Information Sciences 298 (2015) 330–343

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

A framework of mapping undirected to directed graphs for community detection

Peng Gang Sun *, Lill Gao

School of Computer Science and Technology, Xidian University, Xi'an 710071, China
Institute of Computational Bioinformatics, Xidian University, Xi'an 710071, China

A R T I C L E I N F O

Article history:
Received 23 November 2013
Received in revised form 26 August 2014
Accepted 20 October 2014
Available online 3 December 2014

Keywords:
Community detection
Modularity optimization

A B S T R A C T

Community detection gives us a new way to understand the organization characteristics of complex systems. However, for simplicity, most methods for community detection always model the systems as an undirected graph. In this paper, we propose a new framework that unifies undirected into directed graphs by a function, which creates a mapping that transforms an undirected graph into its two directed modes for community detection. We take the method of modularity optimization as an example and apply it on our framework to uncover community structure in complex networks. Compared with the original modularity optimization, we find that the modularity optimization on our framework not only achieves better results on the LFR benchmark, but also a more detailed subdivision on the Zachary club network and the American college football network. Based on the results above, we confirm that our framework provides a good platform for community detection.

1. Introduction

Most complex networks are characterized with groups of nodes, called communities or modules. Communities have more links between nodes of same groups and less links between nodes of different groups [8,11,31,33]. Community detection plays a very important role to understand the organization characteristics of complex networks as well as uncover the hidden correlations among their members [15]. In social networks, communities correspond to groups of people, who share similar characteristics such as hobbies, occupations [8,11–13,31,33]. In biological networks, communities represent groups of molecules, which have same biological functions or participate in same cellular processes [8,11,33]. Therefore, community detection in the past several years has attracted a good deal of attention of scholars in a variety of fields [8,12,13].

Over the years, many algorithms have been proposed to uncover community structure in complex networks [3,4,6–8,19,20,25–27,37–39,41,42]. Newman and Girvan [11,22,23] put forward an interesting work that solves the problem of community detection based on objective optimization. They maximized a measure, *modularity* to determine the best partition of a network [11,22,23]. Later, Rosvall and Bergstrom [29] considered the problem of community detection of a network as a communication process. They minimized the full description length of the network to decide the best partition, which corresponds to an optimal compression of topological structure of the network. Li and Zhang proposed [18] modularity density by considering the sizes of communities instead of total number of links of a network in its definition. They revealed community structure in complex networks by maximizing the measure. Just as objective optimization methods, divisive

* Corresponding author at: School of Computer Science and Technology, Xidian University, Xi'an 710071, China. Tel./fax: +86 29 88202669.
E-mail address: sun198268@163.com (P.G. Sun).

http://dx.doi.org/10.1016/j.ins.2014.10.069
0020-0255/

algorithms also provide a new insight to understand community detection in complex networks [7,8,18,20,30]. For example, Girvan and Newman [11,24] divided a network into smaller parts by iteratively removing the edges with the maximum of betweenness. Finally, the network is broken into smaller subgraphs by the removal process. In the same way, Radicchi et al. [30] and Fortunato et al. [10] iteratively removed the edges of a network with the minimum of edge clustering coefficient and edge information centrality respectively to detect community structure in complex networks. Of course, other techniques such as fuzzy clustering [21,35] also have been used to uncover community structure in complex networks. We can extract more information for community detection methods in a review work by Fortunato [8].

Community structures as very important patterns indeed play a vital role in the study of complex systems, since they always correspond to organization structures or functional units of complex systems. Although the approaches above can help us understand community detection, real systems or networks are more complicated than we can image so that we cannot only consider the real systems as a simple undirected graph. However, most methods in reality for community detection always model the systems as an undirected graph. Actually, a directed graph offers us a more accurate model for the systems as well as a symmetrical metric to extract community structure in complex systems. In this paper, we propose a new framework that unifies undirected into directed graphs by a function, which creates a mapping that transforms an undirected graph into its two directed modes for community detection. We take the modularity optimization as an example and apply it on our framework to evaluate the framework's feasibility and capability on the method for community detection. We also discuss the resolution scale of modularity optimization on our framework and test on the Lancichinetti, Fortunato and Radicchi (LFR) benchmark [16] as well as real-world networks such as the Zachary network [43] and the American college football network [23,27] for community detection.

The rest of the paper is organized as follows. In Section 2, we describe the related works for community detection. In Section 3, we give an introduction on our framework. Section 4 discusses the resolution scale problem of modularity optimization on our framework. Section 5 presents the experimental results on random networks and real-world networks. The conclusion is provided in Section 6.

2. Related works

Community structure widely exists in most real-world networks, and we can describe it as a dense subgraph by graph theory that is because nodes have more links in same communities than that of between different communities. Since community detection can help us understand the organization characteristics of complex networks [8,33], a large number of algorithms for community detection have been put forward in the past few years [3,4,6–8,19,20,25–27,37–39,41,42].

Here, we can classify the algorithms into several categories: Objective optimization methods give us a new insight to understand community detection, e.g. Newman and Girvan [11,22,23] first used an objective optimization approach to uncover community structure in complex networks. They considered a quantitative function, *modularity* or Q, defined as the fraction of links within communities in the actual network, minus the expected fraction of links within comparable communities in a random network [9,29]. The best partition for the actual network is the one with the maximum of modularity. Furthermore, Rosvall and Bergstrom [29] developed an information-theoretic method that considers the process of community detection for a network as a communication process by finding an optimal compression of its topology structure. They used the full description length of a partition for a network as an objective function and minimized this function for community detection. This method achieves better performance in networks with built-in asymmetric communities in different sizes and density. In addition, Li and Zhang put forward [18] modularity density, or D value by considering the sizes of communities instead of total number of links of a network. The D value method achieves better performance for community detection. Objective optimization seems to be an effective approach for community detection [2,5,23,28,34,37]. However, there is a serious resolution scale problem for this strategy. Fortunato and Barthélemy [9] first showed that the modularity optimization approach has the resolution scale problem, and the size of identified communities is dependent on the size of interactions in a network, e.g. complete graphs connected by a single link may not be uncovered correctly [9,18,29], since two or more adjacent cliques may be merged as one. In addition, Rosvall and Bergstrom [29] also pointed out that this problem mainly because the modularity does not contain any information on the number of vertices in a community, and the partition is highly dependent on the total number of edges in the network. Li and Zhang [18] tried to solve the resolution scale problem of modularity optimization by considering the sizes of communities in their objective function instead of total number of links of a network. However, Zhang [44] showed that the modularity density [18] also suffers from a serious resolution scale problem, i.e. some uncovered communities do not satisfy the weak definition of community. For this issue, many works tried to alleviate the resolution scale problem by properly modifying the definition of modularity [1,14,17,32,40], however, no one can solve this problem completely. Therefore, the resolution scale becomes a very important problem for community detection methods.

Divisive algorithms are also well used to uncover community structure in complex networks, since their procedure for revealing community structure is easily understood. Such methods remove the edges of a network in a certain order [8,10,24,36] so that the edges that are the most between nodes each from different communities are deleted firstly. By doing the removal process iteratively, we finally can break the network into smaller subgraphs. A hierarchical tree [8,10,24,36] is a better demonstration of community detection for divisive algorithms. Of course, how to distinguish the edges are within communities from that of between communities are of great importance for divisive algorithms. Girvan and Newman

[11,24] provided edge betweenness, and the measure for an edge is defined as the number of shortest paths between vertices that run along it. Their findings showed that the edges that connect different communities have greater edge betweenness values. Further, Radicchi et al. [30] put forward the edge clustering coefficient, a new measure of edge centrality, since intuitively, the edges within communities are to expect to form cycles, and edges between communities are hardly to be part of cycles. The results showed that low values of the measure are likely to correspond to edges between communities [30]. In addition, Fortunato et al. [10] defined edge information centrality based on the efficient propagation of information over the network. They found that the edges that lie between communities are those with high information centrality values, while those inside communities have low information centrality values. Although divisive algorithms play a vital role for community detection, how to determine the best partition for divisive algorithms is a difficult problem in community detection, e.g. Girvan and Newman [11,24] chose the partition with the maximum of modularity as the best.

In recent years, fuzzy clustering is used to uncover community structure in complex networks [21,35], e.g. fuzzy transitive method is used to detect overlapping and non-overlapping modules [35]. The method treated a network as a fuzzy similarity relation. Finally, each community is mapped as an equivalence class by fuzzy transitive closure. In addition, one equivalence is considered as a core of an overlapping community, and the core can be expanded by adding border nodes, which may be shared by adjacent overlapping communities. Of course, how to measure the similarity between nodes is of great importance to define the fuzzy similarity relation for a network. Edge centralities have been used to evaluate the similarity between nodes [36]. For more information on the categories of community detection algorithms can refer to a review work of Fortunato [8].

From the discussion above, we can see that for simplicity, most methods for community detection always model the complex systems as an undirected graph. Actually, a directed graph offers us a more accurate model for the understanding of the complex systems. In the following section, we will introduce our directed framework to model the complex systems.

3. A framework of mapping undirected to directed graphs

In this section, we propose a new framework that unifies undirected into directed graphs by a function, which creates a mapping that transforms an undirected graph into its two directed modes for community detection.

Here, $G(V,E)$ is often used to denote an unweighted and undirected graph with no self-loops, where V, E are the sets of nodes and of links for the graph respectively, and $|V| = n$, $|E| = l$. We map $G(V,E)$ to its two directed modes, an asymmetrical one, $G^{asym}(V^{asym},E^{asym})$ and a symmetrical one $G^{sym}(V^{sym},E^{sym})$ by the following function respectively.

$$\Re : G \rightarrow G^{asym}, \Re[G(\phi(V), \varphi((E), p(E)))] = G^{asym}(V^{asym}, E^{asym}), \tag{1}$$

$$\begin{aligned}
&\Re' : G \rightarrow G^{sym}, \quad \Re'[G(\phi(V), \varphi'(E))] = G^{sym}(V^{sym}, E^{sym}), \\
&\qquad \phi(V) = V^{asym} = V^{sym}, \\
&\qquad \varphi((E), \quad p(E)) = E^{asym} \\
&\varphi'(E) = E^{sym}
\end{aligned} \tag{2}$$

$$\phi : V \rightarrow V^{asym} \text{ or } V \rightarrow V^{sym}, \qquad \forall v_i \in V, \quad \phi(v_i) = v_i. \tag{3}$$

$$\begin{aligned}
&\varphi : E \rightarrow E^{asym}, \qquad \forall v_i, \quad v_j \in V, \\
&\qquad \text{If } v_i \neq v_j, \quad v_i < v_j \text{ and } (v_i, v_j) \in E, \\
&\qquad\quad \text{then } \varphi((v_i, v_j), p(v_i, v_j)) = \begin{cases} \langle v_i, v_j \rangle & p(v_i, v_j) \\ \langle v_j, v_i \rangle & 1 - p(v_i, v_j) \end{cases} \\
&\qquad\quad \text{where } \varphi((v_i, v_j), p(v_i, v_j)) \sim B(\langle v_i, v_j \rangle, p(v_i, v_j)), \\
&\text{and } \varphi((v_i, v_j), p(v_i, v_j)) \in \{\langle v_i, v_j >, < v_j, v_i \rangle\}. \\
&\qquad \text{If } v_i = v_j \text{ and } (v_i, v_j) \notin E, \quad \text{then } \varphi((v_i, v_j), 1) = \langle v_i, v_j \rangle.
\end{aligned} \tag{4}$$

$$\begin{aligned}
&\varphi' : E \rightarrow E^{sym}, \qquad \forall v_i, \quad v_j \in V, \\
&\qquad \text{If } v_i \neq v_j \text{ and } (v_i, v_j) \in E, \quad \text{then } \varphi'(v_i, v_j) = \langle v_i, v_j \rangle \\
&\qquad \text{If } v_i = v_j \text{ and } (v_i, v_j) \notin E, \text{ then } \varphi'(v_i, v_j) = \langle v_i, v_j \rangle.
\end{aligned} \tag{5}$$

(v_i, v_j) indicates that there is an undirected link between v_i and v_j, and $\langle v_i, v_j \rangle$ denotes a directed link that starts with v_i and ends with v_j. $\varphi((v_i, v_j), p(v_i, v_j)) \sim B(\langle v_i, v_j \rangle, p(v_i, v_j))$ follows "0–1" probability distribution.

The illustration of our framework is showed in Fig. 1. In the figure, an undirected network with two cliques connected by a single link is mapped into its two directed modes based on our framework. We know that the adjacency matrix of G^{asym} is an upper triangular matrix, when $p(v_i, v_j) \equiv 1$, and a lower triangular matrix, when $p(v_i, v_j) \equiv 0$, $\forall v_i, v_j \in V$, while the adjacency matrix of G^{sym} is symmetrical matrix.

4. Modularity optimization on the framework

The modularity optimization is one of the well-known methods and provides us a new insight for the understanding of community detection. However, it has a serious resolution scale problem [9]. In this section, we take the method as an

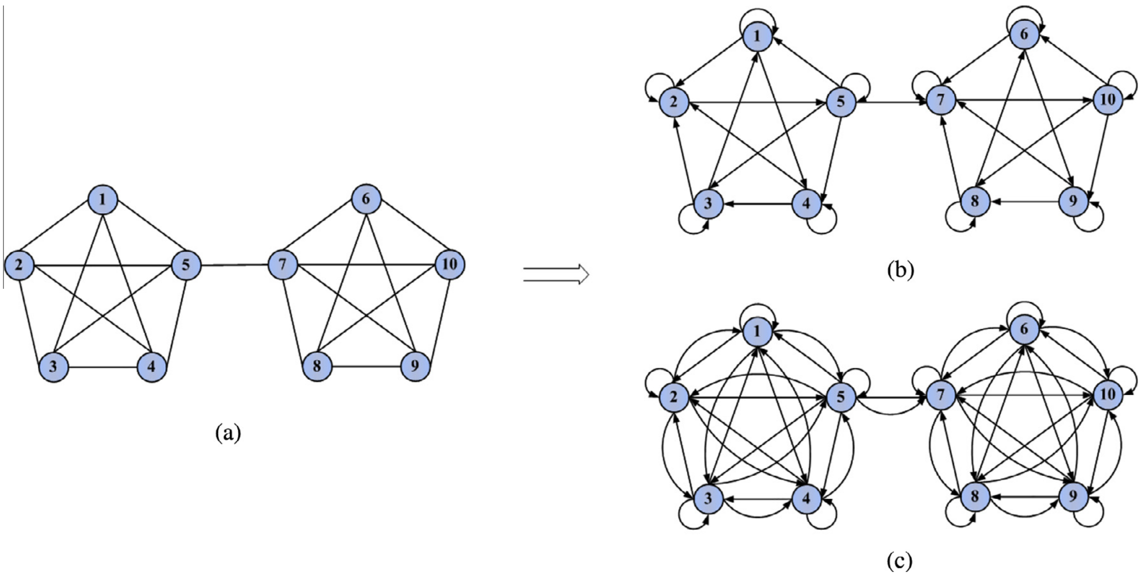

Fig. 1. Illustration of our framework of mapping undirected to directed graphs. (a) An undirected graph consists of two cliques connected by a single link. (b) and (c) Correspond to the asymmetrical and the symmetrical directed modes of the undirected graph respectively.

example and apply it on our framework for community detection. Firstly, we introduce the definition of modularity on our framework based on the two modes, and then provide a detailed discussion on the resolution scale problem.

4.1. Definition of modularity on the framework

The modularity of a partition for a network with n nodes and l links can be written as [11,22,23],

$$Q = \sum_{i=1}^{m}\left[\frac{l_i}{l} - \left(\frac{2l_i + l_i^{ext}}{2l}\right)^2\right] \tag{6}$$

where n_i and l_i are the numbers of nodes and of links in community i respectively, l_i^{ext} is the number of links connecting community i to the rest of the network, $i = 1, 2, \ldots, m$.

From the definition of modularity above, we know that the modularity is the sum of the m terms, each is the fraction of edges within communities in the actual network, minus the expected fraction of edges within comparable communities in a random network [9,11,22,23,29]. Since the modularity does not consider the directions of edges in a network, the modularity on our framework can be described by the following function,

$$Q^{asym} = \sum_{i=1}^{m}\left[\frac{l_i + n_i}{l+n} - \left(\frac{2(l_i + n_i) + l_i^{ext}}{2(l+n)}\right)^2\right] \tag{7}$$

$$Q^{sym} = \sum_{i=1}^{m}\left[\frac{2l_i + n_i}{2l+n} - \left(\frac{2(2l_i + n_i) + 2l_i^{ext}}{2(2l+n)}\right)^2\right] = \sum_{i=1}^{m}\left[\frac{2l_i + n_i}{2l+n} - \left(\frac{2l_i + n_i + l_i^{ext}}{2l+n}\right)^2\right] \tag{8}$$

From Eqs. (6)–(8), we find that Q^{asym} and Q^{sym} not only contain the information in Q, but also the number of nodes within communities and the number of nodes in the whole network. We also find that if we do not consider the directions of edges, then Q^{asym} is a special case in Arenas et al. [1]. The directions of edges are dependent on the method that we apply on the framework.

4.2. The resolution scale of modularity optimization on the framework

Fortunato and Barthélemy [9] pointed out that the modularity optimization method has a serious resolution scale problem. The size of revealed communities is dependent on the total size of links in the network, e.g. complete graphs connected by a single link may not be identified correctly [9,18,29], i.e. two or more adjacent cliques may be merged as one. Rosvall and Bergstrom [29] also showed that it is mainly because the modularity does not contain any information on the number of vertices in a community, and the partition is highly dependent on the total number of links in the network. Therefore, it is necessary for us to give a detailed discussion on the resolution scale problem of Q^{asym} and Q^{sym} respectively.

Here, we follow the approach of Fortunato and Barthélemy [9] that analyze the resolution scale of Q to study that of Q^{asym} and Q^{sym}. All the following discussions are benefited from [9,18,29].

According to the weak definition of communities proposed by Radicchi et al. [30], the number of links within communities is higher than expected, so in Q^{asym}

$$\frac{l_i+n_i}{l+n}-\left(\frac{2(l_i+n_i)+l_i^{ext}}{2(l+n)}\right)^2>0$$

$$\frac{l_i+n_i}{l+n}-\left(\frac{(2+a)(l_i+n_i)}{2(l+n)}\right)^2>0$$

where $l_i^{ext}=a(l_i+n_i)$.

$$l_i+n_i<\frac{4(l+n)}{(2+a)^2}$$

Therefore, we gain $l_i+n_i<\frac{l+n}{4}$ in Q^{asym} compared with $l_i<\frac{l}{4}$ in $Q(a<2)$ [9]

In Q^{sym},

$$\frac{2l_i+n_i}{2l+n}-\left(\frac{2l_i+n_i+l_i^{ext}}{2l+n}\right)^2>0$$

$$\frac{2l_i+n_i}{2l+n}-\left(\frac{(2l_i+n_i)(1+a')}{2l+n}\right)^2>0$$

where $l_i^{ext}=a'(2l_i+n_i)$.

$$2l_i+n_i<\frac{2l+n}{(1+a')^2}$$

Thus, we also gain $2l_i+n_i<\frac{2l+n}{4}$ in Q^{sym} $(a'<1)$ compared with $l_i<\frac{l}{4}$ in $Q(a<2)$ [9]

From the discussion above, we find that the number of internal links within a community is less than one quarter of the total number of links of a network based on the weak definition of communities [30] for Q, Q^{asym} and Q^{sym} respectively.

Here, we also use the network that consists of m cliques [9] to analyze the scales of Q^{asym} and Q^{sym} respectively. In each clique, there are $l_i = l/m$ links and $n_i = n/m$ nodes, $i = 1, 2, \ldots, m$. In this case, Q^{asym} and Q^{sym} are maximal and equal the sum of m identical terms respectively.

$$Q^{asym}=\sum_{i=1}^{m}\left[\frac{l_i+n_i}{l+n}-\left(\frac{2(l_i+n_i)}{2(l+n)}\right)^2\right]=\sum_{i=1}^{m}\left[\frac{l/m+n/m}{l+n}-\left(\frac{2(l/m+n/m)}{2(l+n)}\right)^2\right]=1-\frac{1}{m}$$

$$Q^{sym}=\sum_{i=1}^{m}\left[\frac{2l_i+n_i}{2l+n}-\left(\frac{2l_i+n_i}{2l+n}\right)^2\right]=\sum_{i=1}^{m}\left[\frac{2l/m+n/m}{2l+n}-\left(\frac{2l/m+n/m}{2l+n}\right)^2\right]=1-\frac{1}{m}$$

In this network, we find $Q^{asym} = Q^{sym} = Q$ and when $m\to\infty$, Q^{asym}, Q^{sym}, $Q\to 1$ [9].

In order to analyze the resolution scale of the modularity optimization further, Fortunato and Barthélemy [9] constructed a connected network with n nodes and l links for a fixed number of communities, m in which the modularity reaches the maximum (see Fig. 2(a)). Here, we also analyze the resolution scales of Q^{asym} and Q^{sym} on the network respectively.

The modularity of the network on the framework is

$$Q^{asym}=\sum_{i=1}^{m}\left[\frac{l_i+n_i}{l+n}-\left(\frac{2(l_i+n_i)+2}{2(l+n)}\right)^2\right]=\sum_{i=1}^{m}\left[\frac{l_i+n_i}{l+n}-\left(\frac{l_i+n_i+1}{l+n}\right)^2\right]Q^{sym}=\sum_{i=1}^{m}\left[\frac{2l_i+n_i}{2l+n}-\left(\frac{2l_i+n_i+2}{2l+n}\right)^2\right]$$

where $\sum_{i=1}^{m}l_i=l-m$.

If the equations want to reach their maximums, then all the communities must contain the same number of edges and the same number of nodes, so, $l_i = (l-m)/m = l/m-1$, $n_i = n/m$, $i = 1, 2, \ldots, m$.

$$Q^{asym}_{\max}=\sum_{i=1}^{m}\left[\frac{l/m-1+n/m}{l+n}-\left(\frac{l/m-1+n/m+1}{l+n}\right)^2\right]=\sum_{i=1}^{m}\left[\frac{l/m-1+n/m}{l+n}-\left(\frac{1}{m}\right)^2\right]=1-\frac{m}{l+n}-\frac{1}{m}$$

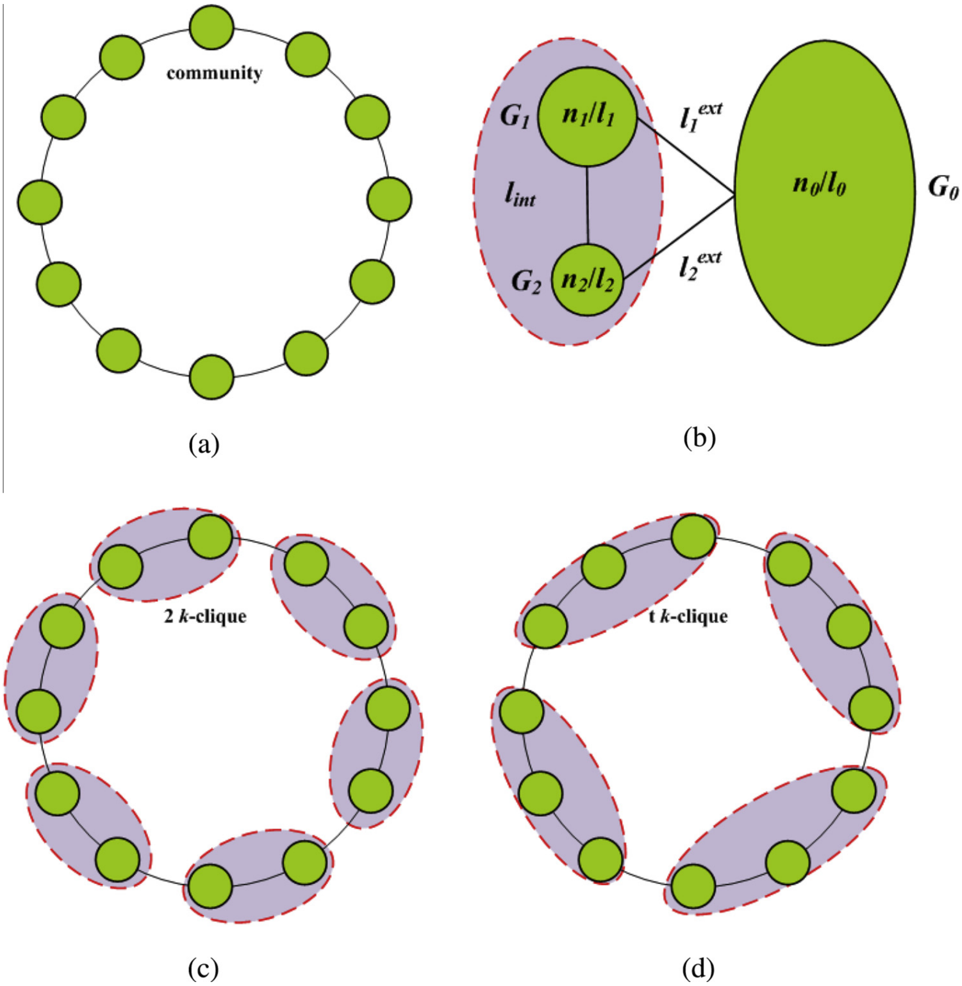

Fig. 2. Examples of networks used for analyzing the resolution scale problem. (a) A network with n nodes and l links for m communities linked by a single edge. (b) A network is partitioned into three or more communities. The two communities, G_1 and G_2 on the left are denoted by circles, the rest of the network G_0 is showed by the oval on the right. G_0, G_1, G_2 are with n_0, n_1, n_2 nodes and l_0, l_1, l_2 links respectively; l_{int} links are between G_1 and G_2; l_1^{ext} links are between G_1 and G_0; l_2^{ext} links are between G_2 and G_0. (c) A network contains m k-cliques connected by single links, and two consecutive cliques are merged as one. (d) t consecutive cliques are merged as one. One node denotes a community or a clique, and the nodes in the circles with dashed red lines show that two or more communities or cliques are merged as one.

$$Q_{\max}^{sym} = \sum_{i=1}^{m}\left[\frac{2l/m-2+n/m}{2l+n} - \left(\frac{2l/m-2+n/m+2}{2l+n}\right)^2\right] = \sum_{i=1}^{m}\left[\frac{2l/m-2+n/m}{2l+n} - \left(\frac{1}{m}\right)^2\right] = 1-\frac{2m}{2l+n}-\frac{1}{m}$$

Compared with

$$Q_{\max} = \sum_{i=1}^{m}\left[\frac{l/m-1}{l} - \left(\frac{l/m}{l}\right)^2\right] = \sum_{i=1}^{m}\left[\frac{l/m-1}{l} - \left(\frac{1}{m}\right)^2\right] = 1-\frac{m}{l}-\frac{1}{m}\,[9]$$

From the discussion above, we find that in this network the expected fraction of links within comparable communities in a random network equals $\left(\frac{1}{m}\right)^2$ for Q, Q^{asym} and Q^{sym}.

Here, we also impose $\frac{\partial Q_{\max}^{asym}}{\partial m} = -\frac{1}{l+n}+\frac{1}{m^2} = 0$, $\frac{\partial Q_{\max}^{sym}}{\partial m} = -\frac{2}{2l+n}+\frac{1}{m^2} = 0$ and gain

$m_{\max}^{asym} = \sqrt{l+n}$, $Q_{\max}^{asym} = 1-\frac{2}{\sqrt{l+n}}$ and $m_{\max}^{sym} = \sqrt{(2l+n)/2}$, $Q_{\max}^{sym} = 1-\frac{2}{\sqrt{(2l+n)/2}}$ compared with $m_{\max} = \sqrt{l}$ and $Q_{\max} = 1-\frac{2}{\sqrt{l}}$ [9].

Based on the discussions above, we know that the resolution scale of Q^{asym} is $m_{\max}^{asym} \sim \sqrt{l+n}$, and that of Q^{sym} is $m_{\max}^{sym} \sim \sqrt{(2l+n)/2}$, and $m_{\max}^{asym} > m_{\max}^{sym}$. Compared with the resolution scale of Q, $m_{\max} \sim \sqrt{l}$, our framework expands the resolution scale of the modularity from $\sqrt{l}$ to $\sqrt{l+n}$ for Q^{asym} [1] or to $\sqrt{(2l+n)/2}$ for Q^{sym} respectively.

Furthermore, we analyze the network in Fig. 2(b). In the figure, the two communities, G_1 and G_2 on the left are denoted by circles, the rest of the network G_0 is showed by the oval on the right. G_0, G_1, G_2 are with n_0, n_1, n_2 nodes and l_0, l_1, l_2 links respectively, and l_{int} links are between G_1 and G_2, l_1^{ext} links are between G_1 and G_0, l_2^{ext} links are between G_2 and G_0. Here, we also consider two partitions like Fortunato and Barthélemy [9]: $A = \{G_0, G_1, G_2\}$ and $A' = \{G_0, (G_1, G_2)\}$.

In Q^{asym},

$$Q_A^{asym} = \frac{l_1 + n_1}{l+n} - \left(\frac{2(l_1+n_1) + l_{int} + l_1^{ext}}{2(l+n)}\right)^2 + \frac{l_2+n_2}{l+n} - \left(\frac{2(l_2+n_2)+l_{int}+l_2^{ext}}{2(l+n)}\right)^2 + Q_{G_0}^{asym}$$

$$Q_{A'}^{asym} = \frac{l_1 + l_2 + l_{int} + n_1 + n_2}{l+n} - \left(\frac{2(l_1+l_2+l_{int}+n_1+n_2) + l_1^{ext} + l_2^{ext}}{2(l+n)}\right)^2 + Q_{G_0}^{asym}$$

where

$$l_{int} = a_1(l_1+n_1) = a_2(l_2+n_2), l_1^{ext} = b_1(l_1+n_1), l_2^{ext} = b_2(l_2+n_2), \text{ and } a_1 + b_1 \leqslant 2, a_2 + b_2 \leqslant 2.$$

$$Q_A^{asym} = \frac{l_1+n_1}{l+n} - \left(\frac{(a_1+b_1+2)(l_1+n_1)}{2(l+n)}\right)^2 + \frac{l_2+n_2}{l+n} - \left(\frac{(a_2+b_2+2)(l_2+n_2)}{2(l+n)}\right)^2 + Q_{G_0}^{asym}$$

$$Q_{A'}^{asym} = \frac{l_1+l_2+a_1(l_1+n_1)+n_1+n_2}{l+n} - \left(\frac{2(l_1+l_2+a_1(l_1+n_1)+n_1+n_2)+b_1(l_1+n_1)+b_2(l_2+n_2)}{2(l+n)}\right)^2 + Q_{G_0}^{asym}$$

$$\Delta Q^{asym} = Q_{A'}^{asym} - Q_A^{asym}$$

$$\Delta Q^{asym} = \frac{[2(l+n)a_1(l_1+n_1) - (a_1+b_1+2)(a_2+b_2+2)(l_1+n_1)(l_2+n_2)]}{2(l+n)^2} < 0$$

$$2(l+n)a_1 - (a_1+b_1+2)(a_2+b_2+2)(l_2+n_2) < 0$$

$$(l_2+n_2) > \frac{2(l+n)a_1}{(a_1+b_1+2)(a_2+b_2+2)}$$

Just as the discussion of Fortunato and Barthélemy [9], we also consider two cases:

For simplicity, we assume that $l_1 = l_2 = l_s$, $n_1 = n_2 = n_s$.

If $a_1 + b_1 = 2$, $a_2 + b_2 = 2$ and $a_1 = a_2 = 2$, $b_1 = b_2 = 0$, then $(l_2+n_2) > \frac{(l+n)}{4}$

$$(l_s+n_s) > \frac{(l+n)}{4}$$

If $a_1 = b_1 = a_2 = b_2 = 1/(l_s + n_s)$, then $(l_2+n_2) > \sqrt{\frac{(l+n)}{2}} - 1$

$$(l_s+n_s) > \sqrt{\frac{(l+n)}{2}} - 1$$

Therefore, if $(l_s+n_s) < \sqrt{\frac{(l+n)}{2}} - 1 < \frac{(l+n)}{4}$, then $Q_{A'}^{asym} > Q_A^{asym}$ that is G_1 and G_2 are merged as one community compared with $l_s < \sqrt{\frac{l}{2}} - 1 < \frac{l}{4}$ in Q [9].

In Q^{sym},

$$Q_A^{sym} = \frac{2l_1+n_1}{2l+n} - \left(\frac{2l_1+n_1+l_{int}+l_1^{ext}}{2l+n}\right)^2 + \frac{2l_2+n_2}{2l+n} - \left(\frac{2l_2+n_2+l_{int}+l_2^{ext}}{2l+n}\right)^2 + Q_{G_0}^{sym}$$

$$Q_{A'}^{sym} = \frac{2l_1+2l_2+2l_{int}+n_1+n_2}{2l+n} - \left(\frac{2l_1+2l_2+2l_{int}+n_1+n_2+l_1^{ext}+l_2^{ext}}{2l+n}\right)^2 + Q_{G_0}^{sym}$$

where

$$l_{int} = a_1'(2l_1+n_1) = a_2'(2l_2+n_2), l_1^{ext} = b_1'(2l_1+n_1), l_2^{ext} = b_2'(2l_2+n_2), a_1' + b_1' \leqslant 1, a_2' + b_2' \leqslant 1.$$

$$Q_A^{sym} = \frac{2l_1 + n_1}{2l + n} - \left(\frac{(a_1' + b_1' + 1)(2l_1 + n_1)}{2l + n}\right)^2 + \frac{2l_2 + n_2}{2l + n} - \left(\frac{(a_2' + b_2' + 1)(2l_2 + n_2)}{2l + n}\right)^2 + Q_{G_0}^{sym}$$

$$Q_{A'}^{sym} = \frac{2l_1 + 2l_2 + 2a_1'(2l_1 + n_1) + n_1 + n_2}{2l + n} - \left(\frac{2l_1 + 2l_2 + 2a_1'(2l_1 + n_1) + n_1 + n_2 + b_1'(2l_1 + n_1) + b_2'(2l_2 + n_2)}{2l + n}\right)^2 + Q_{G_0}^{sym}$$

$$\Delta Q^{sym} = Q_{A'}^{sym} - Q_A^{sym}$$

$$\Delta Q^{sym} = \frac{[2(2l + n)a_1'(2l_1 + n_1) - 2(a_1' + b_1' + 1)(a_2' + b_2' + 1)(2l_1 + n_1)(2l_2 + n_2)]}{(2l + n)^2} < 0$$

$$(2l + n)a_1' - (a_1' + b_1' + 1)(a_2' + b_2' + 1)(2l_2 + n_2) < 0$$

$$(2l_2 + n_2) > \frac{(2l + n)a_1'}{(a_1' + b_1' + 1)(a_2' + b_2' + 1)}$$

If $a_1' + b_1' = 1,\ a_2' + b_2' = 1$ and $a_1' = a_2' = 1,\ b_1' = b_2' = 0$, then $(2l_2 + n_2) > \frac{(2l+n)}{4}$

$$(2l_s + n_s) > \frac{(l + n)}{4}$$

If $a_1' = b_1' = a_2' = b_2' = 1/(2l_s + n_s)$, then $(l_2 + n_2) > \sqrt{2l + n} - 2$

$$(l_s + n_s) > \sqrt{2l + n} - 2$$

Therefore, if $(l_s + n_s) < \sqrt{2l + n} - 2 < \frac{(2l+n)}{4}$, then $Q_{A'}^{sym} > Q_A^{sym}$ that is G_1 and G_2 are merged as one community compared with $l_s < \sqrt{\frac{l}{2}} - 1 < \frac{l}{4}$ in Q [9].

In addition, we also discuss the network in Fig. 2(c) as Fortunato and Barthélemy [9]. The network contains m k-cliques connected by single links. $l = mk(k-1)/2 + m$, and $n = mk$. Here, we consider two partitions: one is that each clique is a community denoted by Q_1^{asym} or Q_1^{sym}; the other is that two consecutive cliques are merged as one community denoted by Q_2^{asym} or Q_2^{sym}.

In Q^{asym},

$$Q_1^{asym} = m\left[\frac{k(k-1)/2 + k}{mk(k-1)/2 + m + mk} - \left(\frac{2(k(k-1)/2 + k) + 2}{2(mk(k-1)/2 + m + mk)}\right)^2\right] = m\left[\frac{k(k+1)}{mk(k+1) + 2m} - \left(\frac{1}{m}\right)^2\right]$$
$$= 1 - \frac{2}{k(k+1) + 2} - \frac{1}{m}$$

$$Q_2^{asym} = \frac{m}{2}\left[\frac{k(k-1) + 1 + 2k}{mk(k-1)/2 + m + mk} - \left(\frac{2(k(k-1) + 1 + 2k) + 2}{2(mk(k-1)/2 + m + mk)}\right)^2\right] = \frac{m}{2}\left[\frac{k(k+1) + 1}{m(k(k+1) + 2)/2} - \left(\frac{2}{m}\right)^2\right]$$
$$= 1 - \frac{1}{k(k+1) + 2} - \frac{2}{m}$$

$$Q_1^{asym} - Q_2^{asym} = \frac{1}{m} - \frac{1}{k(k+1) + 2} > 0$$

$$m < k(k+1) + 2 = \sqrt{2(l + n)}$$

We gain $m < \sqrt{2(l + n)}$ in Q^{asym} compared with $m < \sqrt{2l}$ in Q [9].

In Q^{sym},

$$Q_1^{sym} = m\left[\frac{2k(k-1)/2 + k}{2mk(k-1)/2 + 2m + mk} - \left(\frac{2k(k-1)/2 + k + 2}{2mk(k-1)/2 + 2m + mk}\right)^2\right] = m\left[\frac{k^2}{mk^2 + 2m} - \left(\frac{k^2 + 2}{mk^2 + 2m}\right)^2\right] = 1 - \frac{2}{k^2 + 2} - \frac{1}{m}$$

$$Q_2^{sym} = \frac{m}{2}\left[\frac{2k(k-1) + 2 + 2k}{2mk(k-1)/2 + 2m + mk} - \left(\frac{2k(k-1) + 2 + 2k + 2}{2mk(k-1)/2 + 2m + mk}\right)^2\right] = \frac{m}{2}\left[\frac{2k^2 + 2}{mk^2 + 2m} - \left(\frac{2k^2 + 2 + 2}{mk^2 + 2m}\right)^2\right]$$

$$Q_2^{sym} = 1 - \frac{1}{k^2+2} - \frac{2}{m}$$

$$Q_1^{sym} - Q_2^{sym} = \frac{1}{m} - \frac{1}{k^2+2} > 0$$

$$m < k^2 + 2 = \sqrt{2l+n}$$

We gain $m < \sqrt{2l+n}$ in Q^{sym} compared with $m < \sqrt{2l}$ in Q [9].

Here, we extend it to a more general case that t consecutive k-cliques are merged as one community in the network (see Fig. 2(d)) and denoted by Q_t^{asym}, Q_t^{sym}.

In Q^{asym},

$$Q_t^{asym} = \frac{m}{t}\left[\frac{tk(k-1)/2+(t-1)+tk}{mk(k-1)/2+m+mk} - \left(\frac{2(tk(k-1)/2+(t-1)+tk)+2}{2(mk(k-1)/2+m+mk)}\right)^2\right] = \frac{m}{t}\left[\frac{t(k(k+1)+2(t-1)/t}{m(k(k+1)+2)} - \left(\frac{t}{m}\right)^2\right]$$
$$= 1 - \frac{2-2(t-1)/t}{k(k+1)+2} - \frac{t}{m}$$

$$Q_1^{asym} - Q_t^{asym} = \frac{t-1}{m} - \frac{2(t-1)/t}{k(k+1)+2} > 0 = \frac{1}{m} - \frac{2/t}{k(k+1)+2} > 0$$

$$m < \frac{t}{2}(k(k+1)+2) = t(l+n)$$

In Q^{sym},

$$Q_t^{sym} = \frac{m}{t}\left[\frac{tk(k-1)+2(t-1)+tk}{2mk(k-1)/2+2m+mk} - \left(\frac{tk(k-1)+2(t-1)+tk+2}{2mk(k-1)/2+2m+mk}\right)^2\right] = \frac{m}{t}\left[\frac{tk^2+2t-2}{mk^2+2m} - \left(\frac{tk^2+2t}{mk^2+2m}\right)^2\right]$$
$$= 1 - \frac{2}{t(k^2+2)} - \frac{t}{m}$$

$$Q_1^{sym} - Q_t^{sym} = \frac{t-1}{m} - \frac{2-2/t}{k^2+2} > 0$$

$$m < \frac{t-1}{2-2/t}(k^2+2) = \sqrt{\frac{t}{2}(2l+n)}$$

5. Results and discussions

After we have discussed the resolution scales of Q^{asym} and Q^{sym}, we in this section test them on the LFR benchmark [16] as well as real-world networks such as the Zachary network [43] and the American college football network [23,27] for community detection.

5.1. The LFR benchmark

Lancichinetti, Fortunato and Radicchi (LFR) [16] proposed a new benchmark that can generate random networks with the heterogeneity in distributions of node degrees and community sizes. Here, we produce a serial of random networks based on the benchmark: (1) $m < \sqrt{l} < \sqrt{(2l+n)/2} < \sqrt{l+n}$, this inequality shows that the produced networks are within the resolution scales of Q, Q^{asym} and Q^{sym}. (2) $\sqrt{l} < \sqrt{(2l+n)/2} < \sqrt{l+n} < m$, this inequality denotes that the produced networks exceed the resolution scales of Q, Q^{asym} and Q^{sym}. (3) $\sqrt{l} < \sqrt{(2l+n)/2} < m < \sqrt{l+n}$, this inequality indicates that the produced networks are within the resolution scale of Q^{asym}, but exceed that of Q and Q^{sym}.

Here, we use the normalized mutual information (NMI) [7] to evaluate the consistency between the real communities and the communities detected by Q, Q^{asym} and Q^{sym}.

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

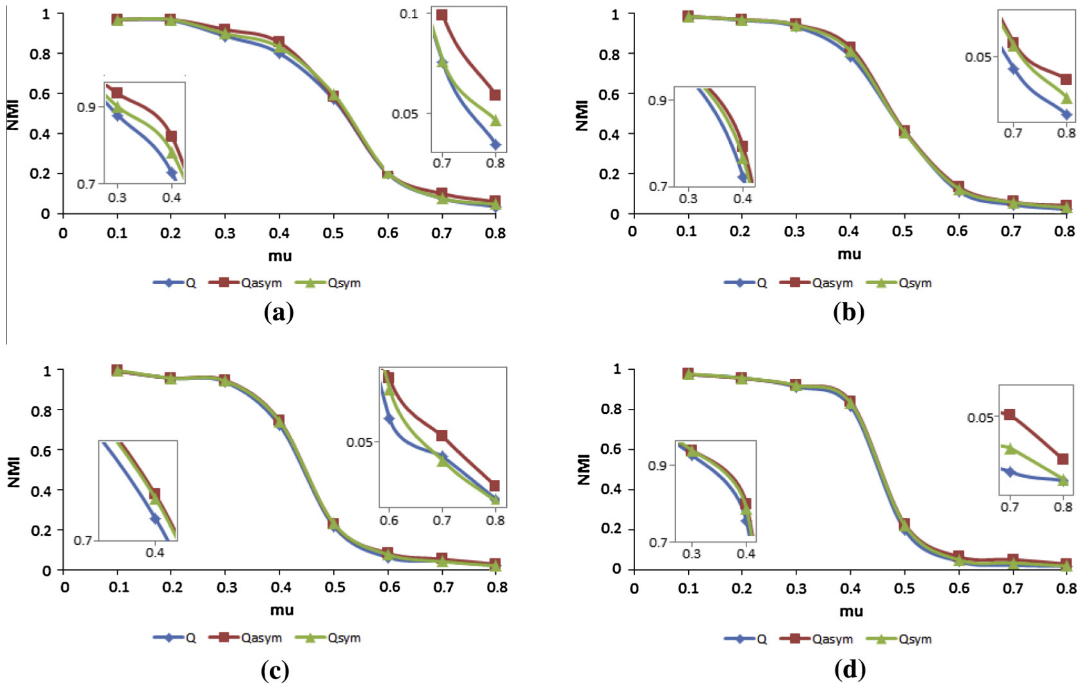

Fig. 3. Testing on the LFR benchmark within the resolution scales of Q, Q^{asym} and Q^{sym} for community detection. (a) $N = 200$, $k = 15$, $maxk = 50$, $minc = 20$, and $maxc = 50$. (b) $N = 200$, $k = 15$, $maxk = 50$, $minc = 30$, and $maxc = 50$. (c) $N = 200$, $k = 15$, $maxk = 50$, $minc = 40$, and $maxc = 50$. (d) $N = 200$, $k = 15$, $maxk = 50$, $minc = 50$, and $maxc = 50$. Each point is on average as a function of *mu*.

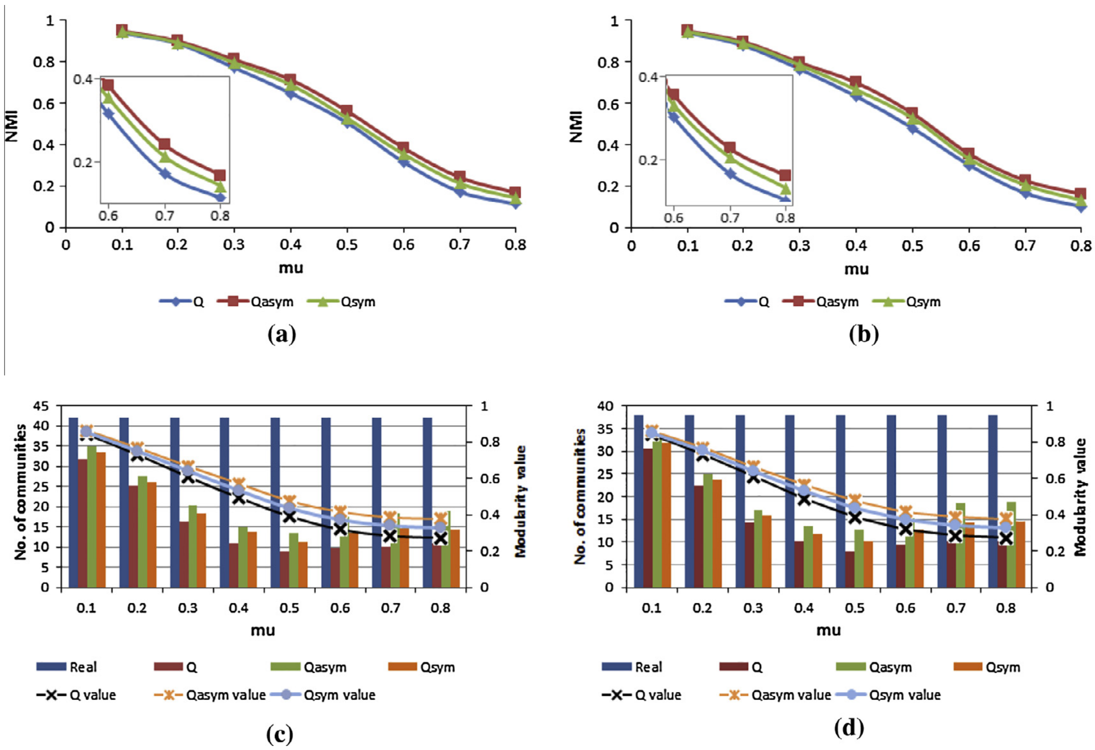

Fig. 4. Testing on the LFR benchmark exceeding the resolution scales of Q, Q^{asym} and Q^{sym} for community detection. (a) The produced networks exceed the resolution scales of Q, Q^{asym} and Q^{sym} such as $N = 1000$, $k = 10$, $maxk = 50$, $minc = 20$, and $maxc = 50$. (b) The produced networks are within the resolution scale of Q^{asym}, but exceed that of Q and Q^{sym} such as $N = 1000$, $k = 10$, $maxk = 50$, $minc = 13$, and $maxc = 50$. (c) and (d) Correspond to (a) and (b) respectively, in which demonstrate the number of communities detected by Q, Q^{asym} and Q^{sym} compared with the real number of communities on the LFR benchmark shown by histogram. The corresponding modularity values are shown by line charts. Each point is on average as a function of *mu*.

where the rows and the columns correspond to the real communities and the detected communities respectively. N_{ij} is simply the number of vertices shared by the real community i and the detected community j. The number of real communities is n_A and the detected communities is n_B, the sum over row i of matrix N_{ij} is $N_{i.}$, and the sum over column j is $N_{.j}$. N is the total number of nodes in the network [7].

Firstly, we produce groups of random networks that are within the resolution scales of Q, Q^{asym} and Q^{sym} such as $N = 200$, $k = 15$, $maxk = 50$, $minc = \{20, 30, 40, 50\}$ and $maxc = 50$. N is the number of nodes in the networks, k is the average node degree, and $maxk$ is the maximum of node degrees. mu is the mixing parameter, i.e. one node shares a fraction, 1-mu of it

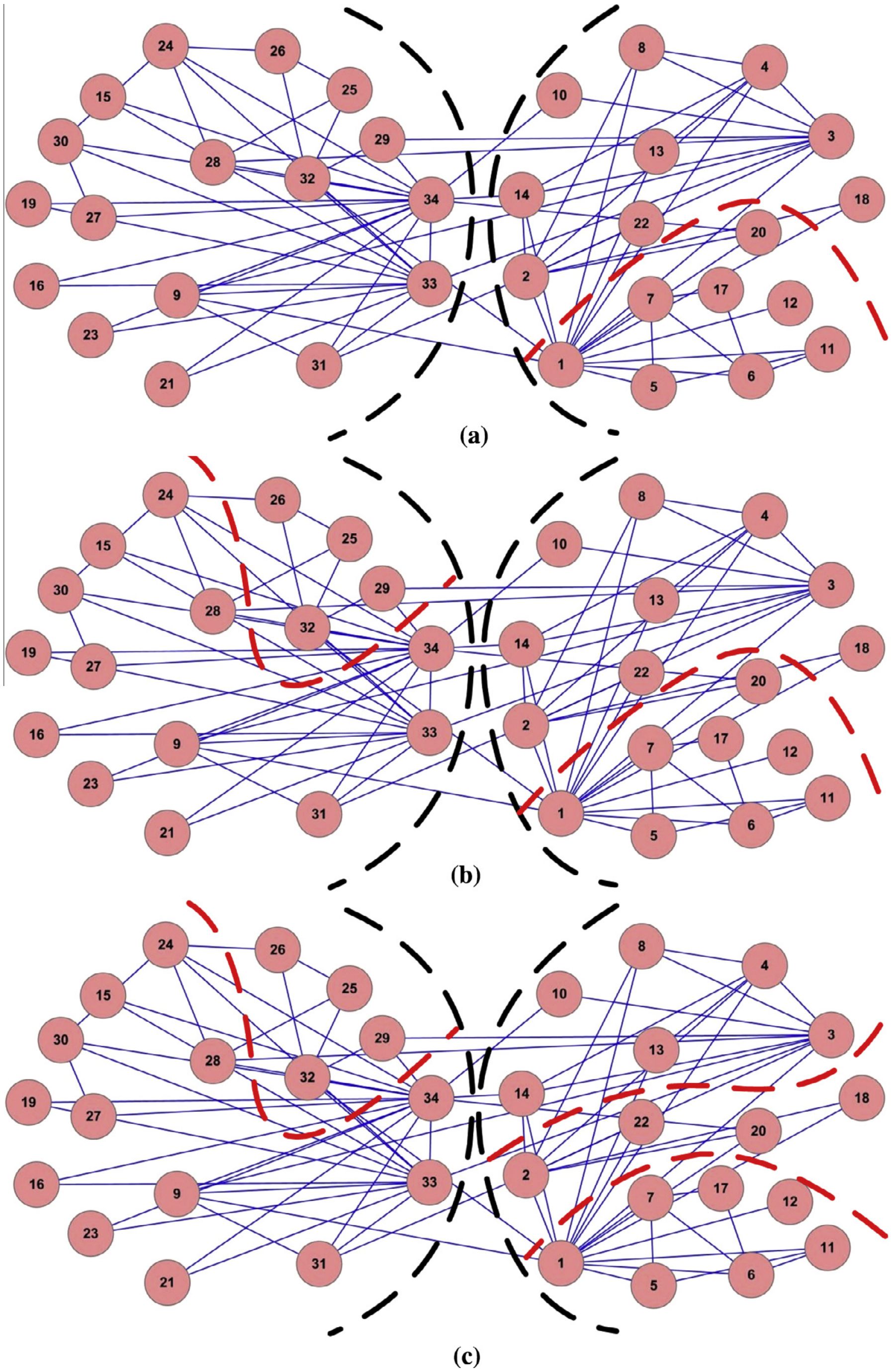

Fig. 5. The communities detected on the Zachary club network. (a), (b) and (c) Are the detected communities by maximizing Q, Q^{sym} and Q^{asym} respectively. The two well-known communities are separated by black dashed lines and can also be subdivided by red dashed lines. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

associated links with the other nodes in its community and a fraction, *mu* with the rest of nodes of the network. *minc* and *maxc* denote the minimum and the maximum of community sizes respectively. If *minc* = *maxc*, then the LFR benchmark is similar to the GN (Girvan and Newman) benchmark [11,18–20].

Fig. 3 shows the results on the LFR benchmark that the produced networks are within the resolution scales of Q, Q^{asym} and Q^{sym} for community detection. We find that Q^{asym} achieves the best results, and Q^{sym} outperforms Q for community detection. With the increase of *mu*, the performance is worse and worse. In particular, when $mu \geqslant 0.5$, the performance declines sharply. In general, the benefits of Q^{asym} and Q^{sym} are not obvious compared with Q for community detection on the random networks that are within their resolution scales.

In view of the results in Fig. 3, we produce groups of random networks that exceed the resolution scales of Q, Q^{asym} and Q^{sym} such as $N = 1000$,$k = 10$, $maxk = 50$, $minc = 20$, and $maxc = 50$. Furthermore, we also generate groups of random networks that are within the resolution scale of Q^{asym}, but exceed that of Q and Q^{sym} such as $N = 1000$, $k = 10$, $maxk = 50$, $minc = 13$, and $maxc = 50$. The results show that Q^{asym} also achieves the best performance on the random networks compared with Q and Q^{sym} for community detection in Fig. 4(a) and (b), and the larger the *mu* is, the better performance Q^{asym} gains.

For the number of detected communities, Q^{asym} is also the best one compared with Q^{sym} and Q. Since $mu = 0.5$ is the hardest case for community detection and the communities on the random networks is the vaguest, we obtain the least number of detected communities by merging several communities as one, which is consistent with the discussion of Fortunato and Barthélemy [9]. We also find that the communities detected by Q^{asym} are closer to that of the real on the random networks, and Q^{asym} obtains greater modularity values compared with Q^{sym} and Q.

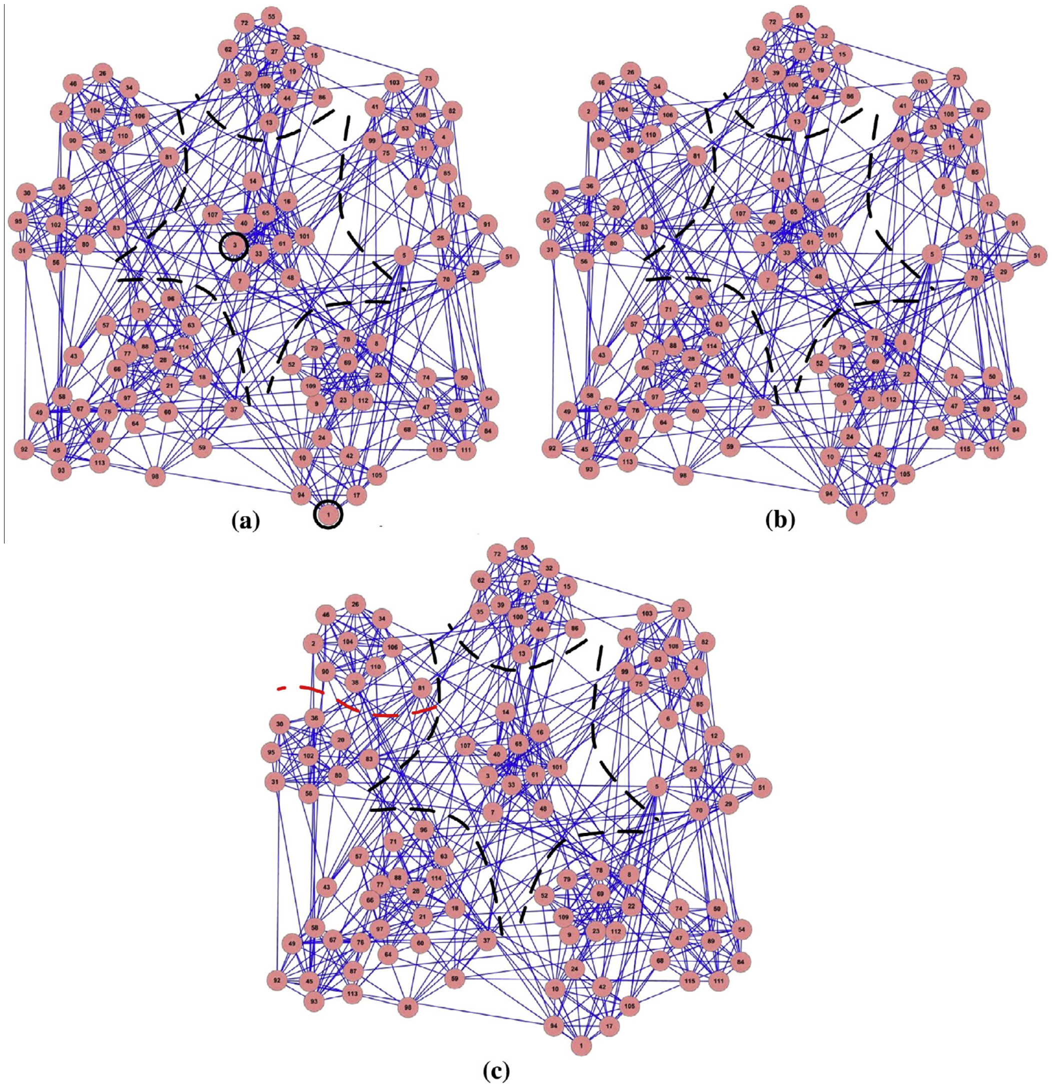

Fig. 6. The communities detected on the American college football network. (a), (b) and (c) correspond to the communities detected by maximizing Q, Q^{sym} and Q^{asym} respectively. The well-known communities are separated by black dashed lines and can also be subdivided by red dashed lines. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5.2. The Zachary club network

In the Zachary club network [43], Zachary considered 34 members of the network for two years. During his experiment, a disagreement occurred between vertex 1, the administrator of the club and vertex 33, the club's instructor, which ultimately lead to a division of the network, since the instructor left and started a new club, taking away nearly a half of members of the original club with him.

Here, we try to uncover the split of club by maximizing Q, Q^{asym} and Q^{sym} respectively. The two well-known communities are detected and separated by black dashed lines and centred with vertex 1 (the administrator of the club) and vertex 33 (the club's instructor) (see Fig. 5). We find that maximizing Q, Q^{asym} and Q^{sym} can further subdivide the two well-known communities into smaller parts respectively. The communities detected by Q^{asym} and Q^{sym} are the subdivisions of Q.

5.3. The American college football network

The network we test in this subsection is the American college football network [23,27]. The nodes in the network indicate the 115 teams, and the links between two nodes denote 616 games played in the course of the year. The teams can be split into 12 conferences and the games generally develop more frequent between members of same conferences than that of different conferences. Fig. 6 shows the identified communities on the American college football network by maximizing Q, Q^{asym} and Q^{sym}. The findings on the network are similar to that of on the Karate club network in the subsection above. Two nodes, 1 and 3 within black circles in Fig. 6(a) are classified wrongly by Q compared with the results by Q^{sym} in Fig. 6(b). In addition, the communities detected by Q^{asym} in Fig. 6(c) are a subdivision of Q^{sym} in Fig. 6(b).

Compared with the original modularity optimization in this section, we find that the modularity optimization on our framework achieves better results on the LFR benchmark as well as a more detailed subdivision on the Zachary club network and the American college football network for community detection. Especially, when test beds exceed the resolution scale, the effectiveness of our framework is much more obviously.

6. Conclusions

In this paper, we introduce a new framework that unifies undirected into directed graphs by a function, which creates a mapping that transforms an undirected graph into its two directed modes for community detection. We apply the method of modularity optimization on our framework to detect community structure in complex networks. We find that our framework indeed gives a remarkable improvement on community detection. Based on the results in this paper, we confirm that our framework provides a good platform for community detection. In the future work, we will apply other quality functions such as the modularity density (D value) on our framework and check the framework's feasibility and capability for community detection.

Acknowledgements

This work is supported by the National Foundation for Studying Abroad, the National Natural Science Foundation of China (Grant Nos. 61202175 and 61202174), the Fundamental Research Funds for the Central Universities (Grant Nos. BDY181417 and BDY021404), and the Research Fund for the Doctoral Program of Higher Education of China (Grant No. 20120203120015).

References

[1] A. Arenas, A. Fernández, S. Gómez, Analysis of the structure of complex networks at different resolution levels, New J. Phys. 10 (2008) 053039.
[2] B. Amiri, L. Hossain, J.W. Crawford, R.T. Wigand, Community detection in complex networks: multi-objective enhanced firefly algorithm, Knowl.-Based Syst. 46 (2013) 1–11.
[3] J.P. Bagrow, Communities and bottlenecks: trees and treelike networks have high modularity, Phys. Rev. E 85 (2012) 066118.
[4] Z. Bu, C. Zhang, Z. Xia, J. Wang, A fast parallel modularity optimization algorithm (FPMQA) for community detection in online social network, Knowl.-Based Syst. 50 (2013) 246–259.
[5] L. Chen, Q. Yu, B. Chen, Anti-modularity and anti-community detecting in complex networks, Inform. Sci. 275 (2014) 293–313.
[6] P. De Meo, E. Ferrara, G. Fiumara, A. Provetti, Enhancing community detection using a network weighting strategy, Inform. Sci. 222 (2013) 648–668.
[7] L. Danon, J. Duch, A. Diaz-Guilera, Comparing community structure identification, J. Stat. Mech: Theory Exp. 09 (2005) P09008.
[8] S. Fortunato, Community detection in graphs, Phys. Rep. 486 (2010) 75–174.
[9] S. Fortunato, M. Barthélemy, Resolution limit in community detection, Proc. Nat. Acad. Sci. 10 (2007) 436–441.
[10] S. Fortunato, V. Latora, M. Marchiori, Method to find community structures based on information centrality, Phys. Rev. E 70 (2004) 056104.
[11] M. Girvan, M.E.J. Newman, Community structure in social and biological networks, Proc. Nat. Acad. Sci. 99 (2002) 7821–7826.
[12] M.S. Granovetter, The strength of weak ties, Am. J. Sociol. 78 (1973) 1360–1380.
[13] M.S. Granovetter, Economic action and social structure: the problem of embeddedness, Sociol. Econ. Life 91 (1985) 481–510.
[14] J.M. Kumpula, J. Saramaki, K. Kaski, J. Kertesz, Limited resolution and multiresolution methods in complex network community detection, Fluctuations Noise Lett. 7 (2007) 209–214.
[15] A. Lancichinetti, S. Fortunato, Consensus clustering in complex networks, Sci. Rep. 2 (2012) 1–7.
[16] A. Lancichinetti, S. Fortunato, F. Radicchi, Benchmark graphs for testing community detection algorithms, Phys. Rev. E 78 (2008) 046110.
[17] A. Lancichinetti, S. Fortunato, Limits of modularity maximization in community detection, Phys. Rev. E 84 (2011) 066122.
[18] Z. Li, S. Zhang, R.S. Wang, X.S. Zhang, L. Chen, Quantitative function for community detection, Phys. Rev. E 77 (2008) 036109.
[19] W. Li, Revealing network communities with a nonlinear programming method, Inform. Sci. 229 (2013) 18–28.
[20] Y. Li, H. Wang, J. Li, H. Gao, Efficient community detection with additive constrains on large networks, Knowl.-Based Syst. 52 (2013) 268–278.

[21] B. Mirkin, S. Nascimento, Additive spectral method for fuzzy cluster analysis of similarity data including community structure and affinity matrices, Inform. Sci. 183 (2012) 16–34.
[22] M.E.J. Newman, Detecting community structure in networks, Eur. Phys. J. B 3 (2004) 8321–8330.
[23] M.E.J. Newman, Modularity and community structure in networks, Proc. Nat. Acad. Sci. 103 (2006) 8577–8582.
[24] M.E.J. Newman, M. Girvan, Finding and evaluating community structure in networks, Phys. Rev. E 69 (2003) 026113.
[25] M.C.V. Nascimento, Community detection in networks via a spectral heuristic based on the clustering coefficient, Discr. Appl. Math. 176 (2014) 89–99.
[26] G.K. Orman, V. Labatut, H. Cherifi, Comparative evaluation of community detection algorithms: a topological approach, J. Stat. Mech: Theory Exp. (2012) P08001.
[27] G. Palla, I. Derenyi, I. Farkas, T. Vicsek, Uncovering the overlapping community structure of complex networks in nature and society, Nature 435 (2005) 814–818.
[28] X. Qi, W. Tang, Y. Wu, G. Guo, E. Fuller, C.-Q. Zhang, Optimal local community detection in social networks based on density drop of subgraphs, Pattern Recogn. Lett. 36 (2014) 46–53.
[29] M. Rosvall, C.T. Bergstrom, An information-theoretic framework for resolving community structure in complex networks, Proc. Nat. Acad. Sci. 104 (2007) 7327–7331.
[30] F. Radicchi, C. Castellano, F. Cecconi, V. Loreto, D. Parisi, Defining and identifying communities in networks, Proc. Nat. Acad. Sci. 101 (2004) 2658–2663.
[31] J. Reichardt, S. Bornholdt, Statistical mechanics of community detection, Phys. Rev. E 74 (2006) 016110.
[32] P. Ronhovde, Z. Nussinov, Local Multiresolution Order in Community Detection, 2012. 1208.5052.
[33] S.H. Strogatz, Exploring complex networks, Nature 410 (2001) 268–276.
[34] P.G. Sun, L. Gao, Y. Yang, Maximizing modularity intensity for community partition and evolution, Inform. Sci. 236 (2013) 82–93.
[35] P.G. Sun, L. Gao, S. Han, Identification of overlapping and non-overlapping community structure by fuzzy clustering, Inform. Sci. 181 (2011) 1060–1071.
[36] P.G. Sun, Methods to find community based on edge centrality, Physica A 9 (2013) 1977–1988.
[37] P.G. Sun, Weighting links based on edge centrality for community detection, Physica A 394 (2014) 346–357.
[38] C. Shi, Y. Cai, D. Fu, Y. Dong, B. Wu, A link clustering based overlapping community detection algorithm, Data Knowl. Eng. 87 (2013) 394–404.
[39] J. Xie, S. Kelley, B.K. Szymanski, Overlapping community detection in networks: the state of the art and comparative study, ACM Comput. Surv. 45 (2013) 1–37.
[40] J. Xiang, K. Hu, Limitation of multi-resolution methods in community detection, Physica A 391 (2012) 4995–5003.
[41] Y. Yang, P.G. Sun, X. Hu, Z.J. Li, Closed walks for community detection, Physica A 397 (2014) 129–143.
[42] B. Yang, J. Di, J. Liu, D. Liu, Hierarchical community detection with applications to real-world network analysis, Data Knowl. Eng. 83 (2013) 20–38.
[43] W.W. Zachary, An information flow model for conflict and fission in small groups, J. Anthropol. Res. 33 (1977) 452–473.
[44] X.S. Zhang, R.S. Wang, Y. Wang, J. Wang, Y. Qiu, L. Wang, L. Chen, Modularity optimization in community detection of complex networks, Europhys. Lett. 87 (2009) 38002.