A Fast Community Detection Algorithm Based on Reconstructing Signed Networks

Chengyi Xia, *Member, IEEE*, Yongping Luo, Li Wang, and Hui-Jia Li, *Member, IEEE*

***Abstract*—Signed networks depict the individual cooperative or hostile relationship in a population, which can help to deeply mine the characteristics of complex networks and predict the potential collaboration between individuals by analyzing their interaction within different groups or communities. In this article, first of all, an improved modularity function for signed networks is proposed on the basis of the existing modularity function. Then, a new community detection algorithm for signed networks has also been devised, and time complexity analysis shows that the time required for the algorithm has a linear relationship with the number of nodes in the sparse networks. Meanwhile, the affinity index that can be used to convert directed signed networks into the corresponding undirected signed networks is come up with. Finally, the current algorithm has been applied into several illustrative and realistic networks. The experimental results indicate that the number of communities given by the proposed algorithm is consistent with that of actual communities, and thus, it can be further conducive to identifying the community structure hidden within the real-world systems.**

***Index Terms*—Affinity index, community detection, community structure, modularity, network reconstruction, signed network.**

I. Introduction

MANY real-world systems can be modeled as a complex network, where the node denotes an element and the link represents the interaction or interconnection between these elements, typical examples include social networks [1], [2], citation networks [3], [4], transportation systems [5], [6], biological and ecological networks [7], [8], etc. Over the past two decades, the research of complex networks has attracted the extensive interests since there exists a great deal of information, which cannot be observed or extracted quickly by the naked eye, hidden inside the networks. However, these information may dominate the whole evolution of the networks and the dissemination of other information such as clustering or community structure [9], [10], influential spreaders [11], [12], and so on.

Manuscript received June 30, 2020; revised October 28, 2020 and February 16, 2021; accepted March 8, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 61773286, Grant 71871233, and Grant 61403280. *(Corresponding author: Chengyi Xia.)*

Chengyi Xia, Yongping Luo, and Li Wang are with the Tianjin Key Laboratory of Intelligence Computing and Novel Software Technology, Engineering Research Center of Learning-Based Intelligent System, Ministry of Education, Tianjin University of Technology, Tianjin 300384, China (e-mail: xialooking@163.com; lyp@zurp.date; wltjut08@126.com).

Hui-Jia Li is with the School of Science, Beijing University of Posts and Telecommunication, Beijing 100876, China (e-mail: hjli@amss.ac.cn).

This article has supplementary downloadable material available at https://doi.org/10.1109/JSYST.2021.3065378, provided by the authors.

Digital Object Identifier 10.1109/JSYST.2021.3065378

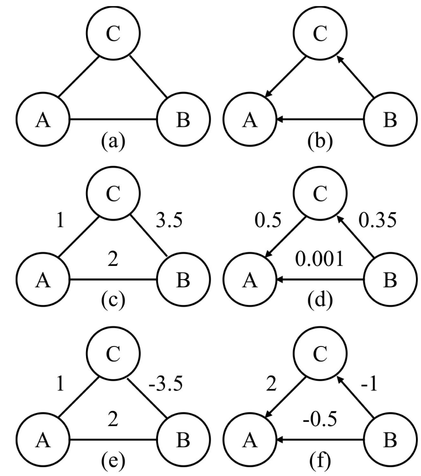

Fig. 1. Undirected networks are presented in the left column [panels (a), (c), and (e)], while directed networks are arranged in the right column [panels (b), (d), and (f)]; the first row [panels (a) and (b)] denotes the unsigned and unweighted networks, the second row [panels (c) and (d)] means the unsigned and weighted networks, and the third row [panels (e) and (f)] stands for the signed and weighted networks.

Based on the link properties or characteristics, the networks can be divided into different types of networks [10]. For instance, the networks are directed or undirected, that is, the interaction between a pair of nodes is asymmetric (directed) or symmetric (undirected); or these networks are weighted or unweighted, where the interaction strength among nodes is identical (unweighted) or different (weighted). However, aforementioned networks can be regarded as the unsigned networks, which are usually represented as a graph that only contains positive links, in the same vein, the one containing both positive and negative links is called a signed network. As an example, in the social network, an individual is viewed as a node and then their interconnection is regarded as a link, in that way, the unsigned network means the frequency of contact, or the similarity between individuals on a certain characteristic (e.g., one person likes Romeo and Juliet and the other person likes it as well), the signed network means the relationship of friends or enemies, or the difference in a particular characteristic between individuals (e.g., another different attitude toward something). In Fig. 1, we present several examples of different types of networks, which may help the readers to understand the slight differences between them.

In a complex network, a community is defined as a set of nodes within which edges are often denser than those connecting them

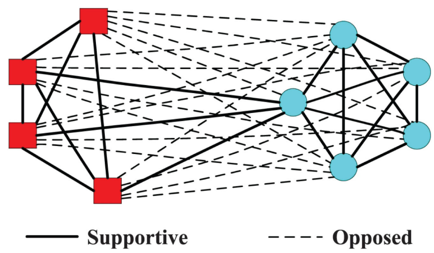

Fig. 2. Community structure of U.S. supreme court justices network.

with other nodes [9]. So far, various algorithms have been presented to detect the community structure hidden in the unsigned network [13]–[16]. Although these algorithms perform well in terms of efficiency and accuracy, most ones are limited to networks with only positive edges, which are difficult to be extended into signed networks. To illustrate the community structure in signed networks, we show a famous real social network in Fig. 2, i.e., the U.S. supreme court justices network, which describes the voting behavior of nine justices in the supreme court of the United States during the period of 2006–2007. The positive links represent that one justice supports the other one, and the negative links denote the opposite meaning. We can see that the U.S. supreme court justices network is divided into two communities, and these two parts also represent the justices with guilty and innocent judgment.

For signed networks, several novel detection algorithms [17]–[19] came up with the identification of communities. However, on the one hand, these algorithms divide the network into two subnetworks—each subnetwork contains only positive or negative edges—and then calculate the modularity for each subnetwork by a weighted rule. On the other hand, the time complexity is higher, which restricts most of them to just handle sparse networks or those with fewer nodes. In addition, the existing modularity function for signed network lacks the meaning that can further explain the nature of the network when evaluating the partition scheme of communities in the signed network [20], [21]. Therefore, in order to probe into the communities within undirected signed networks, starting from the modularity advanced by Newman *et al.* [15], [16], we design a new function of modularity for undirected signed networks (i.e., being weighted or not). Meanwhile, a new algorithm for community detection based on the modularity function is proposed and its time cost is just a linear scale with the number of edges within networks. Taking together, the main contributions in this article can be summarized as the following three aspects.

1) New modularity function is proposed, which can be applied into signed networks for community detection, and it takes the impact of the positive and negative community on the formation of the entire community into account.
2) A network conversion algorithm is presented so that the unsigned network can evolve into the signed network.
3) An affinity index is designed to guide the process of network conversion in order to facilitate the generation of different types of networks.

The rest of this article is structured as follows. First, Section II introduces the related works of community detection on signed networks and briefly describes the problem of community detection, and then further analyzes the feasibility to identify the community structure hidden inside signed network as the end. In Section III, the related improvements are presented to deal with the community detection for signed networks, including the new function of modularity and the affinity index. Then, in Section IV, two algorithms are designed to convert the directed signed networks into the undirected signed networks and detect the community hidden in the undirected signed network. After that, Section V demonstrates the results of the algorithm in some illustrative networks and real-world networks, respectively. Finally, Section VI concludes this article.

II. Backgrounds

A. Preliminaries

In the recent 15 years, the research of community detection has been one of the hot topics in the analysis of complex networks. Numerous works have focused on the community detection algorithms of unsigned networks, howbeit, fewer works are devoted to the community mining on signed networks. In particular, more and more attention is paid to the interpersonal relationship [22], [23] and rumor transmission [24], [25], which attracts quite a few scholars to further explore the community properties of signed networks, and then a number of community detection algorithms for signed networks are advanced. Taking an example, Doreian and Mrvar [19] had proposed a local algorithm that utilized the local search method to divide nodes into different groups on signed networks. Then, Li *et al.* [26] proposed a novel GN-H algorithm, which combines the GN algorithm [15] with the hierarchical clustering algorithm [27]. In order to find the best partition scheme, the GN-H algorithm needs to be iterated many times, and each iteration is further subdivided into two steps. In the first step, the network with all positive edges will be considered as a temporary community. In the second one, the result of the first step and all negative edges are used to define the index of similarity measure to determine the hierarchical clustering, and then the obtained results at this step are considered as the input of the first step in the next iteration. However, the time complexity of the GN-H algorithm is as high as $O(m^2n)$, where m represents the number of edges and n is the total number of nodes in the network.

Along the way of improving the function of modularity, Gómez *et al.* [20] and Traag *et al.* [21] also put forward two new algorithms, which mainly split the network into two parts, that is, a positive and a negative subnetworks. For each subnetwork after division, the value of modularity is calculated independently and the final modularity is determined by weighting each part, and then the experimental results show that the proposed function of modularity can effectively measure the quality of the partition scheme.

After the dynamic mechanism was successfully introduced into unsigned networks [28], Wu *et al.* further utilized the clustering dynamics of signed networks, which highly improved their original work dealing with the overlapping community via dynamics of positive networks [29]. Likewise, Maia *et al.* [30] considered the community detection in complex networks through adapted Kuramoto dynamics, where each individual is represented as a phase oscillator and then the phase of the two oscillators with the positive relationship is gradually approached through the dynamic process; however, the phase of two oscillators with the negative link is gradually far away from each other. Eventually, the phase of oscillators will be synchronized into several different communities. Meanwhile, the number of communities can be determined by fixing the attribute of domain in advance, that is, various community structures can be obtained by setting different values of attribute domain.

In addition to the aforementioned algorithms, for signed networks, the problem of community detection in complex networks has been explored from various angles, such as an iterative algorithm to optimize the local maximum [31], an algorithm to investigate the evolution and memetic [32] by optimizing two objective functions, and so on. Furthermore, the current works are also compared with the FEC algorithm proposed in [17], and the myopic best solutions are given through the Pareto optimality theory [33]. Also, the classifications and evaluations for most of existing community detection algorithms are performed in a recent survey [34].

B. Problem Formulation

Consider an undirected and unweighted signed network as a simple graph $G = (V, E)$, where $V = \{v_1, \ldots, v_n\}$ represents the set of nodes, and E indicates the set of edges connecting different nodes (that is, self-loop or repeated edges are prohibited). Thus, the adjacency matrix $A = [A_{ij}]$ of the graph G can be defined as $A_{ij} = 1$ iff there is an edge with positive interaction between node i and j; Analogously, $A_{ij} = -1$ iff there is an edge with negative interaction between them; otherwise, $A_{ij} = 0$. For further dealing with a weighted network, W denotes its adjacency matrix, the node i and the node j are connected by an edge with the weight of w_{ij}, where the sum of the absolute values of the weights of all edges, the sum of the absolute values of the weights of all positive edges, and the sum of the absolute values of weights of all negative edges can be represented as w, w^+, and w^-, respectively. Then, one can further obtain the following equation $w = w^+ + w^-$. Based on the adjacency matrix A, $A_{i.}$ represents all elements at row i, $A_{.j}$ expresses all elements of the column j, $A^+_{i.}$ ($A^-_{i.}$) denotes all positive (negative) elements of the row i, $A^+_{.j}$ ($A^-_{.j}$) stands for all positive (negative) elements of the column j, and A^+_{ij} (A^-_{ij}) indicates the positive (negative) element at row i and column j. Thus, the number of positive or negative edges, the positive or negative degree of the node i are defined, respectively, as follows:

$$m_i^+ = \sum_j |A^+_{ij}|,\ m_i^- = \sum_j |A^-_{ij}| \tag{1}$$

$$k_i^+ = \sum_j [1 - \delta(A^+_{ij}, 0)],\ k_i^- = \sum_j [1 - \delta(A^-_{ij}, 0)] \tag{2}$$

where the Kronecker delta function $\delta(i, j)$ is a piecewise function of variables i and j, which is assumed to be one if the variables are equal, or zero otherwise.

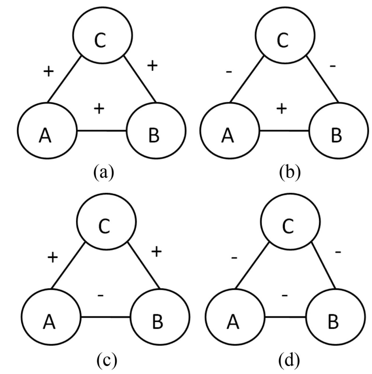

Fig. 3. Different combinations of relationship among three different individuals in the undirected signed network: the positive or negative signs along various edges represent whether the relationships between individuals are friends (+) or enemies (-).

Accordingly, the signed community on the signed network is defined in the following way.

Definition 1 (Signed Community): Assume that $V = V_1 \bigcup V_2 \bigcup \cdots \bigcup V_k$ is a union of k disjoint node clusters on the signed network G, and the number of positive or negative edges inside the same cluster is greater than the number of positive or negative edges between the neighboring clusters, then these k disjoint node clusters are termed as k signed communities of the network G.

Henceforth, the current challenge is to mine the community structure $C = \{C_1, \ldots, C_c\}$, where $C_i \bigcap C_j = \emptyset$ and $\bigcup_{i=1}^{c=|C|} C_i = V$.

C. Social Structure Balance

According to the above-mentioned definition, it is necessary to consider the balance state between these communities, which also means the balance between the nodes. Moreover, we need to study under what kind of case the signed network is in a state of balance, and whether the balance state is strong or weak. Actually, the signed network can be regarded as a generalization of the theory of structural balance, which was proposed by Heider [35]. Henceforth, the signs are used in the signed network to represent the positive and negative relationship among individuals, while the absolute value of corresponding weights characterizes the degree of bias toward these two relationships (but the range of the weights is often limited in the network). The theory of structural balance indicates that the balance among three different individuals can create the relationship of friends or enemies, which typically assumes the following four types of relationships:

1) a friend of my friend is my friend;
2) an enemy of my friend is my enemy;
3) a friend of my enemy is my enemy;
4) an enemy of my enemy is my friend.

But in fact, it is natural to find that the relationship among three individuals can be expressed as an undirected signed network, in which the edge is marked with a symbol or sign, as shown in Fig. 3(a)–(d). At the same time, by analyzing the relationship between individuals based on the theory of structural balance, psychology, and social science, it is easy to know that Fig. 3(c)–(d) is in a state of imbalance, and so there is a tendency to be converted into Fig. 3(a) and (b).

However, these simple networks are often embedded inside a large and real social network, which means that more than two subgroups or communities can appear in a large network. In this way, the balance of structure proposed by Heider [35] may not be applicable since only two subcommunities may be found under some specific constraints. Therefore, the weak structure balance theory and the k-balance network was proposed by Davis [36], and then the nodes in the network can be divided into k different subsets, which can render that the edges within nodes of the same subset are all positive, but the edges between different subsets are all negative. Henceforth, taking the related works on the balance theory into account, we can define the similar concepts on signed networks as follows.

Definition 2 (Structural Balance of Signed Network): Assume that the graph G is an undirected signed network, one can conclude the following.

1) If there exists an integer k such that the node cluster V_i is one of k node clusters satisfying the definition of signed community. Meanwhile, there only have the edges with positive weights within each cluster, and only have the edges with negative weights between different clusters. Thus, the network G is called a signed network with strong balance, that is, G holds a strong balance structure.
2) If there exists an integer k such that the node cluster V_i is one of k node clusters fulfilling the definition of signed community. Besides, the edges inside one cluster are all positive or a small number of edges are negative, but the edges between the different clusters are all those with negative weights or a very few ones with positive weights. Similarly, the network G can be termed as a signed network with weak balance, which means that the network G owns a weak balance structure.
3) If the network G is a strong balance network, the network is said to be a full balance signed network iff $k = 2$, and the network G has a full balance structure under this case.

Thus, the definition of signed community and structural balance for the signed network can be considered as the basis of detecting the communities hidden inside the signed network, in which the network will can be divided into several different communities.

III. Improvements

In the field of community detection, after one partition scheme is given, it is necessary to make the evaluation on the division results via a reasonable measure index. Among them, the validity, optimality, and accuracy are three frequently used quantities to judge the performance of one specific algorithm. Herein, starting from the modularity, some important concepts, measure indicators, and improvements are introduced in detail as follows.

A. Modularity

Regarding the evaluation of community detection algorithms for unsigned networks, several important indicators have been proposed, such as the modularity [15], [16], overlapping modularity [29], normalized mutual information [37], adjusted rand index [38], permanence [39], and so on. Among them, the most commonly used one is the function of modularity defined by Newman *et al.* [15], [16], in which one can calculate the modularity of a network by comparing the difference between the internal link density within the same group and the inter-link density between groups.

The first version of modularity function introduced by Newman and Girvan in [15], can be written as follows:

$$Q = \sum_{i \in C} (e_{ii} - a_i^2) = \sum_{i \in C} \left[\frac{e_i}{m} - \left(\frac{\sum_{j \in i} k_j}{2\,m} \right)^2 \right] \quad (3)$$

where e_{ii} describes the fraction of edges within community i over all the network edges; a_i indicates the fraction of edges connected with any node inside community i in the whole network; e_i denotes the total number of edges within community i; $\sum_{j \in i} k_j$ represents the sum of the degree of all nodes within community i, which means that every edge will be counted two times inside the community; m is the total number of edges in the network.

In [16], in order to simplify the matrix operation and further analyze the spectral characteristics, Newman redefined the function of modularity as

$$Q = \frac{1}{2\,m} \sum_{i,j \in V} \left[A_{ij} - \frac{k_i k_j}{2\,m} \right] \delta(C_i, C_j). \quad (4)$$

where $A = [A_{ij}]$ denotes the adjacent matrix of the whole network, k_i and k_j mean the degree of node i and j, $\delta(C_i, C_j) = 1$ if node i and j belong to the same community; Otherwise, $\delta(C_i, C_j) = 0$. It should be noted that the value of modularity often lies within the interval [−0.5,1], and the community structure becomes obvious when this value is between 0.3 and 0.7.

However, the modularity function cannot be applied into the signed network directly due to the existence of negative edges. Starting from the modularity, Gómez *et al.* [20] put forward an improved function of modularity for the signed network, which can be formulated as

$$Q = \frac{1}{2\,w^+ + 2\,w^-} \sum_i \sum_j \times \left[w_{ij} - \left(\frac{w_i^+ w_j^+}{2\,w^+} - \frac{w_i^- w_j^-}{2\,w^-} \right) \right] \delta(C_i, C_j) \quad (5)$$

where Q means a weighted sum of two parts of the value of modularity, w_{ij} represents the weight of the edge between node i and j, w_i^+, or w_i^- describes the sum of positive or negative edge weights for node i, and w^+ or w^- denotes the total sum of positive or negative weights in the whole network.

Although this modified function of modularity is applicable to the evaluation of community for signed networks, the function of modularity for the signed network just gives the size of one value, and does not give more information that is consistent with the signed network. In order to characterize the information through the value of modularity, it is natural to combine the value and sign of modularity to further improve this quantity. To this end, a new function of modularity is proposed, which can measure the community division results given by the different algorithms on signed networks. Here, it is assumed that the network G is undirected, unweighted, and signed. At this scenario, apart from taking the intracommunity and intercommunity edges into account, it still needs to consider the existence of positive and negative edges as follows:

$$e_{ij} = e_{ij}^{+} + e_{ij}^{-} \tag{6}$$

$$a_i = a_i^{+} + a_i^{-} \tag{7}$$

where e_{ij}^{+} and e_{ij}^{-} represent the proportion of the total number of positive edges and negative edges between community i and j to the total number of edges within the network, respectively; a_i^{+} or a_i^{-} denote the ratio of the total number of positive or negative edges connecting to the internal nodes within community i to the total number of edges for the whole network.

Therefore, the formula of modularity for undirected, unweighted, and signed networks can be expressed as

$$\begin{aligned} SQ &= \sum_{i\in C}\{[e_{ii}^{+} - (a_i^{+})^2] - [e_{ii}^{-} - (a_i^{-})^2]\} \\ &= \sum_{i\in C}\left\{\left[\frac{e_i^{+}}{m} - \left(\frac{\sum_{j\in i} k_j^{+}}{2m}\right)^2\right]\right. \\ &\quad \left. - \left[\frac{e_i^{-}}{m} - \left(\frac{\sum_{j\in i} k_j^{-}}{2m}\right)^2\right]\right\}. \end{aligned} \tag{8}$$

It should be noted that (8) is not derived by substituting (6) and (7) into (3) because all of negative edges are regarded as being infinitely close to the zero edge, but not to the null edge when we calculate the positive modularity. Similarly, when the negative modularity is considered, all of positive edges are thought to infinitely approach the zero edge, but not the null edge. In addition, the null edges indicate no existence, but the zero edges appear in the network only that they own the minimal weights.

According to the above-mentioned function of modularity, one can measure the stand or fall of community division results, that is, we need to judge what kind of values of modularity are excellent or poor. Thus, we will discuss the possible ranges of value of modularity as follows.

1) *Only one community:* The community division results show that the network can only be divided into one community, that is, $|C| = 1$. At this time, it is obvious that $e_{ii}^{\pm} = a_i^{\pm}$, then $SQ = 0$.
2) *The weight of all edges is positive:* If $\forall$ sign (A_{ij}) is larger than zero in the network, the following equation can be derived:

$$SQ = \sum_{i\in C}[e_{ii}^{+} - (a_i^{+})^2] \in [-0.5, 1]. \tag{9}$$

3) *The weight of all edges is negative:* In (8), if $\forall \text{sign}(A_{ij}) \le 0$, one can be rewritten and simplified as

$$SQ = -\sum_{i\in C}[e_{ii}^{-} - (a_i^{-})^2] \in [-1, 0.5]. \tag{10}$$

4) *Each node as a single community:* In this case, the number of communities is equal to the number of nodes in the network, that is, each node can be regarded as an independent community, and it is easy to know $e_{ii}^{+} = 0$ and $a_i^{\pm} = \frac{k_i^{\pm}}{2m} = \frac{w_i^{\pm}}{2w}$, SQ can be computed in the following way:

$$\begin{aligned} SQ &= \sum_{i\in C}\frac{k_i(k_i^{-} - k_i^{+})}{4m^2} \\ &= \sum_{i\in C}\frac{w_i(w_i^{-} - w_i^{+})}{4w^2} \in [-1, 1]. \end{aligned} \tag{11}$$

5) *More than one but less than n communities:* That is, it can be mined that there are multiple communities in the network, but it is not the case where each node is considered as an independent community. Thus, SQ can be deduced as follows:

$$SQ = \sum_{i\in C}[(e_{ii} - a_i^2) - 2(e_{ii}^{-} - a_i a_i^{-})] \in [-1, 1]. \tag{12}$$

From the aforementioned discussions, the possible range of the value of modularity SQ is $[-1, 1]$, and SQ determines the polarity of network. The more the value of SQ is, the more positive the network is. Likewise, the less the SQ value is, the more negative the networks tends to be. As an example, if SQ is close to 1, the network becomes more positive, which means there are a lot of positive communities in the network. Therefore, based on the SQ, we could mine much more information through the calculation of modularity and then easily draw the following conclusions.

1) In the case of only one community, the calculation result is the same as the value of modularity proposed in [15] and [16], i.e., SQ is equal to zero.
2) In the case of all weights being positive in the signed network, the function of modularity can be easily transformed into the classical function of modularity in [15] and [16].
3) In the case of all weights being negative in the signed network, the proposed function differs from the function in [15] and [16] by a negative sign, and it is obvious since the sign of value of modularity indicates that whether the network as a whole is biased toward fully positive network or not.
4) In the other cases, the authors in [15] and [16] cannot acquire the modularity of the signed network, but (8) can calculate the corresponding modularity.

According to the above-mentioned conclusions, the proposed function of modularity is more generalized and can be further adapted in the directed or weighted networks. Meanwhile, although there are some evaluation indicators to gauge the division results, it still lacks a specific index to evaluate the local characteristics. Therefore, a new indicator to evaluate the local characteristics of the node is put forward in the following section.

B. Affinity

Assuming that there is a directed signed graph G standing for a social network, one node i can have two types of edges: injection and ejection ones, where the weight of edges can be positive and negative. For a connection to an individual i, it can be considered as a link established by other individuals after i becomes attractive to them, and then the strength of connection is the recognition that other individuals attach to node i, in fact, it could be called the level of rejection when the strength of the connection is negative. Hence, the greater the attraction of individual i to others, the more likely it is to become a hub node. The neighbors around it may become a community around it, and making it easier for the information dissemination and diffusion [40].

Furthermore, it can be found that a social network will evolve in time or interpersonal relationship, which will lead to changes in connections within the network [41]. These changes will first appear as a change in the strength of connection, that is, with the decrease or increase of the attractiveness of the individual to others in the network, the strength of connection will also vary accordingly. Then, it will further lead to the variation of the number of links, that is, a new link will be created when the attractiveness of an individual to others is increased up to a certain extent. In the same way, when its attractiveness to others is reduced to a lower level, the existing connection between them will be destroyed and the number of links will be decreased in the network. Therefore, based on the injection edge, ejection edge, and attractive forces, two related concepts are built as follows.

Definition 3 (Affinity Edge): In a directed signed network G, if there exists an edge $E < i, j, w_{ij} >$, then the edge E is defined as the affinity edge of node j. If $w_{ij} > 0$, then the edge is termed as a positive affinity edge; otherwise, it is a negative affinity edge.

Definition 4 (Affinity Index): The proportion of the sum of weight for all positive affinity edges of node i to the sum of weight of the affinity edges of node i can be defined as the affinity index of the node i.

According to Definition 3 and 4, the calculation of affinity index is given as follows:

$$F_i = \frac{w_{i,\text{in}}^{+}}{w_{i,\text{in}}} \tag{13}$$

where $w_{i,\text{in}}$ denotes the sum of weight of the affinity edges of node i, and $w_{i,\text{in}}^{+}$ describes the sum of weight for all positive affinity edges of node i.

Combined with the index of affinity, a dynamic evolutionary process can be performed on directed signed networks, which will be eventually converted into undirected signed networks. First, the temporary relationship matrix of the directed signed network G is described as $W(t) = [W_{ij}(t)]$ at time t, where $W_{ij}(t)$ denotes the relationship—friends or enemies—between node i and node j at time t. Then, there are inconsistencies between the sign of different $W_{ij}(t)$ since there are different opinions among individuals in the network, that is, there are positive (friends) and negative (enemies) relationship in the network. Accordingly, a dynamic process over time can be used to describe the process of changing for the social network as mentioned earlier. Meanwhile, in order to ensure that the network is a signed network, the following definition is necessary for the dynamic evolutionary process.

Definition 5 (Dynamic Consistency State of Signed Networks): At time t, if the sign of relationships among all nodes converges into the same sign state in the directed network G, then the state is called the dynamic consistency state of signed networks, which means that $\forall\text{sign}(W_{ij}(t))$ is not less than zero or $\forall\text{sign}(W_{ij}(t))$ is not greater than zero.

Furthermore, one can perform a process of dynamic evolution in directed signed network based on the above-mentioned definition. To ensure that the process of evolution can be ended, and it is necessary to set up the maximum number of iterations T. At the same time, the process needs to perform the dynamic evolution according to the following two rules.

1) *Consistency Rule:* If $W(t)$ converges into the dynamic consistency state of signed network or the number of iterations reaches the maximum iteration number T (i.e., $t = T$), it needs to follow:

$$W_{ij}(t) = W_{ji}(t) = \frac{F_j}{F_i + F_j} W_{ij}(t-1) + \frac{F_i}{F_i + F_j} W_{ji}(t-1). \tag{14}$$

2) *Inconsistency Rule:* If $W(t)$ has not converged into the dynamic consistency state of signed network or the number of iterations has not reached the maximum iteration number T (i.e., $t < T$), the following condition needs to be satisfied:

$$W_{ij}(t) = F_j W_{ij}(t-1) + (1 - F_j) W_{ji}(t-1). \tag{15}$$

On the basis of these two rules, the dynamic evolutionary algorithm seems to be well executed. However, it is also necessary to consider an additional situation, where the node i does not have any positive affinity edge and has only negative affinity edges in the signed network since the limitation of the definition of affinity. Under this case, the node i should have a negative affinity, which needs to be further considered during the process of dynamic evolution of directed signed networks. Henceforth, it will cause $W_{ij}(t-1)$ to be zero when the edge $E < i, j, w_{ij} >$ does not exists in the process of inconsistency stage, and this is why it still needs to assume $W_{ij}(t-1)$ to take some default values at $\{-0.1, 0, 0.1\}$ in the process of evolution.

As a further step, different directed signed networks can be obtained by setting different values of T, and it can be predicted that the final evolved directed signed network, which is different from the original directed signed network, will come from the partial set of nodes and edges with negative edges in the original network. According to the evolved signed network given by the current method, we can convert the original directed signed network into an undirected signed network by using the consistency rule or the inconsistency rule. (The readers are recommended to refer to the Supporting Information for more details.)

So far, the evolutionary processes mentioned above are actually in line with the principle of human communication in social networks, which shows that the proposed dynamic evolution rules for the directed signed network are correct and valid. To be noted that, the newly proposed evolutionary process just requires: an initial directed signed network, the affinity index of the node, which is used to update the weight of edges and two rules that govern the evolutionary process.

After the relevant concepts and formulae have been introduced, the specific algorithms to implement these ideas will be described and detailed analyses will be carried out for each algorithm in Section IV.

IV. Algorithms

A. Description of Conversion Algorithm

The conversion algorithm mainly aims to convert the directed signed network into an undirected signed network, and it can be divided into the following three main stages: in the first stage, the affinity of each node is calculated to evaluate their attractiveness; during the second phase, the weight of edges will be updated, which will simulate the evolution of connection strength between individuals in the social network; in the last period, multiple edges between the same nodes in a directed signed network that achieves the dynamic consistency state are merged into one edge without any direction.

It should be pointed out that there are some special details to be paid more attention at each stage: in the first stage, the temporary relationship matrix for calculating the affinity of each node is provided by the directed signed network after the initial or the second phase, which means that the algorithm will be executed iteratively; during the second phase, according to the node affinity provided by the first stage, the weight of the nonzero edges in the temporary relationship matrix of the directed network is updated so as to mimic the evolution of the social network through the dynamic evolution of the directed signed network; in the last period, if the number of iterations reaches the maximum T or the sign of the weight of all edges in the signed network are already consistent, then it will not be evolved any more.

Based on the aforementioned description, the detailed execution steps of the conversion algorithm can be summarized as Algorithm 4.1.

The detailed analyses of time and space complexity will be conducted here for Algorithm 4.1. In step 2, it takes $O(m)$ time to construct its temporary relationship matrix by reading the list of edges for the directed signed network; it is necessary to traverse all nodes in order to calculate the affinity of each node, and actually each edge is traversed twice since one edge is associated with two nodes, therefore, the time is consumed from step 4 to 6 is $O(2\,m)$; At each process, we need to update the weight of all edges in the network to achieve the goal of dynamic evolution of directed signed networks, and then one process of time required for step 7–9 is $O(m)$; moreover, the process of dynamic evolution may not only be performed once, and thus, the actual time complexity of step 3–11 is $O(tm)$—in fact it can also achieve the purpose of evolution only once by setting $T = 1$; at step 13, the time complexity is $O(m)$ since it is necessary to construct an undirected network.

Algorithm 4.1

Input: directed signed network G_D;
Output: undirected signed network G_U;
step 1: initialize the maximum T and current count $t = 1$;
step 2: build an incidence matrix by the edges of network G_D;
step 3: **while** $W(t)$ is not satisfied with Def.5 and $t < T$ **do**
step 4: **while** the affinity of node i is not calculated in G_D **do**
step 5: utilize Eq.(13) to calculate affinity of the node i;
step 6: **end while**;
step 7: **while** $W_{ij}(t)$ is not equal to zero **do**
step 8: calculate the $W_{ij}(t+1)$ by Eq.(15) and $W(t)$;
step 9: **end while**;
step 10: $t = t + 1$;
step 11: **end while**;
step 12: update $W_{ij}(t+1)$ and $W_{ji}(t+1)$ by Eq.(14);
step 13: construct an undirected signed network G_U by $W(t+1)$;
step 14: return undirected signed network G_U;

During the execution of the algorithm, all operations are performed on the existing directed network, and it just needs to update the existing weights to implement the evolution. For the final step that converses the directed signed network into an undirected signed network, it is implemented by modifying the existing incidence matrix, that is, the algorithm requires a space complexity of $O(n^2)$.

According to the above-mentioned analyses, the worst time complexity of the current Algorithm 4.1 is $O(Tm)$, and the space complexity is $O(n^2)$, where T describes the maximum number of iterations, m represents the number of edges, and n denotes the number of nodes in the network, respectively.

B. Description of Community Detection Algorithm

In this section, the specific community detection algorithm for signed networks is introduced in detail, which is divided into two phases and multiple iterations. The next iteration will be continued on condition that the number of communities is changed after the previous iteration.

In the first phase, each node is individually initialized as a single community, and thus, there are many communities in the initial network. Then, all communities are iterated so that the community i is merged into its neighboring community j in order to obtain the larger value of modularity, which will be increasingly improved. This iteration process will be ended until no improvement can be made to the value of modularity by this action.

After the community i is merged into the neighboring community j, the value of modularity will be changed. Therefore, we can calculate the increment of value of modularity by using (16), which decides whether community i will be merged. That

Algorithm 4.2

Input: undirected signed network G;
Output: detected communities indicated by C;
step 1: read the nodes and edges of the network G;
step 2: **while** the number of nodes has been changed **do**
step 3: initialize each node as an independent community;
step 4: initialize the traversal list $order$;
step 5: **while** there exist moveable nodes in $order$ **do**
step 6: utilize Eq.(16) to move untraversed communities;
step 7: **end while**;
step 8: reconstruct network;
step 9: record the best result into C;
step 10: **end while**;
step 11: return C;

is, community i is merged into the neighboring community j when the increment of value of modularity is positive; if the increment of value of modularity is negative, community i will maintain its existing status

$$\Delta SQ = (e_{ij}^{+} - 2a_i^{+}a_j^{+}) - (e_{ij}^{-} - 2a_i^{-}a_j^{-}). \qquad (16)$$

In the second phase, all nodes within the same community given by the first phase will be folded as a new node, and then all new nodes will be used to reconstruct the network. Henceforth, on the one hand, the edge weight between new nodes is the algebraic sum of weight of edges among two communities, that is, the positive relationship will be neutralized by the negative relationship when there exist positive and negative relationships between communities. On the other hand, if the new node has no self-loop, and then the construction of new network is already a stable global community structure. The new network associated with these subcommunities has no longer been divided, and only associated with the current remaining nodes, and hence, the current algorithm is inclined to merge the network into a larger community. However, if the new node has a self-loop, the algorithm tends to divide the network into smaller communities.

The concrete iteration process is described in Algorithm 4.2.

Likewise, the complexity analyses for Algorithm 4.2 can be carried out as follows. In step 1, it takes $O(m)$ time to read the list of edges for the undirected signed network; for step 3, the time of $O(n)$ is required to initialize the community; in step 4, the time of $O(n)$ is necessary to disrupt the traversal order of the node; from steps 5 to 7, each edge needs to be traversed to find the best merging community, and then the time complexity is $O(m)$; at step 8, in order to reconstruct the network, we need to traverse all edges to calculate the edge weight among new nodes so that the time of $O(|C|)$ is necessary to establish this type of relationship between them. In addition, the algorithm will be repeated L times until it cannot increase the value of modularity through moving or merging communities.

During the running period, the $O(n)$ space is used to save the best partition results and $O(|C|)$ space is needed to maintain the result at the current iteration. Furthermore, the process of network reconstruction requires the $O(m+n)$ space to store the new network.

Taking together, the time complexity of Algorithm 4.2 is $O(L(n+m+|C|))$ and the space complexity is also $O(n+m+|C|)$. In reality, during the simulation experiment, it is found that L is often a smaller constant independent of the scale of network. Comparing with the complexity of many famous existing algorithms is shown in Table I, the computational time of the proposed method is pretty low.

TABLE I
COMPLEXITY OF EXISTING FAMOUS COMMUNITY DETECTION METHODS ON SPARSE GRAPHS

Algorithms	Time
Louvain [13]	$O(N \log N)$
CNM [42]	$O(N \log^2 N)$
DA [43]	$O(N^2 \log N)$
OCR-HK [44]	$O(N^2)$
Bayesian inference [45]	$O(N \log^{\alpha} N)$
Variational Bayian [46]	$O(N^{1.44})$
RN Potts model [47]	$O(N^{1.3})$

V. Experimental Results

In order to evaluate the effectiveness and accuracy of the proposed algorithm, the above-mentioned algorithm is first applied into several illustrative networks. Then, the effect of various parameters on the community division results is explored on some actual networks, which are converted from the directed weighted signed networks.

A. Illustrative Networks

First, two illustrative networks will be used for experiments, and similar experimental data or results can also be obtained in [17] and [48]. In Fig. 4, the network with 28 nodes and 42 edges, which include 12 negative edges, is presented here. On the basis of Fig. 4, seven negative edges are added into the network, as illustrated in Fig. 5. In fact, these two networks can be divided into the following three communities: {1, 2, 3, 10, 11, 12, 19, 20, 21, 28}, {8, 9, 17, 18, 26, 27}, {4, 5, 6, 7, 13, 14, 15, 16, 22, 23, 24, 25}; where there exist some negative edges between the first two communities and the third community in Fig. 4, but there are no negative edges among the first two communities. However, there all exist some negative edges among three communities in Fig. 5. In Figs. 4 and 5, solid lines represent the positive edges, the dashed lines denote the negative ones.

Unlike the multiobjective optimization algorithm [48], the overlapping communities are not considered here because the proposed algorithm will give a division result with the higher value of modularity. It is also different from the compressed sensing algorithm by Jiang *et al.* [6] that mainly explains the reconstruction network from the perspective of matrix, and combines algorithms such as K-SVD with low computation complexity. In Figs. 4 and 5, it can be found that the network is divided into multiple communities at the beginning of the algorithm, where Fig. 4 has nine communities and Fig. 5 has ten communities. However, before the network is reconstructed at

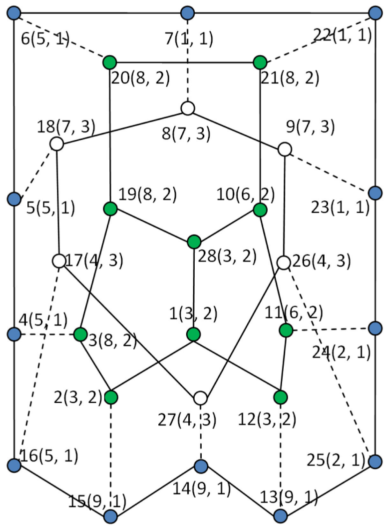

Fig. 4. Topological structure and community partition results of the first illustrative network: the division process is distinguished by the digits in the parentheses after the node, where the second digit denotes the final community numbering that the current node belongs to and the first one represents the community in which the current node lies before the final division is rightly arrived at.

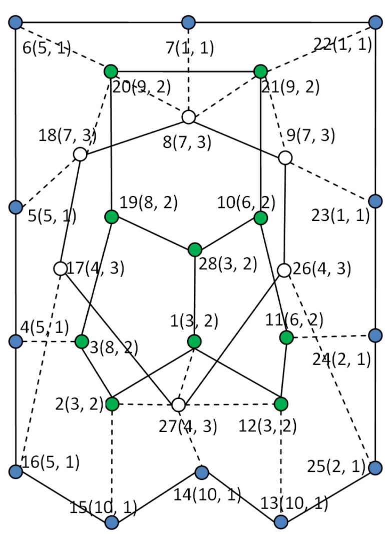

Fig. 5. Topological structure and community partition results of the second illustrative network: the division process is distinguished by the digits in the parentheses after the node, where the second digit denotes the final community numbering that the current node belongs to and the first one represents the community in which the current node lies before the final division is reached.

each iteration, the algorithm merges several small communities into a larger community, which eventually renders the final three communities. The current results indicate that the partitioning scheme given by the proposed algorithm is consistent with the actual situation, and small communities hidden within large communities can be mined during the running process of Algorithm 2. From the final topology of the network in Figs. 4 and 5, it can also be observed that although there are negative edges in the network, these negative edges are located between communities and positive edges can only lie inside the community. Thus, these

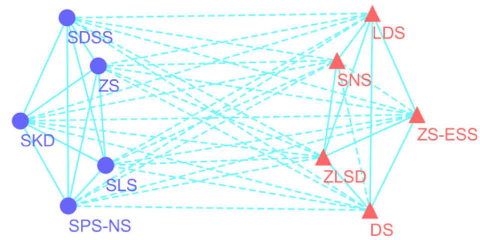

Fig. 6. Division result and the topological structure for the Slovene Parliamentary Parties Network: the blue circles and the red triangles denote two different communities, respectively.

two networks are signed ones in accordance with the definition of strong balance as described in Definition 2.

B. Real-World Networks

Here, several real-world signed networks including the Slovene Parliamentary Parties Network [49], the Gahuku-Gama Subtribes Network [50], and the Bitcoin Alpha [51] are used to test the performance of the proposed algorithm. Among them, the source data of these networks can be found online.[1,2,3]

The Slovene Parliamentary Parties Network describes the degree of association and similarity among ten coalitions, where if there is a link between two alliances, it means that they will have edges inside the network, or else they have no edges in the network. In [49], the edge weight ranges from -3 to 3, which denotes the different cases that are very dissimilar, relatively dissimilar, dissimilar, neither dissimilar nor similar, similar, relatively similar, and very similar. Several works have considered the case that only contains the positive and negative edges with the absolute value of weight 1, such as works in [17] and [52], but the proposed algorithm is also suitable for weighted signed networks. To this end, we take use of the online data[1] signed weighted network, as shown in Fig. 6 where the solid line is used to represent the positive edges and the dashed line denote the negative edges. It is obvious that the network is divided into two communities, that is, {SKD, SDSS, ZS, SLS, SPS-NS} and {ZLSD, LDS, ZS-ESS, DS, SNS}, which is fully consistent with the results obtained in [49]. In addition, from the topological structure shown in Fig. 6, we can find that although the network is divided into two communities, the network does not satisfy the definition of the full balance structure given in Definition 2 since there exists a negative edge between the node SNS and the node DS within the second community, and thus, the network is a weak balance signed network.

The Gahuku-Gama Subtribes Network is a social network of tribes of the Gahuku-Game alliance structure in the Eastern Central Highlands of New Guinea, which contains 58 pairs of relationships between 16 tribes.[2] In Fig. 7, we can clearly identify the topology of the network, where 28 solid lines represent the alliance relationship between tribes while the hostile relationship is pointed out by 28 dashed lines. Based on the relationship among 16 tribes, the network can be divided into three

[1]http://vlado.fmf.uni-lj.si/pub/networks/data/
[2]http://konect.unikoblenz.de/
[3]http://snap.stanford.edu/data

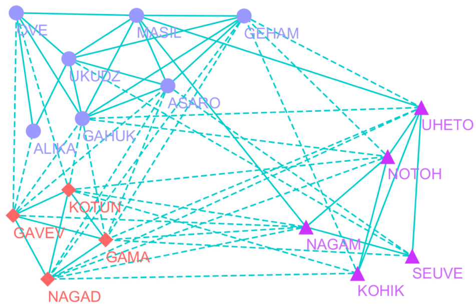

Fig. 7. Division result and the topological structure for the Gahuku-Gama Subtribes Network: three communities denoted by blue circles, red diamonds, and purple triangles correspond to three communities of realistic scenarios.

communities—{ASARO, GEHAM, OVE, UKUDZ, GAHUK, MASIL, ALIKA}, {NAGAD, GAMA, GAVEV, KOTUN}, and {NAGAM, UHETO, KOHIK, NOTOH, SEUVE}. Moreover, it is worth noting that [53] divides the network into four communities in order to achieve the balanced network topology, that is, {NAGAM, SEUVE} will be independent and become a new community. Contrary to Fig. 6, it is easy to find that there are two positive edges between the first community and the third community in Fig. 7, which also indicates that the network is a weak balance network as illustrated in Definition 2.

By observing the topological structure of these two actual networks, it can be found that there are positive or negative edges between different communities or internal communities, which indicates that the opinions from external or internal groups on a specific issue are not consistent for each individual within the realistic social networks [54]. Thus, these inconsistences reflect the human behavior or characteristics, and may have an impact on the results of a specified community detection algorithm, but our proposed algorithm can handle this situation very well. As an example, although the first and third communities in Fig. 7 own the positive edges between individuals, the proposed algorithm can still correctly classify these individuals into the communities, which they should belong to.

The Bitcoin Alpha network is a directed, weighted, and signed network of the who-trusts-whom network of people who trade by using Bitcoin. As Bitcoin users are anonymous, there will exist some risk for the trading between strangers. By marking each user with credit, it can effectively avoid the occurrence of fraudulent transactions. Therefore, the weights of edges fall within the range $[-10, 10]$, which represents the situation from the complete distrust to the full trust. The network constructed by data[4] a directed signed network; however, the proposed algorithm cannot effectively handle the directed signed networks. Therefore, it is necessary to convert the directed signed network into the undirected signed network by utilizing the aforementioned affinity index, and then execute the conversion algorithm on the directed signed network to obtained the corresponding undirected ones. In Fig. 8, the curve shows the change of number of the communities obtained after running the algorithm ten times on the Bitcoin Alpha network. As you can see from Fig. 8, the number of iterations of the algorithm is smaller,

[4]http://snap.stanford.edu/data

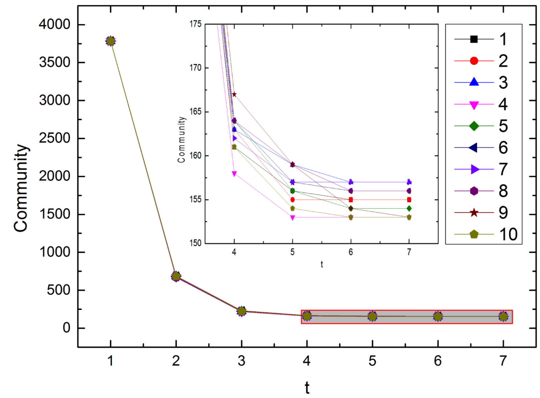

Fig. 8. Changing curve of number of communities for the Bitcoin Alpha network: the ordinate represents the number of communities; the abscissa is the tth iteration of the algorithm; the inset is a partially enlarged graph which showing the number of detected communities is close to 155, where the value displayed on the y-axis is 150–175, and the value displayed on the x-axis is 4–7.

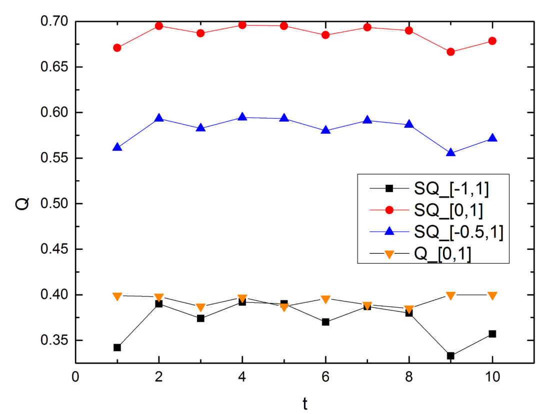

Fig. 9. Modularity comparison of community division on the Bitcoin Alpha network: SQ_$[-1, -1]$ represents the actual value calculated by (8); Q_$[0, 1]$ is the value calculated from (5); SQ_$[-0.5, 1]$ and SQ_$[0, 1]$ denote the value of SQ_$[-1, -1]$ mapped into the space of $[-0.5, 1]$ and $[0, 1]$, respectively.

which is consistent with the results given by the analysis of the algorithm. In order to check the convergence speed, the number of communities has been reduced to around 680 from the initial 3783 communities after just one iteration, and the number of communities inside the network will continuously decrease in the subsequent iteration. Accordingly, it is obvious to see that the number of communities does not change from the fourth iteration, which shows that the algorithm can quickly converge to several stable communities, in which a higher value of modularity can be acquired.

Fig. 9 displays the difference between the values of two kinds of modularity, one of which is (8), and the other is (5). It can be clearly seen that the value of modularity given by (8) is higher than that of modularity calculated by (5) when all values are mapped to the same value space. Meanwhile, the value of modularity is close to 1, which means that the network is biased toward positive communities. Moreover, by analyzing the network structure, it can be found that most positive edges are present in the network, which demonstrates the current prediction. These results further prove that the proposed function of modularity is not only suitable for evaluating the

TABLE II
STRUCTURAL CHARACTERISTICS OF ORIGINAL NETWORKS AND CONVERTED ONES

No.		Scale(Nodes/Edges)	Pst. Edge(%)	A.D.	W.D.
1	D	3,783/24,186	0.93	6.39	9.36
	U	3,783/14,124	0.90	7.47	17.87
2	D	82,140/549,202	0.77	6.69	3.66
	U	82,140/500,481	0.7	12.19	4.13
3	D	5,881/35,592	0.89	6.05	6.13
	U	5,881/21,492	0.86	7.31	14.47
4	D	131,828/841,372	0.85	6.38	4.51
	U	131,828/711,783	0.75	5.40	2.74

1. Pst. Edge represents positive edge; A.D. is the abbreviation of the average degree; W.D. is the abbreviation of the weighted degree.
2. D means that this is a directed network; U denotes an undirected signed network converted from directed signed network.
3. No.1 is the Bitcoin Alpha network; No.2 is user relationships network on the Slashdot obtained in February 2009, where the node set {11 680, 73 373, 78 482, 79 625} is lost; No.3 is the Bitcoin OTC network; No.4 is the epinions social network.

result of community division for undirected signed network, but also it can determine whether the network as a whole is biased toward the positive or negative community based on the sign of values of modularity. Apart from testing three actual networks mentioned above, some other directed networks are also converted into undirected networks and the related features of converted networks are shown in Table II.

As can be seen from Table II, there are also some differences between the converted networks and the original ones. For example, the number of edges decreases since there are edges connected to each other in the original networks, but only one edge can be remained in the converted ones. The proportion of positive edges decreases since a positive edge with a smaller weight may be converted to a negative edge when this node has a lower affinity (whether performing the conversion or not depends on the affinity of two nodes on the corresponding edges). Furthermore, more algorithm tests and performance evaluations are provided in the supplementary materials, and we will omit them in the main text for the brevity.

VI. CONCLUSION

In summary, an extended function of modularity has been presented for the undirected signed network based on the classical modularity, which can be adapted to unweighted and weighted networks. Meanwhile, this function can be further converted into the function of modularity for unsigned networks by removing the second part in the function. Thus, starting from the extended function of modularity in the signed or unsigned networks, the division results can be evaluated for different algorithms. Meanwhile, a fast algorithm of detection communities is proposed on basis of reconstructing networks to optimize the function of modularity and can also handle the disconnected network very well. After analyzing the execution process of algorithm in detail, it can be found that the time complexity can also reach $O(m)$ in a denser network, which can be attributed to the fast convergence speed of the algorithm. Furthermore, an affinity index is also introduced to convert a directed signed network into an undirected signed one. After carefully comparing the characteristics of these two networks, although it can be found that they have minor differences in some properties such as the proportion of the positive edges, key indices can keep the substantially same levels. Thus, the current results demonstrate that the proposed fast algorithms can effectively perform the community division on signed complex networks.

Although the current algorithm exhibits the better performance in several prototype and realistic networks with smaller size, it is still confronting the great challenges when handling the community detection within large-scale online social networks with up to billions of nodes. In addition, the current approach focuses on the optimization of community modularity and the community detection algorithms can be further optimized by combining the modularity with other significant indicators (e.g., number or size of communities).

REFERENCES

[1] J. Fumanal-Idocin, A. Alonso-Betanzos, O. Cordón, H. Bustince, and M. Minárovác, "Community detection and social network analysis based on the Italian wars of the 15th century," *Future Gener. Comput. Syst.*, vol. 113, pp. 25–40, Dec. 2020.
[2] L. Backstorm, P. Boldi, M. Rosa, J. Ugander, and S. Vigna, "Four degrees of separation," in *Proc. 4th ACM Int. Conf. Web Sci.*, Evanston, Il., USA, 2012, pp. 45–54.
[3] S. Datta, P. Basuchowdhuri, S. Acharya, and S. Majumder, "The habits of highly effective researchers: An empirical study," *IEEE Trans. Big Data*, vol. 3, no. 1, pp. 3–17, Mar. 2017.
[4] M. Safar, K. Mahdi, and L. Jammal, "Cyclic entropy of collaborative complex networks," *IET Commun.*, vol. 6, no. 12, pp. 1611–1617, Aug. 2012.
[5] L. Zhu, F. R. Yu, Y. G. Wang, B. Ning, and T. Tang, "Big data analytics in intelligent transportation systems: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 1, pp. 383–398, Jan. 2019.
[6] D. Jiang, W. Wang, and L. Shi, "A compressive sensing-based approach to end-to-end network traffic reconstruction," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 1, pp. 507–519, Jan. 2020.
[7] J. Bascompte, "Disentangling the web of life," *Science*, vol. 325, no. 5939, pp. 416–419, Jul. 2009.
[8] J. Bascompte, "Structure and dynamics of ecological networks," *Science*, vol. 329, no. 5993, pp. 765–766, Aug. 2010.
[9] S. Fortunato, "Community detection in graphs," *Phys. Rep.*, vol. 486, no. 3, pp. 75–174, Jun. 2009.
[10] A. Aleta and Y. Moreno, "Multilayer networks in a nutshell," *Ann. Rev. Condens. Matter Phys.*, vol. 10, pp. 45–62, Jan. 2019.
[11] D. Wang, A. Musaev, and C. Pu, "Information diffusion analysis of rumor dynamics over a social-interaction based model," in *Proc. IEEE 2nd Int. Conf. Collaboration Internet Comput.*, Pittsburgh, PA, USA, 2016, pp. 312–320.
[12] C. Li, L. Wang, S. Sun, and C. Xia, "Identification of influential spreaders based on classified neighbors in real-world complex networks," *Appl. Math. Comput.*, vol. 320, pp. 512–523, 2018.
[13] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *J. Statist. Mech. Exp.*, vol. 2008, no. 10, Oct. 2008, Art. no. 10008.
[14] U. N. Raghavan, R. Albert, and S. Kumara, "Near linear time algorithm to detect community structures in large-scale networks," *Phys. Rev. E*, vol. 76, no. 3, Oct. 2007, Art. no. 036106.
[15] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Phys. Rev. E*, vol. 69, no. 2, Feb. 2004, Art. no. 026113.
[16] M. E. J. Newman, "Modularity and community structure in networks," *Proc. Nat. Acad. Sci. USA*, vol. 103, no. 23, pp. 8577–8582, Jun. 2006.
[17] B. Yang, W. Cheung, and J. Liu, "Community mining from signed social networks," *IEEE Trans. Knowl. Data Eng.*, vol. 19, no. 10, pp. 1333–1348, Nov. 2007.
[18] J. Kunegis, A. Lommatzsch, and C. Bauckhage, "The slashdot zoo: Mining a social network with negative edges," in *Proc. 18th Int. Conf. World Wide Web*, 2009, pp. 741–750.
[19] P. Doreian and A. Mrvar, "A partitioning approach to structural balance," *Social Netw.*, vol. 18, no. 2, pp. 149–168, Apr. 1996.

[20] S. Gomez, P. Jensen, and A. Arenas, “Analysis of community structure in networks of correlated data,” *Phys. Rev. E*, vol. 80, no. 2, Aug. 2009, Art. no. 016114.
[21] V. A. Traag and J. Bruggeman, “Community detection in networks with positive and negative links,” *Phys. Rev. E*, vol. 80, no. 2, Sep. 2009, Art. no. 036115.
[22] M. Yu, W. Si, G. Song, Z. Li, and J. Yen, “Who were you talking to—Mining interpersonal relationships from cellphone network data,” in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining*, Beijing, China, 2014, pp. 485–490.
[23] C. Zhang, B. Wang, B. Lv, and G. Peng, “Social network characteristics of online shopping interpersonal relationship in real and virtual communities,” in *Proc. IEEE 14th Int. Conf. Commerce Enterprise Comput.*, Hangzhou, China, 2012, pp. 101–106.
[24] Y. Suzuki, R. Sugawara, and S. Okuhara, “Exploratory study of the recognition of complex double network model,” in *Proc. 4th Int. Conf. Ind. Eng. Appl.*, Nagoya, Japan, 2017, pp. 322–326.
[25] C. J. Kuhlman, V. S. A. Kumar, M. V. Marathe, S. S. Ravi, and D. J. Rosenkrantz, “Inhibiting diffusion of complex contagions in social networks: Theoretical and experimental results,” *Data Mining Knowl. Discovery*, vol. 29, no. 2, pp. 423–465, Mar. 2015.
[26] X. Li, H. Chen, and S. Li, “Exploiting emotions in social interactions to detect online social communities,” in *Proc. Pac. Asia Conf. Inf. Syst.*, Taipei, Taiwan, 2010, pp. 1426–1437.
[27] S. C. Johnson, “Hierarchical clustering schemes,” *Psychometrika*, vol. 32, no. 3, pp. 241–254, Sep. 1967.
[28] H.-J. Li, Z. Bu, A. Li, Z. Liu, and Y. Shi, “Fast and accurate mining the community structure: Integrating center locating and membership optimization,” *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 9, pp. 2349–2362, May 2016.
[29] J. Wu *et al.*, “Overlapping community detection via network dynamics,” *Phys. Rev. E*, vol. 85, no. 2, Jan. 2012, Art. no. 016115.
[30] D. M. Maia, J. E. de Oliveira, M. G. Quiles, and E. E. Macau, “Community detection in complex networks via adapted Kuramoto dynamics,” *Commun. Nonlinear Sci. Numer. Simul.*, vol. 53, pp. 130–141, Dec. 2017.
[31] A. Mahmood, M. Small, S. A. Al-Maadeed, and N. Rajpoot, “Using geodesic space density gradients for network community detection,” *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 4, pp. 921–935, Apr. 2017.
[32] Y. Li, J. Liu, and C. Liu, “A comparative analysis of evolutionary and memetic algorithms for community detection from signed social networks,” *Soft Comput.*, vol. 18, no. 2, pp. 329–348, Feb. 2014.
[33] M. Ehrgott, *Multicriteria Optimization*, 2nd ed. Berlin, Germany: Springer-Verlag, 2005.
[34] J. Tang, Y. Chang, C. Aggarwal, and H. Liu, “A survey of signed network mining in social media,” *ACM Comput. Surv.*, vol. 49, no. 3, pp. 42:1–42:37, Nov. 2015.
[35] F. Heider, “Attitudes and cognitive organization,” *J. Psychol.*, vol. 21, no. 1, pp. 107–112, Jan. 1946.
[36] J. A. Davis, “Clustering and structural balance in graphs,” *Human Relation*, vol. 20, no. 2, pp. 181–287, May. 1967.
[37] A. Strehl and J. Ghosh, “Cluster ensembles—A knowledge reuse framework for combining multiple partitions,” *J. Mach. Learn. Res.*, vol. 3, no. 3, pp. 583–617, Mar. 2003.
[38] J. M. Santos and M. Embrechts, “On the use of the adjusted rand index as a metric for evaluating supervised classification,” in *Proc. 19th Int. Conf. Artif. Neural Netw.:* Part II, Berlin, Germany, Sep. 2009, pp. 175–184.
[39] T. Chakraborty, S. Srinivasan, N. Ganguly, A. Mukherjee, and S. Bhowmick, “On the permanence of vertices in network communities,” in *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2014, pp. 1396–1405.
[40] A. A. Alshamsi, F. Pinheiro, and C. Hidalgo, “Optimal diversification strategies in the networks of related products and of related research areas,” *Nature Commun.*, vol. 9, no. 1, Apr. 2018, Art. no. 1328.
[41] C. Castellano, S. Fortunato, and V. Loreto, “Statistical physics of social dynamics,” *Rev. Modern Phys.*, vol. 81, no. 2, pp. 591–646, Oct. 2007.
[42] A. Clauset, M. E. J. Newman, and C. Moore, “Finding community structure in very large networks,” *Phys. Rev. E*, vol. 70, no. 6, Dec. 2004, Art. no. 066111.
[43] J. Duch and A. Arenas, “Community detection in complex networks using extremal optimization,” *Phys. Rev. E*, vol. 72, no. 2, Oct. 2005, Art. no. 027104.
[44] S. Boccaletti, M. Ivanchenko, V. Latora, and A. Pluchino, “Detecting complex network modularity by dynamic clustering,” *Phys. Rev. E*, vol. 75, no. 4, Apr. 2007, Art. no. 045102.
[45] M. B. Hastings, “Community detection as an inference problem,” *Phys. Rev. E*, vol. 74, no.3, Sep. 2006, Art. no. 035102.
[46] J. M. Hofman and C. H. Wiggins, “Bayesian approach to network modularity,” *Phys. Rev. Lett.*, vol. 100, no. 25, Jun. 2008, Art. no. 258701.
[47] P. Ronhovde and Z. Nussinov, “Local resolution-limit-free Potts model for community detection,” *Phys. Rev. E*, vol. 81, no. 4, Apr. 2010, Art. no. 046114.
[48] C. Liu, J. Liu, and Z. Jiang, “A multiobjective evolutionary algorithm based on similarity for community detection from signed social networks,” *IEEE Trans. Cybern.*, vol. 44, no. 12, pp. 2274–2287, Dec. 2014.
[49] S. Kropivnik and A. Mrvar, “An analysis of the slovene parliamentary parties network,” *Develop. Data Anal.*, vol. 12, pp. 209–216, Jan. 1996.
[50] H. J. Li, Z. Bu, Z. Wang, and J. Cao, “Dynamical clustering in electronic commerce systems via optimization and leadership expansion,” *IEEE Trans. Ind. Informat.*, vol. 16, no. 8, pp. 5327–5334, Aug. 2020.
[51] S. Kumar, F. Spezzano, V. S. Subrahmanian, and C. Faloutsos, “Edge weight prediction in weighted signed networks,” in *Proc. IEEE Int. Conf. Data Mining*, Barcelona, Spain, Dec. 2017, pp. 221–230.
[52] J. Q. Jiang, “Stochastic block model and exploratory analysis in signed networks,” *Phys. Rev. E*, vol. 91, no. 6, Jun. 2015, Art. no. 062805.
[53] Q. Cai, M. Gong, S. Ruan, Q. Miao, and H. Du, “Network structural balance based on evolutionary multiobjective optimization: A two-step approach,” *IEEE Trans. Evol. Comput.*, vol. 19, no. 6, pp. 903–916, Dec. 2015.
[54] V. Amelkin, F. Bullo, and A. K. Singh, “Polar opinion dynamics in social networks,” *IEEE Trans. Autom. Control*, vol. 62, no. 11, pp. 5650–5665, Nov. 2017.

Chengyi Xia (Member, IEEE) received the Ph.D. degree in control theory and control engineering from Nankai University, Tianjin, China, in 2008.

Since 2013, he has been a Professor with the School of Computer Science and Engineering, Tianjin University of Technology, Tianjin, China. His research interests include complex system modeling and analysis, complex networks, epidemic propagation, and evolutionary game theory.

Yongping Luo received the M.E. degree from the Tianjin University of Technology, Tianjin, China, in 2019.

His research interests include complex network and community detection.

Li Wang received the Ph.D. degree in control theory and control engineering from Nankai University, Tianjin, China, in 2009.

She is currently an Associate Professor with the School of Computer Science and Engineering, Tianjin University of Technology. Her research interests include complex network and multiagent system.

Hui-Jia Li (Member, IEEE) received the Ph.D. degree in the operational research and control theory from the Chinese Academy of Sciences, Beijing, China, in 2013.

He is currently a Research Fellow with the School of Science, Beijing University of Posts and Telecommunications, Beijing, China. His current research interests include complex systems, data mining, pattern recognition, complex network, and control theory.