Topological phase transition in network spreading

Fuzhong Nian(年福忠)[†] and Xia Zhang(张霞)

School of Computer & Communication, Lanzhou University of Technology, Lanzhou 730050, *China*

(Received 26 April 2022; revised manuscript received 30 May 2022; accepted manuscript online 27 June 2022)

This paper investigates information spreading from the perspective of topological phase transition. Firstly, a new hybrid network is constructed based on the small-world networks and scale-free networks. Secondly, the attention mechanism of online users in information spreading is studied from four aspects: social distance, individual influence, content richness, and individual activity, and a dynamic evolution model of connecting with spreading is designed. Eventually, numerical simulations are conducted in three types of networks to verify the validity of the proposed dynamic evolution model. The simulation results show that topological structure and node influence in different networks have undergone phase transition, which is consistent with the phenomenon that followers and individual influence in real social networks experience phase transition within a short period. The infection density of networks with the dynamic evolution rule changes faster and reaches higher values than that of networks without the dynamic evolution rule. Furthermore, the simulation results are compared with the real data, which shows that the infection density curve of the hybrid networks is closer to that of the real data than that of the small-world networks and scale-free networks, verifying the validity of the model proposed in this paper.

Keywords: social network, information spreading, network structure, phase transition

PACS: 89.75.Hc, 89.75.Fb, 64.60.aq **DOI:** 10.1088/1674-1056/ac7bfa

1. Introduction

Computational communication is a promising field of research in complex networks,[1–6] and the study of information spreading in social networks[7–11] is another critical sub-field of computational communication.[12–15] At present, social networks are getting widely used and developed as an increasing number of people get to know and post information through social media platforms. However, there exists a huge amount of information in social networks, in which false information and maliciously oriented opinions tend to be widely spread as well. Hence, it is of great importance to study information spreading in social networks. This study allows us to quickly grasp the laws of information spreading and disseminate effective information throughout the networks.

Numerous scholars have studied information spreading in social networks. Nian *et al.*[16–19] considered the influence of various effects on the information spreading process and proposed the concept of a hybrid spreading model. Wu *et al.*[20] investigated the dual-channel effect of information spreading based on the susceptible acceptable recovery (SAR) model, and explored the effect of both channels on the spreading of explosive information. Some scholars have conducted studies related to the phenomenon of rumor spreading on the internet. Mheidly *et al.*[21] put forward the idea that only the rapid and widespread spreading of proven accurate information among the public can be effective in the combat against the new COVID-19. A heterogeneous network-based infectious disease model was proposed by He *et al.*[22] to describe the spreading process of rumors in the network. In addition, two methods were designed to suppress the rumor spreading. Phase transition[23] is a concept in physics, but it also occurs in social systems.[24] Most scholars have conducted research only on the aspect of information spreading modeling, and few scholars have studied phase transition[25–28] in networks. Boccaletti *et al.*[29] presented a thematic review of mainstream literature related to phase transition, summarizing the existing research results and indicating possible future research orientations. Li *et al.*[30] demonstrated that the von Neumann entropy can well quantify the system phase transition in rumor spreading systems. Xie *et al.*[31] proved that information spreading on social media is a pervasive process, similar to the definition of that in physics. Some scholars have studied the phase transition in information spreading. Lim *et al.*[32] analyzed dynamic thresholds in epidemic spreading based on which control was performed to prevent epidemic outbreaks. Nian *et al.*[33] analyzed the threshold values of information spreading in a high-level network model with three types of closeness, and respectively proposed two information phase transitions from the perspective of global network threshold and individual information entropy threshold. Davis *et al.*[34] demonstrated that there is a critical phase transition point for the avalanche evolution of information spreading range and dissemination efficiency in social networks. Some scholars have also conducted studies on the phase transition caused by recovery factors. Zhang *et al.*[35] and Xiao *et al.*[36] considered both internal and external recovery in the model which proposed to study the recovery process in the network. And by stimulating the recovery process of network nodes, Xiao *et al.*

[†]Corresponding author. E-mail: gdnfz@lut.edu.cn

© 2023 Chinese Physical Society and IOP Publishing Ltd

found that the system is only characterized by first-order phase transition when external recovery is considered, while internal recovery fails to induce this transition. In addition, some scholars have carried out research on opinion phase transition and emotion phase transition. Mansouri *et al.*[37] proposed the phenomenon of phase transition in the opinion model based on social influence theory. Holme *et al.*[38] suggested that a continuous phase transition is experienced from majority opinion diversity to majority sharing the same opinion. Yu *et al.*[39] investigated the phenomenon of group emotion phase transition from two aspects: without network structure and based on the synergistic evolution of emotion and network structure.

Among the above-mentioned literature, most of the studies have focused on identifying phase transition points to control disease outbreaks and opinion and emotion phase transition. Research on phase transition mainly focuses on the improvement of spreading models, while the research on the network structure is not intensive enough and the outcomes are limited. Therefore, this paper studies information spreading from the perspective of topological phase transition. Firstly, a hybrid network model which is more in line with the real social network is constructed by drawing on the ideas of Liu *et al.*,[40–42] who studied image hybrid encryption algorithms. Secondly, the dynamic evolution model of connecting with spreading is proposed. The small-world networks, scale-free networks, and hybrid networks are respectively conducted as basic network models to simulate the continuous network evolution as information spreading.

This study reveals that social networks undergo topological phase transition during the event fermentation, which in turn leads to more rapid spreading, and causes phase transition in some individuals' influence at a certain moment, eventually resulting in more and more intense event spreading, which has great significance for controlling information spreading.

The rest of the paper is organized as follows. In Section 2, various network models and the dynamic evolution model of connecting with spreading are introduced. In Section 3, numerical simulations based on proposed models are carried out to investigate the phase transition during network evolution, and the infection density curves of the experimental networks under different parameters are compared with the real data to verify the validity of models constructed in this paper. In Section 4, a summary about the whole paper is concluded.

2. Model

2.1. Network models

Experiments are conducted by using the SI model in small-world networks,[43,44] scale-free networks,[45,46] and hybrid networks. The hybrid network modeling process is described in the following section.

In real-world social networks, there exist various different sizes social groups, including small-sized, medium-sized, and large-sized groups, in which not only members inside the groups would communicate with each other but also would communicate with members outside the groups. And the social network in the real world is not characterized only by scale-free features or small-world ones. Hence, hybrid networks consisting of small-world networks and scale-free networks are constructed, based on the properties of real networks.

Definition 1 (hybrid network model) $N_{\mathrm{S}} \in (10,15)$ represents the node number of the small-sized groups, $N_{\mathrm{M}} \in (200,300)$ being the node number of the medium-sized groups, $N_{\mathrm{L}} \in (1000,2000)$ being the node number of the large-sized groups. Assuming that the social networks contain U social groups of different sizes, the hybrid network modeling algorithm is as follows.

(i) Randomly generate the number of social groups U, where the number of small-sized groups is U_1, the number of medium-sized groups is U_2, and the number of large-sized groups is U_3, satisfying $U_1+U_2+U_3=U$.

(ii) Suppose the total number of nodes in the hybrid networks is N, where the total number of nodes in the small-sized groups is μN, the total number of nodes in the medium-sized groups is υN, and the total number of nodes in the large-sized groups is $(1-\mu-\upsilon)N$.

(iii) As the small-sized groups conform to the small-world property, U_1 small-world networks are generated randomly in turn, and the number of nodes of U_1 small-world networks is $X_1,X_2,X_3,\ldots,X_{U_1} \in (10,15)$, satisfying $\sum_{i=1}^{U_1} X_i = \mu N$. Internal connections of small-world networks with different numbers of nodes are randomly reconnected starting from the rule graph and following the small-world connection rules.

(iv) As the medium-sized groups conform to the scale-free property, U_2 medium-sized scale-free networks are randomly generated in turn, and the number of nodes of U_2 scale-free networks is $Y_1,Y_2,Y_3,\ldots,Y_{U_2} \in (200,300)$, satisfying $\sum_{j=1}^{U_2} Y_j = \upsilon N$. Internal connections of scale-free networks with different numbers of nodes are connected by preferential connection method and following scale-free connection rules.

(v) As the large-sized groups conform to the scale-free property, U_3 large-sized networks are randomly generated in turn, and the number of nodes of U_3 is $Z_1,Z_2,Z_3,\ldots,Z_{U_3} \in (1000,2000)$, satisfying $\sum_{k=1}^{U_3} Z_k = (1-\mu-\upsilon)N$. The connection method of large-sized scale-free networks is the same as that in step (iv).

(vi) In three groups of different sizes, small, medium, and large, a certain number of nodes are randomly selected and respectively placed in the lists L_i,L_j,L_k, and the nodes in the three lists are randomly connected.

Figure 1 depicts the modeling process of hybrid networks,

where SW represents the small-world networks and SF represents the scale-free networks.

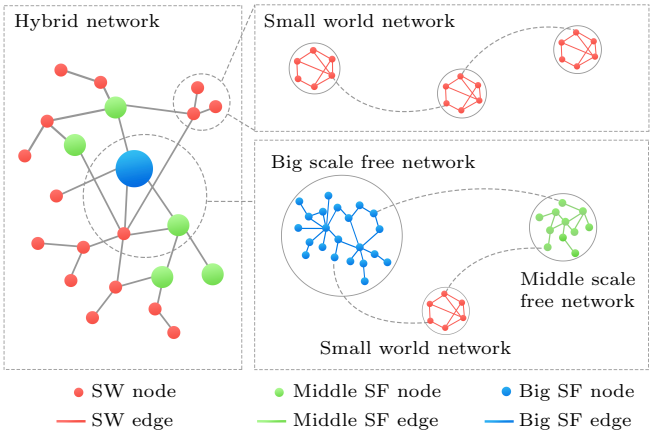


Fig. 1. Hybrid network modeling process.

2.2. Connecting with spreading dynamic revolution model

Viewing information spreading in social networks as a short-time process, its immune state is temporarily not considered. Thus, a disease transmission model similar to the SI is used in the simulation experiments to classify individuals who conduct information spreading in social networks into two types.

(i) The unknown (S): Individuals do not know the information, i.e., individuals will probably be informed by the known.

(ii) The known (I): Individuals know the information, i.e., individuals will inform the unknown with probability β, making him the known.

In the information spreading model above, the unknown (S) or the known (I), all individuals in the social networks will remain in the known state forever once they are informed of the information. The spreading dynamics of the information spreading model in the social network is formulated as

$$\frac{\mathrm{d}S(t)}{\mathrm{d}t} = -\beta\frac{S(t)I(t)}{N}, \quad \frac{\mathrm{d}I(t)}{\mathrm{d}t} = \beta\frac{S(t)I(t)}{N}, \tag{1}$$

where $S(t)$ and $I(t)$ represent the number of the unknown and the known in the network at time t, respectively. N is the population of individuals in the network, satisfying $S(t)+I(t)=N$. Information is spread to the unknown with probability β only when the connection is made between the known and the unknown. As t grows, the connection between the unknown and the known leads to an increase in the number of known individuals.

In the real social networks, as an event spreads, under influence of various factors such as the curiosity of onlookers, the newly known will establish new connections with some of the previously knowns with a certain probability, thus making the newly known gets the news quickly when the event has a new development. Therefore, this paper develops a dynamic evolution model of connecting with spreading, the core idea of which is that the nodes of the newly infected state I (the newly known individuals) will establish new connections with a certain probability with the non-neighbor nodes of the previously infected state I (the previously spreading non-connecting individuals). The connection rule is shown in Fig. 2.

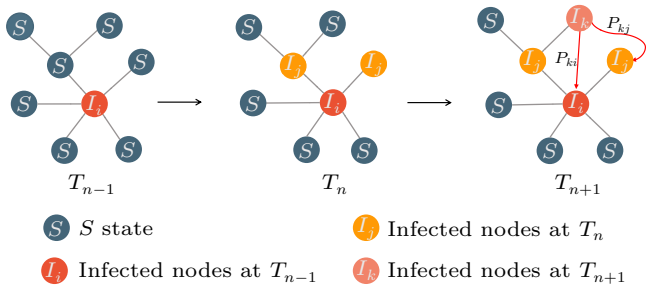


Fig. 2. Connection revolution rule.

(i) At time T_{n-1}, I_i represents the earliest known individuals in the social network, where $i \in (1,n)$.

(ii) At time T_n, I_i successfully spreads the information to I_j, turning them from the unknown to the known. I_j represents individuals changing to the known at T_n, where $j \in (1,n)$. I_j and I_i are connected. Thus, I_j is not required to determine whether to establish a connection with I_i.

(iii) At time T_{n+1}, I_j successfully spreads the information to I_k, turning them from the unknown to the known. I_k represents individuals changing to the known at T_{n+1}, where $k \in (1,n)$. I_k is driven by the desire to know the information source or to get the new information as soon as possible in the next round of spreading to establish connections with the knowns who are not previously connected. Therefore, connection propensity values of I_k and I_i, P_{ki}, are compared with that of I_k and I_j, P_{kj}, to establish connections with the known with higher connection propensity values.

This paper further explores factors that determine how individuals establish attention relationships with other ones in the process of information spreading, under real social networks. Results show that individuals in the networks are more likely to follow individuals from friends of friends as well as individuals with high authority, the ability to create rich content, and high frequency to updating new content.

Definition 2 (connection propensity values P_{ij}) It depends on four factors: social distance factor between two nodes, influence factor of nodes, content value factor of nodes, activeness factor of nodes. Connection propensity values connecting node j and node i are defined as

$$P_{ij} = \omega_1 \times \varepsilon_1(i,j) + \omega_2 \times \varepsilon_2(i) + \omega_3 \times \varepsilon_3(i) + \omega_4 \times \varepsilon_4(i), \tag{2}$$

where node j represents the newly infected nodes as I state, node i represents previously infected nodes as I state excluding j. $\varepsilon_1(i,j), \varepsilon_2(i), \varepsilon_3(i), \varepsilon_4(i)$ represent distance factor of j from i, content value factor of i, activity factor of i, influence

factor of i, respectively. $\omega_1, \omega_2, \omega_3, \omega_4$ represent the weight of connection probability P_{ij} of $\varepsilon_1(i,j)$, $\varepsilon_2(i), \varepsilon_3(i), \varepsilon_4(i)$, respectively, satisfying $\sum_{i=1}^{4} \omega_i = 1$.

In social reality, individuals are more prone to acquaint themselves with friends of friends. According to Gary Charness *et al.*[47] and Marián Boguñá *et al.*,[48] the probability of individuals building social connections will drop with their relative social distance.

Definition 3 (distance factor) Let $\varepsilon_1(i,j)$ represents the distance factor between node i and node j,

$$\varepsilon_1(i,j) = \frac{1}{d_{ij}+1}, \tag{3}$$

where d_{ij} represents the distance between node i and node j.

Definition 4 (content value) In the real social network, the ability of different individuals to create content is different. In this paper, the content value of an individual is defined as the sum of individual likes, comments, and retweets. Let $\Pr(c)$ represents the content value distribution of all individuals in the network, and the analysis of empirical data shows that the content value of all individuals in the social network under this definition obeys a power-law distribution

$$\Pr(c) \propto c^{-\gamma}. \tag{4}$$

Definition 5 (content value factor) Studies have shown that individuals tend to be attracted to ones who have a high ability to create content in order to keep up with new information. Therefore, let $\varepsilon_2(i)$ represents the content value factor of node i,

$$\varepsilon_2(i) = \frac{C_i}{C_{\max}}, \tag{5}$$

where C_i represents the content value of node i, while $C_{\max}$ represents the node with the highest content value in the network.

Definition 6 (activity) Each individual's activity level on social media platforms is different, and in order to better quantify the size of an individual's activity level. In this paper, the activity of an individual is defined as the total number of posts made by an individual. Let $\Pr(a)$ represents the activity distribution of all individuals in the network, and the analysis of empirical data shows that the activity of all individuals in the social network under this definition also obeys a power-law distribution

$$\Pr(a) \propto a^{-\gamma}. \tag{6}$$

Definition 7 (activity factor) In social networks, individuals who post frequently are prone to gaining followers. Therefore, let $\varepsilon_3(i)$ represents the activity factor of node i,

$$\varepsilon_3(i) = \frac{A_i}{A_{\max}}, \tag{7}$$

where A_i represents the activity of node i, while $A_{\max}$ represents the node with the highest activity in the network.

Definition 8 (influence) In real social networks, the individuals' influence is constantly changing as information spreads. Therefore, two components, initial influence and spreading influence are considered in the definition of influence. The influence of node u at time $t+1$ is defined as

$$I_{t+1}(u) = I_t(u) + \sum_{v \in T_1(u)} \frac{d_v}{d_u} A_v, \tag{8}$$

where $T_1(u)$ represents the set of neighbor nodes that are infected as I status by node u during the time $t \sim t+1$. v represents each node in time $T_1(u)$ in traversal, d_v being the degree of node v, d_u being the degree of node u, A_v being the activity of node v.

Equation (8) introduces that the influence of node u at time $t+1$ is equal to that at time t plus the spread influence gained at time $t \sim t+1$. Nodes with a higher degree and activity will be infected as I state by node u, thus resulting in greater spreading influence gained by node u.

Definition 9 (influence factor) As the Matthew effect,[49] the greater the degree is, the greater the influence is. And the greater the influence is, the more it will attract new nodes to connect with it, forming a state of getting more attention. Therefore, let $\varepsilon_4(i)$ represents the influence factor of node i, and it considers both degree and influence,

$$\varepsilon_4(i) = \frac{I_i}{I_{\max}} + \frac{d_i}{d_{\max}}, \tag{9}$$

where I_i represents the influence of node i, while $I_{\max}$ represents the node with the highest influence in the network. d_i represents the degree of node i, while $d_{\max}$ represents the node with the highest degree in the network.

Therefore, by deriving Eqs. (2), (3), (5), (7), and (9), we get

$$P_{ij} = \frac{\omega_1}{d_{ij}+1} + \omega_2 \times \left(\frac{C_i}{C_{\max}}\right) + \omega_3 \times \left(\frac{A_i}{A_{\max}}\right) + \omega_4 \times \left(\frac{I_i}{I_{\max}} + \frac{d_i}{d_{\max}}\right). \tag{10}$$

3. Simulation and experiments

3.1. Phase transition in node degree

Based on the evolution model of connecting with spreading, the experimental networks in different rounds in the spreading process are visualized as shown in Fig. 3.

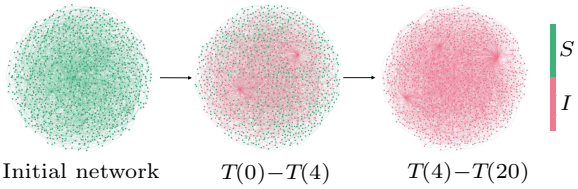


Fig. 3. Spreading process.

The green color on the right side of Fig. 3 represents the S state (the unknown) and the red color represents the I state (the known). From Fig. 3, it can be observed that based on the evolution model of connecting with spreading, nodes with a higher degree are generated in the networks.

Definition 10 (phase transition function of node degree values) The phase transition function of the node degree values is defined as

$$P=\begin{cases} P_1, & d_t(i)\le \rho_1(i,t), \\ P_2, & \rho_1(i,t)<d_t(i)\le \rho_2(i,t), \\ P_3, & d_t(i)>\rho_2(i,t), \end{cases} \tag{11}$$

which describes the differences of node phases under different degree ranges. $d_t(i)$ represents the degree value of node i at time t. $\rho_1(i,t)=\max\{3d_0(i),d_\Sigma(t),d_{\rm M}(t)\}$ is the threshold value changing from phase P_1 to phase P_2, while $\rho_2(i,t)=\max\{(3n)^2d_0(i),n^2d_\Sigma(t),n^2d_{\rm M}(t)\}$is the threshold value changing from phase P_2 to phase P_3 and n takes the value of 10. $d_0(i)$ represents the degree value of node i in the initial networks, while $d_\Sigma(t)=\sum_{i=1}^{N}d(i,t)/N$ is the average value of the degree values of N nodes in the network at the time t. $d_{\rm M}(t)$ is the median of the degree values of N nodes in the network at the time t, where $d_1(t)<d_2(t)<\cdots<d_n(t)$

$$d_{\rm M}(t)=\begin{cases} d_{(n+1)/2}(t), & \text{when } n \text{ is odd number}, \\ \dfrac{d_{n/2}(t)+d_{(n/2)+1}(t)}{2}, & \text{when } n \text{ is even number}. \end{cases} \tag{12}$$

The degree value changes of some nodes in the experimental network under different rounds are shown in Fig. 4.

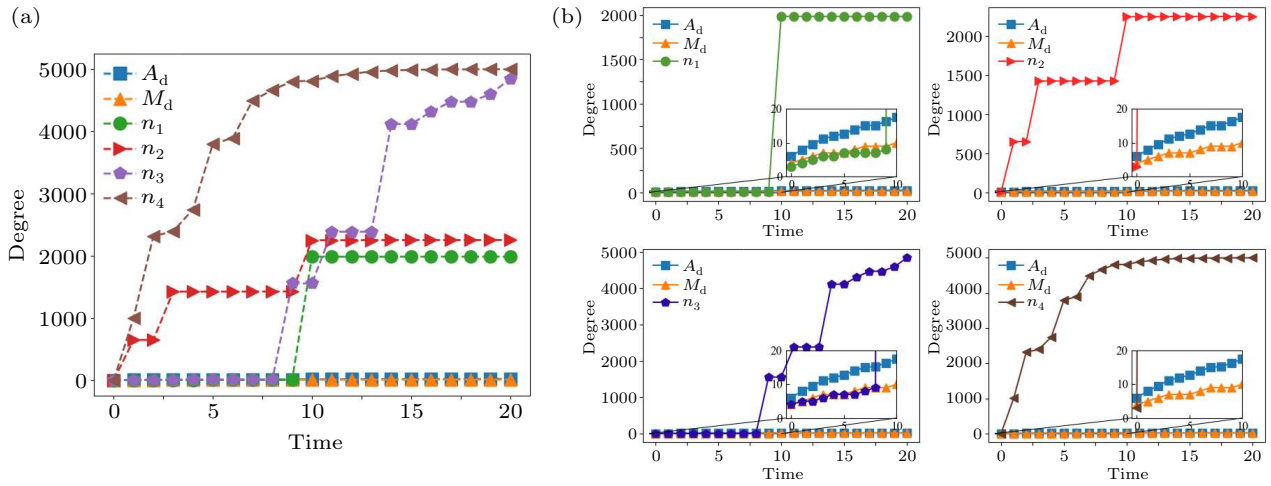


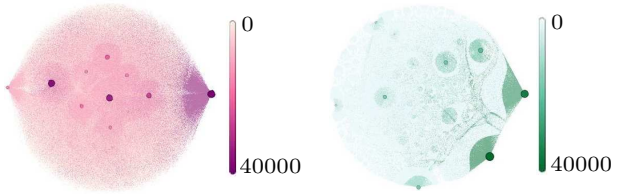
Fig. 4. Changes of degree values. (a) Change of degree value of four nodes. (b) The detailed evolution process of four node degree phase transition points.

In Fig. 4, $A_{\rm d}$ represents the average of the overall degree values of experimental networks, while $M_{\rm d}$ represents the median of the overall degree values of experimental networks. n_1, n_2, n_3, n_4 represent node 3421, 4997, 4094, 4999, respectively in the networks.

The P_1, P_2, and P_3 are divided according to Eq. (11). From Fig. 4, it can be observed that in the time $T(0)$–$T(9)$, the value of node n_1 satisfies $d_t(n_1)\le\rho_1(n_1,t)$. At time $T(10)$, the value of node n_1 satisfies $\rho_1(n_1,t)<d_t(n_1)\le\rho_2(n_1,t)$. Therefore, node n_1, changing from phase P_1 to phase P_2 at $T(9)$–$T(10)$, becomes sub-super node. Likewise, node n_2, changing from phase P_1 to phase P_2 at $T(0)$–$T(1)$, becomes sub-super node. Node n_3, changing from phase P_1 to phase P_2 at $T(8)$–$T(9)$, becomes sub-super node, and when changing from phase P_2 to phase P_3 at $T(13)$–$T(14)$, becomes super node. Node n_4, when changing from phase P_1 to phase P_2 at $T(0)$–$T(1)$, becomes sub-super node, and when changing from phase P_2 to phase P_3 at $T(4)$–$T(5)$, becomes super node.

Analysis shows that the degree values of different nodes undergo phase transition at different times, which is consistent with the fact that different individuals become the central source of information at different times in the spreading process, thus obtaining the attention of a large number of followers.

The degree distributions of the simulated experimental networks and the real network are visualized as shown in Fig. 5.



(a) Experimental network (b) Real network

Fig. 5. Degree distribution comparison of experimental network and real network. (a) Degree distribution of experiment network. (b) Degree distribution of real network.

In Fig. 5, the right color bar from lighter to darker represents the size of the degree value, and the lighter the color is and the smaller the node degree value is. Real network is

built from real data, which is calculated by weibo hotspot system and provided by Sina, with the searching keyword Xi'an Mercedes–Benz Oil Spill from April 1, 2019, to May 8, 2019.

As shown in Fig. 5, there are super nodes and sub-super nodes in the experimental networks, and phase transition has been experienced in node degree value, which is consistent with the phenomenon that the followers of the parties or individuals related will undergo phase transition as the popularity of the event increases in real social networks. Therefore, the evolution processes of the experimental networks are consistent with that of the real networks.

3.2. Phase transition in node influence

Different members are known to have different influences on social networks, and there are some very influential ones, based on which experiments are conducted to simulate how influential members appear in social networks.

Definition 11 (phase transition function of influence) The phase transition function of node influence is defined as

$$\psi = \begin{cases} \psi_1, & I_t(i) \le \vartheta_1(i,t), \\ \psi_2, & \vartheta_1(i,t) < I_t(i) \le \vartheta_2(i,t), \\ \psi_3, & I_t(i) > \vartheta_2(i,t), \end{cases} \tag{13}$$

which describing the different node phases under different influence values. $I_t(i)$ represents the influence value of node i at the time t. $\vartheta_1(i,t) = \max\{3I_0(i), I_\Sigma(t), I_M(t)\}$ is the threshold value of phase ψ_1 changing to phase ψ_2. $\vartheta_2(i,t) = \max\{3nI_0(i), nI_\Sigma(t), nI_M(t)\}$ is the threshold value of phase ψ_2 changing to phase ψ_3, and n takes the value of 10. $I_0(i)$ represents the influence value of node i in the initial network, while $I_\Sigma(t) = \sum_{i=1}^{N} I(i,t)/N$ is the average of the influence values of N nodes at time t. $I_M(t)$ is the median of the influence values of N nodes at time t, where $I_1(t) < I_2(t) < \cdots < I_n(t)$

$$I_M(t) = \begin{cases} I_{(n+1)/2}(t), & \text{when } n \text{ is odd number}, \\ \dfrac{I_{n/2}(t) + I_{(n/2)+1}(t)}{2}, & \text{when } n \text{ is even number}. \end{cases} \tag{14}$$

The influence change of nodes in the experimental network is shown in Fig. 6.

In Fig. 6, n_1, n_2, n_3, n_4 represent node 422, 4525, 2112, 4494 in the networks, respectively. In the color bar on the right, the influence from small to large corresponds to the change from pink to red.

From Fig. 6(a), the influence value of node n_1 satisfies $I_t(n_1) \le \vartheta_1(n_1,t)$ at time $T(0)$–$T(3)$, while satisfies $\vartheta_1(n_1,t) < I_t(n_1) \le \vartheta_2(n_1,t)$ at time $T(4)$. Therefore, the influence of node n_1 changes from phase ψ_1 to phase ψ_2 at $T(3)$–$T(4)$. Likewise, the influence of node n_2 changes from phase ψ_1 to phase ψ_2 at $T(3)$–$T(4)$. The influence of node n_3 changes from phase ψ_1 to phase ψ_2 at $T(3)$–$T(4)$, while changes from phase ψ_2 to phase ψ_3 at $T(4)$–$T(5)$. The influence of node n_4 changes from phase ψ_1 to phase ψ_2 at $T(5)$–$T(6)$. The reason for this is that the known continues to gain spreading influence as it spreads information to the newly unknowns. Therefore, the influence value of the known undergoes phase transition at a certain time.

From Fig. 6(b), it can be observed that the number of nodes with small to great influence in the experimental networks decreases sequentially, which is consistent with the phenomenon that most individuals in real social networks have low influence while only a very small number of individuals have high influence.

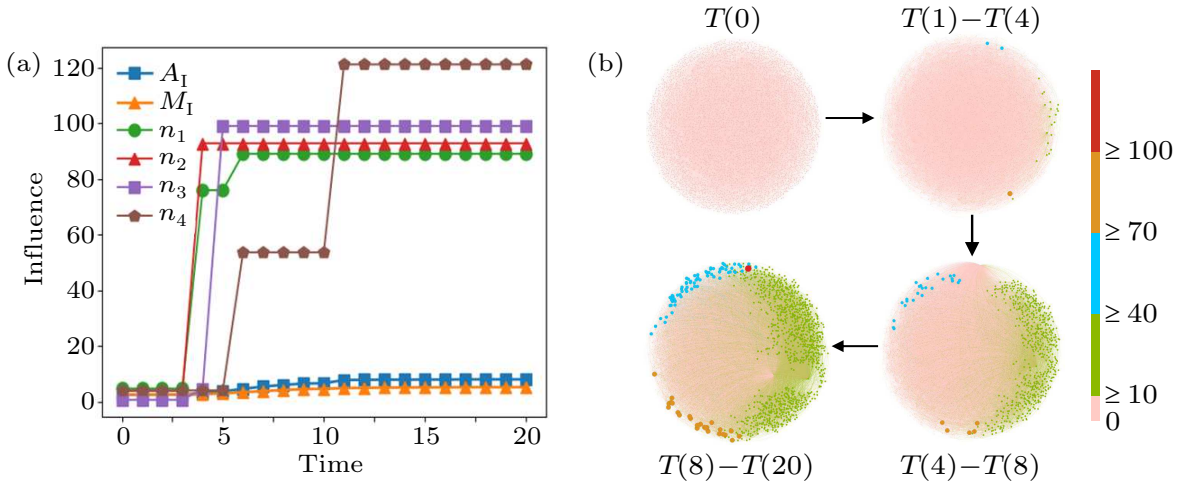


Fig. 6. Node influence changes. (a) Changes of 4 nodes' influence. (b) Changes of all nodes' influence in the networks.

3.3. Phase transition in infection density

Simulations are conducted in small-world networks, scale-free networks, and hybrid networks with the proposed connecting with spreading dynamic evolution model. Different degrees of phase transition in infection density are found. Parameters of the simulation experiments are shown in Ta-

ble 1.

Table 1. Parameters in the experimental model.

Parameter	Definition
T	network type
S	small world network
B	scale-free network
H	hybrid network
N	number of nodes
K	average degree
α	initial infection density
β	rate of infection

To better measure the difference in the degree of phase transition in the infection density under different parameters, the phase transition index of the infection density $\lambda(t)$ is defined as

$$\lambda(t) = i'(t). \tag{15}$$

The N is the total number of individual in the network in Eq. (1), with $S(t)$ and $I(t)$ respectively representing the number of the unknown and the known. $s(t)$ and $i(t)$ represent the infection density of the unknown and the known at time t, respectively, satisfying $N \times s(t) = S(t)$, $N \times i(t) = I(t)$, $s(t)+i(t) = 1$. The ratio of the known at the initial time ($t = 0$) is α. To sum up, we can obtain

$$\frac{\mathrm{d}i(t)}{\mathrm{d}t} = \beta(1-i(t))i(t), \quad i(0) = \alpha. \tag{16}$$

Using separate variables to solve Eq. (16) we obtain

$$i(t) = \frac{1}{1+\left(\frac{1}{\alpha}-1\right)\mathrm{e}^{-\beta t}}. \tag{17}$$

The two most basic derivation rules are

$$(f(x)g(x))' = f'(x)g(x)+f(x)g'(x), \tag{18}$$

$$\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)g(x)-f(x)g'(x)}{(g(x))^2}. \tag{19}$$

From the above equations, the phase transition index $\lambda(t)$ can be derived as

$$\lambda(t) = \frac{\beta\left(\frac{1}{\alpha}-1\right)\mathrm{e}^{-\beta t}}{\left(1+\left(\frac{1}{\alpha}-1\right)\mathrm{e}^{-\beta t}\right)^2}. \tag{20}$$

The maximum value of $\lambda(t)$ without the dynamic evolution rule of connecting with spreading under the corresponding parameters is defined as λ_1. The maximum value of $\lambda(t)$ with the dynamic evolution rule of connecting with spreading under the corresponding parameters is defined as λ_2. The influence of the dynamic evolution rule on the degree of phase transition in infection density is defined as

$$\Delta\lambda = |\lambda_2 - \lambda_1|. \tag{21}$$

3.3.1. Different average degrees

The average degree K is an essential parameter in the networks. Therefore, it is of great importance to study the influence on the degree of phase transition in infection density under different average degrees.

In this section, simulation experiments are conducted in small-world networks and scale-free networks at $N = 20000$, $\alpha = 0.0005$, $\beta = 0.4$ as well as K is equal to 3, 5, and 7, respectively. Simulation results are shown in Fig. 7.

In the following Figs. 7–11, the icon o represents the network without the dynamic evolution rule of connecting with spreading under the corresponding parameters, while the icon f represents the network with the dynamic evolution rule of connecting with spreading under the corresponding parameters.

Figure 7 shows the phase transition differences of infection density when all other conditions are the same but the average degrees are different. From Figs. 7(a) and 7(b), networks with the connecting with spreading dynamic evolution rule show a faster increase in the infection density curve and eventually reach a higher infection density value than that of networks without the connecting with spreading dynamic evolution rule. From Figs. 7(c) and 7(d), it is observed that the changes of $\lambda(t)$ are faster and have higher peaks in the networks with the dynamic evolution rule. From Fig. 7(c), in the small-world networks, when $SK_{\mathrm{o}} = 3$, maximum value 0.0018 is reached at time $T(33)$, while when $SK_{\mathrm{f}} = 3$, maximum value 0.275 is reached at time $T(2)$, thus $\Delta\lambda = 0.2732$ when $SK = 3$ according to Eq. (21). Similarly, the values of $\Delta\lambda$ under other parameters are calculated as shown in Table 2.

Table 2. Values of $\Delta\lambda$ at different network average degrees.

	$SK=3$	$SK=5$	$SK=7$	$BK=3$	$BK=5$	$BK=7$
$\Delta\lambda$	0.2732	0.1802	0.1432	0.09862	0.0716	0.05105

Analyzing $\Delta\lambda$ in Table 2, it is observed that greater phase transition in infection density and faster spreading process are undergone with the dynamic evolution rule of connecting with spreading in the networks with a small average degree.

3.3.2. Different network scales

In real life, social networks are made up of different social communities. And different information is not necessarily spread by the same number of individuals in social networks. Therefore, it is significant to study the influence on the degrees of phase transition in infection density under different network scales.

In this section, simulation experiments are conducted in the small-world networks, scale-free networks, and hybrid networks at $\alpha = 0.0005$, $\beta = 0.4$ while N is 5000 or 20000, respectively. Simulation results are shown in Fig. 8.

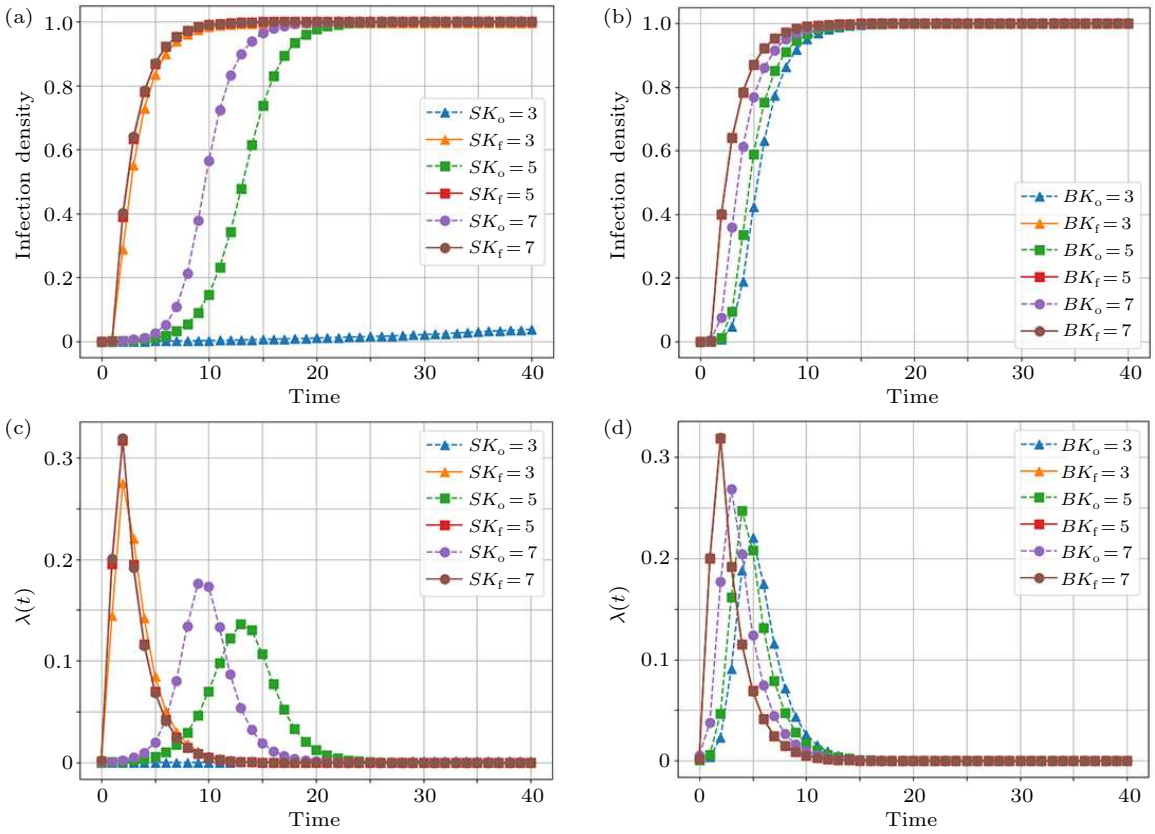


Fig. 7. Changes of infection densities and phase transition indexes under different average degrees. (a) With and without the dynamic evolution rule in small-world networks of different average degrees. (b) With and without the dynamic evolution rule in scale-free networks of different average degrees. Panels (c) and (d) represent the change of phase transition indexes under corresponding conditions of panels (a) and (b), respectively.

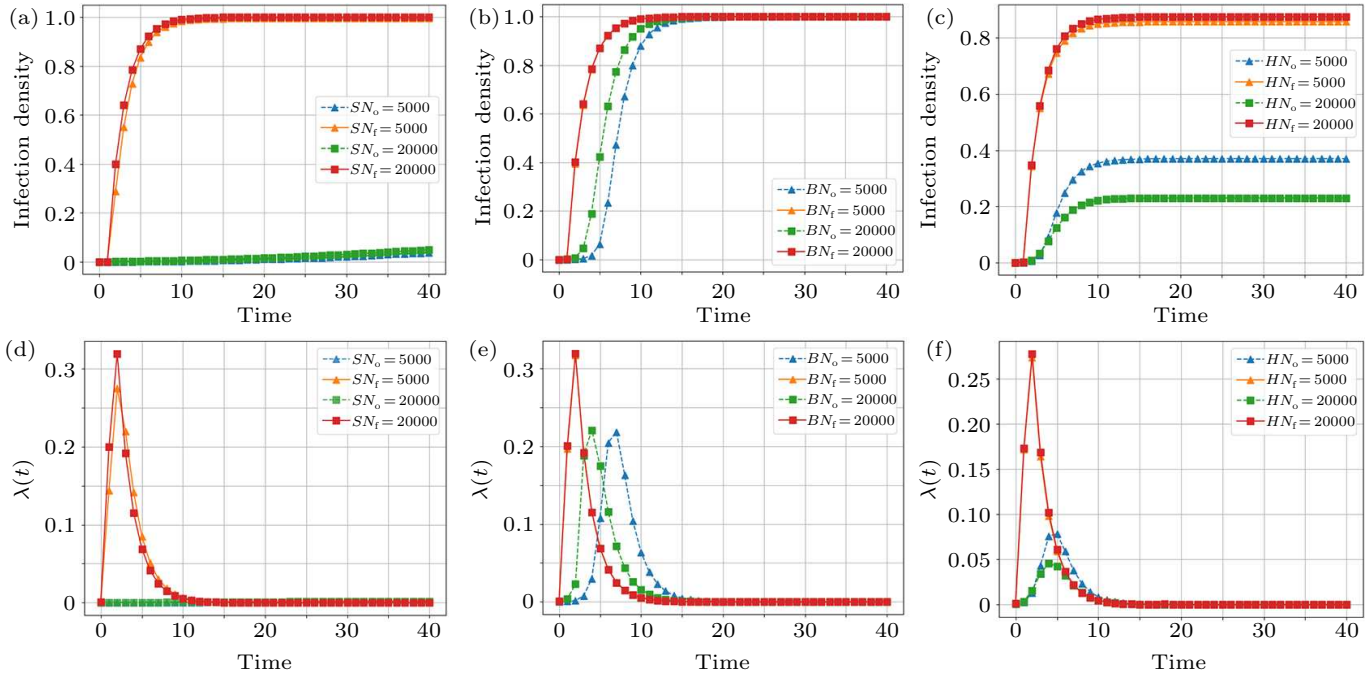


Fig. 8. Changes of infection densities and phase transition indexes under different network scales. (a) With and without the dynamic evolution rule under the small-world networks of different scales. (b) With and without the dynamic evolution rule under the scale-free networks of different scales. (c) With and without the dynamic evolution rule under the hybrid networks of different scales. Panels (d), (e), and (f) represent changes of phase transition indexes under corresponding conditions of panels (a), (b), and (c), respectively.

Table 3. Values of $\Delta\lambda$ at different network scales.

	$SN = 5000$	$SN = 20000$	$BN = 5000$	$BN = 20000$	$HN = 5000$	$HN = 20000$
$\Delta\lambda$	0.2732	0.31753	0.0985	0.09862	0.1952	0.23234

Figure 8 shows phase transition differences in infection density when all other conditions are the same but the network scales are different. From Figs. 8(a)–8(c), it can be observed that infection density curves of networks of different scales increase faster and eventually higher infection density values can be reached with dynamic evolution rule of connecting with spreading. From Figs. 8(d)–8(f), $\lambda(t)$ changes faster and reaches a higher peak with the dynamic evolution rule. From Fig. 8(d), in the small-world networks, when $SN_{\rm o} = 5000$, maximum value 0.0018 is reached at time $T(33)$, while when $SN_{\rm f} = 5000$, maximum value 0.275 is reached at time $T(2)$, thus $\Delta\lambda = 0.2732$ when $SN = 5000$ according to Eq. (21). Similarly, the values of $\Delta\lambda$ under other parameters are calculated as shown in Table 3.

Analyzing $\Delta\lambda$ in Table 3, it is observed that greater phase transition in infection density and faster spreading process is undergone in the networks of relatively large scale with connecting with spreading dynamics evolution rule. In real social networks, information of all kinds often does not spread to the whole network when it is first spread in the social networks, and it needs to be spread several times so that most individuals can get the information. Moreover, the larger social communities are less likely to be aware of the information than the smaller ones at the first time. However, as individuals in the social networks establish new attention relationships, a higher percentage of individuals in the larger social networks is informed of the information in the subsequent spreading process. Therefore, the simulated network spreading processes are more consistent with the real network one.

3.3.3. Different ratios of infection

In this section, the influence of different infection rates on phase transition degrees in infection density is discussed, and simulation experiments are conducted in small-world, scale-free, and hybrid networks with $\alpha = 0.0005$, $N = 20000$ while β is 0.2 and 0.4, respectively. Simulation results are shown in Fig. 9.

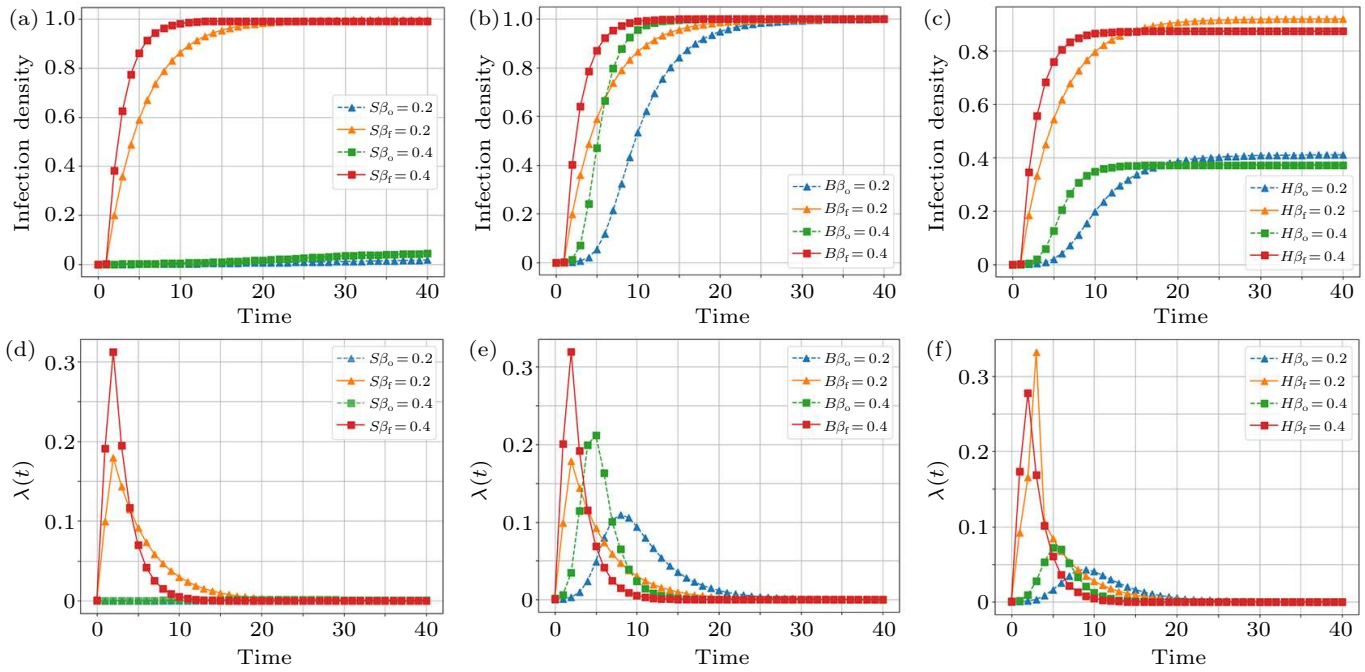


Fig. 9. Changes of infection densities and phase transition indexes under different infection ratios. (a) With and without the dynamic evolution rule under the small-world networks of different infection ratios. (b) With and without the dynamic evolution rule under the scale-free networks of different infection ratios. (c) With and without the dynamic evolution rule under the hybrid networks of different infection ratios. Panels (d), (e), and (f) represent changes of phase transition indexes under corresponding conditions of panels (a), (b), and (c), respectively.

Figure 9 shows phase transition differences in infection density when all other conditions are the same but infection ratios are different. From Figs. 9(a)–9(c), it can be observed that infection density curves under networks of different infection ratios increase faster and eventually higher infection density values can be reached with the dynamic evolution rule of connecting with spreading. From Figs. 9(d)–9(f), $\lambda(t)$ changes faster and reaches a higher peak with the dynamic evolution rule. From Fig. 9(d), in the small-world networks, when $S\beta_{\rm o} = 0.2$, maximum value 0.0006 is reached at time $T(26)$, while when $S\beta_{\rm f} = 0.2$, maximum value 0.1793 is reached at time $T(2)$, thus $\Delta\lambda = 0.1787$ when $S\beta = 0.2$ according to Eq. (21). Similarly, the values of $\Delta\lambda$ under other parameters are calculated as shown in Table 4.

Table 4. Values of $\Delta\lambda$ at different infection ratios.

	$S\beta = 0.2$	$S\beta = 0.4$	$B\beta = 0.2$	$B\beta = 0.4$	$H\beta = 0.2$	$H\beta = 0.4$
$\Delta\lambda$	0.1787	0.3103	0.069	0.10779	0.28887	0.2056

Analyzing $\Delta\lambda$ in Table 4, it is observed that greater phase transition in infection density is undergone with the dynamic evolution rule of connecting with spreading in the small-world networks and the scale-free networks with higher infection ratios, and greater phase transition in infection density is undergone with the dynamic evolution rule of connecting with spreading in hybrid networks with slower infection ratios. The reason for this is that the degrees of evolution are not exactly the same in different types of networks with the dynamic evolution rule of connecting with spreading. The influence in small-world networks and scale-free networks with the connecting with spreading dynamic evolution rule are less than that of the influence of increasing the infection ratio, while in hybrid networks with the connecting with spreading dynamic evolution rule, a larger network evolution process is reached under a relatively smaller infection ratio, thus resulting in larger phase transition in infection density. The latter situation is consistent with real social networks where the infection ratio is small but when attention relationships are larger, the number of individuals knowing the information will also increase more rapidly.

3.3.4. Different initial infection densities

In this section, the influence of different infection densities on phase transition degree in infection density is discussed, and simulation experiments are conducted in small-world, scale-free, and hybrid networks with $\beta = 0.4$, $N = 20000$ while α is equal to 0.01 and 0.0005, respectively. Simulation results are shown in Fig. 10.

Figure 10 shows phase transition differences in infection density when all other conditions are the same but initial infection densities are different. From Figs. 10(a)–10(c), it can be observed that infection density curves under networks of different initial infection densities increase faster and eventually higher infection density values can be reached with the dynamic evolution rule of connecting with spreading. From Figs. 10(d)–10(f), $\Delta\lambda(t)$ changes faster and reaches a higher peak with the dynamic evolution rule. From Fig. 10(d), in the small-world networks, when $S\alpha_o = 0.01$, maximum value 0.0148 is reached at time $T(24)$, while when $S\alpha_f = 0.01$, maximum value 0.3092 is reached at time $T(2)$, thus $\Delta\lambda = 0.2944$ when $S\alpha = 0.01$ according to Eq. (21). Similarly, the values of $\Delta\lambda$ under other parameters are calculated as shown in Table 5.

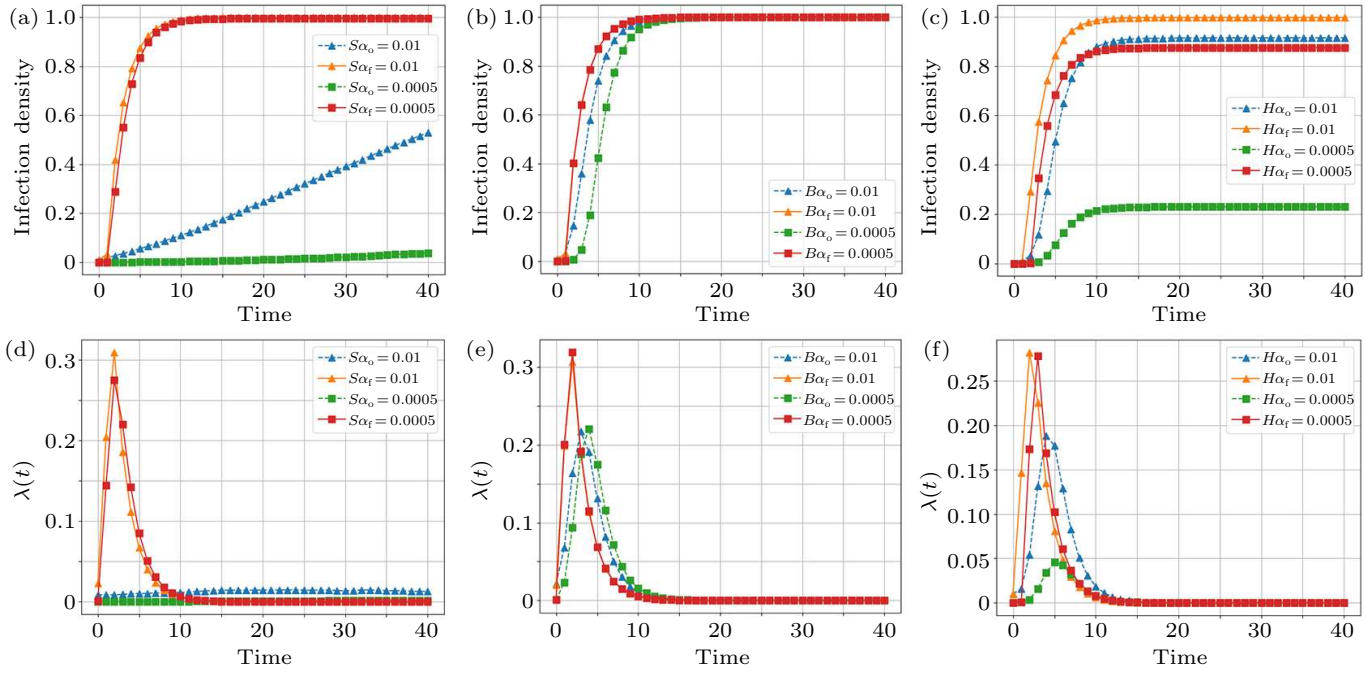


Fig. 10. Changes of infection densities and phase transition indexes under different initial infection densities. (a) With and without the dynamic evolution rule under the small-world networks of different initial infection densities. (b) With and without the dynamic evolution rule under the scale-free networks of different initial infection densities. (c) With and without the dynamic evolution rule under the hybrid networks of different initial infection densities. Panels (d), (e), and (f) represent changes in phase transition indexes under corresponding conditions of panels (a), (b), and (c), respectively.

Table 5. Values of $\Delta\lambda$ at different initial infection densities.

	$S\alpha = 0.01$	$S\alpha = 0.0005$	$B\alpha = 0.01$	$B\alpha = 0.0005$	$H\alpha = 0.01$	$H\alpha = 0.0005$
$\Delta\lambda$	0.2944	0.2732	0.08935	0.09862	0.09418	0.23274

Analyzing $\Delta\lambda$ in Table 5, it is observed that greater phase transition in infection density is undergone with the dynamic rule of connecting with spreading in small-world networks with a larger initial infection density, and greater phase transition in infection density is undergone with the dynamic rule of connecting with spreading in the scale-free networks and

hybrid networks with a smaller initial infection density. This is caused by the differences in the evolution processes under different types of networks. The influence in small-world networks with the connecting with spreading dynamic evolution rule is less than that of the influence of increasing the initial infection density, while in the scale-free networks and the hybrid networks with the connecting with spreading dynamic evolution rule, a larger network evolution process is reached under a relatively smaller initial infection density, thus resulting in larger phase transition in infection density. The latter situation is consistent with real social networks where the initial infection density is small but when attention relationships are larger, the number of individuals knowing the information will also increase more rapidly.

Through the above experiments, it is observed that the infection densities are undergone relatively larger phase transition under different experimental parameters with the dynamic evolution rule of connecting with spreading. In order to better verify the consistency of simulated networks according to this evolution rule with the real social networks, simulation results and real data are compared with each other.

3.3.5. Comparison between simulation results and real data

Based on the above analyses, small-world networks, scale-free networks, and hybrid networks under two different experimental parameters are conducted in this experiment. Spreading processes under these six situations are analyzed to identify spreading process in which situation is closer to that in the real networks. The infection density curves of networks under six different experimental parameters are compared with data of real social networks as shown in Fig. 11.

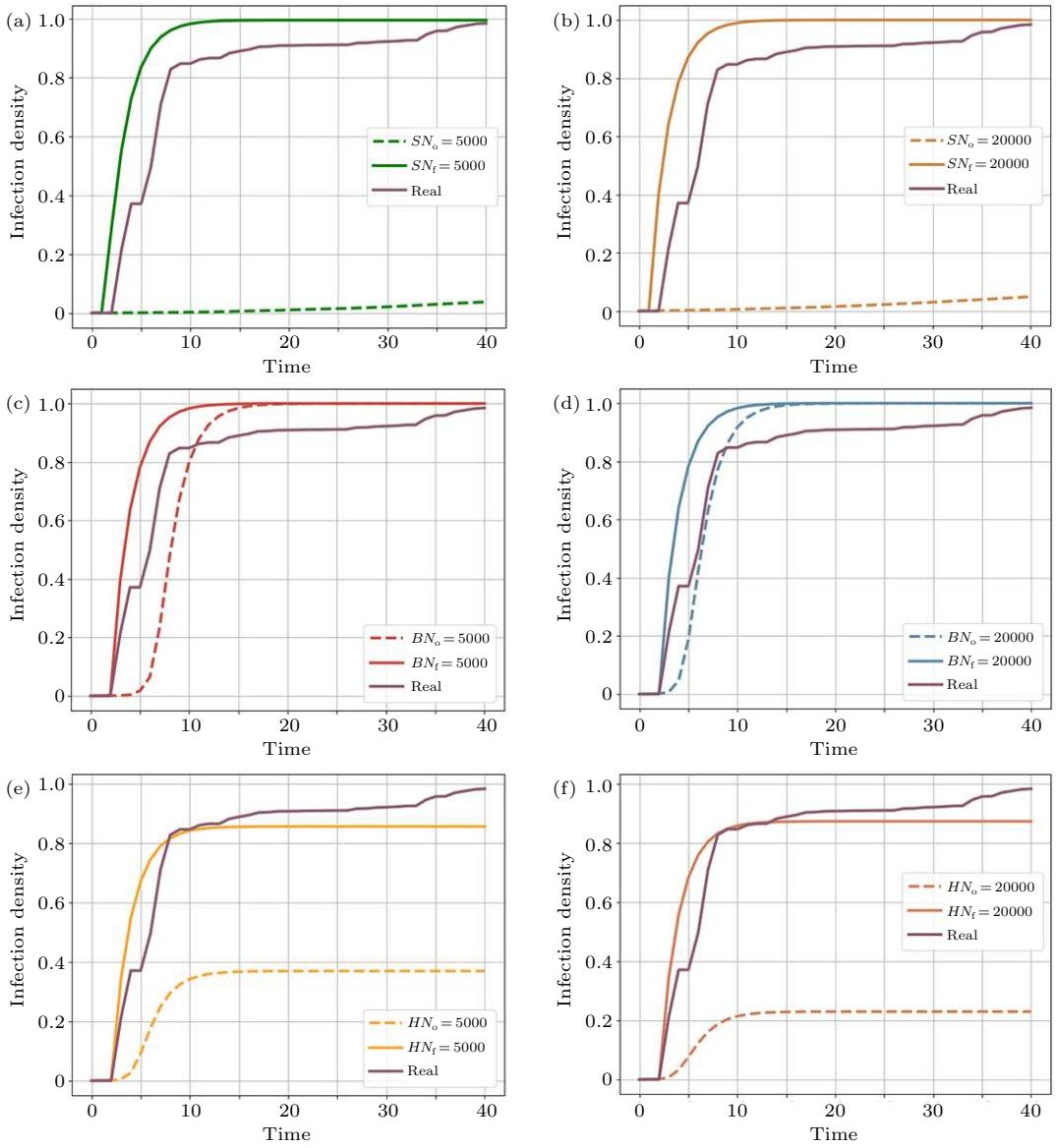


Fig. 11. Comparison between experimental data and real data. Panels (a) and (b) are the comparison between small-world networks with network sizes of 5000 and 20000 respectively and real data. Panels (c) and (d) are the comparison between scale-free networks with network sizes of 5000 and 20000 respectively and real data. Panels (e) and (f) are the comparison between hybrid networks with network sizes of 5000 and 20000 respectively and real data.

In Fig. 11, the initial infection densities under six types of networks are $\alpha = 0.0005$ and infection ratios are $\beta = 0.4$, while real (purple curve) is real data, which is calculated by weibo hotspot system and provided by sina, with the searching keyword Xi'an Mercedes–Benz Oil Spill from April 1, 2019, to May 8, 2019.

From Figs. 11(a) and 11(b), it can be observed that there are certain differences in phase transition points between the infection density curve of the small-world networks with the dynamic evolution rule of connecting with spreading and that of real social networks. And its infection density curve rises too quickly after the phase transition point, which differs from that of real social networks in the latter part of spreading process. From Figs. 11(c) and 11(d), it can be observed that there is little difference in phase transition points between the infection density curve of scale-free networks with the dynamic evolution rule of connecting with spreading and that of real social networks. However, its infection density curve rises too quickly after the phase transition point, which differs from that of real social networks in the latter part of the spreading process. From Figs. 11(e) and 11(f), there is no difference not only in phase transition points between the infection density curve of hybrid networks after evolution and that of real social networks, but also in the latter part of spreading process. In addition, an interesting phenomenon can be found in Fig. 11. Namely, the improvement of adding connecting with spreading rule into the scale-free networks is not as obvious as that in the small world networks and the hybrid networks. This is because the connecting with spreading rule proposed in this paper makes the initial network evolving into a non-uniform network. Therefore, adding this rule leads to the greatest improvement of the small-world network with uniform network characteristics, as well as the smallest improvement of the scale-free network with non-uniform network characteristics, and the improvement of hybrid network consisting of uniform and non-uniform networks is between small-world networks and scale-free networks.

To measure the similarity between experimental data trend and real data trend, the similarity is defined as

$$S(l_1, l_2) = 1 - \frac{\int_0^T |l_1(t) - l_2(t)| \, \mathrm{d}t}{\max\left\{\int_0^T l_1(t)\,\mathrm{d}t, \int_0^T l_2(t)\,\mathrm{d}t\right\}}, \tag{22}$$

where T represents the maximum time after unifying the time axis of curve l_1 and curve l_2. From Eq. (22), the infection density curve similarities between the experimental data and the real data are computed separately in six subplots in Fig. 11, where $S(SN_{\mathrm{f}} = 5000, \mathrm{Real}) = 0.8709$ in Fig. 11(a), $S(SN_{\mathrm{f}} = 20000, \mathrm{Real}) = 0.8589$ in Fig. 11(b), $S(BN_{\mathrm{f}} = 5000, \mathrm{Real}) = 0.8833$ in Fig. 11(c), $S(BN_{\mathrm{f}} = 20000, \mathrm{Real}) = 0.8827$ in Fig. 11(d), $S(HN_{\mathrm{f}} = 5000, \mathrm{Real}) = 0.9712$ in Fig. 11(e), $S(HN_{\mathrm{f}} = 20000, \mathrm{Real}) = 0.9901$ in Fig. 11(f).

Experimental results show that based on the dynamic evolution model of connecting with spreading, the infection density curve trend of the hybrid networks after evolution is closer to that of real social networks than that of the small-world networks and the scale-free networks after evolution.

4. Conclusion

In this paper, experiments on social network evolution processes as information spreading in small-world networks, scale-free networks and hybrid networks, are conducted respectively based on the dynamic evolution model of connecting with spreading. Results show that node degree and influence of experimental networks after evolution undergo phase transition, which is consistent with the phase transition in the individuals' followers and influence under continuous information spreading in real social networks. Moreover, the network topology undergoes phase transition, resulting in a rapid jump in the infection density of experimental networks after evolution compared to that of the initial experimental networks in a short period of time, and reaching phase transition in the infection density. Furthermore, the influence of three types of networks under different experimental parameters with the evolution rule on the degree of phase transition in infection density is investigated. The infection density curves of the experimental networks are compared with the infection density curve of the real data, which shows that the infection density curve trend under hybrid networks with the dynamic evolution rule of connecting with spreading is closer to that under the real social networks than that under the small-world and the scale-free networks with the dynamic evolution rule of connecting with spreading.

By comparing the results of experimental networks with the data of real social networks, the feasibility and rationality of the proposed hybrid networks and the dynamic evolution model of connecting with spreading are demonstrated. Further, it is an essential topic in the field of complex networks to study the evolution of network topology in information spreading and to explore the influence of this evolution mechanism on information spreading in turn. The study on this topic is a great opportunity to master the laws of information spreading so that effective information can be spread throughout the network quickly. However, there still have room for improvement in the dynamic evolution model of connecting with spreading proposed in this paper, as individuals in social networks are influenced by many factors when paying attention to other ones in information spreading. In future undertakings, phase transition in network spreading will be continued to study and explore its general formation mechanism.

Acknowledgments

Project supported by the National Natural Science Foundation of China (Grant Nos. 61863025 and 62266030), Program for International S & T Cooperation Projects of Gansu Province of China (Grant No. 144WCGA166), and Program for Longyuan Young Innovation Talents and the Doctoral Foundation of LUT.

References

[1] Newman M E 2003 *SIAM Rev.* **45** 167
[2] Boccaletti S, Latora V, Moreno Y, Chavez M and Hwang D U 2006 *Phys. Rep.* **424** 175
[3] Wang X F and Chen G 2003 *IEEE Circ. Syst. Mag.* **3** 6
[4] Gallos L K, Song C and Makse H A 2007 *Physica A* **386** 686
[5] Pastor-Satorras R, Castellano C, Van Mieghem P and Vespignani A 2015 *Rev. Mod. Phys.* **87** 925
[6] Arenas A, Díaz-Guilera A, Kurths J, Moreno Y and Zhou C 2008 *Phys. Rep.* **469** 93
[7] Pei S, Muchnik L, Andrade Jr J S, Zheng Z and Makse H A 2014 *Sci. Rep.* **4** 5547
[8] Stieglitz S and Dang-Xuan L 2013 *J. Manage. Inform. Syst.* **29** 217
[9] Karsai M, Kivelä M, Pan R K, Kaski K, Kertész J, Barabási A L and Saramäki J 2011 *Phys. Rev. E* **83** 025102
[10] Li C, Wang L, Sun S and Xia C 2018 *Appl. Math. Comput.* **320** 512
[11] Wang T, Zhou M Y and Fu Z Q 2020 *Chin. Phys. B* **29** 058901
[12] Sun H, Saad D and Lokhov A Y 2021 *Phys. Rev. X* **11** 011048
[13] Leung K, Wu J T and Leung G M 2021 *Nat. Commun.* **12** 1501
[14] Zhao J, Cheng J and Gao H 2014 *Seventh International Joint Conference on Computational Sciences and Optimization* **2014** 325
[15] Lu M, Zhang Z, Qu Z and Kang Y 2018 *IEEE Trans. Knowl. Data Eng.* **31** 1736
[16] Nian F and Dang Z 2018 *Int. J. Mod. Phys. B* **32** 1850106
[17] Nian F, Shi Y and Cao J 2021 *J. Comput. Sci.* **55** 101438
[18] Nian F, Luo L, Yu X and Guo X 2021 *Int. J. Mod. Phys. B* **35** 2150119
[19] Nian F and Liu X 2021 *Appl. Intell.* **52** 889
[20] Wu J, Zheng M, Xu K and Gu C 2020 *Nonlinear Dyn.* **99** 2387
[21] Mheidly N and Fares J 2020 *J. Public Health Pol.* **41** 410
[22] He Z, Cai Z, Yu J, Wang X, Sun Y and Li Y 2017 *IEEE Trans. Veh. Technol.* **66** 2789
[23] Wei X K, Prokhorenko S, Wang B X, Liu Z, Xie Y J, Nahas Y, Jia C L, Dunin-Borkowski R E, Mayer J and Bellaiche L 2021 *Nat. commun.* **12** 1
[24] Perc M 2016 *Phys. Lett. A* **380** 2803
[25] Zurek W H, Dorner U and Zoller P 2005 *Phys. Rev. Lett.* **95** 105701
[26] Kosterlitz J M 2017 *Rev. Mod. Phys.* **89** 040501
[27] Canabarro A, Fanchini F F, Malvezzi A L, Pereira R and Chaves R 2019 *Phys. Rev. B* **100** 045129
[28] Wang W, Liu Q H, Liang J, Hu Y and Zhou T 2019 *Phy. Rep.* **820** 1
[29] Boccaletti S, Almendral J A, Guan S, Leyva I, Liu Z, SendiA-Nadal I, Wang Z and Zou Y 2016 *Phys. Rep.* **660** 1
[30] Li C, Liu F and Li P 2018 *Discrete Dyn. Nat. Soc.* **2018** 1
[31] Xie J, Meng F, Sun J, Ma X, Yan G and Hu Y 2021 *Nat. Hum. Behav.* **5** 1161
[32] Lim S, Jung K and Lui J C 2014 *ACM SIGMETRICS Performance Evaluation Review* **41** 31
[33] Nian F, Yu X, Cao J and Luo L 2020 *Int. J. Mod. Phys. B* **34** 2050203
[34] Davis J T, Perra N, Zhang Q, Moreno Y and Vespignani A 2020 *Nat. Phys.* **16** 590
[35] Zhang S, Wang W, Wu T and Lin T 2019 *Physica A* **534** 122218
[36] Wen X, Yang C, Yang Y P and Chen Y G 2017 *Chin. Phys. Lett.* **34** 058901
[37] Mansouri A and Taghiyareh F 2021 *J. Inform. Syst. Tele.* **9** 1
[38] Holme P and Newman M E 2006 *Phys. Rev. E* **74** 056108
[39] Yu X, Nian F, Yao Y and Luo L 2021 *IEEE Trans. Comput. Soc. Syst.* **8** 1143
[40] Liu H, Zhang Y, Kadir A and Xu Y 2019 *Appl. Math. Comput.* **360** 83
[41] Liu H, Kadir A and Xu C 2020 *Int. J. Bifurcat. Chaos* **30** 2050173
[42] Liu H, Xu Y and Ma C 2020 *Optik* **216** 164925
[43] Watts D J and Strogatz S H 1998 *Nature* **393** 440
[44] Wang X F and Chen G 2002 *Int. J. Bifurcat. Chaos* **12** 187
[45] Barabási A L 2009 *Science* **325** 412
[46] Holme P and Kim B J 2002 *Phys. Rev. E* **65** 026107
[47] Charness G, Haruvy E and Sonsino D 2007 *J. Econ. Behav. Organ.* **63** 88
[48] Boguná M, Pastor-Satorras R, Díaz-Guilera A and Arenas A 2004 *Phys. Rev. E* **70** 056122
[49] Perc M 2014 *J. R. Soc. Interface* **11** 20140378