



Identifying Propagation Source in Temporal Networks Based on Label Propagation

Lilin Fan^{1,2}, Bingjie Li^{1,2}, Dong Liu^{1,2(✉)}, Huanhuan Dai^{1,2},
and Yan Ru^{1,2}

¹ School of Computer and Information Engineering, Henan Normal University,
Xinxiang, Henan, China

liudong@htu.cn

² Big Data Engineering Laboratory for Teaching Resources and Assessment
of Education Quality, Xinxiang, Henan, China

Abstract. The spread of rumors and diseases threatens the development of society, it is of great practical significance to locate propagation source quickly and accurately when rumors or epidemic outbreaks occur. However, the topological structure of online social network changes with time, which makes it very difficult to locate the propagation source. There are few studies focus on propagation source identification in dynamic networks. However, it is usually necessary to know the propagation model in advance. In this paper the label propagation algorithm is proposed to locate propagation source in temporal network. Then the propagation source was identified by hierarchical processing of dynamic networks and label propagation backwards without any underlying information dissemination model. Different propagation models were applied for comparative experiments on static and dynamic networks. Experimental results verify the effectiveness of the algorithm on temporal networks.

Keywords: Source identification · Propagation model · Label propagation

1 Introduction

With the rapid development of social networks represented by Facebook, twitter, microblog and We-Chat, social network has become a very important information dissemination platform. At the same time users share information, they also have to face the adverse effects of rumors and other harmful information. In addition, respiratory infectious diseases such as tuberculosis and 2019-nCoV are highly infectious and seriously threaten human health [1–3]. Therefore, it is of great practical significance to locate source quickly and accurately when rumors or epidemic outbreaks occur.

The spread of rumors or diseases depends on specific networks. For example, the spread of rumors depends on social networks, and the spread of diseases depends on human interaction networks. If the infectious source of the disease is determined in a short period of time, it can better control the spread of the whole disease, narrow its transmission range, and reduce the threat of infectious diseases to people's life safety. If we can quickly locate the source of rumor in the critical period after spreading rumor, we can reduce the negative impact of rumor on politics, people's vital interests and

social stability. Generally, these networks have the characteristics of large scale and complex connections between nodes. Therefore, how to locate a source of infection or rumor in network has become a very meaningful and challenging problem.

The methods of propagation source location in static network are mainly divided into four categories: Propagation source centrality algorithm [4, 5, 9, 10], Belief propagation algorithm [11, 12], Monte Carlo algorithm [6], and Net-Sleuth algorithm [6]. The basic idea of propagation source centrality algorithm is directly obtaining central nodes as infection propagation source; Belief propagation algorithm identify the propagation source through message-passing on a factor graph; Monte Carlo algorithm identifies sources based on simulating the spread of disease; And the feature of the Net-Sleuth algorithm based on the minimum description length is that it does not need the number of sources. The above algorithms are all directed to solve the problem of identifying propagation source in static network. However, most real networks are dynamic and temporal.

In recent years, researchers have begun to focus on the problem of identifying propagation source in dynamic network. The algorithm based on the shortest temporal path effectively identifies the propagation source in temporal network [14]; A method of locating propagation sources based on reverse dissemination strategy was proposed by Jiang et al. [7]; The algorithm based on the time information of all nodes by Antulov-fantulin et al. [6]; and sparse signal reconstruction theories has been applied to locate propagation sources in dynamic network [19]. Although the above algorithm effectively solves the problem of source identification on dynamic networks, the precondition is to know the underlying propagation model in advance. Obviously, it is difficult to obtain model information in practice, and it is hard to obtain the true value of parameters in the preset basic propagation model.

As far as we know, only a few source identification studies have been carried out without any underlying information propagation model. The label propagation algorithm (LPSI) [8] is the first attempt to solve this problem, which is based on the idea of source prominence, namely the nodes surrounded by larger proportions of infected nodes are more likely to be source ones. The main idea of this algorithm is to identify the propagation source by propagating integer labels in the network. It effectively solves the problem that the propagation model cannot be known in advance in practice. However, it is only studied on static networks.

In this paper, we uses the label propagation algorithm to locate the source on the temporal network, the propagation source is identified by hierarchical processing of dynamic networks and iterative propagation of the label. Firstly, dividing dynamic network into multiple single layer static networks. Secondly, label propagation backwards from the last layer of the network to the first layer. Finally, the propagation source is derived based on the node labels of the first layer network. The major contributions of this paper can be summarized as follows:

- use the label propagation algorithm to locate the propagation source on the temporal network;
- identify propagation sources without any underlying propagation model information;
- comparative experiments on static and dynamic networks using different propagation models

The rest of this paper is organized as follows. In Sect. 2, we introduce the related work of propagation source identification in static and temporal networks. Section 3 gives the introduction of label propagation method. Section 4 analyzes the algorithm in detail. Section 5 compares the accuracy of locating propagation source on static network and temporal network under different propagation models. Section 6 summarizes this paper.

2 Related Work

In the past few years, researchers propose a series of methods to identify propagation sources in network. According to the time state of the network, they are divided into propagation source identification methods in static and dynamic networks.

For static networks, Shah et al. [4] put forward the first algorithm to solve the source identification problem for the SI propagation model on the regular tree network. Luo et al. [5] used Jordan center estimation (JCE) algorithm to solve the problem of single source estimation on tree network through Jordan infection center. Fioriti et al. [9] proposed a propagation source identification algorithm called DA by calculating the dynamic age of each node in the infection graph. Comin et al. [10] studied an unbiased betweenness algorithm for computing the centrality of unbiased mediations of all nodes in an infection graph. Lokhov et al. [11] established a dynamic message-passing algorithm for single source identification. Altarelli et al. [12] put forward a belief propagation algorithm (BP) based on factor graph on SIR propagation model. Antulov-fantulin et al. [6] proposed a source location algorithm based on Monte Carlo with single source and known states of all nodes. Prakash et al. [6, 13] recommended the Net-Sleuth algorithm based on the minimum description length in the source identification problem of multi-source and SI propagation model. Wang et al. [8] proposed a source identification method based on label propagation, which is the first attempt to solve the unknown problem of the propagation model in reality.

For temporal networks, Antulov-fantulin et al. [6] proposed a method to locate the source based on the time information of all nodes. Jiang et al. [7] established a method of locating propagation sources based on reverse dissemination strategy. Huang et al. [14] developed a reverse diffusion method to locate the propagation source by calculating the shortest time distance between nodes. Hu et al. [19] a general framework locate the source by bridging structural observability and sparse signal reconstruction theories.

The research goal of this paper is to locate the propagation sources on the dynamic network. However, the current methods of source identification on the dynamic network depend on the specific propagation model, which is not in line with the characteristics of real dissemination. Therefore, this paper presents a dynamic network source identification method based on label propagation. Firstly, dividing dynamic network into multiple single layer static networks. Secondly, label propagation backwards from the last layer of the network to the first layer. Finally, the propagation source is derived based on the node labels of the first layer network.

3 Methodology

The problem statement and algorithm description are shown in this section. Frequently used notations are summarized in Table 1.

Table 1. Notation summarization.

Notation	Definition
G	The topological graph of temporal network
N	Number of nodes
L	Number of network layers
$T = (t_1, t_2, \dots, t_L)$	Time series of temporal network G
$V = (v_1, v_2, \dots, v_N)$	Set of vertexes in G
$E = (E_1, E_2, \dots, E_L)$	Set of edge sets of L -layer temporal network
$Y = (Y_1, Y_2, \dots, Y_L)$	Set of infection node vectors of L -layer temporal network
$Y_m = (y_1^m, y_2^m, \dots, y_N^m)$	Infection node vector at layer m in G , $m \in \{1, 2, \dots, L\}$
A^m	The adjacency matrix at layer m in G , $m \in \{1, 2, \dots, L\}$
P^m	The label propagation probability at layer m in G , $m \in \{1, 2, \dots, L\}$
$g_i^t\{m\}$	i -th node label value of layer m at iteration t , $m \in \{1, 2, \dots, L\}$
S_m	Source estimate set at layer m in G , $m \in \{1, 2, \dots, L\}$
s^*	The propagate source

3.1 Problem Statement

The problem of dynamic network source identification studied in this paper can be formulated as follows: Given a dynamic network $G = (V, E, T)$, According to the time series $T = (t_1, t_2, \dots, t_L)$, the dynamic network is divided into m -layer static network. A node set $V = (v_1, v_2, \dots, v_N)$ with a set E_m at layer $m \in \{1, 2, \dots, L\}$. An infection node vector $Y_m = (y_1^m, y_2^m, \dots, y_N^m)$ at layer m where $y_i^m = 1$ indicates node i is infected and $y_i^m = -1$ otherwise, the goal is to find the original infection source $s^* \subset V$. Because the algorithm used in this paper does not depend on the specific propagation model, it is not necessary to make assumptions on the propagation model.

3.2 Label Propagation Algorithm in Temporal Network

We formally present our method in Algorithm 1. One of the important characteristics of the label propagation algorithm is that it can identify the source independent of the propagation model. In this paper, the label propagation algorithm is used to identify the propagation source in dynamic network, which can be realized through the following three steps.

Assign Labels. We divide the dynamic network into L -layer static network. Known the infection node vector $Y = (Y_1, Y_2, \dots, Y_L)$ when propagating to the last layer network. Notably, the underlying information propagation model for generating Y_L is

unknown in our method. Based on the infected node vector Y_L to assign positive label (+1) and negative label (-1) to infected and uninfected nodes in the last layer network, and put the assigned label into $g_i^0\{L\}$.

Label Propagation. We need to obtain the adjacency matrix A^m at layer m , where $A_{ij}^m = 1$ if there is a connection between nodes i and j , otherwise $A_{ij}^m = 0$. Generating the label propagation probability P^m at layer m from node j to node i based on adjacency matrix. We start from the last network and reversely execute the label propagation algorithm. For each layer network, each node obtains a part of label information from its neighborhood, and retains some label information in its initial state. Therefore, the label value of node i of layer m at iteration $t + 1$ becomes:

$$g_i^{t+1}\{m\} = a \sum_{j:j \in N(i)} P_{ij}^m g_j^t\{m\} + (1 - a)y_i^m \quad (1)$$

where matrix $P^m = D^{(-1/2)}A^mD^{(-1/2)}$, matrix A^m is the adjacency matrix, D is a diagonal matrix with its (i, i) -element equal to the sum of the i -th row of A , and $0 < a < 1$ is the fraction of label information that node i gets from its neighbors. $N(i)$ represents the set of neighbors of node i .

In addition to the first layer network, each layer carries out label propagation with iteration $t = 1$. For the label of nodes in $(m - 1)$ -layer network, if node i belongs to the set of source estimation obtained from m -layer network, assign the positive label (+1), it inherits the label of m -layer network otherwise. Namely,

$$g_i^0\{m - 1\} = \begin{cases} 1, & i \in S_m \\ g_i^0\{m\}, & otherwise \end{cases} \quad (2)$$

For each layer network, it is identified as a suspicious propagation source if a node i meets the following two conditions:

- Node i is initially an infected node;
- The final label value of node i is larger than those of its neighbors.

The first condition of the algorithm is to avoid the interference of uninfected nodes on the accuracy of source identification. The second condition is based on the principle of source prominence [10], namely the nodes surrounded by larger proportions of infected nodes are more likely to be propagation source.

Source Identification. When the label propagates back to the first layer network, the iterative propagation stops until the node label reaches convergence by Eq. 1. Here, "convergence" means that the label value of a node does not change in several consecutive iterations of the label propagation. It is worth mentioning that for a single-layer static network, only the node labels need to be iterated to convergence.

Finally, obtain the node with the highest label value from the suspect node as the propagation source s^* .

Algorithm 1 Input Generation Algorithm

Input: The infected temporal network $G = (V, E, T)$, parameter a ; The initial infection node vector Y_L in the last layer network.

Output: The propagation source s^*

```

1: for each layer network do
2:   Form the weight matrix  $A^m$  defined by  $A_{ij}^m = 1$  if there exists an edge connect-
   ing nodes  $i$  and  $j$ ;
3:   Construct the matrix  $P^m = D^{(-1/2)}A^mD^{(-1/2)}$ , where  $D$  is a diagonal matrix with
   its  $(i, i)$ -element equal to the sum of the  $i$ -th row of  $A_m$ ;
4:   endfor
5:    $g_i^0\{L\} \leftarrow Y_L$ 
6:   for  $m$  from  $L$  to  $1$  do
7:     if  $m \neq 1$  then
8:       for each node  $i$  do
9:          $g_i^1\{m\} = a \sum_{j:j \in N(i)} P_{ij}^m g_j^0\{m\} + (1 - a)y_i^m$ 
10:        endfor
11:       for each node  $i$  do
12:         if  $g_i^1 >$  all  $i$ 's neighbors'  $g_i^1$  then
13:            $S_m = S_m \cup i$ 
14:         endif
15:       endfor
16:        $g_i^0\{m - 1\} = \begin{cases} 1, & i \in S_m \\ g_i^0\{m\}, & \text{otherwise} \end{cases}$ 
17:     else
18:       while  $g_i^t$  dose not reach the convergence  $g^*$  do
19:         for each node  $i$  do
20:            $g_i^{t+1}\{m\} = a \sum_{j:j \in N(i)} P_{ij}^m g_j^t\{m\} + (1 - a)y_i^m$ 
21:         endfor
22:          $t = t + 1$ 
23:       endwhile
24:     endif
25:   endfor
26:   for the first layer network ( $m=1$ ) do
27:     for each node  $i$  do
28:       if  $g_i^t\{1\} >$  all  $i$ 's neighbors'  $g_i^t$  then
29:          $S_1 = S_1 \cup i$ 
30:       endif
31:     endfor
32:   endfor
33:   select node  $i$  with  $\max(g_{i \in S_1}^t\{1\})$ 
34:    $s^* \leftarrow i$ 
35:   return  $s^*$ 

```

3.3 Algorithm Diagram

In order to further understand the algorithm in the section, Fig. 1 shows the process of propagation source identification in multi-layer dynamic network. The data set used in the example is a L-layer small world network with $N = 10$ nodes and average degree of $\langle k \rangle = 4$. Randomly select a source node to propagate through the SI model. We divide the dynamic network into L-layer static networks by chronological order. As shown in Algorithm 1, the green node is the suspect sources of each layer, the white point is the uninfected nodes and the orange one is the propagation source node. First, a set of source estimation nodes is obtained by Eq. 1 on the last layer network. Then, transfer labels to node i between layers in reverse by Eq. 2 and label propagation on the new network layer. Finally, in the first layer network, the label propagation stops until the label converges. The node with the largest label value is output as the propagation source.

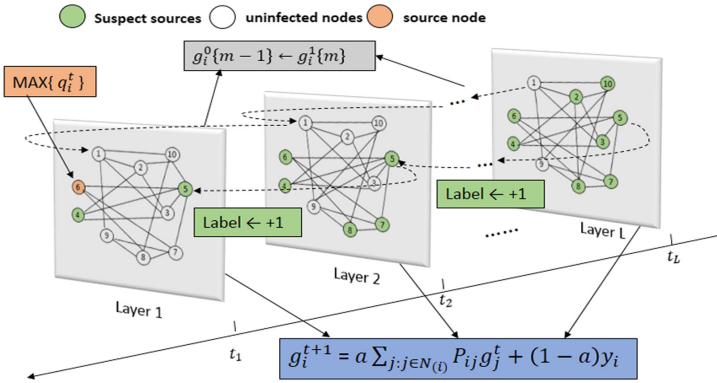


Fig. 1. The process of propagation source identification in multi-layer dynamic network.

4 Algorithm Analysis

Theoretical analysis [8] and time complexity analysis of the algorithm are shown in this section.

4.1 Convergence Analysis

The equation for calculating the label of each node at iteration t can be transformed into another writing method:

$$g^{t+1} = aPg^t + (1-a)y \tag{3}$$

By the initial condition that $g^0 = y$, we get:

$$g^t = (aP)^t + (1 - a) \sum_{i=0}^{t-1} (aP)^i y \quad (4)$$

According to $\lim_{t \rightarrow \infty} (aP)^t = 0$ and $\lim_{t \rightarrow \infty} \sum_{i=0}^{t-1} (aP)^i = (I - aP)^{-1}$, where I is an $n \times n$ identity matrix. Therefore, the label value will converge to:

$$g^* = (1 - a)(I - aP)^{-1}y \quad (5)$$

It can be seen that the label value will reach convergence after iterative propagation finally.

4.2 Source Prominence

Source Prominence, namely the node surrounded by larger proportions of infected nodes are more likely to be source ones. We can understand this idea in two ways, On the one hand, at the edge of the infected area, nodes tend to have fewer infected neighbors. On the other hand, in the center of the infected area, nodes tend to have more infected neighbors. This property has nothing to do with the propagation model, so we can apply this idea to source identification. The experimental results of this paper also verify this idea.

4.3 Time Complexity

For a single-layer network label propagation, the time complexity is $O(t * J * N)$, where J is the average number of neighbors per node and t is the number of iteration. Therefore, For a dynamic network, the time complexity is $O(L * t * J * N)$, where L is the number of network layer.

5 Experiments

5.1 Experiments Setup

Datasets. In this paper, two kinds of networks are selected for experiments, the first type are small-world networks, random networks, and scale-free networks. Number of nodes $N = 100, 500$, average degree $\langle k \rangle = 6, 8, 10$. The second type is the benchmark adopted by Lin et al. in [15], this data set [16] is a dynamic network of a fixed number of communities (named SYN-FIX), analogously to the classical benchmark proposed by Girvan and Newman in [17]. The network consists of 128 nodes divided into four communities of 32 nodes each.

Propagation Models. Existing propagation models can be divided into infection models and influence models. In order to verify that the algorithm can locate rumor sources without any underlying propagation model information, the experimental part of this paper mainly uses three models, namely SI model, SIR model and LT model [20].

For SI model, as shown in Fig. 2(a), each node has only two possible states: S and I. At the initial time of propagation, only one or a few nodes in the network are in state I, while other nodes are in state S. At each time step of the propagation process, any nodes in state I infects each neighbor in state S with the same probability p .

For SIR model, as shown in Fig. 2(b), each node has three possible states: susceptible (S), infected (I), and recovered (R). Every infected node tries to infect its susceptible (uninfected) neighbors independently with probability p . The infected node changes from probability q to R state, and the node will always be in R state, and will not be infected. The infection probability p is chosen uniformly from (0, 1) for the SI model and SIR model. For the SIR model, an extra recovery probability q is chosen uniformly from (0, p).

For LT model, there are only two possible states for each node: active state and inactive state, in which the active state represents that a node receives information in the propagation process, otherwise it is in the inactive state. As shown in Fig. 2(c), each node in the inactive state receives information only when the number of nodes in the active state around it exceeds a certain threshold. Until the sum of the influence of any active node in the network cannot activate the neighbor node in the inactive state, the propagation process ends. Assuming that nodes u, v have degrees d_u and d_v , the infection weight between edges (u, v) is $1/d_v$, and the infect weight between edges (v, u) is $1/d_u$. In addition, the threshold of each node is uniformly selected from a small interval $[0, 0.5]$, in order to infect most of the network.

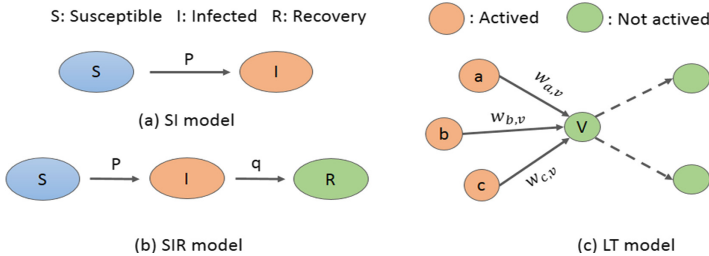


Fig. 2. Propagation Models. (a) SI model with the infection probability p ; (b) SIR model with the infection probability p and the recovery probability q ; (c) LT Models with a threshold θ .

Evaluation Metrics. We use F-score to evaluate the performance of algorithm.

F-score is a suitable metric for measuring multi-labeled classification problems, since it is a combination of precision and recall. The definition of F-score is following:

$$F - \text{score} = \frac{(1 + b^2) * Precision * Recall}{(b^2 * Precision) + Recall} \tag{6}$$

In this article, setting parameter b to 1 means that accuracy and recall are treated equally.

5.2 Results

In order to verify the effectiveness of the algorithm in locating propagation sources on dynamic networks, 100 independent experiments are carried out on four kinds of networks. We set the parameter $a = 0.5$. In each simulation, randomly select a node as the propagation source, and then use SI model, SIR model and LT model to simulate the propagation. In different propagation models, this paper studies the comparison of F-score between dynamic network and static network under different average degree ($\langle k \rangle = 6, 8, 10$) and different network layers ($L = 4, 6, 8$). In order to verify the effectiveness of the algorithm when the network scale is expanded, we increase the number of nodes to 500 for simulation experiments. In addition, the parameter a in Eq. 1 controls the influence of neighbor nodes in the process of label propagation. Therefore, this paper studies the optimal value of a by analyzing the influence of parameter a on the experimental results.

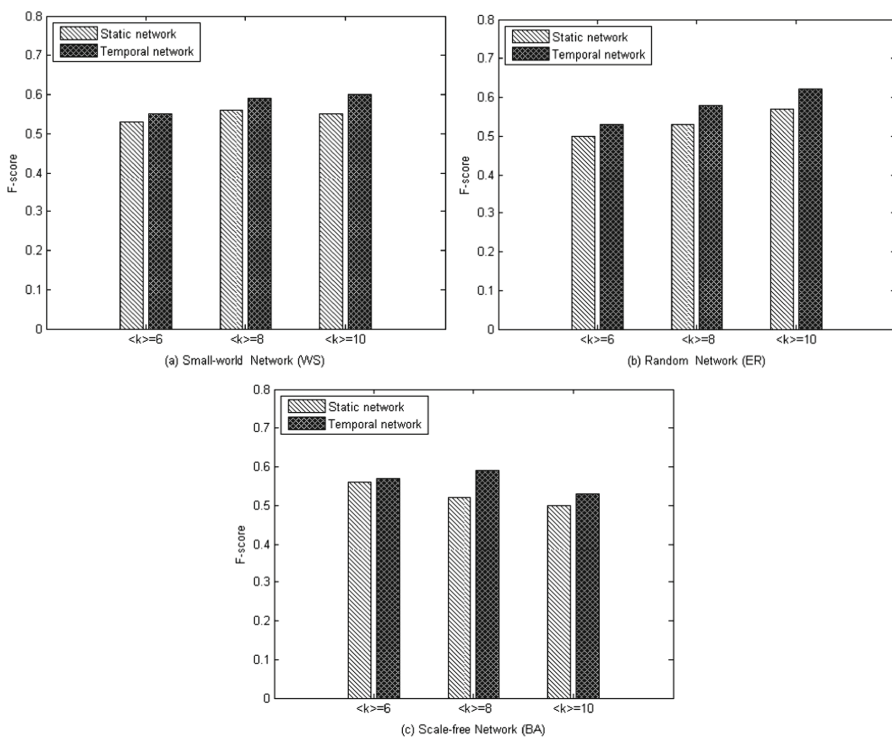


Fig. 3. The comparison of F-score of different average degrees ($\langle k \rangle = 6, 8, 10$) in dynamic networks and static networks under SI model. (a) Small-world network; (b) Random network; (c) Scale-free network.

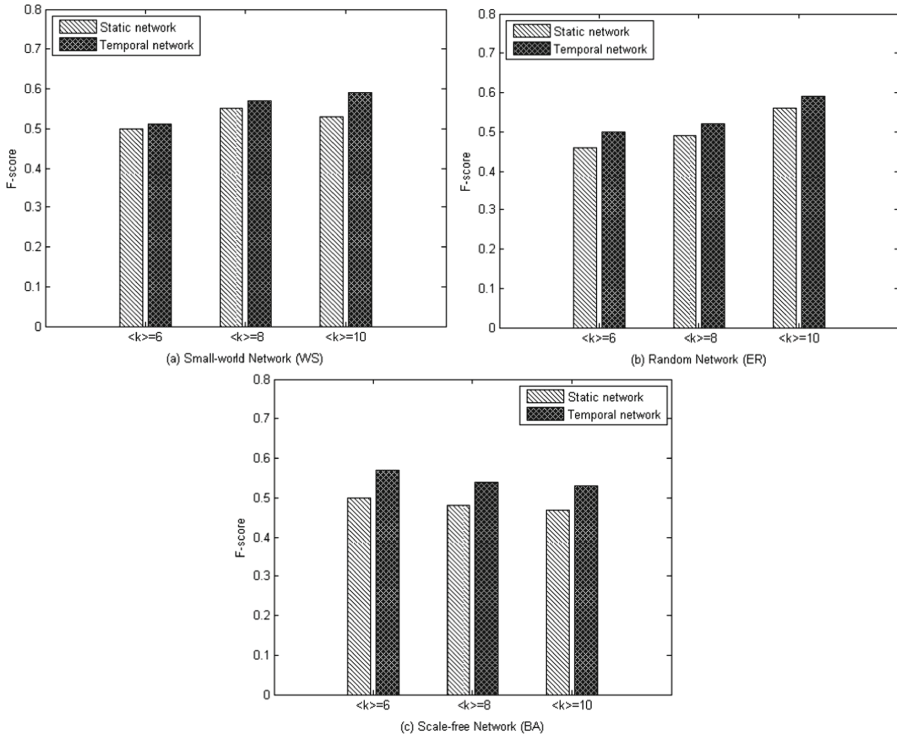


Fig. 4. The comparison of F-score of different average degrees ($\langle k \rangle = 6, 8, 10$) in dynamic networks and static networks under SIR model. (a) Small-world network; (b) Random network; (c) Scale-free network.

The comparison of F-score of different degrees of dynamic network and static network source in SI model, SIR model and LT model can be seen in Fig. 3, Fig. 4, Fig. 5. Figure 6 shows the comparison of F-score of temporal network and static network source under different propagation models for the data set called SYN-FIX. It can be seen that the location accuracy of dynamic source location algorithm is significantly higher than that of static location algorithm. The result proves the effectiveness of the algorithm to identify the propagation source on a temporal network.

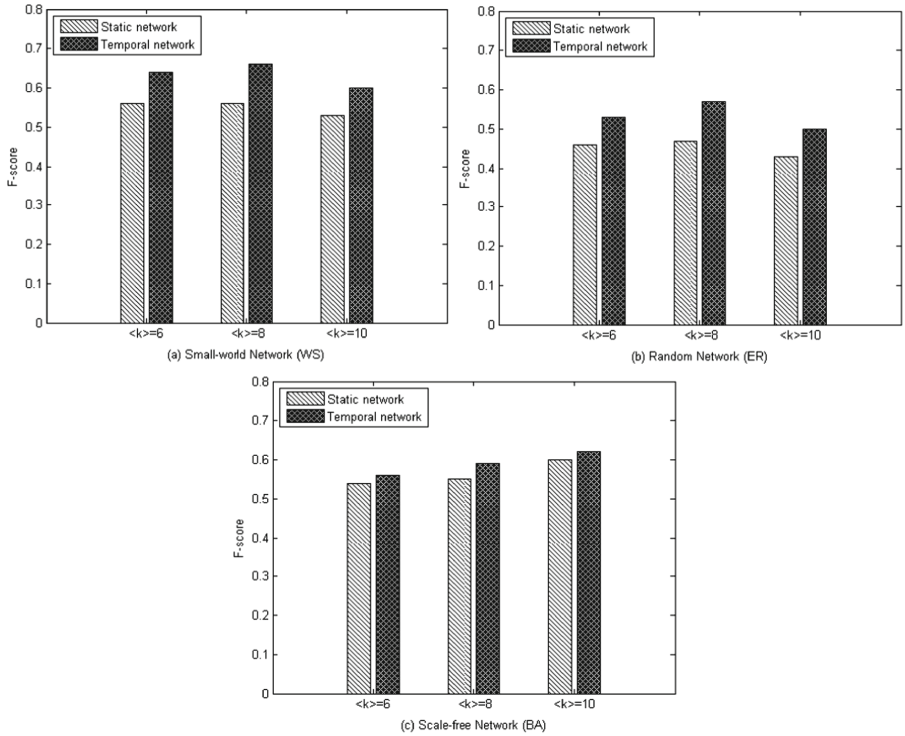


Fig. 5. The comparison of F-score of different average degrees ($\langle k \rangle = 6, 8, 10$) in dynamic networks and static networks under LT model. (a) Small-world network; (b) Random network; (c) Scale-free network.

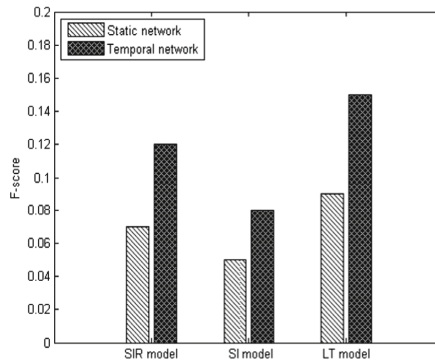


Fig. 6. The comparison of F-score of different propagation models in dynamic network and static network for SYN-FIX.

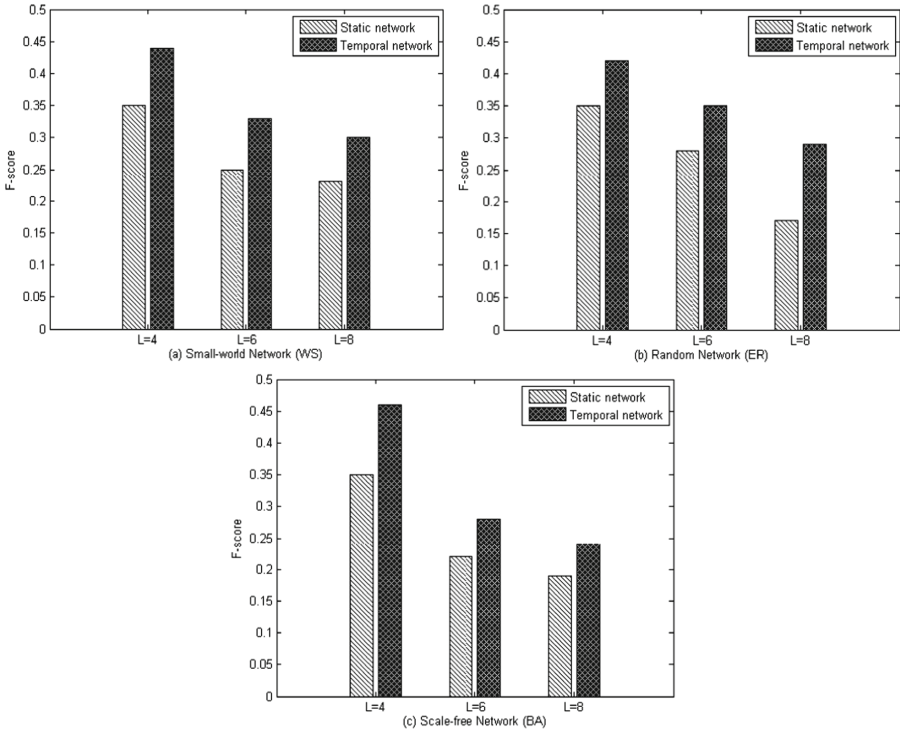


Fig. 7. The comparison of F-score of different network layers ($L = 4, 6, 8$) in dynamic networks and static networks under SI model. (a) Small-world network; (b) Random network; (c) Scale-free network.

Figure 7, Fig. 8, Fig. 9 shows the comparison of F-score of different network layers of dynamic network and static network source under SI model, SIR model and LT model respectively. The location probability of dynamic source location algorithm is significantly higher than that of static location algorithm. In addition, it can be seen that the accuracy of source location decreases with the increase of network layers. The main reason is that when the transmission time is long, information is widely spread in the network, and the topology will change greatly, which will reduce the accuracy.

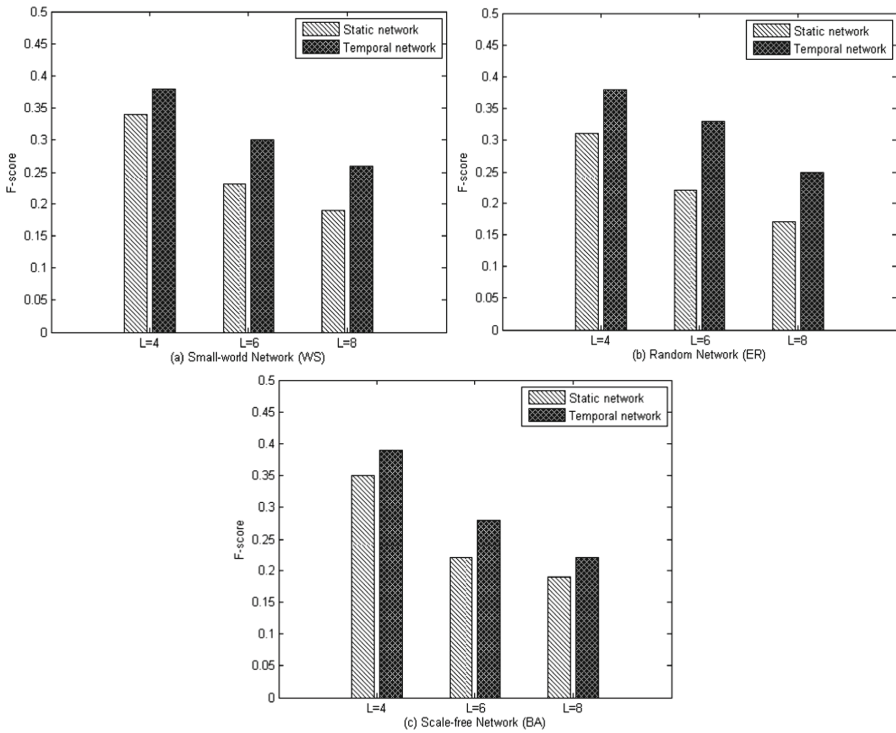


Fig. 8. The comparison of F-score of different network layers ($L = 4, 6, 8$) in dynamic networks and static networks under SIR model. (a) Small-world network; (b) Random network; (c) Scale-free network.

Figure 10 shows the comparison of F-score of different propagation models of dynamic networks and static networks with $N = 500$. The results show that when the network scale is expanded, the algorithm is still applicable.

Figure 11 is the result of the change of F-score on small world network with parameter a under SI model, SIR model and LT model. It is observed that the F-score decreases when parameter a approaches 0 or 1, regardless of the propagation model setting. On the other hand, the performances with $a \in [0.2, 0.6]$ are always stable and preferred. The result confirm with the intuition that we should consider both the initial infection status and the effects from neighbors for source detection.

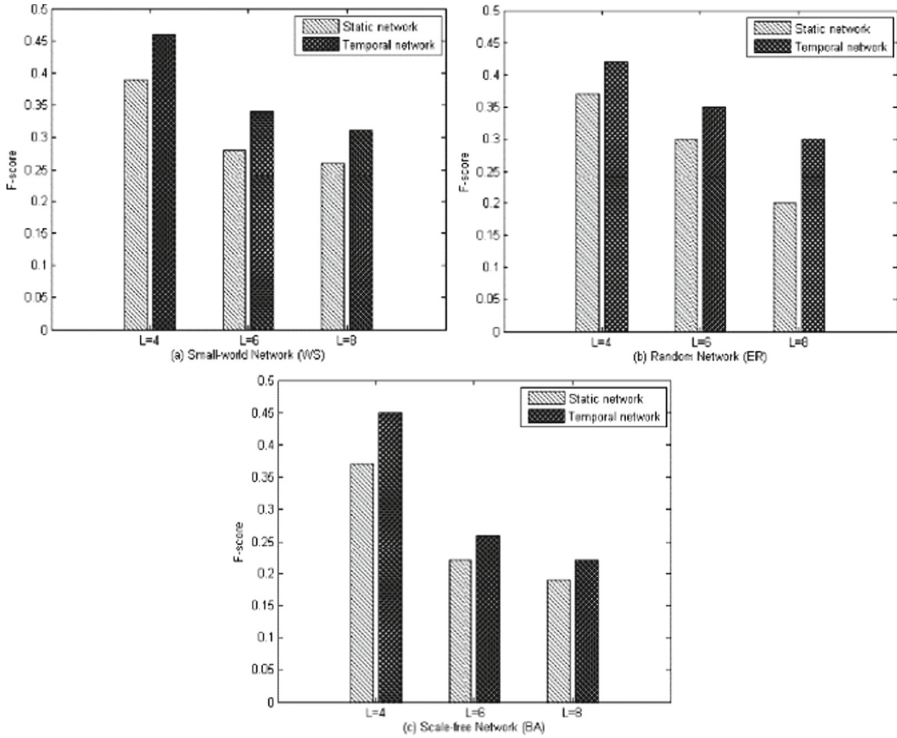


Fig. 9. The comparison of F-score of different network layers ($L = 4, 6, 8$) in dynamic networks and static networks under LT model. (a) Small-world network; (b) Random network; (c) Scale-free network.

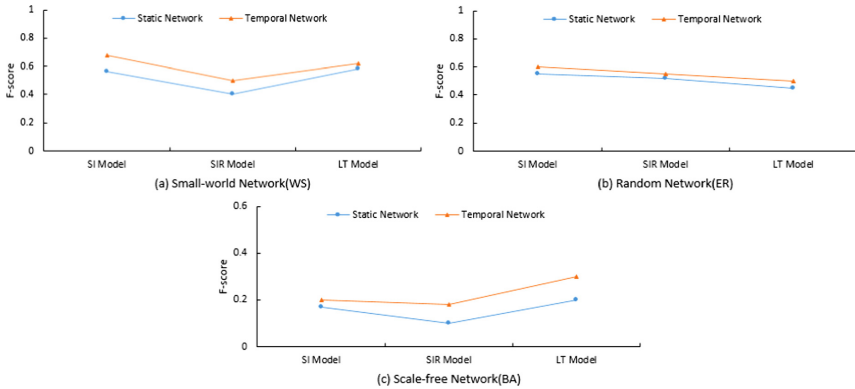


Fig. 10. The comparison of F-score of different propagation models of dynamic network and static network source with $N = 500$.

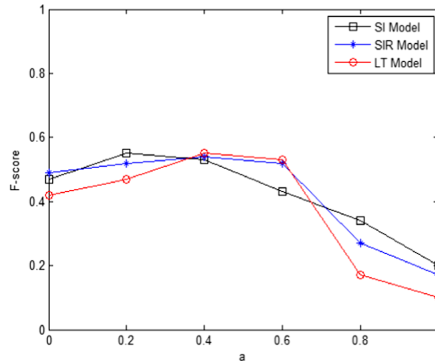


Fig. 11. The change of F-score with parameter a.

6 Conclusion

Locating the source in temporal network is a practical research issue. In this paper, label propagation algorithm is used to locate the source in dynamic network, and the simulation results on static and temporal network are given under SI model, SIR model and LT model. The experimental results show that the location accuracy of the temporal network is higher than that of the static network without considering the dynamic topology changes. In addition, the number of layers of temporal network has a certain impact on the accuracy of source location.

The algorithm in this paper has a premise that dynamic social networks only consider edge changes. However, for the online social network represented by microblog, it is in line with the actual situation in a short time. For the case of node and edge changing at the same time, it will be discussed in the future work.

References

1. Williams, B.G., Granich, R., Chauhan, L.S., et al.: The impact of HIV/AIDS on the control of tuberculosis in India. *Proc. Nat. Acad. Sci. U.S.A.* **102**(27), 9619–9624 (2005)
2. Smith, R.D.: Responding to global infectious disease outbreaks: lessons from SARS on the role of risk perception, communication and management. *Soc. Sci. Med.* **63**(12), 3113–3123 (2006)
3. World Health Organization: Global tuberculosis report 2013. World Health Organization, Geneva, Switzerland (2013)
4. Shah, D., Zaman, T.: Detecting sources of computer viruses in networks: theory and experiment. In: *Proceedings of the ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems*, New York, USA, pp. 203–214 (2010)
5. Luo, W., Tay, W.P., Leng, M.: How to identify an infection source with limited observations. *IEEE J. Sel. Top. Sign. Process.* **8**(4), 586–597 (2014)
6. Antulovfantulin, N., Lancic, A., Šmuc, T., Štefančić, H., Šikić, M.: Identification of patient zero in static and temporal networks: robustness and limitations. *Phys. Rev. Lett.* **114**(24), 248701 (2015)

7. Jiang, J., Wen, S., Yu, S., et al.: Rumor source identification in social networks with time-varying topology. *IEEE Trans. Dependable Secure Comput.* **99**, 1 (2016)
8. Wang, Z., Wang, C., Pei, J., Ye, X.: Multiple source detection without knowing the underlying propagation model. In: *AAAI*. AAAI Press, pp. 217–223 (2017)
9. Fioriti, V., Chinnici, M.: Predicting the sources of an outbreak with a spectral technique. arXiv preprint [arXiv:1211.2333](https://arxiv.org/abs/1211.2333) (2012)
10. Comin, C.H., Da Fontoura, C.L.: Identifying the starting point of a spreading process in complex networks. *Phys. Rev. E* **84**(5), 56105 (2011)
11. Lokhov, A.Y., Mzard, M., Ohta, H., et al.: Inferring the origin of an epidemic with a dynamic message-passing algorithm. *Phys. Rev. E* **90**(1), 12801 (2014)
12. Altarelli, F., Braunstein, A., Dall Asta, L., et al.: Bayesian inference of epidemics on networks via belief propagation. *Phys. Rev. Lett.* **112**(11), 118701 (2014)
13. Prakash, B.A., Vreeken, J., Faloutsos, C.: Spotting culprits in epidemics: how many and which ones? In: *IEEE 12th International Conference on Data Mining (ICDM)*, Brussels, Belgium, vol. 2012, pp. 11–20 (2012)
14. Huang, Q.: Source locating of spreading dynamics in temporal networks. In: *The 26th International Conference. International World Wide Web Conferences Steering Committee* (2017)
15. Lin, Y.-R., Zhu, S., Sundaram, H., Tseng, B.L.: Analyzing communities and their evolutions in dynamic social networks. *ACM Trans. Knowl. Discov. Data* **3**(2), 18 (2009)
16. Folino, F., Pizzuti, C.: An evolutionary multiobjective approach for community discovery in dynamic networks. *IEEE Trans. Knowl. Data Eng.* **26**(8), 1838–1852 (2014)
17. Girvan, M., Newman, M.E.J.: Community structure in social and biological networks. *Proc. Nat. Acad. Sci. U.S.A.* **99**, 7821–7826 (2002)
18. Zhou, D., Bousquet, O., Lal, T.N., Weston, J., Scholkopf, B.: Learning with local and global consistency. *Adv. Neural Inform. Process. Syst.* **16**(16), 321–328 (2004)
19. Hu, Z.L., Shen, Z., Cao, S., et al.: Locating multiple diffusion sources in time varying networks from sparse observations. *Sci. Rep.* **8**(1) (2018)
20. Dong, M., Zheng, B., Hung, N., Su, H., Guohui, L.: Multiple rumor source detection with graph convolutional networks. 569–578 (2019). <https://doi.org/10.1145/3357384.3357994>