

Shortest path of temporal networks: An information spreading-based approach

Yixin Ma(马一心), Xiaoyu Xue(薛潇雨), Meng Cai(蔡萌), and Wei Wang(王伟)
Citation: Chin. Phys. B, 2020, 29 (**12**): 128902. DOI: 10.1088/1674-1056/abb302
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Influential nodes identification in complex networks based on global and local information

Yuan-Zhi Yang(杨远志), Min Hu(胡敏), Tai-Yu Huang(黄泰愚)
Chin. Phys. B, 2020, 29 (**8**): 088903. DOI: 10.1088/1674-1056/ab969f

Generation of optimal persistent formations for heterogeneous multi-agent systems with a leader constraint

Guo-Qiang Wang(王国强), He Luo(罗贺), Xiao-Xuan Hu(胡笑旋)
Chin. Phys. B, 2018, 27 (**2**): 028901. DOI: 10.1088/1674-1056/27/2/028901

Optimized routing strategy for complex network with multiple priorities

Shi-Bao Li(李世宝), Zong-Xing Sun(孙宗星), Jian-Hang Liu(刘建航), Hai-Hua Chen(陈海华)
Chin. Phys. B, 2016, 25 (**8**): 088902. DOI: 10.1088/1674-1056/25/8/088902

Improved routing strategy based on gravitational field theory

Song Hai-Quan, Guo Jin
Chin. Phys. B, 2015, 24 (**10**): 108901. DOI: 10.1088/1674-1056/24/10/108901

Traffic resource allocation for complex networks

Ling Xiang, Hu Mao-Bin, Long Jian-Cheng, Ding Jian-Xun, Shi Qin
Chin. Phys. B, 2013, 22 (**1**): 018904. DOI: 10.1088/1674-1056/22/1/018904

Shortest path of temporal networks: An information spreading-based approach*

Yixin Ma(马一心)[1,2], Xiaoyu Xue(薛潇雨)[1,2], Meng Cai(蔡萌)[3], and Wei Wang(王伟)[2,†]

[1] College of Cybersecurity, Sichuan University, Chengdu 610065, China
[2] Cybersecurity Research Institute, Sichuan University, Chengdu 610065, China
[3] School of Humanities and Social Sciences, Xi'an Jiaotong University, Xi'an 710049, China

(Received 14 May 2020; revised manuscript received 20 July 2020; accepted manuscript online 27 August 2020)

The shortest path is a widely studied network science problem and has attracted great attention. Nevertheless, it draws little attention in temporal networks, in which temporal edges determine information dissemination. In this paper, we propose an information spreading-based method to calculate the shortest paths distribution in temporal networks. We verify our method on both artificial and real-world temporal networks and obtain a good agreement. We further generalize our method to identify influential nodes and found an effective method. Finally, we verify the influential nodes identifying method on four networks.

Keywords: temporal network, shortest path, information spreading

PACS: 89.75.Hc, 87.19.X–, 87.23.Ge **DOI:** 10.1088/1674-1056/abb302

1. Introduction

Networks can be implemented in numerous systems in the real world,[1] such as logistics transportation, World Wide Web, social contact.[2–6] Previous researches mainly focused on static network, in which the network structures do not change with time. Extensive empirical studies indicated that the network structures are time-varying, which should be described by using the framework of temporal networks.[7–10]

Information and diseases can be disseminated through nodes in temporal networks. For temporal network propagation, the scope of dissemination can be directly studied by simulating the interaction between groups within a period.[11] And it can be predicted by individual-based (IB) models for research[12–15] as well. IB model overcomes the shortcomings of the homogeneous and heterogeneous mean-field approaches. Aims to decrease the complexity of the solution from exponential $\mathcal{O}(3^N)$ to polynomial $\mathcal{O}(N)$, it is inspired by the continuous-time Markov chain SIR model.[16] However, compared to the individual-based model, the contact-based (CB) model proposed by Koher *et al.*[17] is derived from the message-passing framework,[18,19] which improves the approximation of existing epidemic thresholds and analytic estimations concerning the IB approach for arbitrary time-evolving graphs.

In the network propagation, information spreads through nodes, hence forms numerous paths between different pairs of nodes. Knowing the path distribution within the networks is significant for many reasons such as the World Wide Web[20] design, epidemic spreading,[21,22] information transmission on social networks,[23] network system structures reconstruction,[24] flood control, and water resources conservation.[25] Among these paths, only the shortest one can be used to define the length between two nodes. Some studies showed interest in the shortest paths distribution (SPD). The authors in Refs. [26–28] studied the statistical characteristics of uncorrelated random networks with an arbitrary degree distribution. Hofstad *et al.* studied the SPD in various random graph models with general degree sequences and extreme modular networks.[29,30] Some developed analytical methods to compute the SPD in Erdős and Rényi and configuration-model networks.[31,32] Another paper has studied the SPD in modular random networks,[33] testing the conditions in which the number of interlinks between two or more modules controls the network topology. SPD was also applied in identifying the diffusion source on networks.[34]

However, the SPD problem in the temporal network has drawn little attention. According to Eitan *et al.*, an one-pass algorithm, as well as a graph transformation approach has been proposed.[35] In this paper, we treat SPD calculating process as a way of Markovian information dissemination and put forward a new method to calculate the SPD in temporal networks. By predicting the Markovian information dissemination process,[17] we compute the SPD of four heterogeneous temporal networks. Besides, we utilize simulation of temporal network message spreading to verify the validity of our method. Some studies have introduced methods to calculate

*Project supported by the National Natural Science Foundation of China (Grant No. 61903266), China Postdoctoral Science Foundation (Grant No. 2018M631073), China Postdoctoral Science Special Foundation (Grant No. 2019T120829), the Fundamental Research Funds for the Central Universities, China, and Sichuan Science and Technology Program, China (Grant No. 20YYJC4001).

†Corresponding author. E-mail: wwzqbx@hotmail.com

© 2020 Chinese Physical Society and IOP Publishing Ltd

the influence maximization in information propagation.[36–38] In this article, we set strategies to study the influence of a node in propagation range in four temporal networks.

2. Model

In this section, we will introduce the basic concepts and models of temporal networks, as well as the property dynamics model.

2.1. Artificial temporal networks

The network in the traditional sense is a static network composed of nodes and edges. The temporal network we are studying is a network that takes the interaction time of nodes into account. The interactions between nodes could be seen as events. Each edge is active only when the event occurs, that is, two nodes solely have connected edge between them when the event occurs, and every activity happens in chronological order. Mathematically, we use $\boldsymbol{G}$ stands for the network at some point in time, and $\boldsymbol{G}_t$ for a sequence of $\boldsymbol{G}$. The adjacency matrix of $\boldsymbol{G}_t$ at time t is $\boldsymbol{A}_t$. If the interaction between node i and node j exists at time t, then $\boldsymbol{A}_t(i,j) = 1$. Otherwise, if no connection between node i and j at time t, $\boldsymbol{A}_t(i,j)$ equals 0.

Here we introduce two types of artificial temporal network models that will be applied in the rest of the paper. The first is the ER temporal network, in which the temporal network $\boldsymbol{G}_t$ is built by using ER[39] network. The network generating method is as follows. (i) Specify the number of nodes N and the node connect probability p of each pair of nodes. (ii) At each time t, generate an ER network $\boldsymbol{G}_t$ with nodes being arbitrarily connected with probability p. The degree distribution of network $\boldsymbol{G}_t$ is $P_t(k) = \binom{N}{k} p^k (1-p)^{N-k}$, and the average degree is $\langle k_t \rangle = p(N-1) \approx pN$, cluster coefficient $c = p$. (iii) At the end of each time step, deleting all edges. (iv) Repeating steps (ii) and (iii) until $t = t_{\max}$. Therefore, the structure of ER temporal network at two adjacent moments has no correlation.

The second artificial network is the activity-driven network.[40] The steps of generating temporal network $\boldsymbol{G}_t$ are as follows. i) Specify the number of nodes N and time T of the network. For each node i, we assign an activity potential value x_i, which follows a power-law distribution $f(x) \sim x^{-\gamma}$, $x \in (\varepsilon, 1)$ and γ is an exponent. The larger value of γ, the more homogeneous of degree distribution. ii) At each time step t, nodes are activated with a probability $n_i = \eta x_i$, where η is a remodeling factor, and each activated node randomly connects to m nodes. iii) At the end of each time step, deleting all edges. iv) Repeating steps ii) and iii) until $t = t_{\max}$.

2.2. Real-world temporal networks

In this paper, we utilize two real-world temporal networks as well. The first is a high school dynamic contact network,[41] which contains the temporal network of contacts between students in a high school in Marseilles, France, during four days in December 2011. The dataset includes the connection between 126 people during 13617-time steps, and each time step in dataset representing the active contacts in 20-second intervals of the data collection.

The second real-world temporal network is a dynamic contact network of hospital ward.[42] This dataset contains the temporal network of contacts between patients, patients, and health-care workers (HCWs) and among HCWs in a hospital ward in Lyon, France, from Monday, December 6, 2010, at 1:00 pm to Friday, December 10, 2010, at 2:00 pm. The study included 46 HCWs and 29 patients during 17375 times, and each time representing the active contacts during 20-second intervals of the data collection. Table 1 shows the details of the data we apply in this paper.

Table 1. Temporal network data.

Network	Node number	Time	Time interval
ER	200	200	1
Activity driven	200	200	1
High school	126	13617	20 s
Hospital ward	75	17375	20 s

2.3. Propagation dynamics model

In this paper, we use the susceptible-informed (SI) model to describe information spreading dynamics. At any time, a node could only have two states: susceptible or informed. If a node is in an informed state, that means it has received the message and has the chance to forward it to others. If a node is in a susceptible state, which means it has not received any messages yet. Each informed node transmits the message to its neighboring nodes that are in a susceptible state, and once neighbors receive, it is in a known state and could never return to a susceptible state. The steps of information dissemination on the static network are as follows. Initially, we set one node i in the network as informed at the first step, which is the source of propagation, and the others are susceptible. Node i then conveys the message to its susceptible neighbors with a particular transmission probability λ. If successful, its neighbors changed from susceptible to known ones. At this point, information initiates to spread on the network. Whereafter, each informed node will notify their susceptible neighbors with the propagation probability λ until all susceptible nodes in the network know the information.

The above is the notification situation of the SI model in a static network. Now we discuss the process of the SI model in a temporal network. Similarly, we set one node i in the network $\boldsymbol{G}_0$ to be informed at the initial time t_0, which is the source of propagation, and the others are susceptible. At the time t_0, the transmission source notifies its susceptible

neighbors with specific transmission feasibility λ. If the notification is successful, its neighbors change from susceptible to informed. Distinct from the static network, the temporal network contains time. Each time after that, only several nodes have active edges, and the ER temporal network and activity-driven temporal network we use have different connection situation at each time. Therefore, not all known nodes can inform their neighbors at t_1. Only when there are informed nodes among nodes that have active edges at this time, can these informed nodes notify susceptible neighbors. Figure 1(a) is an example of the propagation process of the SI model in a temporal network. Each node in the informed state only has neighbors at certain moments, so it can only propagate at these moments.

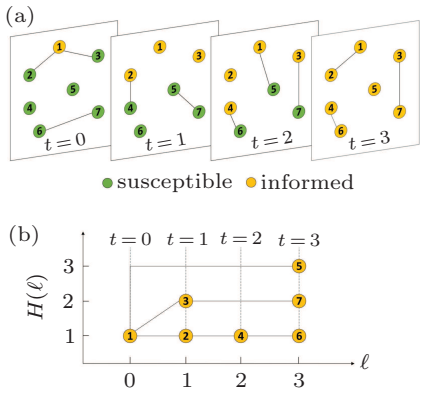


Fig. 1. Illustration of the relationship between spreading and shortest path. (a) An example of the information propagation procedure of the SI model in a temporal network. (b) The number $H(\ell)$ of path length ℓ versus ℓ on the temporal network above, and it shows the exact time of each node being notified and their message source. Meanwhile, it reveals the similarity between the information dissemination process and the SPD calculation process so that we can compute the path length and distribution in the procedure of information propagation. From the figure, it is shown that the distance between the node and the initial source is equal to the time it is notified.

2.4. Relationships between the shortest path and spreading dynamics

For temporal networks, the shortest path is time-dependent. Mathematically, the shortest path between nodes i and j represents that a walker can reach node j from node i through a temporal path $\{e_{i,1}(t_a), e_{2,3}(t_b), \ldots, e_{\ell,j}(t_x)\}$, and the path length ℓ is the minimum value for any paths between the two nodes, where $e_{m,n}(t_b)$ means that nodes m and n are connected at time t_b.

To calculate the shortest path of temporal network, we can see that the procedure of each reached point reaching its neighbors is similar to the information propagation process in the SI model mentioned above. The former is that a node changes from an unreached state to a reached state until all nodes become reached, while the latter is from an unknown state to a known state until all nodes become informed. Figure 1 exhibits the relationship between the SPD and the message passing process. We pick out the successful process of notification in panel (a) and show it in panel (b), and mark the path length from each node to the propagation source (node 1) as well as the number of paths length. For example, the path length from node 5 to the source is 3. Some may have doubts that this path length seems to be 1, however, because the connection of nodes in the temporal network changes with time, there is no connection between node one and node 5 in the first two steps while they are connected in the third step, so the path length between them is 3. From the figure, there are three nodes with a path length of 3 to the source, which means $H(3) = 3$. Each step of the notification process requires a time unit, so the shortest path length from each node to the source and the notified time are numerically equal, which can be seen in the figure. Therefore, by performing information dissemination, we calculate the time each node is notified to obtain the shortest path length.

3. Simulations

In numerical simulation, we set the network size $N = 200$ and maximum time step $t_{\max} = 200$ for ER temporal networks. For activity-driven networks, we set $N = 200$, $t_{\max} = 200$, $\varepsilon = 10^{-3}$, $m = 20$, and $\eta = 10$. In both networks, the propagation probability $\beta = 1$.

To compute the shortest path distribution of a temporal network, we should use a fast Monte-Carlo simulation method to study the SI model. From Subsection 2.3, it is illustrated that each node in the network is set as the source node to calculate the shortest path distribution, so it causes a large amount of calculation. The time complexity is $\mathcal{O}(N^2)$, n represents the number of nodes. Inspired by Ref. [11], we use a generalized a matrix-based approach with time complexity $\mathcal{O}(LT)$. We get a set of formulas to express its calculation process $\widetilde{\boldsymbol{S}}_{t+1} = 1 - \widetilde{\boldsymbol{I}}_{t+1}$ and $\widetilde{\boldsymbol{I}}_{t+1} = \widetilde{\boldsymbol{I}}_t + \beta\boldsymbol{\Phi}_t$, where $\widetilde{\boldsymbol{S}}_{t+1}$ represents the susceptible node vector at time $t+1$, $\widetilde{\boldsymbol{I}}_{t+1}$ for the informed node vector at $t+1$, and $\boldsymbol{\Phi}_t$ denotes the edge connection situation among all the nodes at time t, and its concrete expression will be discussed in the next paragraph. Both vectors are $N \times 1$ in size. In the $\widetilde{\boldsymbol{S}}$ matrix, the value is nonzero if the node is in a susceptible state, otherwise it is 0. Similarly, we know the meaning of the element in the vector $\widetilde{\boldsymbol{I}}$, *i.e.*, if the node is informed, its value is nonzero, otherwise, it is 0. On the connection $\boldsymbol{\Phi}$ between a susceptible node and an informed node, the susceptible node is notified with β probability. To compute the shortest path distribution, we should set $\beta = 1$, thus $\widetilde{\boldsymbol{I}}_{t+1} = \widetilde{\boldsymbol{I}}_t + \boldsymbol{\Phi}_t$.

Subsequently, we will discuss the specific expression of $\boldsymbol{\Phi}$. We use a common and convenient way to represent temporal networks: a set of adjacency matrices $\boldsymbol{A}_t$ with dimension $N \times N$. These adjacency matrices represent the connected edges of each node at time $t = 1, \ldots, t_{\max}$ Therefore, we can get $\boldsymbol{\Phi}_t = \boldsymbol{A}_t\widetilde{\boldsymbol{I}}_t$. The informed node at each moment notifies

its neighbors according to the active edge status at that moment. At the initial time, only one node has a value of 1 in $\widetilde{\boldsymbol{I}}_0$, which indicates the source of notification, and the remaining values are all 0. After analysis, we get the final expression that simulates propagation

$$\boldsymbol{S}_{t+1} = 1 - \boldsymbol{I}_{t+1}, \tag{1}$$

$$\boldsymbol{I}_{t+1} = \boldsymbol{I}_t + \boldsymbol{A}_t \boldsymbol{I}_t. \tag{2}$$

From this method, we can get the known status of each node in the temporal network at time t_i, so that we can calculate the number of informed nodes n_{ti} that increases at each time compared to the previous time. For the statistical value, since the notification time can be regarded as the length of the temporal network path, n_{ti} represents the number of paths whose length is t_i, and the value of the time t_i corresponding to the last non-zero newly increased informed node is the shortest path of the temporal network.

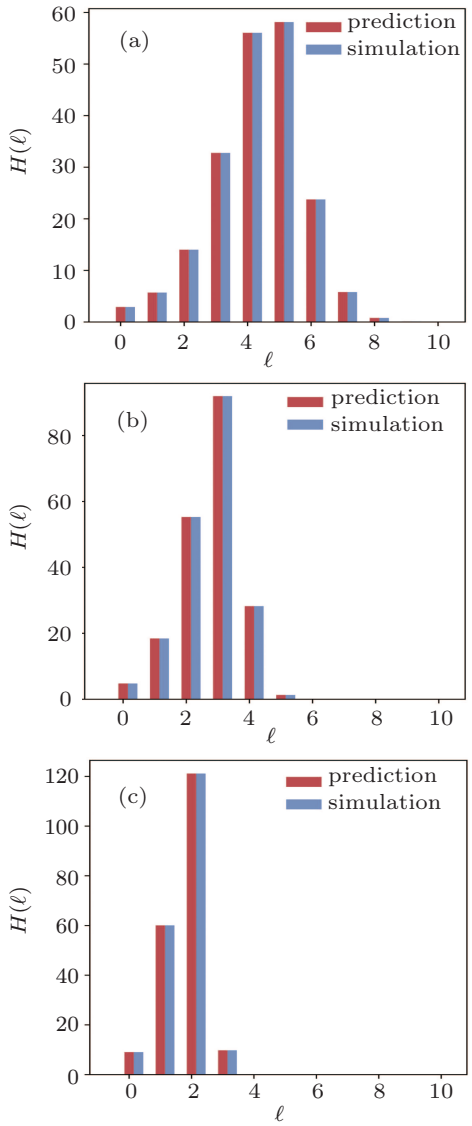


Fig. 2. The number $H(\ell)$ of path length ℓ *versus* ℓ on ER temporal networks. Nodes connection probability is $p = 0.01$ in panel (a), $p = 0.02$ in panel (b) and $p = 0.04$ in panel (c), respectively.

In what follows, we calculate the path length distribution for artificial and real-world temporal networks. Figure 2 exhibits the path distribution of ER temporal network with connection probabilities of $p = 0.01$, 0.02, and 0.04. From the figure, we can intuitively see that as the probability of continuous edges in the ER graph increases, the time for information to propagate to the entire network becomes shorter, so the shortest path will be shortened, and the distribution of the shortest path will increase. Simultaneously, it can be seen clearly that the theoretical prediction results (see details in Section 4) are consistent with the simulation results, indicating that our theory can well predict the numerical results.

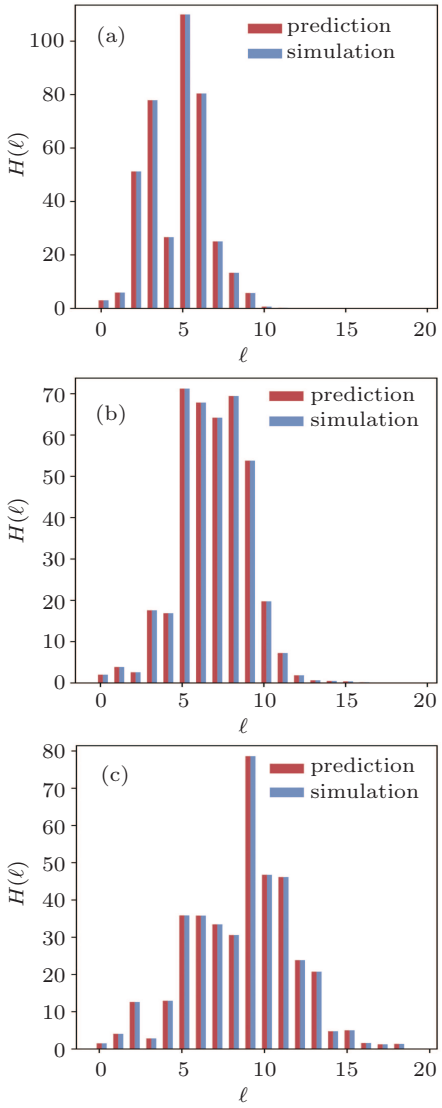


Fig. 3. The number $H(\ell)$ of path length ℓ *versus* ℓ on activity-driven temporal networks. $\gamma = 2.3$ in panel (a), $\gamma = 3.0$ in panel (b), and $\gamma = 3.5$ in panel (c), respectively.

Figure 3 shows the temporal network path distribution of the activity-driven model. The degree distribution is determined by γ. The smaller the γ, the stronger heterogeneous the degree distribution. When $\gamma = 2.3$, the degree distribution is more significant than $\gamma = 3.0$ and 3.5, the time required to inform all nodes is less, and the path is shorter as well. Through

a large number of simulation experiments, it is found that other values of γ will not qualitatively change our results. We further study the shortest path distribution on two real-world networks. Figure 4 shows the high school dynamic contact network and the hospital ward dynamic contact network. In high school data, $N = 126$, $t = 13617$, $\beta = 1$, and in hospital data, $N = 75$, $t = 17375$, $\beta = 1$. The x axis is the prediction result and the y axis is the simulation result. It can be seen that the slope of the function in the figure is 1, which indicates that the simulation is consistent with the prediction result.

We further propose strategies to identify the influential nodes. We propose five strategies. For strategies 1 to 4, the source is selected with the widest propagation scope in the first 1 to 4 steps, respectively. In strategy 5, the source node is randomly selected. The spreading scope of these strategies over time is shown in Fig. 5, with the propagation probability $P = 0.1$. We study the spread of the different sources of information dissemination in the four networks. In Fig. 5(a), we set the SI information spreading on ER temporal network with nodes connection probability $p = 0.01$. In Fig. 5(b), the message spreading on activity-driven temporal network $\gamma = 2.3$, $\varepsilon = 10^{-3}$, $m = 20$, and $\eta = 10$. For Figs. 5(c) and 5(d), the information dissemination on high school and hospital ward dynamic contact networks. To make Figs. 5(c) and 5(d) more intuitive, we sample the data according to a certain length for display. Conducting five experiments on each network, we use different shapes of scattered points to represent the mean of each strategy in five experiments and use error bars to represent the standard deviation of each strategy.

The spreading patterns of the four networks in Fig. 5 differ from each other. Since figures 5(a) and 5(b) are artificial temporal networks, the connection probability of nodes is fixed, so their information propagation range is relatively stable. At the same time, figures 5(c) and 5(d) are real-world temporal networks, the connection between nodes and nodes

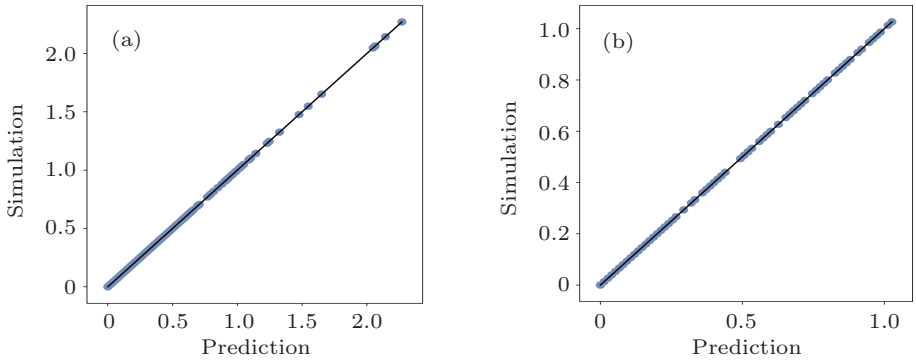


Fig. 4. Simulation results *versus* prediction results in the high school dynamic contact network (a) and the hospital ward dynamic contact network (b), respectively. The scatter points stand for true value, and the line is an auxiliary line, which is drawn to see the comparison of simulation and prediction results more intuitively.

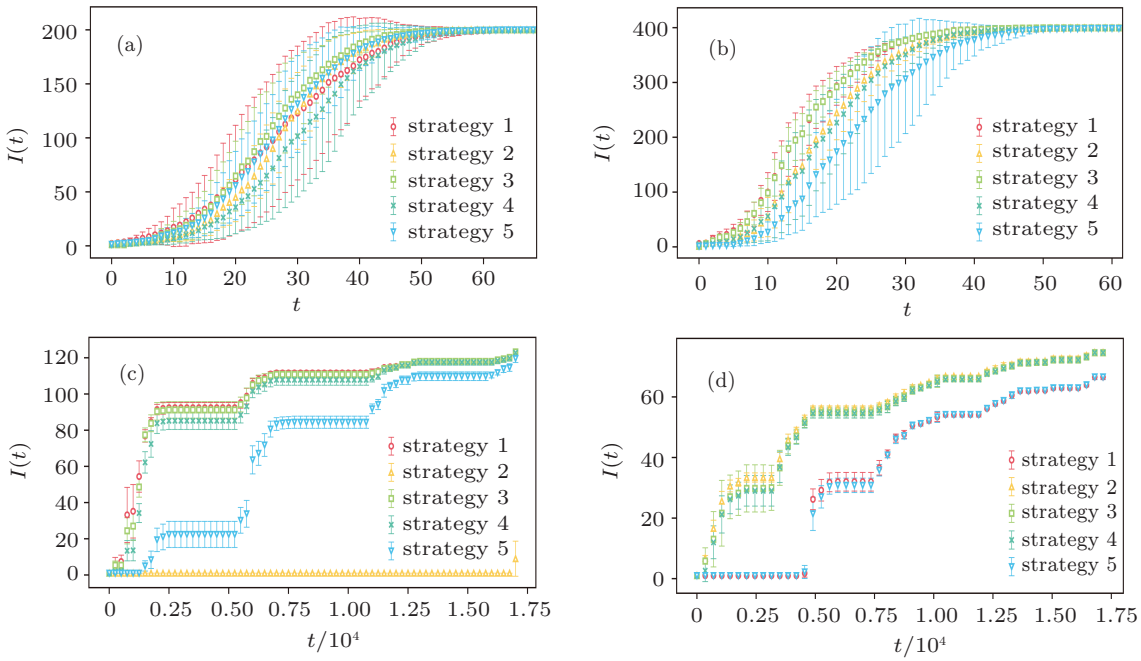


Fig. 5. The information spread range $I(t)$ *versus* time t on four networks with 5 strategies. Panels (a)–(d) are ER, activity-driven, high school contact, and hospital ward contact temporal networks, respectively.

is affected by practical factors, so they have uncertainty, which makes the growth of the scope of information spread unstable. At the same time, the propagation range is also affected by the spreading probability $P = 0.1$ and degree distributions of different graphs. As can be seen from the figure, nodes with a wide range of propagation have a higher probability of spreading fast, and the propagation speed of random node 5 is slower than that of nodes with high degree distribution. According to the performance of the five strategies, we can conclude that strategy 3 has the best effect in the range of dissemination, and strategy 5 has the worst impact.

4. Theory

In this section, we will describe mathematically how to predict SPD in a temporal network. When calculating the SPD of a temporal network, we use a contact-based (CB) model to predict. In some related works, temporal networks use individual-based (IB) models [12–16] to research the propagation of temporal networks. The authors in Refs. [18,19] proposed accurate message-passing frameworks on dynamic networks. Still, according to the research by Koher,[17] the CB model is derived from them and inherits their accuracy on the loop-free topology, and improves analytical estimates concerning the IB approach for arbitrary temporal graphs.

Importantly, the accuracy of the CB model improves the approximation of existing thresholds, which is a crucial measure for public information dissemination. Therefore, we use the CB model for propagation prediction to improve the accuracy of the SPD calculation. The CB model requires iteration through all edges at each time step, and thus the time complexity scales with $\mathcal{O}(LT)$, in which L is the number of edges.

In the SI model, nodes have only two possible states: the susceptible state (S) or the informed state (I), and at each moment, and informed state node can only propagate to the susceptible state neighbors with a distance of 1 from them. First, we consider the probability that node l is susceptible at time t, because of which, we assume that the probability of node l being susceptible at the initial time $S_l(0) = p_l$, it shows that l is not notified by any neighbor nodes before time t, so the probability of node l being susceptible $S_l(t) = p_l\Psi_l(t)$.

Next, we discuss $\Psi_l(t)$. If the probability of node l at time t is calculated from the informed source at the initial time, the known status of each node from the initial time to time t must be considered. The edge states are very complicated and vary from the time, so this method is not applicable. Therefore, we can start by considering the neighbors of node l at time t, and rewrite $\Psi_l(t)$ as

$$S_l(t) = p_l \prod_{k\in N_l} \Theta_{k\to l}(t), \tag{3}$$

where N_l represents neighbor nodes of l at time t. $\Theta_{k\to l}(t)$ denotes the probability that no message will propagate from k node to l node from the initial time to time t. Since there is no information transmission at the initial time, $\Theta_{k\to l}(0) = 1$. What is more, $\Theta_{k\to l}(t+1)$ is determined by two factors: (I) k and l have temporary edges at time t, the probability is represented by $u_{k\to l}(t)$, and (II) k is at informed state at time t and does not transmit the message to l before time t, which is denoted by $I_{k\to l}(t)$. We set β as the propagation probability, then the chance of l being notified by k at time t is $\beta u_{k\to l}(t)I_{k\to l}(t)$, from which the formula can be derived as

$$\Theta_{k\to l}(t+1) = \Theta_{k\to l}(t) - \beta u_{k\to l}(t)I_{k\to l}(t). \tag{4}$$

The probability $I_{k\to l}(t+1)$ that k is in the informed state and did not transmit the message to l before time $t+1$ is also determined by two factors: I) k is in the known state before time t and does not inform l, which is $I_{k\to l}(t)$, and k still does not notify l at t with the propagation probability of β and that is $1-\beta u_{k\to l}(t)$. II) k is notified by its neighbor nodes other than l at time t $\Delta S_{k\to l}(t) = S_{k\to l}(t) - S_{k\to l}(t+1)$, since notification can occur only one step at a time, k cannot inform l at this time. From this we can get $I_{k\to l}(0) = 1 - p_l$ at the initial moment, and derive the formula

$$I_{k\to l}(t+1) = [1-\beta u_{k\to l}(t)]I_{k\to l}(t) + \Delta S_{k\to l}(t). \tag{5}$$

Then we discuss the feasibility $S_{k\to l}(t)$ of the neighbor node k of l at time t. Since k is not notified by l, then if k is not notified by all its neighbors except l $j \in N_k \setminus l$ at time t, k is susceptible. We know that the corresponding probability[43] is $\Psi_{k\to l}(t) = \prod_{j\in N_k\setminus l}\Theta_{j\to k}(t)$, which is similar to Eq. (3). So $S_{k\to l}(t+1) = p_k\Psi_{k\to l}(t+1)$, which is

$$S_{k\to l}(t+1) = p_k \prod_{j\in N_k\setminus l} \Theta_{j\to k}(t+1). \tag{6}$$

The progression of message transmission in the CB model is mainly described by Eqs. (5) and (6). Finally, we discuss the node-centric formula. $S_l(t)$ has been shown in Eq. (3). Moreover, SI model has only two states of susceptible state S and informed state I, so the node-centric formula can be expressed as

$$S_l(t+1) = p_l \prod_{k\in N_l} \Theta_{k\to l}(t+1). \tag{7}$$

$$I_l(t+1) = 1 - S_l(t+1). \tag{8}$$

The CB model is accurate for temporal networks and allows us to consider the dynamic correlation of the research objects fully.

5. Conclusion

In this article, we proposed a method for calculating the shortest path in temporal networks. By utilizing the SI model of information propagation and simulating the propagation process of the Markovian chain, we calculated the SPD. And

using the method of predicting the whole process of information dissemination to compute SPD, we performed calculations on artificial and real-world temporal networks. The results are compared to our method to verify the accuracy of ours. By displaying the histogram and scatter plot of the calculation results of path distribution, we can see the exactitude of our proposed method. Moreover, we put forward five strategies to study the influence of a node in propagation in four temporal networks. By comparing the information dissemination range of the strategy, we get the conclusion which is the best or the worst strategy.

References

[1] Newman M E J 2003 *SIAM Rev. Soc. Ind. Appl. Math.* **45** 167
[2] Hufnagel L, Brockmann D and Geisel T 2004 *Proc. Natl. Acad. Sci. India Sect. B Biol. Sci.* **101** 15124
[3] Hidalgo C, Klinger B, Barabasi A L and Hausmann R 2007 *Bull. Am. Phys. Soc.* **52** No. 1
[4] Liljeros F, Edling C R, Amaral L A N, Stanley H E and Åberg Y 2001 *Nature* **411** 907
[5] Cohen R and Havlin S 2010 *Complex networks: structure, robustness and function* (Cambridge: Cambridge University Press)
[6] Newman M 2018 *Networks* (Oxford: Oxford University Press)
[7] Holme P and Saramäki J 2012 *Phys. Rep.* **519** 97
[8] Sekara V, Stopczynski A and Lehmann S 2016 *Proc. Natl. Acad. Sci. USA* **113** 9977
[9] Stopczynski A, Sekara V, Sapiezynski P, Cuttone A, Madsen M M, Larsen J E and Lehmann S 2014 *PloS one* **9**
[10] Casteigts A, Flocchini P, Quattrociocchi W and Santoro N 2012 *International Conference on Ad-Hoc Networks and Wireless* **27** 387
[11] Koher A, Lentz H HK, Hövel P and Sokolov I M 2016 *PloS One* **11**
[12] Valdano E, Ferreri L, Poletto C and Colizza V 2015 *Phys. Rev. X* **5** 021005
[13] Chakrabarti D, Wang Y, Wang C, Leskovec J and Faloutsos C 2008 *ACM Trans. Inf. Syst. Secur.* **10** 1
[14] Mieghem P V, Omic J and Kooij R 2008 *IEEE ACM Trans. Netw.* **17** 1
[15] Wang Y, Chakrabarti D, Wang C, and Faloutsos C 2003 *Proceedings of 22nd International Symposium on Reliable Distributed Systems* 25
[16] Youssef M and Scoglio C 2011 *J. Theor. Biol.* **283** 136
[17] Koher A, Lentz H H K, Gleeson J P and Hövel P 2019 *Phys. Rev. X* **9** 031017
[18] Karrer B and Newman M EJ 2010 *Phys. Rev. E* **82** 016101
[19] Lokhov A Y, Mézard M, Ohta H and Zdeborová L 2014 *Phys. Rev. E* **90** 012801
[20] Zhang Z K, Liu C, Zhan X X, Lu X, Zhang C X and Zhang Y C 2016 *Phys. Rep.* **651** 1
[21] Zhan X X, Liu C, Zhou G, Zhang Z K, Sun G Q, Zhu J J and Jin Z 2018 *Appl. Math. Comput.* **332** 437
[22] Kan J Q and Zhang H F 2017 *Commun. Nonlinear Sci. Numer. Simul.* **44** 193
[23] Liu C, Zhan X X, Zhang Z K, Sun G Q and Hui P M 2015 *New J. Phys.* **17** 113045
[24] Ma C, Chen H S, Li X, Lai Y C and Zhang H F 2020 *SIAM J. Appl. Dyn. Syst.* **19** 124
[25] Gong Y, Liu P, Cheng L, Chen G, Zhou Y, Zhang X and Xu W 2020 *J. Flood Risk Manag.* **13** e12586
[26] Newman M EJ, Strogatz S H and Watts D J 2001 *Phys. Rev. E* **64** 026118
[27] Dorogovtsev S N, Mendes J F F and Samukhin A N 2003 *Nucl. Phys. B* **653** 307
[28] van der Hofstad R, Hooghiemstra G and Van Mieghem P 2005 *Random Structures & Algorithms* **27** 76
[29] van der Hofstad R and Hooghiemstra G 2008 *J .Math. Phys.* **49** 125209
[30] Asher E E, Sanhedrai H, Panduranga N K, Cohen R and Havlin S 2020 *Phys. Rev. E* **101** 022313
[31] Katzav E, Nitzan M, ben Avraham D, Krapivsky P L, Kühn R, Ross N and Biham O 2015 *Europhys. Lett.* **111** 26006
[32] Nitzan M, Katzav E, Kühn R and Biham O 2016 *Phys. Rev. E* **93** 062309
[33] Dorogovtsev S N, Mendes J F, Samukhin A N and Zyuzin A Y 2008 *Phys. Rev. E* **78** 056106
[34] Xu S, Teng C, Zhou Y, Peng J, Zhang Y and Zhang Z K 2019 *Physica A* **527** 121267
[35] Wu H, Cheng J, Huang S, Ke Y, Lu Y and Xu Y 2014 *Proceedings of the VLDB Endowment* **7** 721
[36] Qiu L Q, Yu J F, Fan X, Jia W and Gao W W 2019 *IEEE Access* **7** 42052
[37] Kumar R and Calders T 2017 *Advances in Database Technology, EDBT* 2017*: Proceedings of the* 20*th International Conference on Extending Database Technology*, March 21–24, 2017, Venice, Italy 270
[38] Chen Y C, Zhu W Y, Peng W C, Lee W C and Lee S Y 2014 *ACM Trans. Intell. Syst. Technol.* **5** 1
[39] Erdős P and Rényi A 1960 *Publ. Math. Inst. Hung. Acad. Sci.* **5** 17
[40] Perra N, Gonçalves B, Pastor-Satorras R and Vespignani A 2012 *Sci. Rep.* **2** 469
[41] Fournet J and Barrat A 2014 *PloS One* **9**
[42] Vanhems P, Barrat A, Cattuto C, Pinton J F, Khanafer N, Régis C, Kim B, Comte B and Voirin N 2013 *PloS One* **8**
[43] Lokhov A Y 2014 *Dynamic cavity method and problems on graphs*, Ph. D. Dissertation (Paris: Paris 11)