Physica A 604 (2022) 127875

Contents lists available at ScienceDirect

Physica A

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

A label propagation-based method for community detection in directed signed social networks

Maryam Hosseini-Pozveh [a,*], Maedeh Ghorbanian [b], Maryam Tabaiyan [a]

[a] Department of Computer Engineering, Shahreza Campus, University of Isfahan, Iran

[b] Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan, Iran

ARTICLE INFO

Article history:
Received 19 April 2021
Received in revised form 26 June 2022
Available online 30 June 2022

Keywords:
Label propagation
Community detection
Directed signed social network

ABSTRACT

Community detection is one of the most essential issues in social networks analysis field. Among the available categories of algorithms, the label propagation algorithm-based (LPA-based) methods, due to their proper time complexity, are of high concern. As all the social networks explicitly or implicitly include signed relationships, the attempt here is to suggest an LPA-based approach for community detection in the directed signed social networks. The direction of edges is not addressed in available LPA-based community detection methods for signed social networks. In this respect, 1) a weighting method is suggested in order to utilize the direction information that converts the network into an undirected weighted signed social network, 2) this weight is combined with a second weight obtained from the sign information of the edges, and 3) the LPA is extended, where the combined weights are applied in label propagation. Moreover, the directed signed modularity and the directed signed flow-based capacity measures are proposed. The findings of the run experiments indicate that the proposed method as to the directed signed modularity, directed signed flow-based capacity, and frustration measures on real-world and synthetic data sets, outperforms its counterparts.

1. Introduction

Community structure is one of the primary properties of the social networks [1]; therefore, community detection has become a major research issue in social networks analysis field [2]. According to a well-known definition, communities are subgroups of nodes where intra-community connections are strong while inter-community connections are weak [3]. In this context, in most studies, social networks with undirected trust relationships are of concern [1]. Consequently, two important pieces of information: (1) the direction of the edges, and (2) the sign of the edges, are ignored in most methods for simplicity, leading to the loss of information throughout the community detection process. There exist methods which consider the edge direction [1] and the edge sign (positive as trust and negative as distrust) [4].

Different categories of community detection methods exist like the clustering-based, modularity-based [4,5], random-walk based, and label propagation algorithm-based (LPA-based) [5,6] methods, among which the methods based on label propagation algorithm (LPA) are of interest due to their lower run time in comparison with others which make them more scalable for large social networks [5,7]. The focus of this study is on LPA-based community detection methods.

* Corresponding author.
E-mail addresses: m.hosseini@shr.ui.ac.ir (M. Hosseini-Pozveh), m.ghorbanian@alumni.iut.ac.ir (M. Ghorbanian), ms.tabaiyan@alumni.iut.ac.ir (M. Tabaiyan).

https://doi.org/10.1016/j.physa.2022.127875
0378-4371/

Community detection has many important applications in social network analysis domain like network summarization, link prediction, public health, criminology, politics, and recommender systems [8]. The signed social networks are subject to these applications as well, as signed relationships exists in all social networks implicitly, where, distrust, foe, or negative relationships insist next to trust, friendship, or positive ones. Consequently, some social networks like the Epinions and Slashdot explicitly model these relationships besides positive relationships [9]. Due to the importance of the implicit distrust relationships, there exist an important research field in the social network analysis domain, which focuses on mining implicit negative relationships from the metadata of the social networks. This research field is named sign prediction or negative link prediction [4,10]. Signed relationships are inherently directed as has been modeled by most real-world signed social networks like the Slashdot and Epinions. The directional essence of signed relationships is considered in Status theory, one of the primary theories about signed social networks, as well [9]. In the Status theory, a directed positive link from node v to node u is interpreted in the sense that node v considers node u having higher status, and a directed negative link from node v to node u is interpreted in the sense that node v considers node u having lower status [9]. Sign and direction information makes simultaneously the community detection results more accurate and the community detection process more complex. Accordingly, community detection in directed signed social networks has been and is essential, requiring more effective and efficient algorithms. The objective of this study is to propose a label propagation-based method for community detection in directed signed social networks. To the best knowledge of the researchers here, there exists no study where the LPA-based community detection incorporates both the direction and the sign. Contributions of this study are outlined as follows:

- The directed signed social network is converted into a weighted undirected signed social network by assigning and averaging the following two weight measures:
  - For applying the edge direction, a weighting method proposed by Li [11], is extended to convert the information of both positive and negative edges' direction into the edge weight.
  - For considering the sign information of edges, a weighting method proposed by Zarei et al. [12] is applied to assign weight to the edges.
- LPA [7] is extended here to find community structure in the obtained weighted undirected signed social networks. The main extension consists of the modified label propagation rule to include weights of the obtained network.
- A modularity measure and a flow-based capacity measure are proposed for directed signed social networks.
- Experiments are conducted on four real-world data sets: (1) the Wikipedia Requests for Adminship [13], (2) the Bitcoin OTC [14], (3) the Epinions [15], (4) the Slashdot [16], and on synthetic data sets, to assess the effectiveness of the proposed method where results indicate that this method outperforms the benchmarks in identifying the community structure by considering modularity, flow-based capacity, and frustration measures. Moreover, the superiority of the flow-based capacity measure in comparison with the modularity measure in detecting communities considering the importance of the information flow direction in directed signed social networks becomes evident.

The rest of this article is organized as follows. The related works are reviewed in Section 2. The modularity and flow-based capacity measures for directed signed social networks are presented in Section 3. In Section 4, the proposed method is discussed. Details of the experiments on datasets are presented in Section 5. Finally, the study is concluded in Section 6.

2. Related works

2.1. Definitions

Social networks model the individuals and the relationships between them and are represented by the graphs, where, nodes are the individuals and edges are the relationships. A social network is expressed as $G = (V, E)$, where V is the set of nodes and E is the set of edges [1]. Three essential properties are determined for the edges of each social network: (1) direction: edges are either directed or undirected [1], (2) weight: edges are either weighted or un-weighted [1], and (3) sign: edges are either signed or unsigned [4].

Directed and undirected social networks: In an undirected graph $G = (V, E)$, every edge $\{i, j\}$ is an unordered pair of nodes specifying a relationship in both directions. In a directed graph $G = (V, E)$, every edge (i, j) is an ordered pair of nodes specifying a relationship from i to j [1]. The adjacency matrix A of the both directed and undirected networks with n nodes is an $n \times n$ matrix, expressed as follows [1]:

$$a_{i,j} = \begin{cases} 1, & \textit{if there is a relation from } i \textit{ to } j \\ 0, & \textit{otherwise} \end{cases} \tag{1}$$

The matrix of an undirected social network is symmetric and the matrix of a directed one is asymmetric [1].

Weighted social networks: In a weighted graph $G = (V, E)$, a weight is assigned to each edge (i, j) [1]. The weighted social networks are either directed or undirected. The adjacency matrix A of a weighted network with n nodes is an $n \times n$

matrix, expressed as follows [1]:

$$a_{i,j} = \begin{cases} w_{i,j}, & \textit{if there is a relation from i to j} \\ 0, & \textit{otherwise} \end{cases} \tag{2}$$

where, $w_{i,j}$ is the assigned weight to the edge *(i, j)*.

Signed social networks: In a signed social network $G = (V, E)$, every edge *(i, j)* is either positive (meaning trust or friendship) or negative (meaning distrust or foe relation). This fact makes the edges to appear in the two sets of E^+ and E^-. The adjacency matrix A of a signed network with n nodes is an $n \times n$ matrix expressed as follows [4]:

$$a_{i,j} = \begin{cases} 1, & \textit{if there is a positive relation from i to j} \\ -1, & \textit{if there is a negative relation from i to j} \\ 0, & \textit{otherwise} \end{cases} \tag{3}$$

A signed social network is either directed or undirected, and weighted or un-weighted simultaneously [4].

2.2. The label propagation-based community detection algorithms in unsigned social networks

There exist various categories of approaches for community detection in social networks which include clustering-based, modularity-based, random-walk based, and Label propagation algorithm (LPA)-based [5,6,17] methods. The LPA-based methods are simple, time efficient and scalable for large social networks. In LPA [7] first, a unique label is assigned to every node, next, a set of iterations begins and in every iteration, in random ordering, every node chooses a label with the highest frequency among its neighbors. Iterations end when no node can change its label, and finally, the nodes with the same label are grouped as the same community.

The main drawback of LPA is the randomness of identified communities in different runs of the algorithm [7,17] addressed in many studies. Updating the labels of nodes by considering an ordering of nodes is one of the main solutions for this issue. For this purpose, different measures like node degree, eigenvector centrality, and clustering coefficient are applicable for ordering [18]. This updating is similar to assigning a preference-based weight to the nodes and changing the update rule of labels based on this weight while labels of the nodes are updated in random ordering [18]. The related studies include assigning the preference based on the K-shell [19] or the core nodes [20] concepts. Assigning weights to the edges is another similar approach which generates robust communities. Lou et al. [21] applied coherent neighborhood propinquity weight which is computed between every pair of nodes to update labels thereof. Chin and Ratnavelu [22] assigned a weight to every edge based on the mutual neighbor score between the nodes. Jokar and Mosleh [23] proposed the strength-based weight between every node and its neighbors according to the number of the edges in the neighborhood of these nodes.

LPA is extended based on different insights [17,24] like extending for bipartite networks [25–27], weighted networks [28] and dynamic networks [28–30]. In some of the LPA-based algorithms the idea of belonging of the nodes to more than one label (community) in each iteration to find an overlapping community structure is provided; consequently, a belonging coefficient (BC) determines the strength of the correlation between the node and the community [31–33].

2.3. Edge direction in community detection

Because in practice many of social networks are of the directed type, edge direction is of high concern in the field of community detection, otherwise in community detection process many useful information will be lost. The challenge here is to determine the viewpoint about the concept of the edge direction and to define the concept of communities and the links among them based on this view.

A basic study in this context is presented by Leicht and Newman [34] where the modularity measure is of concern. They establish that if a node i has high out-degree and a node j has high in-degree, an edge from j to i should contribute more to the modularity than an edge from i to j, because it is a more unexpected edge than the other, consequently, they presented the directed modularity measure and an algorithm which detects communities through maximizing this measure. According to this viewpoint about edge direction, the flow of information in the network is of concern. Kim et al. [35] revealed that the mentioned directed modularity measure cannot distinguish the information flow importance among some directed edges. They proposed a measure named LinkRank extending PageRank to determine the importance of directed edges. According to the similar viewpoint about the direction of edges, Li [11] applied the idea of converting the directed graph into an undirected weighted graph in the label propagation algorithm, where, the following equation is applied to assign a weight to each edge based on the number of in-coming and out-coming edges of the nodes [11]:

$$W_{ij} = 1 - \frac{k_i^{out} * k_j^{in}}{d_i * d_j} \tag{4}$$

where, k_i^{out}, k_j^{in}, d_i, and d_j are the out-degree of the node i, the in-degree of the node j, the degree of the node i, and the degree of the node j, respectively.

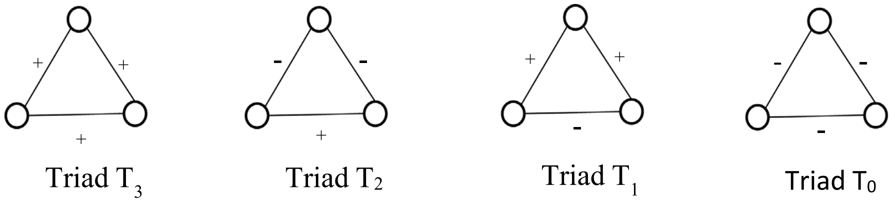

Fig. 1. The triads in an undirected signed social network. Triads T_0 and T_1 are unbalanced; Triads T_2 and T_3 are balanced [9].

Following, the labels are selected by nodes through aggregating over the labels of their neighbors considering the weight of the edges. According to them, this equation outperforms the directed modularity in identifying the importance of the edges in terms of local information flow direction [11].

Yang et al. [36] proposed a probabilistic generative model for community detection in directed networks, followed by an efficient Expectation–Maximization algorithm to compute the maximum likelihood solutions to their proposed model. Yang et al. [37] presented a community-affiliation network model for community detection in directed networks, where, the contribution to the directed networks is that nodes can both send and receive edges form communities in the affiliation model. Ning et al. [38] suggested a method based on local community extraction concept for community detection in directed social networks. The viewpoint about the edge direction in their study is based on the direction consistency of links from the nodes of a community to the other parts of the network. They proposed an MCMC-based approach to detect local communities. Although such direction consistency concept is appropriate for some networks, it is not proper for social networks because in the latter, the direction of edges are information flow-related. Jonnalagadda and Kuppusamy [39] proposed a game theory-based approach for detecting communities, where, first, a weight based on the common neighbors is assigned to every directed edge, and next, a cooperative game is applied to model the interactions among the nodes. In this process, a greedy algorithm detects the stable coalitions made among the nodes. Chang et al. [40] presented a probabilistic framework by sampling the network to generate bivariate distributions with the same marginal distributions. Based on this framework, the three algorithms for community detection in undirected networks are extended to directed network. Long et al. [41] presented a method for community detection in directed networks by formulating edge intensity for directed graphs, where the skeletons of networks is extracted by splitting the skeleton chain iteratively to detect communities based on the intensity-based modularity.

Direction in community detection in signed social networks is assessed by some researches as well. Traag and Bruggeman [42] extended an existing Potts model to incorporate negative links in order to detect communities in directed signed networks. In their study, the modularity is extended to signed networks and an optimization-based method is applied to detect communities. Chang et al. [43] proposed a method for community detection in signed networks based on error-correcting codes. Their objective is to minimize the sum of the number of positive edges between the two different communities and the number of negative edges within communities as the two sources of errors. Three different approaches are applied in accomplishing their objective. Chang et al. [44] presented a sampling-based probabilistic framework for centrality analysis and community detection in attributed networks including signed networks. The path measures are introduced and applied to determine how the sampling method should be twisted from the original sampling method in the underlying graph. Li et al. [45] proposed a spectral clustering based algorithm for directed signed graph partitioning. In [46] a greedy search is applied to optimize the value of the exact integrated complete data likelihood (ICLex) derived for the signed stochastic block model to find communities in directed signed social networks. Hu et al. [47] proposed a random walk-based method for community detection in directed signed social networks.

2.4. Label propagation-based community detection algorithms in signed social networks

The sign of the edges is an essential issue of concern in the community detection field in social networks. Accordingly, communities consist of disjoint subgroups of nodes, where in the communities, positive relationships are strong and negative relationships are weak while between communities, positive relationships are weak and negative relationships are strong [4,48]. This definition is subject to structural balance (SB) theory [9] of the signed social networks. Two types of triads are introduced in SB theory: the balanced and the unbalanced. With no concern on edge direction, there exist four possible triads in the signed social networks which are presented in Fig. 1. [9].

As observed in Fig. 1, based on the number of their positive and negative links, triads T_2 and T_3 are balanced and triads T_0 and T_1 are unbalanced [9]. As to community detection, an entirely balanced network (which includes only balanced triads) is divided into two communities where all the links within communities are positive and all the links between communities are negative. In practice, this does not hold in signed social networks, thus, detecting even a k-balanced community structure is impossible. A k-balanced community structure is defined as having *k* communities where all the links within are positive and all the links between communities are negative. Accordingly, a proper community detection approach seeks to find a structure as close as possible to a *k*-balanced structure [42,49].

According to the knowledge of the researchers here, only two studies have customized LPA for undirected signed social networks. In the first, two algorithms, signed Network Label Propagation Algorithm (SLPA), and signed network label propagation algorithm with structural balance degree (SBDSLPA) [50], are presented. In SLPA, after assigning a unique label to every node, a set of iterations begins where at every iteration, each node i chooses a label l_i according to the following equation:

$$l_i = arg\ \max_{l}(|N_i^+(l)| - |N_i^-(l)|) \tag{5}$$

where, N_i^+ is the set of the positive neighbors of i, N_i^- is the set of the negative neighbors of i, $|N_i^+(l)|$ is the number of positive neighbors with label l, and $|N_i^-(l)|$ is the number of negative neighbors with label l.

Iterations continues until no node can change its label; consequently, the nodes with same label are grouped as the same community. The main drawback of the SLPA is the randomness of identified communities in different runs of algorithm. To overcome this drawback, SBDSLPA assigns a degree, k_{ij}, to each edge $\{i, j\}$, named structural balance degree of the edge based on the structural balance theory. The values of k_{ij} s are computed as follows. First, for every negative edge, k_{ij} is set to -1 and for every positive edge to $+1$, next, if the edge is negative, for any balanced triads it belongs to, k_{ij} is decreased by 1, while if the edge is positive, for any balanced triads it belongs to, k_{ij} is increased by 1. After computing the weights of the edges, first, a unique label is assigned to every node, and next, a set of iterations begins and in every iteration, each node i chooses a label l_i based on the following equation:

$$l_i = arg\ \max_{l} \sum_{l_j = l, v_j \in N_i} k_{ij} \tag{6}$$

where N_i is the set of the neighbors of the node i.

Iteration continues until no node can change its label. Finally, the nodes with the same label are grouped as the same community [50].

In the second study [12], a method named weighted label propagation algorithm (WLPA) is proposed, where, a weight is assigned to every edge based on the Jacquard similarity measure. This measure is computed through the following equation:

$$sim(u, v) = \frac{S^+(u, v) - S^-(u, v)}{|\{N(u), u\} \cup \{N(v), v\}|} \tag{7}$$

where:

$$S^+(u, v) = |N^+(u) \cap |N^+(v)| + |N^-(u) \cap |N^-(v)| \tag{8}$$

and:

$$S^-(u, v) = |N^+(u) \cap |N^-(v)| + |N^-(u) \cap |N^+(v)| \tag{9}$$

where in both, $|N^+(u)|$, $|N^+(v)|$, $|N^-(u)|$ and $|N^-(v)|$ are the number of positive neighbors of u, the number of positive neighbors of v, the number of negative neighbors of u, and the number of negative neighbors of v, respectively.

After computing the weights of the edges, first, a unique label is assigned to every node. Then a set of iterations begins and in each, each node i chooses a label l_i according to the following equation:

$$l_i = arg\ \max_{l} \sum_{l_j = l, v_j \in N_i} sim(i, j) \tag{10}$$

where N_i is the set of the neighbors of i.

The iterations continue until no node can change its label. Finally, the nodes with the same label are grouped as the same community. In their study [12], the WLPA is not assessed in comparison with SLPA [50], and SBDSLPA [50].

Based on the content presented above, this study is the first with the objective of presenting an LPA-based method for community detection in directed signed social networks.

3. Directed signed measures

Two measures are proposed for evaluating the effectiveness of the detected communities in directed signed social networks. First one is the extended version of modularity for directed signed social networks. The second one is the directed signed flow-based capacity measure, where the information flow through the link structure of the social networks in forming the communities is of concern.

3.1. Directed signed modularity

Modularity, proposed by Newman and Girvan [51], is an essential and highly applied measure in assessing the quality of detected communities in the un-weighted undirected social networks. This measure is extended by Arenas et al. [52], for un-weighted directed social networks. The extended version of this measure for un-weighted undirected signed social networks is presented by Li et al. [53].

To the best knowledge of the researchers here, there exists no modularity definition for the directed signed social networks; consequently, in this study, it is defined as follows:

$$Q = \frac{m^+}{m^+ + m^-}Q^+ - \frac{m^-}{m^+ + m^-}Q^- \tag{11}$$

where, Q is the modularity symbol, Q^+ is the positive modularity symbol, Q^- is the negative modularity symbol, m^+ is the number of the positive edges of the social network, and m^- is the number of the negative edges of the social network. Q^+ is defined as:

$$Q^+ = \frac{1}{m^+}\sum_{ij}({A_{ij}}^+ - \frac{k_i^{out+}k_j^{in+}}{m^+})\delta(i,j) \tag{12}$$

where, ${A_{ij}}^+$ is the number of positive edges with i to j orientation, k_i^{out+} is the positive out-degree of node i, k_j^{in+} is the positive in-degree of node j, and $\delta(i,j)$ is 1 if i and j belong to the same community, otherwise 0. Q^- is defined as:

$$Q^- = \frac{1}{m^-}\sum_{ij}({A_{ij}}^- - \frac{k_i^{out-}k_j^{in-}}{m^-})\delta(i,j) \tag{13}$$

where, ${A_{ij}}^-$ is the number of negative edges with i to j orientation, k_i^{out-} is the negative out-degree of node i, k_j^{in-} is the negative in-degree of node j, and $\delta(i,j)$ is 1 if i and j belong to same community, otherwise 0.

3.2. Directed signed flow-based capacity measure

The concept of community based on the information flow in directed networks is discussed by Barroso et al. [54], where, a strong and cohesive community is defined as the one where communication among its members is high. They proposed a directed flow-based capacity measure to assess the quality of detected communities. For the group of nodes, S, this measure is computed as following equation:

$$\mu = \frac{\sum_{i,j\in S} f_{i,j}}{\sum_{i,j\in V} f_{i,j}} \tag{14}$$

where $f_{i,j}$ is the flow from i to j for the directed social networks.

Here, this measure is redefined based on information diffusion in communities for directed signed social networks. The information diffusion is modeled based on the information diffusion models. Community structure influences on how information flows in networks. Information tends to flow in dense communities; therefore there exist many community-based methods to find the most influential nodes of social networks [55].

The signed directed flow-based capacity measure is defined as follows:

$$\mu = \frac{\sum_{c\in C}\sum_{i\in c} f_{i,c}}{\sum_{i\in V} f_i} \tag{15}$$

where, $f_{i,c}$ is the overall influence of node i that flows in its community c and f_i is the overall influence of node i that flows throughout the social network. This influence is computed based on the SC-B model presented by Hosseini-Pozveh et al. [56]. To compute the influence of node i based on this model, all nodes are considered in one of the active, inactive, or blocked states. In this computation, first, node i is considered active while all nodes of community or social network are inactive. Then, the activation process for inactive nodes begins and proceeds in the discrete steps as follows:

In every step t, when a node v becomes active, it tries to activate each one of its inactive incoming neighbor nodes once. In this attempt, if its neighbor w trusts it, v tries to make w active based on $p^+_{v,w}$, but if there is a distrust, v tries to block w based on $p^-_{v,w}$. If w is surrounded by more than one newly activated neighbors who are trying to activate or block it, they follow a random permutation order to do, so If v succeeds, then w would be active or blocked in the next step. This process continues until no other activation is possible. The whole process continues for R rounds and the average number of activated nodes in these R rounds are considered as the f_i.

4. Proposed method

The existing LPA-based methods for signed social networks, SLPA [50], SBDSLPA [50], and WLPA [12] are not only designed for undirected ones, but also have their limitations. SLPA faces the randomness problem in detected communities in different runs [50]. SBDSLPA applies its proposed structural balance degree to identify communities. This degree is assigned to every edge and is calculated subject to the structural balance theory. The limitation in computing this degree is schemed in Fig. 2, where as observed, only the balanced triads to which the edge belongs, are involved in computations. For example in Fig. 2., the negative edge $\{a, d\}$ belongs to one unbalanced triad and zero balanced one so its degree remains

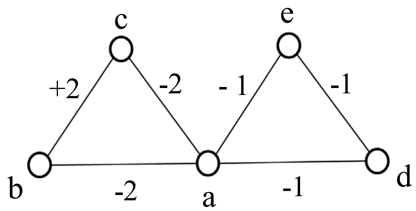

Fig. 2. A part of an assumed signed social network. Triad a-b-c is balanced and the a-d-e is unbalanced. The degrees of edges are computed through the SBDSLPA [50].

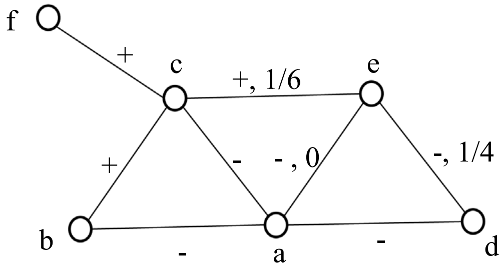

Fig. 3. A part of an assumed signed social network. The weights of edges $\{e, a\}$, $\{e, c\}$ and $\{e, d\}$ are computed through the WLPA [12].

-1 as its initial value. But, the degree of negative edge $\{a, b\}$ changes from -1 to -2 as it belongs to one balanced triad and zero unbalanced one. Consequently, node a prefers to accept a label from node d rather than b, while node a is in more balance with node b than node d.

According to the Eq. (7) of computing the similarity weights in WLPA algorithm [12], a node may prefer to choose a label from its negative neighbor than from its positive neighbor while that negative edge is unbalanced and that positive edge is balanced. As observed, in Fig. 3, the weights of edges $\{e, a\}$, $\{e, d\}$, and $\{e, c\}$ are computed as WLPA; consequently, node e chooses the label of its negative neighbor d rather than its positive neighbor c while edge $\{e, c\}$ is balanced and edge $\{e, d\}$ is unbalanced.

In this study, a new measure is suggested to be assigned to every edge and be incorporated in an LPA-based method for community detection in directed signed social networks. The proposed method (Fig. 4), named Directed Balanced LPA (DBLPA), is described in details as follows.

(1) Converting the direction of the edges into the weight:

A well-accepted approach for incorporating the direction of the edges into community detecting, is to first, convert the information of direction into the weight for the edges, and next, consider the graph as an undirected weighted graph to which the existing community detection methods can be applied on [1]. As described in Section 2, the method presented by Li [11] has applied this idea in LPA method. Here, this method is extended by considering both the positive and negative edges in social networks:

- The direction weight for every positive edge, directed from node i to node j, is calculated through the following equation:

$$W_{1,ij} = 1 - \frac{k_i^{out+} * k_j^{in+}}{d_i^+ * d_j^+} \tag{16}$$

  where, k_i^{out+} is the positive out-degree of the source node i, k_j^{in+} is the positive in-degree of the target node j, d_i^+ is the positive degree of the source node i, and d_j^+ is the positive degree of the target node j.

  This weight represents the importance of local positive flow of the edge among other local positive edges.
- The direction weight for every negative edge, directed from node i to node j, is calculated through the following equation:

$$W_{1,ij} = 1 - \frac{k_i^{out-} * k_j^{in-}}{d_i^- * d_j^-} \tag{17}$$

  where, k_i^{out-} is the negative out-degree of the source node i, k_j^{in-} is the negative in-degree of the target node j, d_i^- is the negative degree of the source node i, d_j^- is the negative degree of the target node j.

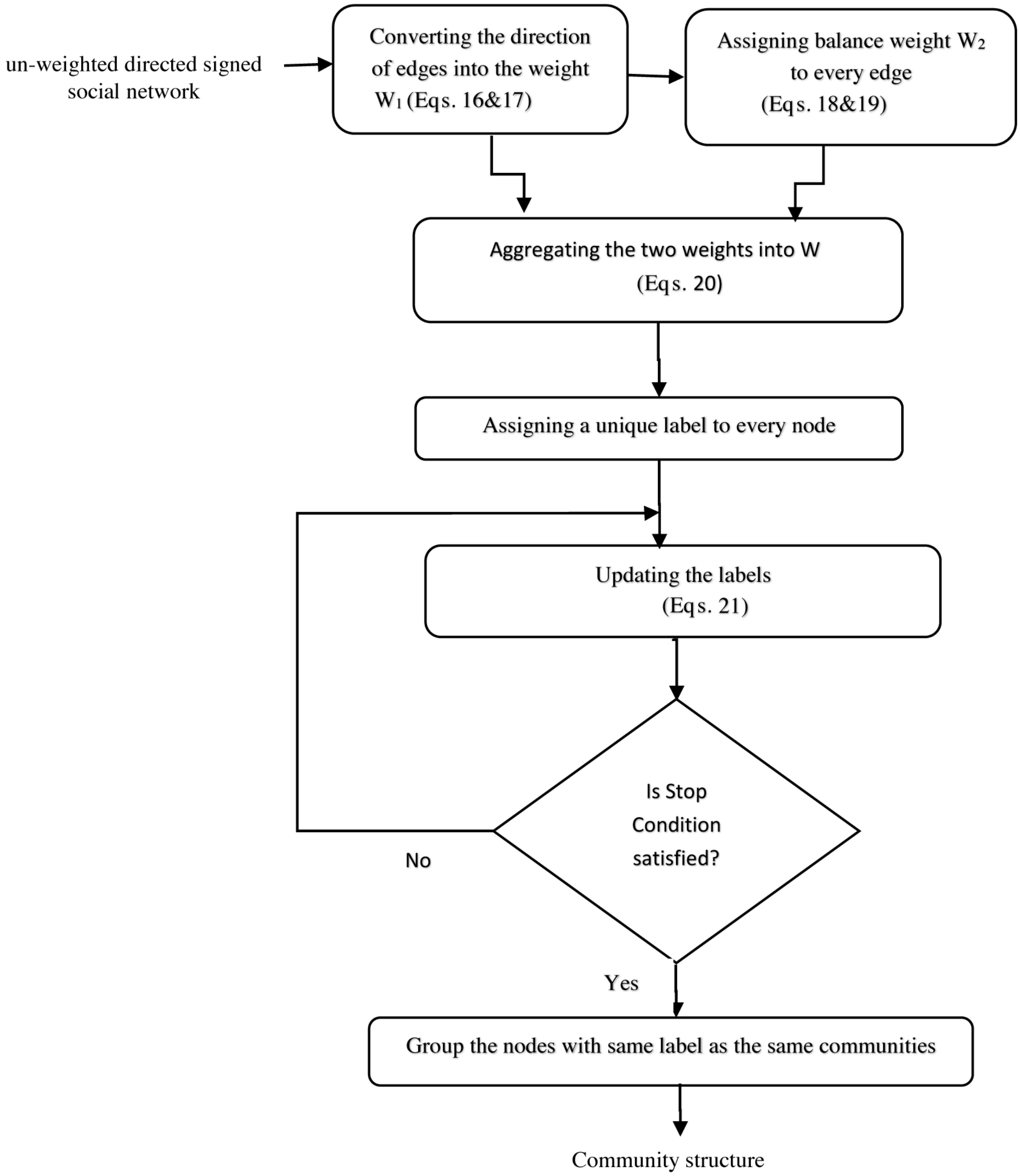

Fig. 4. The system diagram of the proposed method.

This weight represents the importance of local negative flow of the edge among other local negative edges.

The contribution of the local negative flow of some negative edges are compared in Fig. 5. (Contribution of the local positive flow follows the same pattern). According to the Eq. (17), $W_{1,b_1,a_1} = 0$, $W_{1,a_2,b_2} = 0$, $W_{1,b_3,a_3} = 0$, $W_{1,b_4,a_4} = 0$, $W_{1,b_5,a_5} = 5/6$, $W_{1,a_6,b_6} = 2/6$, $W_{1,a_7,b_7} = 15/16$, and $W_{1,a_8,b_8} = 7/16$. The higher this weight for positive edges, the more the edge's contribution in forming the communities. On the contrary, the higher this weight for negative edges, the less the edge's contribution in forming the communities. As observed in Fig. 5-a to -d, a negative edge is surrounded by positive edges. A negative edge among many positive edges, despite its direction, could contribute in forming a community as the weights of all of these edges are computed as 0. When the number of local negative edges increases, the contribution of negative edges in forming communities becomes more complex. In this regard, edge (a_5, b_5) has more negative contribution in comparison to Edge (a_6,b_6), because node b_5 which has more incoming negative edges has negative link toward a_5 and it is a more unexpected edge in comparison with the link from a_6 toward b_6. In a similar manner, edge (a_7,b_7) is more unexpected than edge (a_8,b_8). What is desired

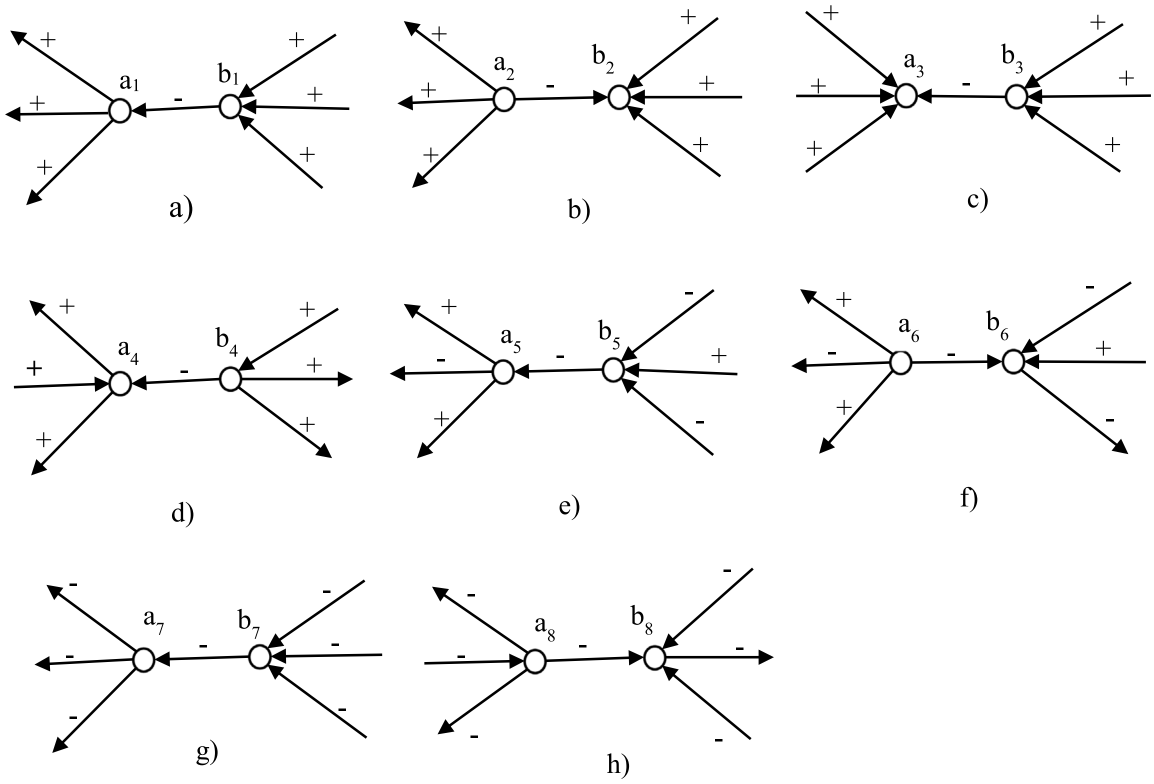

Fig. 5. The dynamics of some edges in the directed signed social networks.

for negative edges is to consider these weights in a manner that edges with smaller weights contribute more in forming communities (as it is evident in Eqs. (16) and (17)).

As it has been discussed, this weight utilizes the information of edges in community detection process properly. However, enriching this weight with the information of balance structure of the social network is of essence, where, in the signed networks, structural balance topology of the network would affect the information flow [57]. As observed in Fig. 5, the weight of edges in Fig. 5-a to -d are all zero as their direction information does not have any superiority over others, but their contribution in balance structure of their neighborhood may be different and should be considered in community forming.

With the available degree of nodes, the time complexity is $O(m_p + m_n) = O(m)$ where m_p is the number of the positive edges, m_n is the number of the negative edges, and m is the number of the edges of the social network.

(2) Assigning the balance weight to every edge considering the balance theory of signed social networks:

Every edge may belong to some balanced and some unbalanced triads. For every edge, the number of unbalanced triads the edge $\{i,j\}$ belongs to, is expressed as U_{ij}, and, the number of balanced triads the edge $\{i,j\}$ belongs to, is expressed as B_{ij}.

The similarity weight in the WLPA algorithm [12], expressed as Eq. (7), is rewritten here as follows:

– For every positive edge, directed from node i to node j, its balance weight is computed as:

$$W_{2,ij} = \begin{cases} \dfrac{B_{ij} - U_{ij}}{|\{N(i), i\} \cup \{N(j), j\}|}, & \mathit{if}\ |\{N(i), i\} \cup \{N(j), j\}| \neq 0 \\ 0, & \mathit{otherwise} \end{cases} \tag{18}$$

– For every negative edge, directed from node i to node j, its balance weight is computed as:

$$W_{2,ij} = \begin{cases} \dfrac{U_{ij} - B_{ij}}{|\{N(i), i\} \cup \{N(j), j\}|}, & \mathit{if}\ |\{N(i), i\} \cup \{N(j), j\}| \neq 0 \\ 0, & \mathit{otherwise} \end{cases} \tag{19}$$

where, in both, B_{ij}, U_{ij}, $|N^{+}(i)|$, $|N^{+}(j)|$, $|N^{-}(i)|$ and $|N^{-}(j)|$ are the number of balanced triads of edge *{i,j}*, the number of unbalanced triads of edge {*i*,*j*}, the number of positive neighbors of *i*, the number of positive neighbors of *j*, the number of negative neighbors of *i*, and the number of negative neighbors of *j*, respectively.

This phase takes the time $O(m.\overline{d}.log\overline{d} + m_p + m_n)$ where m is the number of the edges, $\overline{d}$ is the average degree of the nodes of the social network, m_p is the number of the positive edges and m_n is the number of the negative edges.

(3) Combining the direction weight and the balance weight for every edge:

The final weight for every edge is computed by averaging the two previously mentioned weights through the following equation:

$$W_{ij} = \frac{1}{2}(W_{1,ij} + W_{2,ij}) \tag{20}$$

Both the local flow importance and the local balance information of the edges are of concern in this measure which is designed to apply the positive essence of the positive edges and negative essence of the negative edges. This phase takes the time $O(m)$.

(4) Initializing label of the nodes:

A unique label is assigned to every node of the given social network, which takes the time $O(n)$.

(5) Updating label of the nodes:

A random order of all the nodes is determined and the label of every node is updated accordingly by considering the following presented rule:

$$l_i = arg\ \underset{l}{max}\Big(\sum_{l_j=l, j\in N_i^{+}} W_{ij} - \sum_{l_j=l, j\in N_i^{-}} W_{ij} \Big) \tag{21}$$

where, l_i and l_j are the labels of nodes i and j respectively, and N_i^{+} is the set of the trusted neighbors of i and N_i^{-} is the set of the distrusted neighbors of i.

This phase takes the time $O(m)$ same as the original LPA.

(6) Evaluating the final condition:

Step 5 repeats until the label of no node is changed any more. At this point, the label of a node determines its community.

4-1- Overall time Complexity

Steps 5 to 6, takes the $O(mt)$ time, where m is the number of the edges, and t is the number of the iterations until the final condition is met. Considering all the steps, the dominant time complexity belongs to steps 2 and 5 to 6, the $O(m.t + m.\overline{d}.log\overline{d})$.

5. Experimental evaluation

The experiments are conducted to address the following main question:

– Can the proposed method (DBLPA) improve the quality of detected communities in the directed signed social networks in comparison with benchmark methods?

All the implementations are performed in Python. Experiments are run on a 2.5 GHz Intel core i7 processor with 16 GB memory.

- Comparison measures

  Next to applying the proposed modularity measure for directed signed social networks (Eq. (11)) and the directed signed flow-based capacity measure (Eq. (15)), the frustration measure [58] is applied here. To calculate frustration [58], the number of negative edges between nodes in the same communities is summed up with the number of positive edges between nodes of different communities as it is presented in the following equation [58]:

$$F = \sum_{i,j\in V} {A_{ij}}^{-}\delta(i,j) + {A_{ij}}^{+}(1-\delta(i,j)) \tag{22}$$

  where, V is the set of the nodes, ${A_{ij}}^{+}$ is the number of positive edges between i and j, ${A_{ij}}^{-}$ is the number of negative edges between i and j, and $\delta(i,j)$ is 1 if i and j belong to the same community, otherwise 0.

  In current study, the normalized version of the frustration measure is applied as:

$$F = \frac{\sum_{i,j\in V} {A_{ij}}^{-}\delta(i,j) + {A_{ij}}^{+}(1-\delta(i,j))}{m} \tag{23}$$

  where m is the number of the edges of the social network.

  Although the time complexity of the presented method is discussed, the run time of the algorithms as another comparison measure is presented, too.

Table 1
Comparison of computed measures for Bitcoin OTC dataset.

	Directed signed modularity	Directed signed flow-based capacity measure	Normalized frustration	Run time (s)
DBLPA	**0.4126**	**0.7711**	**0.0693**	41.66
UBLPA	0.3749	0.6318	0.0784	36.07
SLPA	0.2962	0.6215	0.0886	**27.12**
SBDSLPA	0.3137	0.6135	0.0763	33.40
WLPA	0.3623	0.6303	0.1021	30.35

Table 2
Comparison of computed measures for Wikipedia dataset.

	Directed signed modularity	Directed signed flow-based capacity measure	Normalized frustration	Run time (s)
DBLPA	**0.3440**	**0.5979**	**0.1894**	86.44
UBLPA	0.2944	0.5035	0.2006	70.32
SLPA	0.2213	0.4158	0.1963	**58.52**
SBDSLPA	0.2812	0.4289	0.2019	62.23
WLPA	0.2851	0.5283	0.2074	71.05

Table 3
Comparison of computed measures for Slashdot data set.

	Directed signed modularity	Directed signed flow-based capacity measure	Normalized frustration	Run time (s)
DBLPA	0.3191	**0.5562**	0.2386	171.34
UBLPA	**0.3412**	0.4418	0.2515	138.03
SLPA	0.2722	0.3989	**0.2311**	**96.79**
SBDSLPA	0.3348	0.4125	0.2533	131.80
WLPA	0.3380	0.4325	0.2860	128.10

Table 4
Comparison of computed measures for Epinions data set.

	Directed signed modularity	Directed signed flow-based capacity measure	Normalized frustration	Run time (s)
DBLPA	**0.5241**	**0.4332**	0.1933	210.33
UBLPA	0.4791	0.3695	**0.1914**	180.56
SLPA	0.3805	0.3306	0.2110	**109.13**
SBDSLPA	0.3942	0.3348	0.1986	135.48
WLPA	0.4205	0.3527	0.2241	159.55

- Comparison methods
  The performance of the proposed method is compared with the three algorithms: SLPA, SBDSLPA [50] and WLPA [12], as there exists no LPA-based method for community detection in directed signed social networks. Because these three methods are presented for undirected signed social networks, the direction of the edges in social network is ignored when applying them on directed data sets.
  Next to these methods, the capability of an undirected version of the proposed method in improving the quality of detected communities is assessed. Accordingly, the first step of DBLPA method (described in Section 4, Eqs. (16) and (17)) is ignored in the case of community detection in undirected signed social networks, and the averaged weight of step three (Eq. (20)) is considered equal to the computed weight of the step two (Eqs. (18) and (19)). This version of proposed method is expressed as UBLPA abbreviated of undirected BLPA.
- The real-world datasets results
  The four real world directed datasets applied in this study are:
  - BitCoin OTC dataset [14]: This dataset contains 5881 nodes and 35 592 distinct voter/votee relationships.
  - Wikipedia Requests for Adminship dataset [13]: This dataset contains 10,835 nodes and 159,388 voter/votee relationships collected through votes assigned by the users to each other within the -1 to $+1$ range.
  - Slashdot dataset [16]: This dataset contains 82 140 nodes and 549 202 signed relationships.
  - Epinions dataset [15]: A signed dataset containing the information of 132 000 nodes and 841 372 trust/distrust edges.

  The results of computed directed signed modularity, signed directed flow-based capacity, normalized frustration, and run-time measures on the detected communities through different algorithms are tabulated in Tables 1, 2, 3, and 4 for Bitcoin OTC, Wikipedia, Slashdot, and Epinions respectively.

Table 5
Features of the six applied synthetic datasets.

Synthetic datasets	N	k_{avg}^{in}	k_{max}^{in}	com_{min}	com_{max}	t_1	t_2	μ_1, p^+, p^-
DS1	1000	10	50	20	50	2	1	{0.0,0.1,0.2,0.3,0.4,0.5}
DS2	1000	20	100	20	100	2	1	{0.0,0.1,0.2,0.3,0.4,0.5}
DS3	5000	20	100	20	100	2	1	{0.0,0.1,0.2,0.3,0.4,0.5}
DS4	5000	40	200	20	200	2	1	{0.0,0.1,0.2,0.3,0.4,0.5}
DS5	10 000	40	100	20	100	2	1	{0.0,0.1,0.2,0.3,0.4,0.5}
DS6	10 000	80	200	20	200	2	1	{0.0,0.1,0.2,0.3,0.4,0.5}

Table 6
Comparison of average directed signed modularity and average directed signed directed flow-based capacity measure in synthetic datasets.

	Average directed Signed modularity					Average directed signed flow-based capacity measure				
	DBLPA	UBLPA	WLPA	SBDSLPA	SLPA	DBLPA	UBLPA	WLPA	SBDSLPA	SLPA
DS1	0.2430	0.2741	**0.2900**	0.2232	0.2150	**0.5519**	0.4846	0.4918	0.4820	0.4732
DS2	0.2695	**0.3011**	0.2212	0.2135	0.1962	**0.5222**	0.4690	0.4561	0.4501	0.4365
DS3	**0.2638**	0.2420	0.2301	0.2188	0.2003	**0.5745**	0.5213	0.5111	0.5170	0.5039
DS4	**0.2743**	0.2276	0.2095	0.2074	0.1871	**0.6120**	0.5880	0.5604	0.5494	0.5400
DS5	0.2886	**0.3206**	0.2796	0.2583	0.2313	**0.6211**	0.5615	0.5510	0.5540	0.5427
DS6	**0.2815**	0.2480	0.2173	0.2328	0.2010	**0.6437**	0.5309	0.5440	0.5186	0.5264

Table 7
Comparison of average normalized frustration and average run time measures in synthetic datasets.

	Average normalized frustration					Average run-time (s)				
	DBLPA	UBLPA	WLPA	SBDSLPA	SLPA	DBLPA	UBLPA	WLPA	SBDSLPA	SLPA
DS1	**0.0951**	0.1271	0.1840	0.1512	0.1576	7.77	6.06	6.98	6.34	**4.35**
DS2	**0.0832**	0.1230	0.1787	0.1467	0.1496	16.90	11.66	9.45	10.89	**8.90**
DS3	0.0911	**0.0810**	0.1709	0.1639	0.1800	33.89	28.12	25.12	20.90	**15.14**
DS4	**0.1072**	0.1189	0.1810	0.1722	0.1606	44.96	39.78	37.37	37.20	**28.73**
DS5	0.1325	0.1338	0.1836	**0.1005**	0.1361	64.55	59.34	59.80	53.66	**49.99**
DS6	0.1185	**0.1131**	0.1679	0.1477	0.1562	97.44	80.22	83.34	73.56	**65.67**

Based on the obtained results, it is revealed that as to directed signed flow-based capacity measure, by considering all directed signed data sets, the DBLPA outperforms its counterparts. As to the directed signed modularity, although DBLPA is among the top two methods, in terms of all the data sets it is not the best yet. This is because in finding the communities, DBLPA emphasizes on the information flow through the links instead of merely the link structure. As to the normalized frustration, as emphasizing on the balance structure may slightly increase frustration in some cases, DBLPA is among the top two.

The DBLPA run time is more than other methods because of its higher computation volume than others, while the values establish the scalability of the method regarding different sizes of applied data sets. This fact is evident in the available time complexity analysis conducted before.

The undirected version of this newly proposed method, UBLPA, yields the second best results among its counterparts. Among three benchmark methods, as to the directed signed flow-based capacity measure and directed signed modularity measures, WLPA yields more acceptable results though its frustration value is high.

- The synthetic datasets results

The synthetic networks are generated by applying the LFR benchmark method. These networks are formed according to the instructions of generating directed social networks [59] and signed social networks [60]. Here, the parameters include N (number of nodes), k_{avg}^{in} (average in-degree), k_{max}^{in} (maximum in-degree), com_{min} (minimum community size), com_{max} (maximum community size), t_1 (exponent of in-degree power low distribution size), t_2 (exponent of community power low distribution size), μ_1 (mixing parameter of in-degree), p^+ (fraction of positive links among the communities), p^- (fraction of negative links inside the communities). Increasing the μ_1 parameter makes the community structure of the network unclear. Features of the six applied synthetic datasets are tabulated in Table 5.

Experiments are run for every data set considering the different values of the μ_1, p^+, p^- parameters. Accordingly, these three parameters are set to different values of {0.0, 0.1, 0.2, 0.3, 0.4, 0.5} in different runs; thus, the experiments are conducted for $6 \times 6 \times 6$ (216) times for each dataset. The average directed signed modularity and average directed signed flow-based capacity measures in these synthetic datasets are compared in Table 6. The average normalized frustration and average run time measures in these synthetic datasets are compared in Table 7.

The results obtained for the synthetic datasets confirm the results on real-world dataset. Accordingly, DBLPA outperforms other methods considering average signed directed flow-based capacity measure while it is not always the top first considering the average directed signed modularity and frustration measures. The findings indicate that this proposed method considers the direction of the information flow based on the balance link structure.

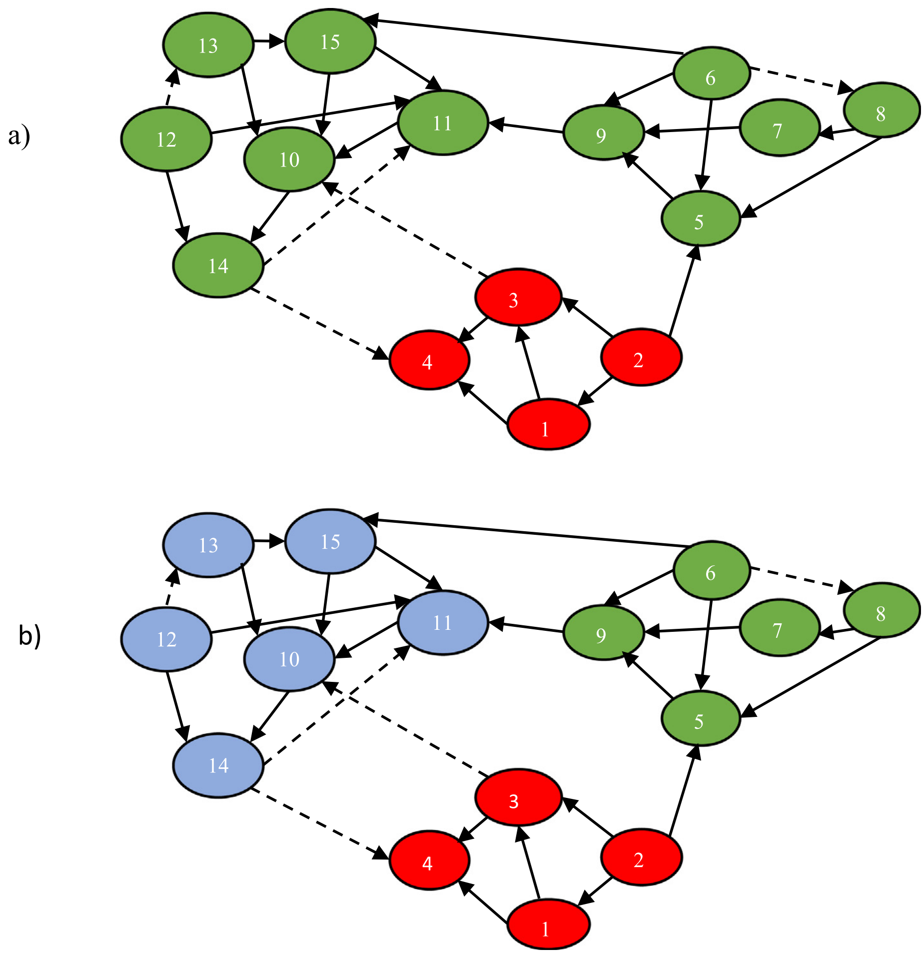

Fig. 6. (a) Communities detected by DBLPA, (b) communities detected by UBLPA. Positive edges are presented in solid lines, and negative edges are presented in dashed lines. Community members are shown in the same color.

Next to the mentioned experiments, the functionality of DBLPA is assessed in comparison with UBLPA through an illustration in Fig. 6. UBLPA determines communities only based on balanced structure of the social network. Therefore, the edge (9,11) is determined as an outer edge between two communities. It is while nodes 5, 7, and 6 are influenced by node 9 and node 9 itself is influenced by node 11. In fact, information flows from node 11 to node 9 and to nodes 5, 6, 7, consequently. This process is determined by DBLPA as it is shown in Fig. 6. In Fig. 6-a, frustration value is 4, directed signed modularity is 0.3340, and directed signed flow-based capacity value is 0.7120 while in Fig. 5-b, frustration value is 6, directed signed modularity is 0.4060, and directed signed flow-based capacity value is 0.6550.

6. Conclusion

An LPA-based method (DBLPA) is proposed to identify the community structure of the directed signed social networks. In this regard, first, an approach is proposed to convert the information of direction into the weights, next, the weights obtained from the edges direction are enriched through utilizing the balance theory, and then, an extended version of the LPA is presented to incorporate the obtained weights into the label propagation phase. Moreover, a version of DBLPA, named the UBLPA, is introduced to be applied in undirected signed social networks.

The proposed method is evaluated by considering the frustration, modularity, and flow-based capacity measures. In this regard, a modularity measure for directed signed social networks, and a directed signed flow-based capacity measure are proposed. Evaluations are conducted on synthetic datasets and four real-world social networks and the obtained results reveal the superiority of this proposed method in community detection of directed signed social networks.

One of the future tasks is to assess the other features and theories of signed social networks like the status theory in LPA-based community detection methods. Another work is to improve the proposed method for overlapping community detection in undirected and directed signed social networks. Assessing more effective variants of modularity measure considering the direction of flow and the resolution limit issue for LPA-based community detection in directed signed social networks is another research path that should be of concern.

Improving run-time efficiency of LPA-based methods for community detection in directed signed social networks is of essence, and requires applying the status theory, and the information diffusion_based concepts, where both may be able to consider the direction and sign simultaneously when transforming the edge and sign information to its weight in order to decrease the computations. Applying the idea of boundary nodes in decreasing the number of nodes with updated labels in each iteration of DBLPA is to be considered as another future task regarding run-time.

CRediT authorship contribution statement

Maryam Hosseini-Pozveh: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition. **Maedeh Ghorbanian:** Software, Validation. **Maryam Tabaiyan:** Methodology, Software, Validation, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The reference to datasets are presented in references part.

References

[1] F.D. Malliaros, M. Vazirgiannis, Clustering and community detection in directed networks: A survey, Phys. Rep. 533 (4) (2013) 95–142.
[2] S.R. Chintalapudi, M.K. Prasad, A survey on community detection algorithms in large scale real world networks, in: 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), IEEE, 2015, pp. 1323–1327.
[3] S. Fortunato, D. Hric, Community detection in networks: A user guide, Phys. Rep. 659 (2016) 1–44.
[4] J. Tang, Y. Chang, C. Aggarwal, H. Liu, A survey of signed network mining in social media, ACM Comput. Surv. 49 (3) (2016) 1–37.
[5] S. Souravlas, A. Sifaleras, M. Tsintogianni, S. Katsavounis, A classification of community detection methods in social networks: a survey, Int. J. Gen. Syst. 50 (1) (2021) 63–91.
[6] M. Plantié, M. Crampes, Survey on social community detection, in: Social Media Retrieval, Springer, London, 2013, pp. 65–85.
[7] U.N. Raghavan, R. Albert, S. Kumara, Near linear time algorithm to detect community structures in large-scale networks, Phys. Rev. E 76 (3) (2007) 036106.
[8] A. Karataş, S. Şahin, Application areas of community detection: A review, in: 2018 International Congress on Big Data, Deep Learning and Fighting Cyber Terrorism (IBIGDELFT), IEEE, 2018, pp. 65–70.
[9] J. Leskovec, D. Huttenlocher, J. Kleinberg, Signed networks in social media, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, 2010, pp. 1361–1370.
[10] M. Ozer, M.Y. Yildirim, H. Davulcu, Negative link prediction and its applications in online political networks, in: Proceedings of the 28th ACM Conference on Hypertext and Social Media, 2017, pp. 125–134.
[11] X. Li, Directed LPA: Propagating labels in directed networks, Phys. Lett. A 383 (8) (2019) 732–737.
[12] B. Zarei, M.R. Meybodi, B. Masoumi, Detecting community structure in signed and unsigned social networks by using weighted label propagation, Chaos 30 (10) (2020) 103118.
[13] [dataset], R. West, H.S. Paskovv, J. Leskovec, P. Christopher, Exploiting social network structure for person-to-person sentiment analysis, Trans. Assoc. Comput. Linguist. 2 (Oct) (2014) 297–310.
[14] [dataset], S. Kumar, F. Spezzano, V.S. Subrahmanian, C. Faloutsos, Edge weight prediction in weighted signed networks, in: Data Mining (ICDM), 2016 IEEE 16th International Conference on, IEEE, 2016, pp. 221–230.
[15] [dataset], P. Massa, P. Avesani, Extended epinions dataset, 2006, http://www.trustlet.org/extended_epinions.html.
[16] [dataset], J. Leskovec, D. Huttenlocher, J. Kleinberg, Signed networks in social media, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, 2010, pp. 1361–1370.
[17] L. Brahim, B. Loubna, I. Ali, A literature survey on label propagation for community detection, in: 2021 Fifth International Conference on Intelligent Computing in Data Sciences (ICDS), IEEE, 2021, pp. 1–7.
[18] L. Šubelj, M. Bajec, Robust network community detection using balanced propagation, Eur. Phys. J. B 81 (3) (2011) 353–362.
[19] Y. Xing, F. Meng, Y. Zhou, M. Zhu, M. Shi, G. Sun, A node influence based label propagation algorithm for community detection in networks, Sci. World J. (2014) (2014).
[20] Q. Gui, R. Deng, P. Xue, X. Cheng, A community discovery algorithm based on boundary nodes and label propagation, Pattern Recognit. Lett. 109 (2018) 103–109.
[21] H. Lou, S. Li, Y. Zhao, Detecting community structure using label propagation with weighted coherent neighborhood propinquity, Physica A 392 (14) (2013) 3095–3105.
[22] J.H. Chin, K. Ratnavelu, Detecting community structure by using a constrained label propagation algorithm, Plos One 11 (5) (2016).
[23] E. Jokar, M. Mosleh, Community detection in social networks based on improved label propagation algorithm and balanced link density, Phys. Lett. A 383 (8) (2019) 718–727.
[24] S.E. Garza, S.E. Schaeffer, Community detection with the label propagation algorithm: A survey, Physica A 534 (2019) 122058.
[25] X. Liu, T. Murata, Community detection in large-scale bipartite networks, Trans. Japanese Soc. Artificial Intell. 25 (1) (2010) 16–24.
[26] Z. Li, R.S. Wang, S. Zhang, X.S. Zhang, Quantitative function and algorithm for community detection in bipartite networks, Inform. Sci. 367 (2016) 874–889.

[27] H. Taguchi, T. Murata, X. Liu, Bimlpa: community detection in bipartite networks by multi-label propagation, in: International Conference on Network Science, Springer, Cham, 2020, pp. 17–31.
[28] J. Xie, M. Chen, B.K. Szymanski, LabelrankT: Incremental community detection in dynamic networks via label propagation, in: Proceedings of the workshop on dynamic networks management and mining, 2013, pp. 25-32.
[29] K. Liu, J. Huang, H. Sun, M. Wan, Y. Qi, H. Li, Label propagation based evolutionary clustering for detecting overlapping and non-overlapping communities in dynamic networks, Knowl.-Based Syst. 89 (2015) 487–496.
[30] M. Sattari, K. Zamanifar, A cascade information diffusion based label propagation algorithm for community detection in dynamic social networks, J. Comput. Sci. 25 (2018) 122–133.
[31] S. Gregory, Finding overlapping communities in networks by label propagation, New J. Phys. 12 (10) (2010) 103018.
[32] S. Liu, F. Zhu, H. Liu, Z. Du, A core leader based label propagation algorithm for community detection, China Commun. 13 (12) (2016) 97–106.
[33] N. Chen, Y. Liu, H. Chen, J. Cheng, Detecting communities in social networks using label propagation with information entropy, Physica A 471 (2017) 788–798.
[34] E.A. Leicht, M.E. Newman, Community structure in directed networks, Phys. Rev. Lett. 100 (11) (2008).
[35] Y. Kim, S.W. Son, H. Jeong, Finding communities in directed networks, Phys. Rev. E 81 (1) (2010).
[36] T. Yang, Y. Chi, S. Zhu, Y. Gong, R. Jin, Directed network community detection: A popularity and productivity link model, in: Proceedings of the 2010 SIAM International Conference on Data Mining, Society for Industrial and Applied Mathematics, 2010, pp. 742–753.
[37] J. Yang, J. McAuley, J. Leskovec, Detecting cohesive and 2-mode communities in directed and undirected networks, in: Proceedings of the 7th ACM international conference on Web search and data mining, 2014, pp. 323–332.
[38] X. Ning, Z. Liu, S. Zhang, Local community extraction in directed networks, Physica A 452 (2016) 258–265.
[39] A. Jonnalagadda, L. Kuppusamy, Mining communities in directed networks: A game theoretic approach, in: International Conference on Intelligent Systems Design and Applications, Springer, Cham, 2017, pp. 826–835.
[40] C.S. Chang, D.S. Lee, L.H. Liou, S.M. Lu, M.H. Wu, A probabilistic framework for structural analysis and community detection in directed networks, IEEE/ACM Trans. Netw. 26 (1) (2017) 31–46.
[41] H. Long, T. Wu, H. Yin, A skeleton-based community detection algorithm for directed networks, in: 2020 IEEE 3rd International Conference on Information Systems and Computer Aided Education (ICISCAE), IEEE, 2020, pp. 118–123.
[42] V.A. Traag, J. Bruggeman, Community detection in networks with positive and negative links, Phys. Rev. E 80 (3) (2009) 036115.
[43] C.S. Chang, D.S. Lee, L.H. Liou, S.M. Lu, Community detection in signed networks: an error-correcting code approach, in: 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), IEEE, 2017, pp. 1–8.
[44] C.H. Chang, C.S. Chang, C.T. Chang, D.S. Lee, P.E. Lu, Exponentially twisted sampling for centrality analysis and community detection in attributed networks, IEEE Trans. Netw. Sci. Eng. 6 (4) (2018) 684–697.
[45] Y. Li, S. Yuan, X. Wu, A. Lu, On spectral analysis of directed signed graphs, Int. J. Data Sci. Anal. 6 (2) (2018) 147–162.
[46] S. Ping, D. Liu, B. Yang, Y. Zhu, H. Chen, Z. Wang, Community detection in signed networks based on the signed stochastic block model and exact ICL, IEEE Access 7 (2019) 53667–53676.
[47] B. Hu, H. Wang, Y. Zheng, Sign prediction and community detection in directed signed networks based on random walk theory, Int. J. Embed. Syst. 11 (2) (2019) 200–209.
[48] P. Anchuri, M. Magdon-Ismail, Communities and balance in signed networks: A spectral approach, in: Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2012), IEEE Computer Society, 2012, pp. 235–242.
[49] P. Anchuri, M. Magdon-Ismail, Community detection in signed networks, 2011.
[50] L. Fang, Q. Yang, J. Wang, W. Lei, Signed network label propagation algorithm with structural balance degree for community detection, in: International Conference on Smart Homes and Health Telematics, Springer, Cham, 2016, pp. 427–435.
[51] M.E. Newman, M. Girvan, Finding and evaluating community structure in networks, Phys. Rev. E 69 (2) (2004) 026113.
[52] A. Arenas, J. Duch, A. Fernández, S. Gómez, Size reduction of complex networks preserving modularity, New J. Phys. 9 (6) (2007) 176.
[53] Y. Li, J. Liu, C. Liu, A comparative analysis of evolutionary and memetic algorithms for community detection from signed social networks, Soft Comput. 18 (2) (2014) 329–348.
[54] M. Barroso, I. Gutiérrez, D. Gómez, J. Castro, R. Espínola, Group definition based on flow in community detection, in: International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Springer, Cham, 2020, pp. 524–538.
[55] S. Banerjee, M. Jenamani, D.K. Pratihar, A survey on influence maximization in a social network, Knowl. Inf. Syst. 62 (9) (2020) 3417–3455.
[56] M. Hosseini-Pozveh, K. Zamanifar, A.R. Naghsh-Nilchi, Assessing information diffusion models for influence maximization in signed social networks, Expert Syst. Appl. 119 (2019) 476–490.
[57] H.J. Li, W. Xu, S. Song, W.X. Wang, M. Perc, The dynamics of epidemic spreading on signed networks, Chaos Solitons Fractals 151 (2021) 111294.
[58] A. Amelio, C. Pizzuti, Community mining in signed networks: a multiobjective approach, in: Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ACM, 2013, pp. 95–99.
[59] A. Lancichinetti, S. Fortunato, Benchmarks for testing community detection algorithms on directed and weighted graphs with overlapping communities, Phys. Rev. E 80 (1) (2009) 016118.
[60] Y. Su, B. Wang, F. Cheng, L. Zhang, X. Zhang, L. Pan, An algorithm based on positive and negative links for community detection in signed networks, Sci. Rep. 7 (1) (2017) 1–12.