Random walk on signed networks

Article *in* Physica A: Statistical Mechanics and its Applications · May 2018

DOI: 10.1016/j.physa.2018.05.139

Jianlin Zhou
Beijing Normal University

Ying Fan
Beijing Normal University

Zengru Di
Beijing Normal University

Physica A 508 (2018) 558–566

Contents lists available at ScienceDirect

Physica A

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

Random walk on signed networks

Jianlin Zhou, Lingbo Li, An Zeng*, Ying Fan*, Zengru Di

School of Systems Science, Beijing Normal University, 100875 Beijing, China

HIGHLIGHTS

- Design a signed random walk model which the agent can walk along the negative link.
- The position and density of negative links will affect the signed random walk model.
- This model could be used to explore the community structure in the signed networks.

ARTICLE INFO

Article history:
Received 25 February 2018
Received in revised form 5 May 2018
Available online 31 May 2018

Keywords:
Random walk
Signed network
Convergence
Community detection

ABSTRACT

Random walks on the traditional networks have achieved a series of research results in many aspects, such as analysis of node centrality, community detection, link prediction, etc. and have a wide range of applications. Actually, random walks can also apply to the signed networks which contain two types of links: positive links and negative links. However, there are few related researches about random walks on signed networks. And also we find that most researches about random walks on signed networks assume that the agent walks only along the positive links in the diffusion process, which loses the effective information of negative links. So in this paper, we propose a signed random walk model which allows that the random walker walks along the negative links with a smaller probability than positive links. We focus on two aspects of the signed random walk as follows: (1) the convergence of transition probability matrix. (2) the application to community detection in signed networks. And we apply the signed random walk to both artificial signed networks and real-world signed networks. The results show that the position and density of negative links in the signed network will affect the convergence rate of transition probability matrix. We also find that the signed random walk can be used to uncover the meaningful community structures in the signed networks.

1. Introduction

Many real-world complex systems such as information systems, social systems, ecological systems and so on can be described and understood by complex networks. This approach provides chances for studying the structure, functions and dynamics of these systems [1–4]. Most edges in traditional complex networks have no sign, but there exists a kind of complex networks called signed networks whose links have positive or negative sign. For example, in cortical neural networks, positive links denote excitatory relationship between neurons while negative links represent inhibitory relationship [5]; in signed social Networks, positive and negative links can represent friendly and hostile relations between the users respectively [6]. In the signed networks, nodes are units that can be genes, individuals, nations, etc. And positive and negative

* Corresponding authors.
E-mail addresses: anzeng@bnu.edu.cn (A. Zeng), yfan@bnu.edu.cn (Y. Fan).

https://doi.org/10.1016/j.physa.2018.05.139
0378-4371/

links represent the positive and negative relationships between a pair of units, respectively. Such opposite relationships are widely existed in social, information, even biological fields and it is of significant value to explore the signed networks.

Scholars in different fields have made a series of achievements in exploring the signed networks with the two basic theories: structural balance theory [7,8] and social status theory [9,10] proposed. These research works mainly involve the network topology analysis (node centrality measures, community detection, the balance of signed network, etc.), link and sign prediction, personalized recommendation, etc [11]. And these results also reveal that the negative links play an important role in the network formation, structural evolution, dynamical process, etc. However, most studies focus on the network structure of the signed networks and there are few studies involving dynamical process [12].

Random walks are one of the classical stochastic processes. Recently, Masuda et al. have made a detailed introduction about the random walks on networks in their review paper [13]. In the paper we can find that the related theories of random walks on networks such as stationary density, relaxation time, exit probability, mean first-passage and recurrence times, cover time have made a series of progress. In addition, variety of applications via random walks on networks have been developed. For example, based on random walks, we could use the ranking methods (PageRank [14], TempoRank [15], Random-walk betweenness centrality [16], etc.) to identify important nodes [17] and detection algorithms (Netwalk [18], Walktrap [19], InfoMap [20], stability [21], etc.) used for finding community. This review paper focuses mainly on the traditional networks such as directed or undirected network, weighted or unweighted network, single-layer network or multi-layer network, etc. and it mentions little about random walks on the signed networks.

Random walks on the signed networks have achieved some research results but it still remains unclear how the structural information of the negative links affects the dynamical behavior. Based on the random walks, researchers perform personalized ranking or community detection in the signed network. Jung et al. designed a new random walk model which introduces a sign into an agent who will change its sign when it travels the negative link and performed personalized rankings in the signed networks based on the model [22]. Yang et al. proposed a random walk based heuristic algorithm called FEC to mine signed network communities by estimating the probability for an agent starting from an arbitrary node of the network to reach the target node [23]. In addition, Wang et al. proposed a Self-Avoiding Pruning (SAP) random walk model on the signed network and investigated how network structure characteristics such as link density and degree distribution influence the properties of the SAP walk [24].

Community detection is an important part of the signed network researches and random walk can be used to explore community structure in the signed network by computing the distance between nodes. Negative links provide some information which could help improve the accuracy of the community detection [25,26]. However, at present, the existing transition rule which the agent's walk is restricted only along the positive links is not appropriate and this rule makes the random walker lose the information of negative links. In addition, since the negative links mainly exist between the communities, if the agent is allowed to walk along the negative links with larger probability, it will be easier for the random walker to enter other communities rather than staying its original community. Therefore, an agent should be allowed to walking along the negative links with a smaller probability compared with positive links.

Negative links are important components of the signed networks. In this paper, we design a signed random walk model (denoted by SRW) on undirected signed networks which the random walker will walk along the positive links with a larger probability compared with negative links. We firstly compare the convergence of the signed random walk and regular random walk (denoted by RRW) on the synthetic signed networks and random signed networks. We find that the location and density of the negative links have a greater impact on the convergence of SRW. Especially, the SRW converges more slowly when it is applied in the synthetic signed networks which the positive links only exist within communities. It indicates that the existence of community structure affects the convergence of SRW. We also find that as the proportion of negative edges in random signed networks decreases, the convergence of the SRW and RRW become faster. Then we apply the SRW in the artificial signed networks and real-world signed networks for community detection. The experimental results show that the SRW model could be used to explore the community structure in the signed networks.

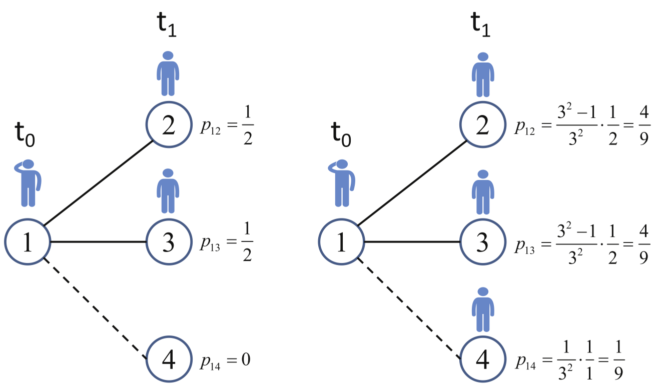
2. The walk rule on the signed network

Consider a discrete-time random walk on the undirected and unsigned networks. This kind of network could be described by adjacency matrix A, where its element $A_{ij} = 1$ if node i and node j are connected and 0 otherwise. We assume that each node is like a place and initially an agent is located at a node selected randomly. At each step the agent moves from the current position to one of its adjacent positions. Mathematically, the transition probability for an agent from node i to node j is defined as

$$T_{ij} = \frac{A_{ij}}{k_i} \tag{1}$$

where k_i is the degree of node i. Based on the transition matrix, we can obtain $T^t = \left\{T_{ij}^t\right\}$ which is the powers of matrix T, showing the probability of going from node i to node j through a random walk of length t.

Given an undirected and signed network with n nodes and m links, we can describe it by an adjacency matrix B with n dimensions. The entries of the adjacency matrix B are defined as follows: $B_{ij} = 1$ if there is a positive link between node i and node j, $B_{ij} = -1$ if there is a negative link between node i and node j, $B_{ij} = 0$ otherwise. In addition, the adjacency matrix B



(a) Regular random walk. (b) Signed random walk.

Fig. 1. (Color Online) The example of two random walks on the signed network. The solid edges denote positive links and the dashed edges represent negative links. At t_0, a random walker stays at the node 1. At t_1, the random walker will stay one of node 2, 3, 4. (a) and (b) provide the random walker with different strategies. (a) indicates that the random walker only walks along the positive links and chooses one of them evenly and randomly. (b) shows that the random walker can travel along the negative link with a smaller probability.

can be split into two parts: the matrices B^+ and B^-, $B = B^+ - B^-$, which $B^+_{ij} = B_{ij}$ if $B_{ij} > 0$ and 0 otherwise, and $B^-_{ij} = -B_{ij}$ if $B_{ij} < 0$ and 0 otherwise. Then the positive and negative degrees of node i can be defined as

$$k_i^+ = \sum_j B^+_{ij} \quad k_i^- = \sum_j B^-_{ij} \tag{2}$$

So the degree $k_i = k_i^+ + k_i^-$.

We also consider a discrete-time random walk on this undirected and signed network. For random walk on the signed network, the regular way for the agent is walking only along the positive links during the process as shown in Fig. 1(a). However, differentiate from the transition rules mentioned above, we propose a signed random walk model which the probability of walking along the positive links for the agent is greater than walking along the negative links as shown in Fig. 1(b). In this research we suppose that each node has at least one positive link. We give different walking rules considering two situations respectively: (1) the number of negative links of node i is more than zero. (2) the number of negative links of node i is equal to zero. For the first situation, the transition probability from node i to node j is defined as

$$T_{ij} = \begin{cases} \dfrac{k_i^2-1}{k_i^2}\dfrac{1}{k_i^+} & B_{ij}=1 \\ 0 & B_{ij}=0 \\ \dfrac{1}{k_i^2}\dfrac{1}{k_i^-} & B_{ij}=-1 \end{cases} \tag{3}$$

We can see that $\sum_j T_{ij} = 1$. For the second situation, the agent will walk along one of positive links which is chosen uniformly.

3. The empirical networks

3.1. Synthetic signed networks

Inspired by Yang et al. [23], in this article we build the synthetic signed networks with controlled community structure. The communities in the signed networks are defined as the groups of nodes within which the positive links are dense and between which negative links are dense. In this synthetic signed networks, most positive links exist within communities and most negative links exist between the communities. The synthetic signed network generated model is controlled by the following parameters: $c, n, k, p_{in}, p_-, p_+$, where c is the number of communities in the signed network, n is the number of nodes in each community, k is the average degree of a node, p_{in} indicates the proportion of edges existing within communities, $1-p_{in}$ indicates the proportion of edges existing between communities, p_- denotes the probability of negative links existing within communities, and p_+ represents the probability of positive links existing between communities. Then the proportion of negative edges in signed network is $p_{in} * p_- + (1-p_{in}) * (1-p_+)$. In this generated model, $c = 4$, $n = 32$, $k = 16$. Once the parameters c, n, k are determined, the structure of signed network only depends on parameters p_{in}, p_-, p_+. In addition, the synthetic signed networks with known community structure can be used to validate and compare the community detection algorithms.

Id (Parties)	1	2	3	4	5	6	7	8	9	10
1 (LDS)	0	173	157	134	23	-241	-254	-89	-142	-203
2 (DS)	173	0	170	57	-6	-184	-191	-170	-97	-109
3 (ZS-ESS)	157	170	0	77	-9	-120	-160	-77	-188	-80
4 (ZLSD)	134	57	77	0	49	-253	-230	-215	-150	-217
5 (SNS)	23	-6	-9	49	0	-132	-164	-210	-106	-174
6 (SLS)	-241	-184	-120	-253	-132	0	235	176	140	177
7 (SPS-SNS)	-254	-191	-160	-230	-164	235	0	117	116	180
8 (SKD)	-89	-170	-77	-215	-210	176	117	0	94	114
9 (ZS)	-142	-97	-188	-150	-106	140	116	94	0	138
10 (SDSS)	-203	-109	-80	-217	-174	177	180	114	138	0

Fig. 2. The adjacency matrix of the Slovene parliamentary party network.

Id (subtribes)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1 (GAVEV)	0	1	1	1	-1	0	0	0	0	-1	-1	-1	0	0	0	-1
2 (KOTUN)	1	0	1	1	-1	-1	-1	0	0	-1	0	-1	0	0	0	0
3 (NAGAD)	1	1	0	1	-1	-1	-1	-1	0	0	0	0	0	0	-1	-1
4 (GAMA)	1	1	1	0	-1	0	0	-1	-1	0	0	-1	0	0	-1	-1
5 (NAGAM)	-1	-1	-1	-1	0	1	0	0	1	0	0	0	1	0	0	0
6 (NOTOH)	0	-1	-1	0	1	0	1	1	0	0	0	-1	0	0	-1	0
7 (KOHIK)	0	-1	-1	0	0	1	0	1	0	0	0	0	0	0	-1	0
8 (UHETO)	0	0	-1	-1	0	1	1	0	1	0	0	-1	1	0	-1	0
9 (SEUVE)	0	0	0	-1	1	0	0	1	0	0	0	0	0	-1	0	-1
10 (OVE)	-1	-1	0	0	0	0	0	0	0	0	1	1	1	1	0	0
11 (ALIKA)	-1	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0
12 (GAHUK)	-1	-1	0	-1	0	-1	0	-1	0	1	0	0	1	1	1	1
13 (MASIL)	0	0	0	0	1	0	0	1	0	1	0	1	0	1	1	1
14 (UKUDZ)	0	0	0	0	0	0	0	0	-1	1	1	1	1	0	1	1
15 (GEHAM)	0	0	-1	-1	0	-1	-1	-1	0	0	0	1	1	1	0	1
16 (ASARO)	-1	0	-1	-1	0	0	0	0	-1	0	0	1	1	1	1	0

Fig. 3. The adjacency matrix of the Gahuku–Gama subtribes network.

3.2. Two real-world signed networks

The Slovene parliamentary party network is a weighted and signed network shown in Fig. 2, which describes the relations between 10 parties of the Slovene Parliamentary in 1994 [27]. In this signed network, the nodes are the parties and the edges represent the similarity between two parties. The relations between Parliamentary political parties were estimated by 72 members of the Slovene National Parliament via questionnaires. The questionnaires were designed by a group of expert on parliament activities to estimate the distances among the ten parties on a scale from -3 to 3. The weights of links in the network were the average values of 72 questionnaires and multiplied by 100. In this paper, we transform this weighted network into unweighted network and perform the following analysis on this unweighted network.

The Gahuku–Gama subtribes network is a binary and signed network shown in Fig. 3, where the nodes are subtribes, the positive links denote the political alliance between subtribes and the negative links represent the enmities between subtribes. It was created based on Read's study on the cultures of highland New Guinea [28].

4. The convergence of the transition probability matrix

The core of investigating the properties of random walks is to study their transition probability matrices. And the convergence of the transition probability matrix is usually as one of its key features. In this section, we focus on the convergence of the SRW's transition probability matrix and investigate how the network structure affects its convergence rate. Firstly, we define $p_i(t)$ as the probability that the random walker reaches node i in the signed network at discrete time t. Then the probability that the random walker reaches each node at discrete time $t+1$ can be calculated by $P(t+1) = P(t)T$, where $P(t) = (p_1(t), p_2(t), \ldots, P_N(t))$, T is the transition probability matrix of the random walk and N is the number of nodes in signed network. At some time t^*, if $P(t^*) = P(t^*+1)$, the transition probability matrix converges. Then we

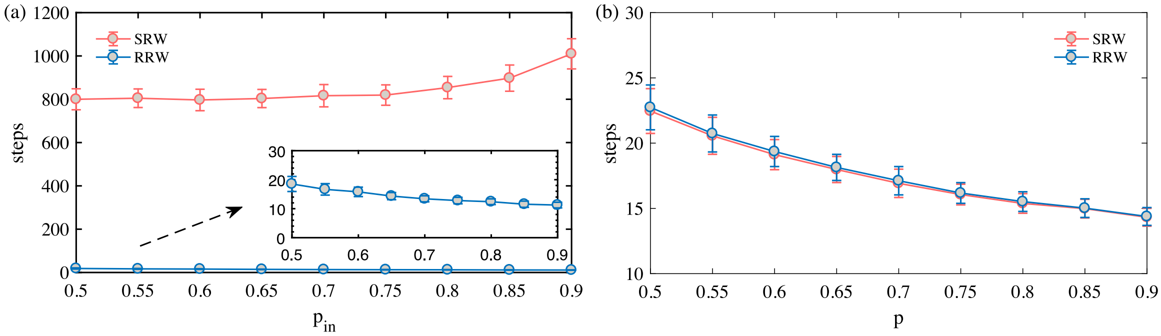


Fig. 4. (Color Online) (a) shows the number of steps before the SRW and RRW reach convergence in different synthetic signed networks generated by varying p_{in} and $p_- = p_+ = 0$. (b) shows the difference between the SRW and RRW on the random signed networks as the proportion of negative links decreases. Each random signed network has 128 nodes and the average degree of a node is 16. The points in the curves are all averaged over 50 realizations of synthetic signed networks or random signed networks.

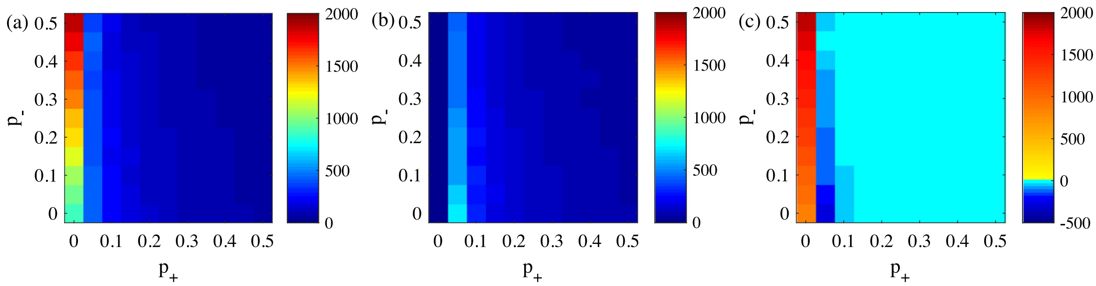


Fig. 5. The number of steps before the SRW and RRW reach convergence in the signed network generated by fixing the $p_{in} = 0.8$ and varying p_+, p_- from 0 to 0.5, are shown in (a) and (b), respectively. Different color represents different number of steps when the random walks converge. (c) shows the difference of the convergence between SRW and RRW on these synthetic signed networks. Each point in the heatmap is averaged over 50 realizations of the synthetic signed networks. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

define t^* as the time that the transition probability matrix achieves convergence. The smaller the time t^* is, the faster the transition probability matrix converges. In this paper, we define the error as $E(t) = \sum_{j=1}^{N} \left|p_j(t+1) - p_j(t)\right|$, where $p_j(t+1) = \sum_{i=1}^{N} p_i(t)T_{ij}$, $p_1(0) = 1, p_2(0) = \cdots = p_N(0) = 0$. And as the discrete time t increases, the error $E(t)$ gradually decreases. When the error is equal to zero, the transition probability matrix converges. To simplify the calculation, in this paper, when the error $E(t^*)$ is less than 10^{-4}, we think that the transition probability matrix approximately achieves convergence and t^* is the convergence time of the transition probability matrix.

To study the convergence of transition probability matrices of the SRW and RRW, we perform numerical simulation on the synthetic signed networks and random signed networks. Firstly, we fix parameters $p_- = 0$, $p_+ = 0$ and adjust p_{in} from 0.5 to 0.9 with the interval 0.05. We generate 50 connected synthetic signed networks under each set of parameters. In these networks, the positive links only exist within communities and negative links only exist between the communities. At this time, p_{in} is the proportion of positive links in the signed networks. Then we apply the SRW and RRW to these signed networks and calculate the average transition steps respectively when their transition probability matrices converge. The analytical result is shown in Fig. 4(a). We can observe that the convergence of SRW is significantly slower than the RRW's when there are only negative links between communities. We also find that as the p_{in} increases, that is, positive links within the communities increase and negative links between the communities decrease, the convergence of SRW becomes slower.

We also generate the random signed networks which are controlled by three parameters: m, k, p, where m is the number of nodes in the signed networks, k is the average degree of a node and p is the proportion of positive edges in the signed networks. The random signed networks have no communities. We generate random signed networks which their parameters: $m = 128$, $k = 16$, and varying p from 0.5 to 0.9 with the interval 0.05. We also apply the SRW and RRW to these random signed networks and compare their convergence time shown in Fig. 4(b). We can see that there are slight differences of the convergence time between SRW and RRW on random signed networks. At the same time, we also find that as the proportion of negative edges decreases, the convergence of these two random walks both become faster, which indicates that the link density of negative links can affect the convergence rate of random walks.

Secondly, we fix the parameter $p_{in} = 0.8$ and vary the other two parameters p_+ and p_- from 0 to 0.5, respectively. We also generate 50 connected synthetic signed networks under each set of parameters. The rate of convergence for RRW and SRW on these signed networks could be seen in Fig. 5. We could find that when $p_+ = 0$, the convergence rate of SRW will be slower as the parameter p_- gradually increases shown in Fig. 5(a), but the convergence rate of the RRW remains relatively

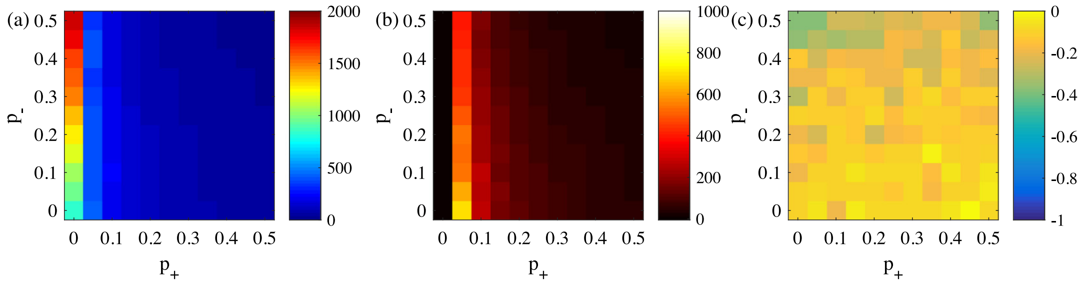


Fig. 6. (Color Online) The difference between SRW and RRW on synthetic signed networks and random signed networks in convergence are shown in (a) and (b), respectively. And (c) shows the difference of the convergence between SRW and RRW on random signed networks. The synthetic signed networks are generated by fixing the $p_{in} = 0.8$ and varying p_+, p_- from 0 to 0.5. The random signed networks have the same proportion as the negative links of the corresponding synthetic signed networks and they are generated 50 times under each pair of different parameters. Each random signed network has 128 nodes and the average degree of a node is 16.

stable shown in Fig. 5(b). And the number of steps of SRW is obviously higher than RRW's. Moreover, when $p_+ = 0.05$, the RRW converges more slowly than SRW under different parameters p_- shown in Fig. 5(c). However, for other parameters p_+ and p_-, there are no obvious differences between the SRW and RRW in convergence. We also compare the convergence time of these two random walks on these synthetic signed networks and their corresponding random signed networks shown in Fig. 6. These generated random signed networks have the same proportion of negative edges with their corresponding synthetic signed networks. We can see that there exists obvious difference about the convergence time of SRW and RRW on synthetic signed networks and random signed networks. And these two random walks on the synthetic signed networks converge slower. But there exists no obvious difference between SRW and RRW on random signed network.

In addition, we also test the convergence of the SRW and RRW on the real-world signed networks. For the Slovene parliamentary party network, there exists no positive link between the communities, the SRW converges more slowly than RRW, which is same as the finding in the synthetic signed networks. And for the Gahuku–Gama subtribes network, the SRW also converges more slowly compared with the RRW. Overall, there are many factors affecting the convergence of SRW, one of which are the position and link density of negative links.

5. The application to community detection in the signed networks

One of important applications for random walks is to explore the community structure of the complex networks via calculating the similarity or distance between nodes. For example, Zhou defined the distance between node i and node j as the average number of steps for a random walker moving from node i to node j [29] and then proposed a dissimilarity index used for finding communities [30]. Latapy and Pons proposed a detection algorithm called Walktrap and calculated the distance between nodes on the basis of the probability which a random walker moves from node i to node j in limited steps [19]. In this section, we combine the signed random walk with the Walktrap method (denoted by SRWT) to detect communities in the signed networks.

The procedure of SRWT is as follows:

(1) using the signed random walk to measure the similarity between pairs of nodes. We can calculate the distance between node i and node j according to the following definition:

$$r_{ij} = \sqrt{\sum_{k=1}^{N} \frac{\left(T_{ik}^{t} - T_{jk}^{t}\right)^2}{d\left(k\right)}} \tag{4}$$

where $d(k)$ is the degree or strength of node k and t is the number of transition steps. In this paper, we set $t = 3$. The probability $T_{G_\alpha i}^{t}$ that the random walker moves to node i from community G_α in t steps is defined as

$$T_{G_\alpha i}^{t} = \frac{1}{N_{G_\alpha}} \sum_{k \in G_\alpha} T_{ki}^{t} \tag{5}$$

where N_{G_α} is the number of nodes in community G_α. And then the distance between community G_i and G_j can be calculated by

$$r_{G_iG_j} = \sqrt{\sum_{k=1}^{N} \frac{\left(T_{G_ik}^{t} - T_{G_jk}^{t}\right)^2}{d\left(k\right)}} \tag{6}$$

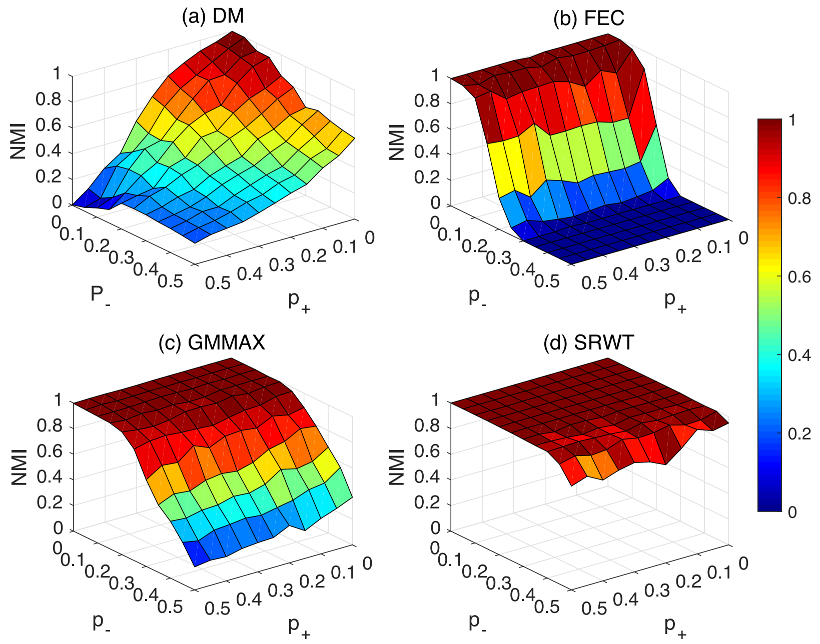


Fig. 7. (Color Online) Evaluation of detection methods on the synthetic signed networks generated by the parameters $p_{in} = 0.8$, and p_-, p_+ changing from 0 to 0.5 respectively. Each point is averaged over 20 realizations of the networks.

(2) classifying the nodes into different communities within hierarchical clustering algorithm. In the process of clustering, one can determine which two communities to merge based on Ward's method and there must exist at least one edge between the two communities merged.

(3) determining the optimal partition on the basis of modularity. In this paper, we apply the following modularity into different partitions [31]:

$$Q = \frac{2w^+}{2w^+ + 2w^-} Q^+ - \frac{2w^-}{2w^+ + 2w^-} Q^- \tag{7}$$

$$Q^+ = \frac{1}{2w^+} \sum_i \sum_j \left(w_{ij}^+ - \frac{w_i^+ w_j^+}{2w^+} \right) \delta(G_i, G_j) \tag{8}$$

$$Q^- = \frac{1}{2w^-} \sum_i \sum_j \left(w_{ij}^- - \frac{w_i^- w_j^-}{2w^-} \right) \delta(G_i, G_j) \tag{9}$$

where G_i and G_j are two communities, $\delta(G_i, G_j) = 1$ if $G_i = G_j$ and zero otherwise, w_i^+ is the positive strengths of node i, w_i^- is the negative strengths of node i, $2w^+ = \sum_i w_i^+$, $2w^- = \sum_i w_i^-$. The best partition will get the maximum of Q value compared to other partitions.

To evaluate the performance of detection algorithms applied in the signed network, we adopt the normalized mutual information (NMI) which is one of most popular measures for the similarity of partitions [32]. Note that $0 \le NMI \le 1$. If $NMI = 0$, it means the two partitions are completely different; if $NMI = 1$, it means the two partitions are exactly the same. The larger the NMI value, the more similar the two partitions are.

We perform experiments on the synthetic signed networks and two real-world signed networks to verify the effectiveness of the SRWT in community detection. We also compare SRWT method with other three classical methods used to detect communities in the signed networks which are DM algorithm [33], GMMAX algorithm [34] and FEC algorithm [23]. DM algorithm finds the optimal communities by optimizing a criterion function which is minimizing the number of negative links within communities and positive links between communities. GMMAX algorithm is through signed modularity optimization to detect communities. FEC algorithm is also a kind of random walks-based methods which the random walker only walks along the positive links. We test these methods on the signed networks which are generated by the parameters $p_{in} = 0.8$, and p_-, p_+ changing from 0 to 0.5. The results are as shown in Fig. 7. As we can see, in general, the performance of the SRWT algorithm outperforms the other three methods in community detection by comparing their NMI values and the SRWT's NMI values are always 1 except for a few parameter intervals. The DM and FEC have the worst performance among the four

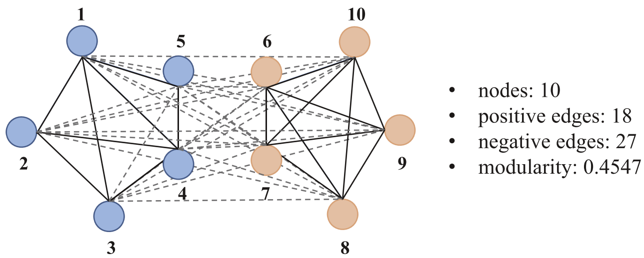


Fig. 8. The community detection of the Slovene parliamentary party network via the SRWT. The solid edges denote positive links and the dashed edges represent negative links. Each color represents different community. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

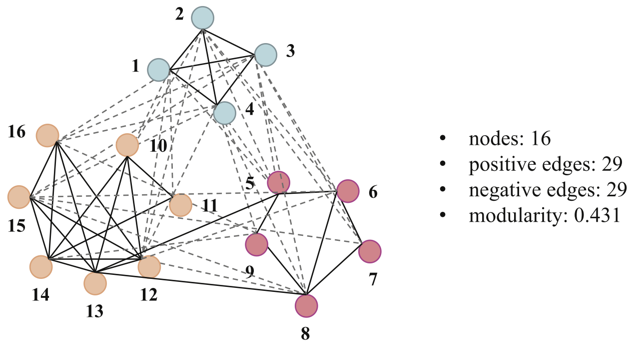


Fig. 9. The community detection of the Gahuku–Gama subtribes network via the SRWT. The solid edges denote positive links and the dashed edges represent negative links. Each color represents different community. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

methods and their NMI values changes greatly. For the FEC, it can have a better performance when $0 \le p_+ \le 0.5$ and $0 \le p_- \le 0.1$, whose NMI values are close to 1. But when $p_- > 0.1$, the performance of FEC begins to drop sharply. For the GMMAX, overall it is better than DM and FEC, but its performance begins to drop when $p_- > 0.2$.

Finally, we apply the SRWT on two real-world signed networks: the Slovene parliamentary party network and Gahuku–Gama subtribes network. These two networks are usually used as benchmarks to test community detection algorithms in the signed networks. Applying the SRWT in the Slovene parliamentary party network can detect two communities: {1, 2, 3, 4, 5} and {6, 7, 8, 9, 10} shown in Fig. 8, which is consistent with the results provided by Kropivnik and Mrvar [27]. And the corresponding modularity of this division calculated by the above equation is 0.4547. We apply the SRWT to the Gahuku–Gama subtribes network and obtain three communities: {1, 2, 3, 4}, {5, 6, 7, 8, 9}, and {10, 11, 12, 13, 14, 15, 16} shown in Fig. 9, which its corresponding modularity is 0.431. This result is identical to the Read's study on this network [28].

6. Conclusion

In this paper, we propose a signed random walk on signed networks in which the agent is allowed to walk along the negative links but with a smaller probability compared with the positive links. We focus on the two aspects of this random walk: the convergence of the transition probability matrix and one of its important application, community detection. The studies of these two aspects are both validated on the synthetic and real-world signed networks. The position and density of negative links will affect the convergence of the signed random walk. We find that the transition probability matrix of the signed random walk converges slowly compared with the regular random walk when all the positive links exist within communities in the synthetic signed networks. It reveals that the convergence of the SRW's transition probability matrix is affected by the existence of community structure in the signed networks. We also apply the SRWT method to explore the community structure of the signed networks and the simulation results show that the SRWT method is effective to detect the communities of signed network compared with several traditional methods.

Although we have proposed a signed random walk model, the influence of negative links on random walk is complex and diverse. It is necessary to design new walking rule according to the position and density of negative links in the signed network. This model also has its own limitation, for example, this model requires that each node has at least one positive link. In addition, the signed random walk model does not apply to directed signed network. In the next research work, we will consider redefining the signed random walk in order to fit directed signed networks.

Acknowledgment

This work is supported by the National Natural Science Foundation of China (Grant Nos. 61573065 and 71731002).

References

[1] S.H. Strogatz, Exploring complex networks, Nature 410 (2001) 268.
[2] R. Albert, A.L. Barabàsi, Statistical mechanics of complex networks, Rev. Modern Phys. 74 (2002) 47.
[3] M.E. Newman, The structure, and function of complex networks, SIAM Rev. 45 (2003) 167.
[4] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.U. Hwang, Complex networks: Structure, and dynamics, Phys. Rep. 424 (2006) 175.
[5] C. Parisien, C.H. Anderson, C. Eliasmith, Solving the problem of negative synaptic weights in cortical models, Neural Comput. 20 (2008) 1473.
[6] N. Girdhar, K.K. Bharadwaj, Signed social networks: A survey, in: International Conference on Advances in Computing and Data Sciences, Springer, Singapore, 2016, pp. 326–335.
[7] F. Heider, Attitudes and cognitive organization, J. Psychol. 21 (1946) 107.
[8] D. Cartwright, F. Harary, Structural balance: a generalization of Heider's theory, Psychol. Rev. 63 (1956) 277.
[9] R. Guha, R. Kumar, P. Raghavan, A. Tomkins, Propagation of trust and distrust, in: Proceedings of the 13th International Conference on World Wide Web, ACM, New York, 2004, pp. 403–412.
[10] J. Leskovec, D. Huttenlocher, J. Kleinberg, Signed networks in social media, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, Atlanta, 2010, pp. 1361–1370.
[11] J. Tang, Y. Chang, C. Aggarwal, H. Liu, A survey of signed network mining in social media, ACM Comput. Surv. 49 (2016) 42.
[12] X. Yuan, Y. Hu, H.E. Stanley, S. Havlin, Eradicating catastrophic collapse in interdependent networks via reinforced nodes, Proc. Natl. Acad. Sci. USA (2017) 201621369.
[13] N. Masuda, M.A. Porter, R. Lambiotte, Random walks and diffusion on networks, Phys. Rep. 716–717 (2017) 1.
[14] S. Brin, L. Page, The anatomy of a large-scale hypertextual Web search engine, Comput. Netw. ISDN Syst. 30 (1998) 107.
[15] L.E.C. Rocha, N. Masuda, Random walk centrality for temporal networks, New J. Phys. 16 (2014) 063023.
[16] M.E.J. Newman, A measure of betweenness centrality based on random walks, Soc. Netw. 27 (2005) 39.
[17] S. Ji, L. Lü, C.H. Yeung, Y. Hu, Effective spreading from multiple leaders identified by percolation in the susceptible–infected–recovered (SIR) model, New J. Phys. 19 (2017) 073020.
[18] H. Zhou, R. Lipowsky, Network brownian motion: A new method to measure vertex-vertex proximity and to identify communities and subcommunities, in: International Conference on Computational Science, Springer, Heidelberg, 2004, pp. 1062–1069.
[19] P. Pons, M. Latapy, Computing communities in large networks using random walks, J. Graph Algorithms Appl. 10 (2006) 191.
[20] M. Rosvall, C.T. Bergstrom, Maps of random walks on complex networks reveal community structure, Proc. Natl. Acad. Sci. USA 105 (2008) 1118.
[21] J.C. Delvenne, S.N. Yaliraki, M. Barahona, Stability of graph communities across time scales, Proc. Natl. Acad. Sci. USA 107 (2010) 12755.
[22] J. Jung, W. Jin, L. Sael, U. Kang, Personalized ranking in signed networks using signed random walk with restart, in: 2016 IEEE 16th International Conference on Data Mining, IEEE, Barcelona, 2016, pp. 973–978.
[23] B. Yang, W. Cheung, J. Liu, Community mining from signed social networks, IEEE Trans. Knowl. Data Eng. 19 (2007) 1333.
[24] H. Wang, C. Qu, C. Jiao, W. Ruszel, Self-Avoiding Pruning Random Walk on Signed Network. arXiv:1710.05744.
[25] P. Esmailian, M. Jalili, Community detection in signed networks: the role of negative ties in different scales, Sci. Rep. 5 (2015) 14339.
[26] I. Mendonça, R. Figueiredo, V. Labatut, P. Michelon, Relevance of negative links in graph partitioning: A case study using votes from the European parliament, in: 2nd European Network Intelligence Conference, ENIC, IEEE, Karlskrona, 2015, pp. 122–129.
[27] S. Kropivnik, A. Mrvar, An analysis of the slovene parliamentary parties network, Dev. Stat. Methodol. 209 (1996).
[28] K.E. Read, Cultures of the central highlands, New Guinea, Southwest. J. Anthropol. 10 (1954) 1.
[29] H. Zhou, Network landscape from a Brownian particle's perspective, Phys. Rev. E 67 (2003) 041908.
[30] H. Zhou, Distance, dissimilarity index, and network community structure, Phys. Rev. E 67 (2003) 061901.
[31] S. Gómez, P. Jensen, A. Arenas, Analysis of community structure in networks of correlated data, Phys. Rev. E 80 (2009) 016114.
[32] L.N.F. Ana, A.K. Jain, Robust data clustering, in: Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE, Madison, 2003, pp. 128–133.
[33] P. Doreian, A. Mrvar, A partitioning approach to structural balance, Soc. Netw. 18 (1996) 149.
[34] J.Q. Jiang, Stochastic block model and exploratory analysis in signed networks, Phys. Rev. E 91 (2015) 062805.