ISSN 1674-1056
http://cpb.iphy.ac.cn
http://iopscience.iop.org/cpb
中国物理 B
Chinese Physics
中国物理
Chinese Physics B

Modeling for heterogeneous multi-stage information propagation networks and maximizing information

Ren-Jie Mei(梅人杰), Li Ding(丁李), Xu-Ming An(安栩明), Ping Hu(胡萍)

Citation: Chin. Phys. B . 2019, 28(2): 028702. **doi:** 10.1088/1674-1056/28/2/028702
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

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

Subtle role of latency for information diffusion in online social networks

Fei Xiong(熊菲), Xi-Meng Wang(王夕萌), Jun-Jun Cheng(程军军)

Chin. Phys. B . 2016, 25(10): 108904. **doi:** 10.1088/1674-1056/25/10/108904

Ising model on evolution networks and its application on opinion formation

Zhu Xiao-Long, Zhang Hai-Tian, Sang Jian-Ping, Huang Sheng-You, Zou Xian-Wu

Chin. Phys. B . 2014, 23(6): 068701. **doi:** 10.1088/1674-1056/23/6/068701

A 2-stage strategy updating rule promotes cooperation in the prisoner's dilemma game

Fang Xiang-Sheng, Zhu Ping, Liu Run-Ran, Liu En-Yu, Wei Gui-Yi

Chin. Phys. B . 2012, 21(10): 108702. **doi:** 10.1088/1674-1056/21/10/108702

Modeling for heterogeneous multi-stage information propagation networks and maximizing information*

Ren-Jie Mei(梅人杰), Li Ding(丁李)†, Xu-Ming An(安栩明), and Ping Hu(胡萍)

School of Electrical Engineering and Automation, Wuhan University, Wuhan 430072, *China*

(Received 5 October 2018; revised manuscript received 19 November 2018; published online 10 January 2019)

In this paper, we propose a heterogeneous multi-stage model to study the effect of social reinforcement on information propagation. Both heterogeneity of network components and the heterogeneity of individual reinforcement thresholds are considered. An information outbreak condition is derived, according to which the outbreak scale and individual density of each state under specific propagation parameters can be deduced. Monte Carlo experiments are conducted in Facebook networks to demonstrate the outbreak condition, and we find that social reinforcement effects generally inhibit the propagation of information though it contributes to the emergence of certain hot spots simultaneously. Additionally, by applying Pontryagin's Maximum Principle, we derive the optimal control strategy in the case of limited control resources to maximize the information propagation. Then the forward–backward sweep method is utilized to verify its performance with numerical simulation.

Keywords: heterogeneous network, social reinforcement, multi-stage, optimal resource allocation

PACS: 87.23.Ge, 89.70.–a, 87.55.de, 89.75.–k **DOI:** 10.1088/1674-1056/28/2/028701

1. Introduction

Information propagation on networks is ubiquitous in our daily life. The network of information circulation is extremely complex due to the diversity of information itself and the multiple ways of its propagation.[1–4] The formation of complex social networks as well as the aggregation and dissemination of various pieces of information on the network have produced many macro-social phenomena,[5,6] which have deeply affected human life and cannot be well described by a simple contagion model.[7] Over decades, a large number of scholars have been keen to explore the dynamic mechanism and key factors of information propagation by investigating various propagation phenomena, such as the spreading of rumors,[8,9] opinions,[10] fads and innovations.[11] A variety of factors that affect information propagation are also explored, such as network topology,[12,13] individual characteristics,[14–17] and social environment.[9,18] These studies continuously enrich our cognition of propagation phenomenon.

Despite this, many macroscopic phenomena of propagation caused by individual complexity have not yet been well modeled and explained. The phenomenon of social reinforcement[19,20] is one of them. Centola[21] conducted experiments on the dissemination of health behavior using real social networks, and the result proved that information propagation did have a social reinforcement effect. It can explain a phenomenon such as that an individual is more inclined to accept and disseminate information after being informed several times rather than accepting the information through just one single contact. This form of dissemination is considered to be complex contagion which has been studied by Byungjoon Min using a statistical model in Ref. [22]. There are also other statistical models such as threshold model[23–26] and continuous probability model[27] that were generated to study the reinforcement effect in propagation. However, the propagation dynamics associated with the social reinforcement effect is far from well studied. An author in Refs. [28]–[30] established a multi-stage information propagation model, in which an individual experienced multiple intermediate states as they were repeatedly informed before eventually becoming a spreading individual. This model not only fuses and explains the social reinforcement effect, but also describes the information dissemination in society as a dynamic system. It is worth noting that social reinforcement, but without the heterogeneity of network components and the heterogeneity of individual reinforcement threshold, are considered in the current research on multi-stage information propagation models.

The fundamental purpose of studying information propagation is to discover which factors influence the spreading process, as well as to find out ways to intervene to control the process. A large number of control strategies have been proposed to stop rumors[31] and expand positive influence.[32,33] In general, heuristic control and optimal control are two types of control strategies that are most commonly adopted. The former enables the system to evolve toward the desired objective by adjusting parameters according to current feedback state with a fixed formula, while the latter can maximize or minimize the desired objective function with the optimal control strategy.[34] However, the control strategy of optimal resource allocation proposed to maximize information dissemination in

*Project supported by the National Natural Science Foundation of China (Grant No. 61873194).

†Corresponding author. E-mail: liding@whu.edu.cn

© 2019 Chinese Physical Society and IOP Publishing Ltd

Ref. [35] is of more practical significance because the budget required for control strategies is usually limited. Therefore, considering the optimal allocation of limited control resources in our heterogeneous multi-stage model is a more novel and meaningful exploration.

Motivated by the above analysis, we propose a heterogeneous multi-stage information propagation model for the information propagation process in this paper. The reinforcement threshold that denotes the number of times an individual needs to successfully receive information from a state of ignorance of the information to the contagious state is different from individual to individual. We divide individuals with the same reinforcement attribute into the same group. Based on this model, we study the outbreak scale of propagation about the transmission parameters and individual's stage attributes, and draw the conclusion that the information propagation on the network is usually not easy to break out, but once it does, there will be a high probability of outbreak on a large scale. Finally, in view of the characteristics of the propagation model, we explore the optimal control strategy for different stage individuals to maximize information propagation when control resources are limited and verify its optimality by applying heuristic control at the same time. Our study not only provides useful guidance for practical application, but also reveals different propagation characteristics of heterogeneous individuals from the perspective of optimal control strategies.

The rest of this paper is organized as follows. Section 2 describes the process of establishing a heterogeneous multi-stage model. Section 3 is the analysis of outbreak scale of the proposed information propagation model. The restrained optimal control strategy is studied in Section 4. Finally, we summarize the full text in Section 5.

2. Heterogeneous multi-stage information propagation model

We consider a social network made up of N individuals that can be classified into m groups according to their reinforcement thresholds. Generally, an empirical way of selecting groups is based on individuals clustering analysis. Once the historical behavior characteristics of individuals are obtained and quantified, individuals can be grouped. In our theoretical study, we only concentrate on the heterogeneity of the individual towards a specific behavior. $\mathbb{N}$ denotes the set of all individuals in the network, and $\mathbb{N}=\cup_{i=1}^{m}\mathbb{N}_i$, $0<i\leq m$. $\mathbb{N}_i$ represents the group in which an individual becomes contagious only when it contacts with a contagious one at least i times. The *SIS* compartmental model is generalized which is widely adopted in the study of epidemic spreading. An individual in the I compartment who believes and transmits the information is considered to be a contagious individual, while individuals in the S compartment are considered to be non-contagious. For the individuals of the $\mathbb{N}_i$ group, we further expand the non-contagious state to i stages to represent the number of times the individual receives information. When an individual in the $\mathbb{N}_i$ group successfully receives information $i-1$ times, the individual will move into stage $i-1$ of non-contagious and then transfer into the I compartment as long as it is informed one more time.

The state transition process of the $\mathbb{N}_i$ group member is shown in Fig. 1.

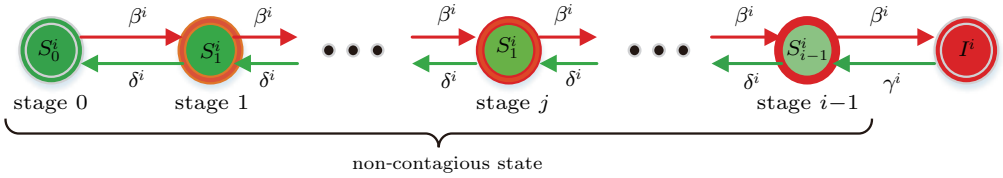


Fig. 1. The state transition process for $\mathbb{N}_i$ group members.

In addition, an individual may also go back to the previous stage with a certain probability due to forgetting or diminishing interest. The state transition dynamics of the $\mathbb{N}_i$ group member is

$$\begin{cases} \dot{S}_0^i(t) = -\beta^i S_0^i(t) I(t) + \delta^i S_1^i(t), \\ \dot{S}_1^i(t) = \beta^i S_0^i(t) I(t) - \beta^i S_1^i(t) I(t) - \delta^i S_1^i(t) + \delta^i S_2^i(t), \\ \quad \cdots, \\ \dot{S}_{i-1}^i(t) = \beta^i S_{i-2}^i(t) I(t) - \beta^i S_{i-1}^i(t) I(t) - \delta^i S_{i-1}^i(t) + \gamma^i I^i(t), \\ \dot{I}^i(t) = \beta^i S_{i-1}^i(t) I(t) - \gamma^i I^i(t), \end{cases} \tag{1}$$

where $S_0^i(t)$ denotes the proportion of $\mathbb{N}_i$ group members in the network that are in the state of stage 0, the superscript of the symbol $S_0^i(t)$ shows individuals' reinforcement thresholds and the subscript shows the exact intermediate stage that individuals are in. β^i denotes the propagation rate that is the probability of receiving the information when a non-contagious individual comes across a contagious one; δ^i denotes the hesitation rate and this means the probability of individuals who are

in non-contagious state tracing back one stage to the previous state. γ^i denotes the attenuation rate and its meaning is similar to δ^i. The difference between δ^i and γ^i is that the former acts on intermediate stages and the latter acts between the contagious state and non-contagious state. $I(t)$ in the model is the sum of all contagious state individuals in the network at time t.

Notice that there exists a group of individuals without intermediate stages ($i=1$ and $\delta^1=0$) in the network, whose state transition process is similar to that of *SIS* model and corresponds to Eq. (2).

$$\begin{cases} \dot{S}_0^1(t) = -\beta^1 S_0^1(t) I(t) + \gamma^1 I^1(t), \\ \dot{I}^1(t) = \beta^1 S_0^1(t) I(t) - \gamma^1 I^1(t). \end{cases} \tag{2}$$

Usually, the extent to which individuals need to be strengthened is different from individual to individual. We assume that there are m groups of individuals that hold different reinforcement attributes and different proportions. Concretely, the proportion of the i-stage group is p_i that satisfies normalized conditions $\sum_{i=1}^m p_i = 1$. Then, the model for heterogeneous multi-stage information propagation networks can be obtained

$$\begin{cases} \dot{S}_0^1(t) = -\beta^1 S_0^1(t) I(t) + \gamma^1 I^1(t), \\ \dot{I}^1(t) = \beta^1 S_0^1(t) I(t) - \gamma^1 I^1(t), \\ \dot{S}_0^2(t) = -\beta^2 S_0^2(t) I(t) + \delta^2 S_1^2(t), \\ \dot{S}_1^2(t) = \beta^2 S_0^2(t) I(t) - \beta^2 S_1^2(t) I(t) - \delta^2 S_1^2(t) + \gamma^2 I^2(t), \\ \dot{I}^2(t) = \beta^2 S_1^2(t) I(t) - \gamma^2 I^2(t), \\ \quad \cdots, \\ \dot{S}_0^m(t) = -\beta^m S_0^m(t) I(t) + \delta^m S_1^m(t), \\ \dot{S}_1^m(t) = \beta^m S_0^m(t) I(t) - \beta^m S_1^m(\mathrm{t}) I(t) - \delta^m S_1^m(t) + \delta^m S_2^m(t), \\ \quad \cdots, \\ \dot{S}_{m-1}^m(t) = \beta^m S_{m-2}^m(t) I(t) - \beta^m S_{m-1}^m(t) I(t) - \delta^m S_{m-1}^m(t) + \gamma^m I^m(t), \\ \dot{I}^m(t) = \beta^m S_{m-1}^m(t) I(t) - \gamma^m I^m(t), \end{cases} \tag{3}$$

where $I(t) = \sum_{i=1}^m I^i(t)$. It tells us that the contagious individual has a chance of transmitting information to any non-contagious individual.

Compared to other models[32,36] of information propagation, such as the threshold model, the continuous probability model and independent cascade model, the multi-stage proposed in our paper between the initial non-contagious state and the contagious state presents us a continuous process of social reinforcement effects and it provides additional details for us to study the intermediate process of propagation. Moreover, the heterogeneous reinforcement and forgetting mechanisms of individuals are also the focus of our study, which contain richer characteristics of human beings and contribute to a more reasonable model for describing and studying information propagation under the social reinforcement effect.

3. Outbreak scale of the information propagation

3.1. Theoretical analysis

From previous study, it is well known that either in virus transmission or in information propagation, the propagation rate needs to be large enough to allow viruses or information to break out. In this section, we explore the outbreak conditions of information propagation based on our proposed model Eq. (3).

Firstly, we provide the following theorem about the steady state of $I(t)$.

Theorem 1 The steady state of $I(t)$ in Eq. (3) is the zero or the positive real solutions sets of the following equation

$$I(\infty) = \frac{p_1}{1+\dfrac{\gamma^1}{\beta^1 I(\infty)}} + \frac{p_2}{1+\left(1+\dfrac{\delta^2}{\beta^2 I(\infty)}\right)\dfrac{\gamma^2}{\beta^2 I(\infty)}} + \cdots + \frac{p_m}{1+\left[\sum\limits_{j=0}^{m-1}\left(\dfrac{\delta^m}{\beta^m I(\infty)}\right)^j\right]\dfrac{\gamma^m}{\beta^m I(\infty)}}, \tag{4}$$

where $\sum_{i=1}^m p_i = 1$.

Proof The propagation dynamics model for individuals in $\mathbb{N}_i$ group is Eq. (1).

Let the derivative of each state in Eq. (1) equal zero. The relationship between the density of contagious individuals and non-contagious ones in the same group can be obtained as

$$\begin{cases} S_j^i(\infty) = \dfrac{\delta^i}{\beta^i I(\infty)} S_{j-1}^i(\infty), \\ S_{i-1}^i(\infty) = \dfrac{\gamma^i}{\beta^i I(\infty)} I^i(\infty), \end{cases} \tag{5}$$

where $j = 1, 2, \ldots, i-1$.

Since the individual density of each state in the group satisfies $S_0^i(t) + S_1^i(t) + \cdots S_j^i(t) + \cdots S_{i-1}^i(t) + I^i(t) = p_i$. Therefore, we derive the proportion of contagious individuals in the i-stage group, which is expressed as follows:

$$I^i(\infty) = \frac{p_i}{1+\left[1+\frac{\delta^i}{\beta^i I(\infty)}+\left(\frac{\delta^i}{\beta^i I(\infty)}\right)^2+\cdots+\left(\frac{\delta^i}{\beta^i I(\infty)}\right)^{i-1}\right]\frac{\gamma^i}{\beta^i I(\infty)}}. \tag{6}$$

There are m groups of individuals in the network, and thus $I(t) = \sum_{i=1}^{m} I^i(t)$. This leads to the expression of Eq. (4). The proof of Theorem 1 is completed.

Obviously, the zero is a trivial solution to Eq. (4), which means that zero is always an equilibrium of system Eq. (3). Besides, the proportion of contagious individuals $I(\infty)$ can be solved. Propagation will break out if the solution of Eq. (4) is a positive real number and the scale of the outbreak is exactly this solution. This result may shed light on revealing the relationship between propagation parameters, the outbreak conditions and the steady state of the outbreak scale.

Next, we provide numerical examples to present the impact of parameters on the scale of propagation.

3.2. Numerical experiments

Firstly, we consider that information propagates on a homogeneous network that is composed of only one single group of individuals, which means $\mathbb{N} = \mathbb{N}_i$, $0 < i \leq m$. We study whether individuals with intermediate stages have influence on the propagation, and how much effect individuals with different stages have on the propagation. Three cases of simulations are given in Table 1.

Table 1. Three cases of information propagation on homogeneous network.

	p_1	p_2	p_3	γ^1	γ^2	γ^3	δ^1	δ^2	δ^3
Case 1	1	0	0	0.06	0	0	0	0	0
Case 2	0	1	0	0	0.06	0	0	0.06	0
Case 3	0	0	1	0	0	0.06	0	0	0.06

As shown in Fig. 2(a), the outbreak scale of propagation continuously changes with the propagation rate β in case 1. However, when the individual has multiple stages, the first order transition occurs between the outbreak scale and the propagation rate shown in the case 2 and case 3. The greater the number of intermediate stages of the individuals, the larger the propagation rate β has to be when the outbreak occurs, and the longer time is also required for the outbreak. In general, it will be more difficult for an outbreak to occur under the effect of social reinforcement.

Next, another numerical experiment is carried out on heterogeneous networks to show the effect of propagation parameters in Eq. (4) on the final scale of outbreak. The network is composed of three groups of individuals with different stage attributes. They are $\mathbb{N} = \sum_{i=1}^{3} \mathbb{N}_i, i = 1,2,3$ respectively and their proportions in the networks correspond to p_1, p_2, p_3. Three sets of proportions of the three groups are adopted in Table 2 to analyze the impact of the heterogeneity of the network on the propagation. For simplicity and without loss of generality, we assume $\beta^1 = \beta^2 = \beta^3$, $\gamma^1 = \gamma^2 = \gamma^3$, $\delta^2 = \delta^3$. The simulation results are shown in the Fig. 3, Fig. 4, and Fig. 5.

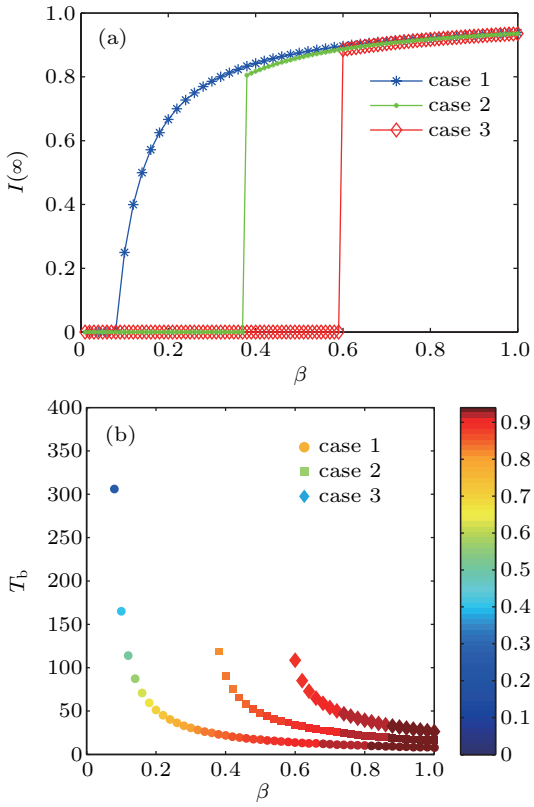


Fig. 2. The outbreak scale and outbreak time are plotted against the propagation rate β in three cases. (b) T_b denotes the time required for the system to reach 95% of the steady state value.

The common causal relationship in the three simulations is that the outbreak scale of the propagation is positively correlated with the transmission rate β and negatively correlated with the attenuation rate γ and hesitation rate δ. From the comparison between Fig. 3, Fig. 4, and Fig. 5, we can derive the influence of the heterogeneity of multi-stage individuals on the propagation. The more individuals with multi-stage attributes (Fig. 4 and Fig. 5 compared with Fig. 3) and the more stages of individuals (Fig. 5 compared with Fig. 4), the smaller the propagation parameter space will be which enables the propagation break out to a medium scale (medium propagation scale refers to the yellow–green area in Fig. 3, Fig. 4 and Fig. 5). We can conclude that the social reinforcement effect leads to the phenomenon of polarization propagation that means the propagation tends to die out, or break out with enormous scale.

Table 2. Three sets of a proportion of three group individuals on the heterogeneous network.

	p_1	p_2	p_3
Set 1	0.6	0.2	0.2
Set 2	0.2	0.6	0.2
Set 3	0.2	0.2	0.6

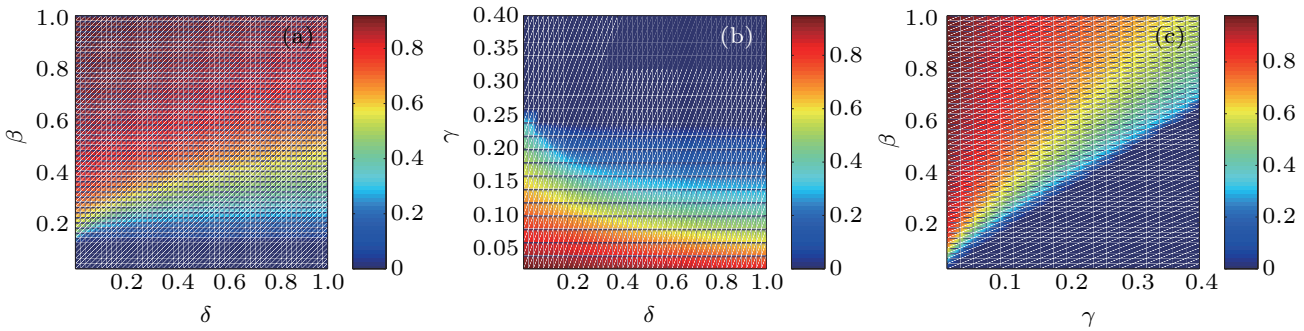


Fig. 3. The outbreak scale as a function of propagation parameters β, δ, and γ under the condition of Set 1. (a) $\gamma^1 = \gamma^2 = \gamma^3 = 0.08$, (b) $\beta^1 = \beta^2 = \beta^3 = 0.4$, (c) $\delta^2 = \delta^3 = 0.08$.

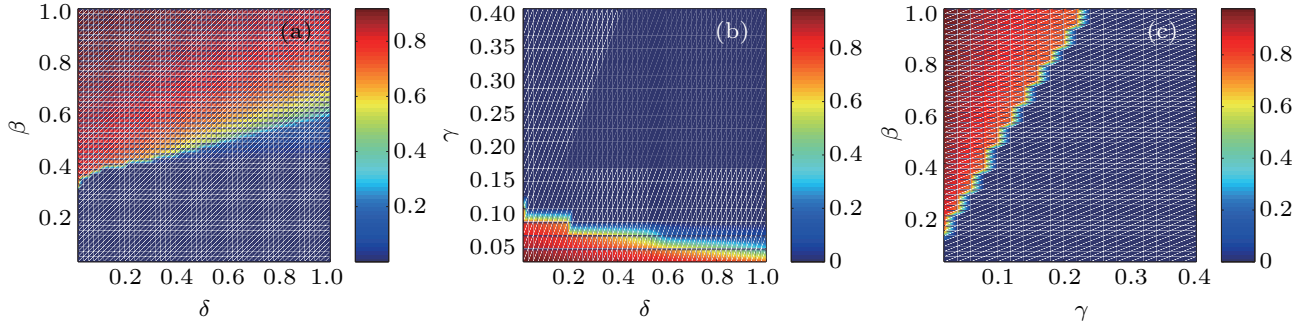


Fig. 4. The outbreak scale as a function of propagation parameters β, δ, and γ under the condition of Set 2. (a) $\gamma^1 = \gamma^2 = \gamma^3 = 0.08$, (b) $\beta^1 = \beta^2 = \beta^3 = 0.4$, (c) $\delta^2 = \delta^3 = 0.08$.

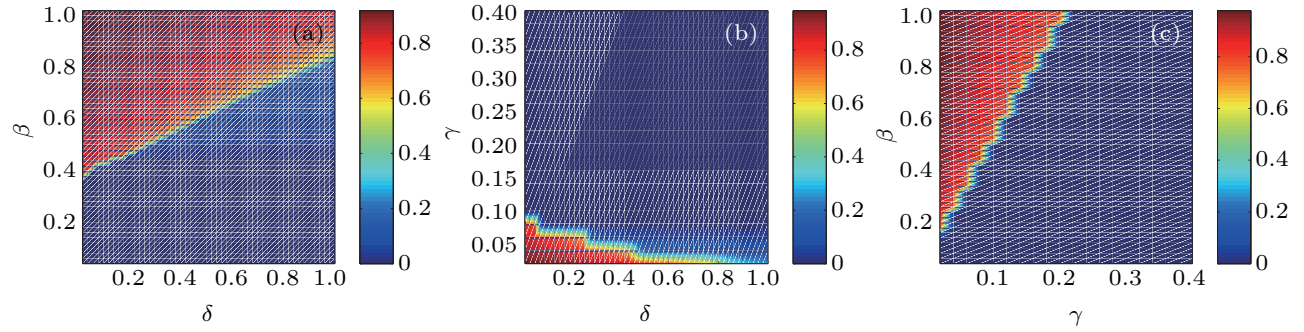


Fig. 5. The outbreak scale as a function of propagation parameters β, δ, and γ under the condition of Set 3. (a) $\gamma^1 = \gamma^2 = \gamma^3 = 0.08$, (b) $\beta^1 = \beta^2 = \beta^3 = 0.4$, (c) $\delta^2 = \delta^3 = 0.08$.

The characteristics of this model provide an explanation of the fact that it is only a few messages that hit the headlines although plenty of messages propagate around. Social reinforcement effects inhibit the spread of information generally though it contributes to the emergence of certain hot spots at the same time.

3.3. Monte-Carlo experiments

In this section, we analyze our multi-stage propagation mechanism on real social networks. We extract the topology of the Facebook network with 4039 nodes and 88234 edges in total.[37] All individuals are randomly divided into three groups and the proportions of the three groups are the same as those in Table 2. The proportion I_0 of initial seed nodes is also randomly selected in each group. At every time, we consider that each node can only communicate with one of its neighbors, and a non-contagious individual will transfer into the next stage with a probability of β if it communicates with a contagious one. When $\delta = \gamma = 0.08$ is adopted, the variation of outbreak scale against propagation rate is shown in Fig. 6. Each result of the Monte Carlo experiment is averaged 30 times run.

As we can see, the experiment's results are similar to the theoretical results and the polarization characteristic of multi-stage propagation mechanism is still obvious although there are some errors when the model is applied to real networks. The causes of errors are roughly divided into two categories. One is the experimental error of simulation, such as the length of simulation time and the limitation of network size. The other is caused by the simplified condition of the model. Our model assumes that the network scale is infinite and does not involve a specific network structure.

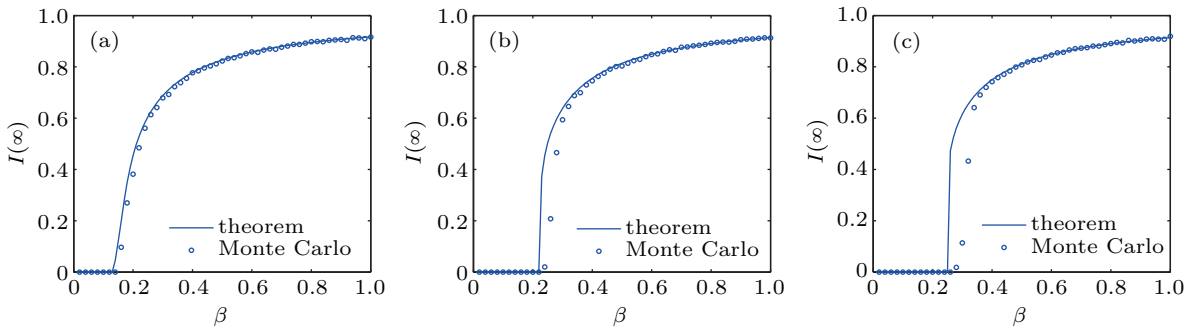


Fig. 6. The outbreak scale is plotted against the propagation rate β in three sets as shown in Table 1. (a) Set 1, (b) Set 2, (c) Set 3. $\gamma^i = 0.08$, $\delta^i = 0.08$. $i = 1, 2, 3$.

4. Maximize the propagation of information

4.1. Optimal control strategy with limited control resources

The variation of propagation parameters or network topology will change the spreading dynamics. From the perspective of control theory, it can be considered that the propagation process is intervened by controlling the propagation parameters or network structure, and thus the propagation result is controlled. Heuristic control and optimal control are two commonly utilized control strategies. While the former is not aimed at a clear objective function, the latter is.

In this section, we consider how to achieve the goal of activating the contagious individuals as much as possible by the optimal control of the attenuation rate under the propagation dynamics of Eq. (3) when the control resources are limited within time interval $[0, T]$.

The problem can be stated as the following optimization program

$$\max J = \int_0^T \sum_{i=1}^{m} \omega_i I^i(t)\,\mathrm{d}t, \tag{7}$$

where $\omega_i > 0$ denotes the relative importance of the i-stage individuals, the value of ω_i is positively related to the importance of the i-stage individuals.

The cost constraint is

$$\int_0^T \sum_{i=1}^{m} f_i\left(\gamma^i(t)\right)\mathrm{d}t \le C_0, \tag{8}$$

where $f_i\left(\gamma^i(t)\right) = \frac{1}{2} h_i p_i \left(\gamma^i(t) - \gamma_0^i\right)^2$ is the control cost function of $\mathbb{N}_i$ group. h_i , $\gamma^i(t)$ and γ_0^i denote the cost coefficient, the control strategy and the natural attenuation rate of that group respectively. C_0 denotes the resource budget.

The state constraints satisfy Eq. (3).

We know that in the context of the above problems, the objective function will reach the maximum when the available control resources are fully used. In other words, when the problem (Eq. (7) and Eq. (8)) is solved, the equation $C = C_0$ must be satisfied, $C = \int_0^T \sum_{i=1}^{m} f_i\left(\gamma^i(t)\right)\mathrm{d}t$.

Next, we refer to the conversion method of the literature,[35] multiply the cost constraint Eq. (8) by a multiplier and add it to the objective function Eq. (7) to form a new objective function.

$$\begin{aligned}\max J = & \int_0^T \sum_{i=1}^{m} \omega_i I^i(t)\,\mathrm{d}t \\ & - \theta\left(\int_0^T \sum_{i=1}^{m} \frac{1}{2} h_i p_i \left(\gamma^i(t) - \gamma_0^i\right)^2 dt - C_0\right),\end{aligned} \tag{9}$$

where $\theta > 0$ is defined as a cost auxiliary variable multiplier.

State constraints satisfy Eq. (3).

Obviously, the objective function Eq. (9) is equivalent to the objective function Eq. (7), and Eq. (8) if and only if $C = C_0$. In the subsequent simulation study, we approximate it as $C = C_0$ when the condition $|C - C_0| \le 10^{-3}$ is satisfied.

Defining

$$\boldsymbol{\Lambda}(t) = \left[\lambda_{1,1}(t)\ \lambda_{1,2}(t)\ \lambda_{2,1}(t)\ \lambda_{2,2}(t)\ \lambda_{2,3}(t)\ \cdots \lambda_{m,1}(t)\ \lambda_{m,2}(t) \cdots \lambda_{m,m}(t)\ \lambda_{m,m+1}(t)\right]^{\mathrm{T}}$$

and

$$\boldsymbol{ST}(t) = \left[S_0^1(t)\ I^1(t)\ S_0^2(t)\ S_1^2(t)\ I^2(t) \cdots S_0^m(t)\ S_1^m(t) \cdots S_{m-1}^m(t)\ I^m(t)\right]^{\mathrm{T}}$$

as the vector of costate variable and system state. By applying the Pontryagin Maximum Principle (PMP), we can derive the following Hamiltonian function of the problem (9)

$$H\left(I^i(t), \lambda_{i,j}(t), \gamma^i(t)\right) = \sum_{i=1}^{m} \omega(i) I^i(t) + \lambda_{1,1}(t)\left(-\beta^1 S_0^1(t) I(t) + \gamma^1 I^1(t)\right) + \sum_{i=2}^{m} \lambda_{i,1}(t)\left(-\beta^i S_0^i(t) I(t) + \delta^i S_1^i(t)\right)$$

$$+\sum_{i=3}^{m}\sum_{j=2}^{i-1}\lambda_{i,j}(t)\left(\beta^i S^i_{j-2}(t)I(t)-\beta^i S^i_{j-1}(t)I(t)-\delta^i S^i_{j-1}(t)+\delta^i S^i_j(t)\right)$$
$$+\sum_{i=2}^{m}\lambda_{i,i}(t)\left(\beta^i S^i_{i-2}(t)I(t)-\beta^i S^i_{i-1}(t)I(t)-\delta^i S^i_{i-1}(t)+\gamma^i I^i(t)\right)$$
$$+\sum_{i=1}^{m}\lambda_{i,i+1}(t)\left(\beta^i S^i_{i-1}(t)I(t)-\gamma^i I^i(t)\right)-\theta\sum_{i=1}^{m}h_i p_i\frac{\left(\gamma^i(t)-\gamma^i_0\right)^2}{2}, \tag{10}$$

where $m>2$.

Adjoint equations are

$$\begin{cases}\dfrac{\mathrm{d}\lambda_{i,1}}{\mathrm{d}t}=-\dfrac{\partial H}{\partial S^i_0}=(\lambda_{i,1}-\lambda_{i,2})\beta^i I(t), & i=1,2,\ldots,m,\\ \dfrac{\mathrm{d}\lambda_{i,j}}{\mathrm{d}t}=-\dfrac{\partial H}{\partial S^i_{j-1}}=\left(\lambda_{i,j}-\lambda_{i,j+1}\right)\beta^i I(t)+\left(\lambda_{i,j}-\lambda_{i,j-1}\right)\delta^i, & i=2,\ldots,m,\ j=2,\ldots,i,\\ \dfrac{\mathrm{d}\lambda_{i,i+1}}{\mathrm{d}t}=-\dfrac{\partial H}{\partial I^i}=-\omega_i+(\lambda_{i,i+1}-\lambda_{i,i})\gamma^i+\displaystyle\sum_{k=1}^{m}\sum_{j=1}^{k}\left(\lambda_{k,j}-\lambda_{k,j+1}\right)\beta^i S^i_{j-1}, & i=1,2,\ldots,m.\end{cases} \tag{11}$$

Denote $\gamma^{i*}(t)$ as the optimal control strategy for the i-stage individuals, $i=1,2,\ldots,m$. Without loss of generality, we assume that control variable $\gamma^{i*}(t)$ is bounded in $[0,\gamma^i_0]$. The corresponding functions of system state and costate variables are denoted as $\boldsymbol{ST}^*(t)$ and $\boldsymbol{\Lambda}^*(t)$ respectively. Equation $J(\gamma^{i*})=\max\{J(\gamma^i)\},i=1,2,\ldots,m$ is satisfied. Then PMP implies that the optimal control at time t satisfies

$$\begin{aligned}&\max_{\gamma^i(t)}H\left(I^{i*}(t),\lambda^*(t),\gamma^i(t),t\right)\\ &=H\left(I^{i*}(t),\lambda^*(t),\gamma^{i*}(t),t\right),\end{aligned} \tag{12}$$

and we derive

$$\begin{aligned}\left.\frac{\partial H}{\partial\gamma^i(t)}\right|_{\gamma^i=\gamma^{i*}}&=-\theta h_i p_i\left(\gamma^{i*}(t)-\gamma^i_0\right)\\ &\quad+(\lambda_{i,i}-\lambda_{i,i+1})I^{i*}=0,\end{aligned}$$

$$\Rightarrow\ \gamma^{i*}(t)=\min\left\{\max\left\{0,\frac{1}{\theta^*\cdot h_i p_i}(\lambda_{i,i}-\lambda_{i,i+1})I^{i*}+\gamma^i_0\right\},\gamma^i_0\right\}, \tag{13}$$

where θ^* denotes the optimal auxiliary variable multiplier.

The transversality conditions

$$\lambda^*_{i,j}(T)=0, \tag{14}$$

where $i=1,2,\ldots,m$, $j=1,2,\ldots,i+1$.

4.2. Simulations

In this section, we will provide a numerical experiment and analysis of the optimal control strategy Eq. (13) in the case of limited control resource.

Table 3. Algorithm: Forward and backward sweep algorithm for optimal control with limited control resources.

Input: $\theta_{\text{low}},\theta_{\text{high}}$, C_0, T, N_{sweep}, p_k, $\boldsymbol{ST}(0)$, $\boldsymbol{\Lambda}(T)$, $\beta^i,\delta^i,\gamma^i_0,i\in\mathrm{M},j=1,\ldots,i+1,t\in[0,T]$
Output: Optimal control strategy $\gamma^{i*}(t)$, optimal auxiliary variable multiplier θ^*.

$\Delta_{\mathrm{C}}\leftarrow\min\{10^{-3}\times C_0,10^{-5}\}$
repeat
  $\theta^*\leftarrow(\theta_{\text{low}}+\theta_{\text{high}})/2$
  repeat
    Calculate $\boldsymbol{ST}^*(t)$ using state equation (3) with initial value $\boldsymbol{ST}(0)$.
    Calculate $\boldsymbol{\Lambda}^*(t)$ with Eq. (12) with transversality Conditions $\boldsymbol{\Lambda}(T)=0$.
    Calculate the optimal control strategy $\gamma^{i*}(t),i=1,2,\ldots,m$, by Eq. (13).
  until error $\le\Delta_{\mathrm{S}}$ or Circle=N_{sweep}
  $C\leftarrow\frac{1}{2}\int_0^{\mathrm{T}}\sum_{i=1}^{m}h_i p_i\left(\gamma^i(t)-\gamma^i_0\right)^2\mathrm{d}t$
  if $C>C_0$ **then**
    $\theta_{\text{low}}\leftarrow\theta^*$
  end if
  if $C<C_0$ **then**
    $\theta_{\text{high}}\leftarrow\theta^*$
  end if
until $|C-C_0|<\Delta_{\mathrm{C}}$

The optimal control strategy obtained above is only the structure of the optimal solution, and the analytical solution of the control function cannot be obtained. We apply the forward–backward sweep algorithm that is shown in Table 3 to make the optimal control strategy play a role in the simulation.

Denote Δ_S and Δ_C as the state error tolerance and cost error tolerance respectively. N_{sweep} is the times of inner loop. The value of optimal auxiliary variable multiplier θ^* is adjusted by the bisection algorithm (initialized by appropriate guesses for minimum and maximum values for θ^*) until the cost constraint Eq. (8) is satisfied with desired accuracy.

When the individual's ability to disseminate information is weak, the propagation cannot break out in a natural way, which can be concluded from Fig. 7(a).

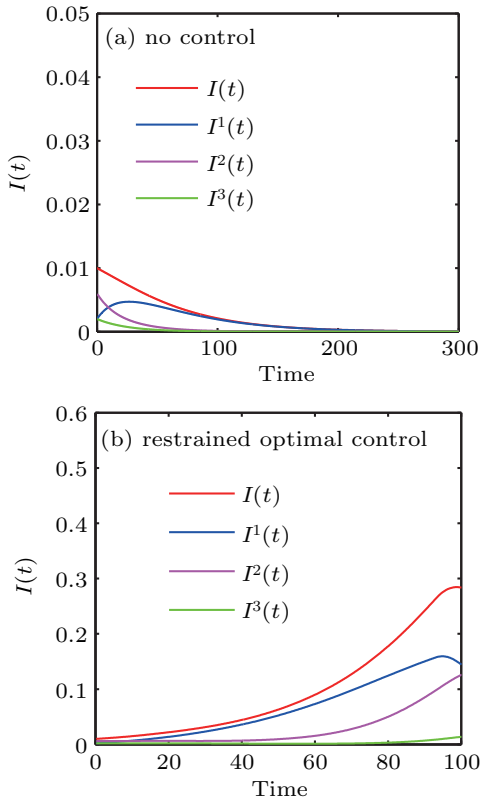


Fig. 7. The evolution of information propagation under no control and restrained optimal control. (a) Propagation eventually stops when no control is applied. Graph (b) shows the evolution of propagation in each group under the optimal control strategy when the control resources $C_0 = 6$ are fully used.

Consider a propagation network $\mathbb{N} = \cup_{i=1}^{3}\mathbb{N}_i$, and let the parameters $p_1 = 0.2$, $p_2 = 0.6$, $p_3 = 0.2$, all individuals are randomly divided into three groups. $h_1 = h_2 = h_3 = 1$ and $\beta^1 = \beta^2 = \beta^3 = 0.2$, $\delta^2 = \delta^3 = 0.08$, $\gamma_0^1 = 0.06, \gamma_0^2 = 0.05, \gamma_0^3 = 0.04$. At the beginning of propagation, one percent of individuals are randomly selected as contagious ones in each group.

In order to study the performance of the optimal control strategy, we introduce a heuristic control strategy that contains direct feedback of the current propagation level $I^i(t)$, *i.e.*,

$$\gamma^i(t) = \gamma_0^i \left(I^i(t)/p_i\right)^{\alpha_i}, i = 1,2,3,$$

where α_i denotes the nonlinear feedback strength. Two sets of parameters of the heuristic control strategy are adopted. Let $\alpha_1 = 0.3, \alpha_2 = 0.27, \alpha_3 = 1$ as the heuristic control strategy 1 (HC 1) and $\alpha_1 = 1.1, \alpha_2 = 1.1, \alpha_3 = 1.1$ as the heuristic control strategy 2 (HC 2). The corresponding propagation evolution of three different control strategies in both numerical experiments and Monte Carlo experiments are shown in Fig. 8.

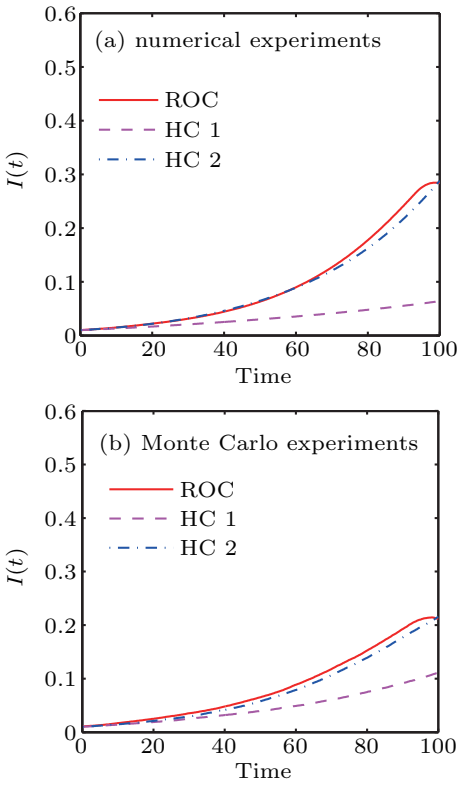


Fig. 8. The evolution of propagation in the network under three different control strategies. Heuristic control strategies 1 and 2 are plotted with dotted line and dash-dot lines. (b) Monte Carlo experiments are conducted on the Facebook network with 4039 nodes and 88234 edges in total. The cause of errors between numerical experiments and Monte Carlo experiments is similar to that in Subsection 3.3.

Table 4. Compare the size of the outbreak and the control resources consumed under different control strategies.

	Restrained optimal control	Heuristic control 1	Heuristic control 2
Consumed resources	6.000	6.1062	10.5932
Numerical $I(T)$	0.2841	0.0639	0.2880
Monte Carlo $I(T)$	0.2115	0.1114	0.2156

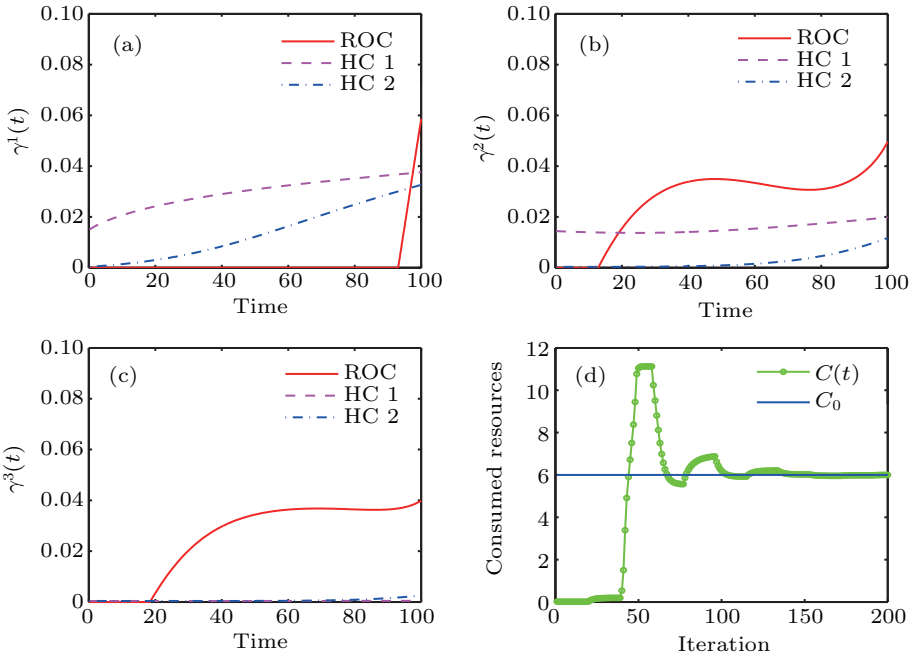


Fig. 9. The optimal and heuristic control strategies are displayed in panels (a)–(c), and panel (d) is the iteration process of adjusting the consumed control resources when we look for an optimal control strategy that satisfies the conditions.

As revealed in Fig. 8, both the optimal control strategy and heuristic control strategy have an important impact on the propagation process and steady state value of contagious individuals. The variation of the attenuation rate $\gamma^{i*}(t)$, $i = 1,2,3$ within the control time $[0,T]$ changes the equilibrium state of the dynamic system. However, the control performance of the two control methods has significant differences. From Fig. 8 and Table 4, it is not surprising to find that a larger propagation scale is achieved by the optimal control strategy than that by the heuristic control under the same mount of control resources. In turn, when the propagation scale is equal, less control resources are required for the optimal control than the other.

From the perspective of the optimal control strategy itself, the optimal control strategies for different groups of individuals reflect the different propagation properties of these groups. Comparing the optimal control strategies of the three groups in Fig. 9, we notice that the control resources are used to the maximum extent for the first group and $\gamma^1 = 0$ almost throughout the control time. However, γ^1 rise rapidly to the natural attenuation rate $\gamma^1 = 0.06$ near the fixed final time T. This is because there is no need for such investment of control when almost all the individuals of $\mathbb{N}_1$ group have been transformed into contagious states and the space for conversion tends to be saturated, as shown in Fig. 7(b). In sharp contrast to the $\mathbb{N}_1$ group, until the fixed final time T, lots of individuals in the $\mathbb{N}_2$ and $\mathbb{N}_3$ groups with less control intensity are still in the non-contagious state, and the space for investment and conversion are large. Such a method of resource allocation may reflect the different importance of the three group individuals from the point of view of propagation efficiency. As there is no intermediate state for the $\mathbb{N}_1$ group individuals, in which the non-contagious individual can transform into the contagious state in at least one step, its conversion is inevitably faster and more efficient than individuals in other groups.

5. Conclusion

This paper considers the social reinforcement effect of different groups of individuals and establishes a heterogeneous multi-stage information propagation model that presents a rich propagation dynamics phenomenon. Through the outbreak condition theorem, we can deduce whether the propagation will break out and its outbreak scale as long as the parameters of the system are given. Another conclusion about the propagation model is that the higher the proportion of individuals with multiple intermediate states, the more difficult it is for an outbreak of propagation. What is more, the scale of the outbreak shows a significant polarization. In other words, either it does not break out or the scale of the outbreak is large with a high probability.

Analysis of the model by applying the Pontryagin's Maximum Theorem, the optimal control strategy obtained under the condition of limited resources has achieved the purpose of promoting and maximizing information propagation. Obviously, the optimal strategy is more efficient in resource utilization compared with the heuristic control strategy. The strategy of optimal resource allocation is of great practical significance for the wide dissemination of opinions, innovations and products. We believe that the outcomes of this paper can provide diverse inspirations for further study of multi-stage propagation problems, and establish the theoretical basis for the uti-

lization and control of information propagation on the social networks.

References

[1] Huang W M, Zhang L J, Xu X J and Fu X 2016 *Sci. Rep.* **6** 23766
[2] Xing Q B, Zhang Y B, Liang Z N and Zhang F 2011 *Chin. Phys. B* **20** 120204
[3] Wu Y H, Deng S and Huang H B 2013 *Commun. Nonlinear Sci. Num. Simul.* **18** 3469
[4] Feng Y, Ding L, Huang Y H and Guan Z H 2016 *Chin. Phys. B* **25** 128903
[5] Onnela J P and Reed-Tsochas F 2010 *Proc. Natl. Acad. Sci. USA* **107** 18375
[6] Doer B, Fouz M and Friedrich T 2012 *Commun. ACM* **55** 70
[7] Czaplicka A, Toral R and San Miguel M 2016 *Phys. Rev. E* **94** 062301
[8] Zhu L and Zhao H 2017 *Int. J. Syst. Sci.* **48** 2064
[9] Li D D and Ma J 2017 *Physica A* **469** 284
[10] Pires M A and Crokidakis N 2017 *Physica A* **467** 167
[11] Krapivsky P L, Redner S and Volovik D 2011 *J. Stat. Mech-Theory E* **2011** P12003
[12] Liu T, Li P, Chen Y and Zhang J 2016 *PLoS One* **11** e0152021
[13] Chen W and Jia Z 2015 *Proceedings of the* 2015 *International Conference on Intelligent Systems Research and Mechatronics Engineering*, April 11–13, 2015, Zhengzhou, China, p. 347
[14] Kang L, and Xiang B B, Zhai S L, Bao Z K and Zhang H F 2018 *Acta Phys. Sin.* **67** 198901 (in Chinese)
[15] Liu H L, Huang Y L, Luo C H and Hu H B 2016 *Acta Phys. Sin.* **65** 158901 (in Chinese)
[16] Xiong F, Wang X M and Cheng J J 2016 *Chin. Phys. B* **25** 108904
[17] Deng S and Li W 2017 *Phys. Rev. E* **95** 042306
[18] Xiao Y P, Li S Y and Liu Y B 2017 *Acta Phys. Sin.* **66** 030501 (in Chinese)
[19] Ma J, Li D and Tian Z 2016 *Physica A* **447** 108
[20] Zheng M, Lü L and Zhao M 2013 *Phys. Rev. E* **88** 012818
[21] Centola D 2010 *Science* **329** 1194
[22] Min B and San Miguel M 2018 *Sci. Rep.* **8** 10422
[23] Wang W, Tang M, Shu P and Wang Z 2016 *NJPh* **18** 013029
[24] Wang W, Tang M, Zhang H F and Lai Y C 2015 *Phys. Rev. E* **92** 012820
[25] Zhou C, Zhao Q and Lu W 2015 *PLoS One* **10** e0140556
[26] Rossi W S, Como G and Fagnani F 2017 *IEEE Transactions on Network Science and Engineering* **1**
[27] Chung K, Baek Y, Kim D, Ha M and Jeong H 2014 *Phys. Rev. E* **89** 052811
[28] Lin Y, Lui J C S, Jung K and Lim S 2013 *9th International Conference on Signal-Image Technology and Internet-Based Systems*, December 2–5, 2013, pp. 501–508
[29] Guo Q T, Jiang X, Lei Y J, Li M, Ma Y F and Zheng Z M 2015 *Phys. Rev. E* **91** 012822
[30] Melnik S, Ward J A, Gleeson J P and Porter M A 2013 *Chaos* **23** 013124
[31] Huo L A, Lin T, Fan C, Liu C and Zhao J 2015 *Advances in Difference Equations* **2015** 54
[32] Kempe D, Kleinberg J and Tardosé 2003 *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, August 24–27, 2003, New York, USA, p. 137
[33] Zhang H Y, Dinh T N, and Thai M T 2013 *IEEE* 33*rd International Conference on Distributed Computing Systems*, July 8–11, 2013, Philadelphia, PA, USA, p. 317
[34] Kandhway K and Kuri J 2017 *IEEE T. Syst. Man. Cy-S* **47** 1099
[35] Kandhway K and Kuri J 2016 *IEEE ACM T. Network* **24** 3204
[36] Lu L Y, Chen D B and Zhou T 2011 *NJPh* **13** 123005
[37] http://cpb.iphy.ac.cn/EN/column/column297.shtml