Subtle role of latency for information diffusion in online social networks*

Fei Xiong(熊菲)[1,2,†], Xi-Meng Wang(王夕萌)[1,2], and Jun-Jun Cheng(程军军)[3]

[1] School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China
[2] Key Laboratory of Communication and Information Systems, Beijing Municipal Commission of Education, Beijing Jiaotong University, Beijing 100044, China
[3] China Information Technology Security Evaluation Center, Beijing 100085, China

(Received 14 April 2016; revised manuscript received 19 May 2016; published online 25 August 2016)

Information diffusion in online social networks is induced by the event of forwarding information for users, and latency exists widely in user spreading behaviors. Little work has been done to reveal the effect of latency on the diffusion process. In this paper, we propose a propagation model in which nodes may suspend their spreading actions for a waiting period of stochastic length. These latent nodes may recover their activity again. Meanwhile, the mechanism of forwarding information is also introduced into the diffusion model. Mean-field analysis and numerical simulations indicate that our model has three nontrivial results. First, the spreading threshold does not correlate with latency in neither homogeneous nor heterogeneous networks, but depends on the spreading and refractory parameter. Furthermore, latency affects the diffusion process and changes the infection scale. A large or small latency parameter leads to a larger final diffusion extent, but the intrinsic dynamics is different. Large latency implies forwarding information rapidly, while small latency prevents nodes from dropping out of interactions. In addition, the betweenness is a better descriptor to identify influential nodes in the model with latency, compared with the coreness and degree. These results are helpful in understanding some collective phenomena of the diffusion process and taking measures to restrain a rumor in social networks.

Keywords: information diffusion, node latency, user behavior, complex networks

PACS: 89.75.–k, 87.23.Ge, 89.75.Fb, 05.10.–a **DOI:** 10.1088/1674-1056/25/10/108904

1. Introduction

The process of information diffusion has attracted wide attention. Many researchers have devoted themselves to investigating how information propagates from several initial spreaders to the population.[1,2] Different models have been put forward to describe local behaviors and analyze the dynamic diffusion process in various kinds of networks.[3,4] These models make a connection between local diffusion actions and the macroscopic state. Analysis and simulations have been conducted to reveal the dependence on model parameters.[5,6] As a typical representative, epidemic models reflect the mechanism of infection and recovery for diseases, and can be used to characterize information diffusion, such as the susceptible-infected (SI) model,[7,8] susceptible-infected-susceptible model (SIS),[9] susceptible-infected-refractory model (SIR),[10–12] etc. Rumor models consider that the rumor may be clarified when two spreaders contact each other.[13] The underlying networks are of great importance for the diffusion process, determining the propagation velocity and extent.[14,15] A propagation model in two-layer multiplex networks was presented, and results proved that interactions between layers enhanced the diffusion process.[16] Moreover, in conjoint social-physical networks,[17] the scale of information diffusion is increased dramatically. In Ref. [18], the effect of community structure was studied in the presence of social reinforcement. Meanwhile, some studies extracted the most influential nodes in the diffusion process.[19,20] In Ref. [21], compared with the degree, the coreness constitutes a good topological descriptor to identify influential spreaders in complex networks, if epidemic models are used. When the diffusion process is characterized by rumor models, the coreness does not determine the spreading ability of nodes, but it reflects whether a node can suppress the rumor.[22]

With the development of Web 2.0, online social networks have become one of the most important ways for people to gain and share information. A lot of researches focused on information diffusion in online social networks.[23,24] Power-law degree distribution in online social networks yields an optimal structure for information diffusion.[25] The spread of online user behaviors was explored, and it was found that clustered-lattice networks accelerated behavior spreading.[26] Besides, weak ties also greatly affect the spreading dynamics. In Ref. [27], selecting weak ties preferentially cannot make information diffuse quickly, but they act as bridges connecting isolated local communities. The environment in online social

*Project supported by the National Natural Science Foundation of China (Grant Nos. 61401015 and 61271308), the Fundamental Research Funds for the Central Universities, China (Grant No. 2014JBM018), and the Talent Fund of Beijing Jiaotong University, China (Grant No. 2015RC013).

†Corresponding author. E-mail: xiongf@bjtu.edu.cn

© 2016 Chinese Physical Society and IOP Publishing Ltd

networks is more complex than real society, so researchers try to characterize the elaborate diffusion process.[28] In Ref. [29], the simultaneous diffusion of negative and positive information was considered, and more states of user behaviors were introduced into the diffusion model. Based on the model, in Ref. [30] two methods of restraining rumor spreading were compared, and it was found that the method of the truth clarification had a long-term performance while blocking rumors at the nodes with a large degree only took effect in the early stage. To describe online user decisions and actions, an evolutionary game theoretic framework[31] was proposed to model the diffusion dynamics, and the model could be explained by differential equations.

Current studies often assume that agents focus their attention on the topic all the time, and they continue to forward information and express their ideas. In current models, agents have the ability to spread information once they are infected by it. In fact, even if agents in online social networks are convinced by spreaders, they have no influence on neighbors until they publish or republish the information. After the forwarding event, susceptible neighbors can contact the information. As depicted in Ref. [29], the mechanism of forwarding information for agents was introduced into the rumor model. Their model is only suitable for the environment where two opposite kinds of information coexist, and the forwarding event is individual-related, so the spreading threshold cannot be derived analytically. Furthermore, users in online social networks do not always keep active. After a diffusion action, they may enter into a waiting period of stochastic length where no further action takes place. Therefore, individual latency generally exists in consecutive user behaviors. In Ref. [32], latency of voters was introduced into opinion dynamic models, and results showed that the macroscopic state was changed by latency. In these models it was assumed that agents have initial opinions towards a topic, and information is expected to reach all the population before the dynamics. However, whether latency in information diffusion affects the spreading scale and velocity still needs further exploring, and this problem is important for understanding the real intrinsic diffusion dynamics. In this paper, we propose a diffusion model with node latency. Spreaders may enter into the latent state and do not spread the information temporarily, but they may be reactivated again. We conduct mean-field analysis and numerical simulations of our model. We find that the node latency does not change the spreading threshold, but determines the final diffusion extent.

The rest of this paper is structured as follows. In Section 2, an information diffusion model with node latency for the online social medium is presented. In Section 3, the mean-field analyses of the model in homogeneous networks and heterogeneous networks are conducted. In Section 4 numerical simulations and a discussion about the model are given. Concluding remarks are given in Section 5.

2. Diffusion model with latency

In online social networks, users spread information to others by sending a message or publishing a post. They are not active all the time. After they take a diffusion action, their activity decreases and they temporarily keep away from the diffusion process. With the time elapsing, latent users may be reactivated, and they will interact with neighbors. Therefore, the activity of users is heterogeneous and dynamic, and it affects the process of information diffusion. We introduce latency of nodes into the epidemic model. After a node takes a diffusion action, it enters into a waiting period of stochastic length.

We introduce our diffusion model. In the model, a node may be in one of four potential states, i.e., the susceptible state, the infected state, the latent state, and the refractory state. The susceptible state means that agents do not see any information, and have no idea about the topic. Susceptible nodes may be infected if they receive the information from their neighbors. The infected state means that nodes are convinced of the information and are willing to propagate it. However, they do not have influence on neighbors until they forward the information. Latent users have accepted the information, but they freeze it and do not take diffusion actions. Refractory nodes absolutely lose their interest in the topic, and drop out of interactions.

Nodes interact with neighbors in terms of their states. Each time, infected nodes may send a message to their neighbors with the probability $p_{\rm lat}$ and then enter into the latent state. Susceptible neighbors receive the message and read it, and they may be infected with the spreading probability λ. Meanwhile, the latent nodes may recover their activity and enter into the infected state again with the probability $p_{\rm inf}$, or else, they may become refractory and withdraw from interactions with the probability δ. We provide the state transition graph for an arbitrary node in Fig. 1.

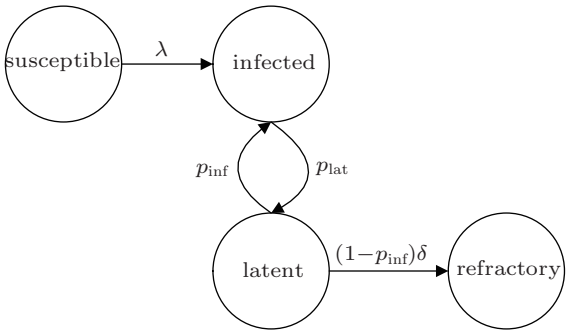


Fig. 1. State transition graph of a node in the topology.

From the above model, after nodes temporarily stop their diffusion actions, they may become active again with the probability $p_{\rm inf}$, and they may also lose their activity permanently

with the probability $(1-p_{\rm inf})\,\delta$. In the model, only the refractory state is stable, and infected and latent nodes will become extinct in the diffusion process. The final system evolves into the regime in which all nodes become refractory or the susceptible and refractory states coexist with different densities. From Fig. 1, the latent state can be treated as an intermediate state between diffusion actions and withdrawal from interactions. Although latent nodes trust the topic, they currently do not diffuse information to others. Therefore, the effect of node latency on the diffusion process cannot be ignored. The threshold of information diffusion is also related to the latent state. We will conduct mean-field analysis of the spreading dynamics.

3. Analytical approach

If the underlying network is a heterogeneous network, node degrees exhibit distribution $P(k)$. The average degree of the network is

$$\bar{k}=\sum_{k=0}^{+\infty}kP(k). \tag{1}$$

The probability that a node with degree k meets a neighbor with degree k' is $P(k'|k)$. The probability represents degree correlations of the network. Apparently, nodes belonging to different connectivity classes have different spreading dynamics. The overall densities of nodes with degree k in the susceptible, infected, latent, and refractory state are defined as $\rho_{\rm s}(k,t)$, $\rho_i(k,t)$, $\rho_{\rm l}(k,t)$, and $\rho_{\rm r}(k,t)$, respectively. We are concerned about the variations of these four densities. $\rho_{\rm s}(k,t)$ decreases all the time since susceptible nodes are infected by receiving the information from neighbors. The probability that a susceptible node with degree k contacts an infected neighbor with degree k' is $P(i_{k'}|s_k)$. In the mean-field approximation, the probability can be written as

$$P(i_{k'}|s_k)\approx\rho_i\left(k',t\right). \tag{2}$$

The degrees of infected neighbors for a susceptible node are heterogeneous, and their densities are often different. For a susceptible node with degree k, the average density of infected neighbors is

$$\begin{aligned}\bar{\rho}_{\rm i}(s_k,t) &= \sum_{k'=0}^{+\infty}P\left(k'|k\right)P(i_{k'}|s_k)\\ &\approx \sum_{k'=0}^{+\infty}P\left(k'|k\right)\rho_{\rm i}\left(k',t\right).\end{aligned} \tag{3}$$

An infected neighbor forwards a message with the probability $p_{\rm lat}$, so the susceptible node is infected by the neighbor in the time interval $[t,t+\Delta t]$ with the probability $\lambda p_{\rm lat}\Delta t$. For a susceptible node with degree k, it receives many messages from neighbors, and therefore, it may be infected with a greater chance. The probability of being infected by k neighbors is defined as $P_{si}(k,t)$, and it is calculated by being multiplied with the probability for a single neighbor, and the resulting expression is as follows:

$$P_{si}(k,t)=1-\left(1-\lambda p_{\rm lat}\Delta t\right)^k. \tag{4}$$

From Eq. (4), the probability for multiple infected sources is the same as that from the independent cascade model that has widely been validated in the process of information diffusion.[33] In the limit $\Delta t\to 0$, $P_{si}(k,t)$ is approached by

$$P_{si}(k,t)=1-\left(1-\lambda k p_{\rm lat}\Delta t\right)=\lambda k p_{\rm lat}\Delta t. \tag{5}$$

Thus, we obtain the transition rate of $\rho_{\rm s}(k,t)$. The variation of $\rho_{\rm s}(k,t)$ during $[t,t+\Delta t]$ is $-\rho_{\rm s}(k,t)\cdot\bar{\rho}_{\rm i}(s_k,t)P_{si}(k,t)$. We have the following differential equation:

$$\begin{aligned}\partial\rho_{\rm s}(k,t)/\partial t &= \frac{\rho_{\rm s}(k,t+\Delta t)-\rho_{\rm s}(k,t)}{\Delta t}\\ &= -\frac{\rho_{\rm s}(k,t)\,\bar{\rho}_{\rm i}(s_k,t)\,P_{si}(k,t)}{\Delta t}\\ &= -\lambda k p_{\rm lat}\rho_{\rm s}(k,t)\sum_{k'=0}^{+\infty}P\left(k'|k\right)\rho_{\rm i}\left(k',t\right).\end{aligned} \tag{6}$$

The density of infected nodes $\rho_{\rm i}(k,t)$ increases because susceptible agents are infected by the information from their neighbors and latent agents recover their activity. $\rho_{\rm i}(k,t)$ decreases as infected agents become latent. After infected nodes forward information, they will enter into the latent state, so the variation of $\rho_{\rm i}(k,t)$ is $p_{\rm lat}\rho_{\rm i}(k,t)$. The variation of latent agents that are reactivated is $p_{\rm inf}\rho_{\rm l}(k,t)$. Therefore, the transition rate of $\rho_{\rm i}(k,t)$ is expressed as

$$\begin{aligned}\partial\rho_{\rm i}(k,t)/\partial t &= \lambda k p_{\rm lat}\rho_{\rm s}(k,t)\sum_{k'=0}^{+\infty}P\left(k'|k\right)\rho_{\rm i}\left(k',t\right)\\ &\quad -p_{\rm lat}\rho_{\rm i}(k,t)+p_{\rm inf}\rho_{\rm l}(k,t).\end{aligned} \tag{7}$$

Analogously, $\rho_{\rm l}(k,t)$ increases when infected agents lose their activity temporarily, but decreases when latent agents become refractory or they are reactivated. The transition rates of $\rho_{\rm l}(k,t)$ and $\rho_{\rm r}(k,t)$ are given by

$$\begin{aligned}&\partial\rho_{\rm l}(k,t)/\partial t=p_{\rm lat}\rho_{\rm i}(k,t)-p_{\rm inf}\rho_{\rm l}(k,t)-(1-p_{\rm inf})\,\delta\rho_{\rm l}(k,t),\\ &\partial\rho_{\rm r}(k,t)/\partial t=(1-p_{\rm inf})\,\delta\rho_{\rm l}(k,t).\end{aligned} \tag{8}$$

We will analyze the threshold of spreading probability below which the information cannot propagate in a network. Homogeneous and heterogeneous networks are used as interaction topology.

3.1. Homogeneous network

First, we consider homogeneous networks to explore basic features of the diffusion model. In the networks, degree fluctuations are very small with the same degree $\bar{k}$, and there are no degree correlations. Therefore, $\rho_{\rm s}(k,t)$, $\rho_{\rm i}(k,t)$, $\rho_{\rm l}(k,t)$

and $\rho_{\rm r}(k,t)$ are simply written as $\rho_{\rm s}(t)$, $\rho_{\rm i}(t)$, $\rho_{\rm l}(t)$, and $\rho_{\rm r}(t)$. Equations (6)–(8) are reduced to

$$\begin{aligned}
&\partial\rho_{\rm s}(t)/\partial t = -\lambda\bar{k}p_{\rm lat}\rho_{\rm s}(t)\rho_{\rm i}(t),\\
&\partial\rho_{\rm i}(t)/\partial t = \lambda\bar{k}p_{\rm lat}\rho_{\rm s}(t)\rho_{\rm i}(t) - p_{\rm lat}\rho_{\rm i}(t) + p_{\rm inf}\rho_{\rm l}(t),\\
&\partial\rho_{\rm l}(t)/\partial t = p_{\rm lat}\rho_{\rm i}(t) - p_{\rm inf}\rho_{\rm l}(t) - (1-p_{\rm inf})\delta\rho_{\rm l}(t),\\
&\partial\rho_{\rm r}(t)/\partial t = (1-p_{\rm inf})\delta\rho_{\rm l}(t).
\end{aligned} \tag{9}$$

Assume that information propagates in the global scope after time τ. When $t<\tau$, information is trapped in the local area. Thus, we have $\rho_{\rm s}(\tau)\approx\rho_{\rm s}(0)$, $\rho_{\rm i}(\tau)\approx\rho_{\rm i}(0)$, and $\rho_{\rm l}(\tau)\approx\rho_{\rm l}(0)$. In the beginning, only one infected node is assigned in the network, so we have the following initial condition:

$$\rho_{\rm s}(0)=1-1/N,\ \ \rho_{\rm i}(0)=1/N,\ \ \rho_{\rm l}(0)=0,\ \ \rho_{\rm r}(0)=0. \tag{10}$$

When information propagates beyond the local region, the density of refractory agents varies beyond 0. In the situation, $\partial\rho_{\rm i}(t)/\partial t>0$ should hold for $t=\tau+\Delta t$. Since $\Delta t\to 0$ and the increase of $\rho_{\rm l}(t)$ is usually later than the increase of $\rho_{\rm i}(t)$, we have $\partial\rho_{\rm l}(\tau+\Delta t)/\partial t\approx 0$. After some elementary manipulations of the third equation in Eq. (9), we have the following equation with the above condition:

$$\rho_{\rm l}(\tau+\Delta t)\approx\frac{p_{\rm lat}}{p_{\rm inf}+(1-p_{\rm inf})\delta}\rho_{\rm i}(\tau+\Delta t). \tag{11}$$

The threshold condition of the second equation in Eq. (9) can also be written as

$$\begin{aligned}
&\partial\rho_{\rm i}(\tau+\Delta t)/\partial t\\
&=\lambda\bar{k}p_{\rm lat}\rho_{\rm s}(\tau+\Delta t)\rho_{\rm i}(\tau+\Delta t) - p_{\rm lat}\rho_{\rm i}(\tau+\Delta t)\\
&\quad + p_{\rm inf}\rho_{\rm l}(\tau+\Delta t)>0.
\end{aligned} \tag{12}$$

Inserting Eq. (11) into Eq. (12), we obtain

$$\lambda\bar{k}p_{\rm lat}\rho_{\rm s}(\tau+\Delta t) - p_{\rm lat} + \frac{p_{\rm lat}p_{\rm inf}}{p_{\rm inf}+(1-p_{\rm inf})\delta}>0. \tag{13}$$

Since $\rho_{\rm s}(\tau+\Delta t)\approx 1$, we find the final condition

$$\lambda\bar{k}>1-\frac{p_{\rm inf}}{p_{\rm inf}+(1-p_{\rm inf})\delta}. \tag{14}$$

From Eq. (14), the threshold of spreading probability correlates with the average node degree $\bar{k}$, and refractory probability δ and the reactivating probability $p_{\rm inf}$. However, the probability of becoming latent $p_{\rm lat}$ does not affect the spreading threshold.

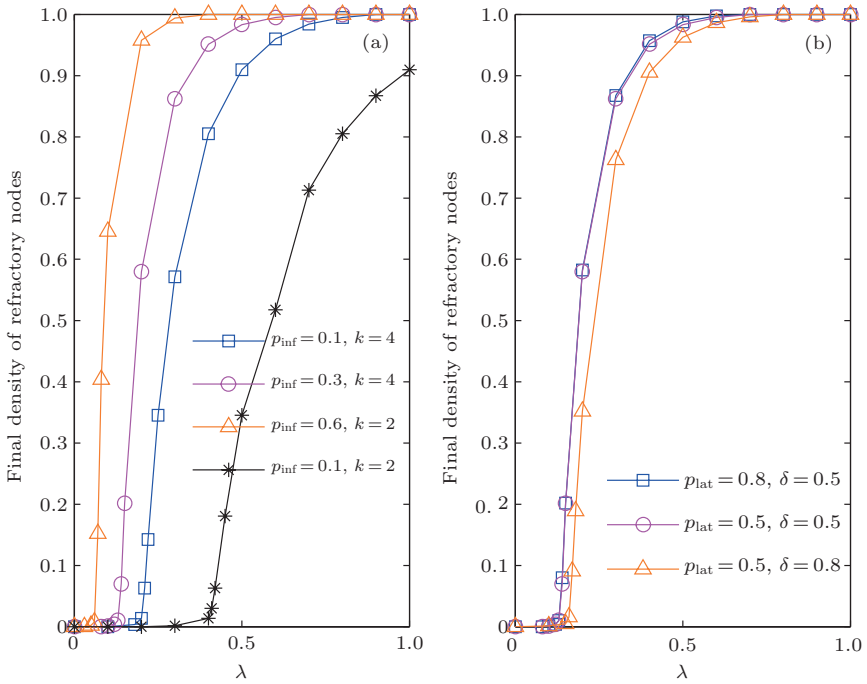


Fig. 2. (color online) Plots of final density of refractory nodes $\rho_{\rm r}(\infty)$ versus λ, given by Eqs. (9) in homogeneous networks, with $\rho_{\rm i}(0)=1/2500$. (a) $\delta=0.5$, and $p_{\rm lat}=0.5$; (b) $k=4$, and $p_{\rm inf}=0.3$.

Figure 2 illustrates the solution of Eq. (9). From Fig. 2, a threshold of spreading probability for information diffusion exists clearly, and the value of the threshold approaches Eq. (14), which is in accordance with our analysis. Above the threshold, $\rho_{\rm r}(\infty)$ increases drastically with λ increasing, leading to a phase transition. In the left panel of Fig. 2, large average degree and reactivating probability $p_{\rm inf}$ advance the diffusion process and reduce the threshold. We notice that when $\bar{k}=4$, the information for $\lambda\geq 0.8$ invades the whole system, but when $\bar{k}=2$, some nodes are still unaware of the information for any λ. As shown in the right panel of Fig. 2, the latency parameter $p_{\rm lat}$ is unrelated to threshold, which is

in consistence with the analysis of Eq. [14]. However, when information can propagate in the system, $\rho_{\rm r}(\infty)$ varies slightly with the latency parameter. The larger refractory probability increases the threshold a little and reduces the diffusion extent, but its effect also depends on $p_{\rm inf}$. When $p_{\rm inf}$ is large enough, δ does not have an obvious influence.

3.2. Heterogeneous network

Now, we consider uncorrelated inhomogeneous networks. The degree correlations of networks can be written as

$$P\left(k'|k\right) = k'P\left(k'\right)/\bar{k}. \tag{15}$$

The overall densities of infected and latent nodes are obtained by averaging the densities over different degrees, and given as

$$\rho_{\rm i}(t) = \sum_{k=0}^{+\infty}\rho_{\rm i}(k,t) \text{and} \rho_{\rm l}(t) = \sum_{k=0}^{+\infty}\rho_{\rm l}(k,t). \tag{16}$$

Integrating Eq. (6) from time 0 to time t, we have

$$\begin{aligned}\rho_{\rm s}(k,t) &= \exp\left(-\lambda k p_{\rm lat}\sum_{k'=0}^{+\infty}P\left(k'|k\right)\int_0^t\rho_{\rm i}\left(k',t'\right)\mathrm{d}t'\right)\\ &= \mathrm{e}^{-\lambda k p_{\rm lat}\varphi(t)},\end{aligned} \tag{17}$$

where

$$\begin{aligned}\varphi(t) &= \sum_{k'=0}^{+\infty}P\left(k'|k\right)\int_0^t\rho_{\rm i}\left(k',t'\right)\mathrm{d}t'\\ &= \frac{1}{\bar{k}}\sum_{k'=0}^{+\infty}k'P\left(k'\right)\int_0^t\rho_{\rm i}\left(k',t'\right)\mathrm{d}t'.\end{aligned} \tag{18}$$

Obviously, the initial value of $\phi(t)$ is $\phi(0)=0$. We multiply the first equation of Eq. (8) with $kP(k)/\bar{k}$ and sum over different k values, and then have the following expression:

$$\begin{aligned}&\frac{1}{\bar{k}}\sum_{k=0}^{+\infty}kP(k)\,\partial\rho_{\rm l}(k,t)/\partial t\\ &= \frac{1}{\bar{k}}\sum_{k=0}^{+\infty}kP(k)\left(p_{\rm lat}\rho_{\rm i}(k,t) - p_{\rm inf}\rho_{\rm l}(k,t) - (1-p_{\rm inf})\,\delta\rho_{\rm l}(k,t)\right)\\ &= p_{\rm lat}\frac{\mathrm{d}\phi(t)}{\mathrm{d}t} - \left(p_{\rm inf}+(1-p_{\rm inf})\,\delta\right)\frac{1}{\bar{k}}\sum_{k=0}^{+\infty}kP(k)\,\rho_{\rm l}(k,t).\end{aligned} \tag{19}$$

We integrate Eq. (19) from time 0 to time t. Noticing that in the final system $\rho_{\rm l}(k,t)=0$ when $t\to\infty$, the integral of Eq. (19) is simplified into

$$\begin{aligned}p_{\rm lat}\phi(\infty) &= \left(p_{\rm inf}+(1-p_{\rm inf})\,\delta\right)\frac{1}{\bar{k}}\\ &\quad\times\sum_{k=0}^{+\infty}kP(k)\int_0^{+\infty}\rho_{\rm l}\left(k,t'\right)\mathrm{d}t'.\end{aligned} \tag{20}$$

Analogously, we do the same sum manipulation of Eq. (7) and integrate it. Considering Eq. (17), we obtain

$$\begin{aligned}\frac{\mathrm{d}\phi(t)}{\mathrm{d}t} &= \frac{\lambda p_{\rm lat}}{\bar{k}}\sum_{k=0}^{+\infty}k^2P(k)\int_0^t\mathrm{e}^{-\lambda k p_{\rm lat}\phi(t')}\frac{\mathrm{d}\phi(t')}{\mathrm{d}t'}\mathrm{d}t' - p_{\rm lat}\phi(t)\\ &\quad+\frac{p_{\rm inf}}{\bar{k}}\sum_{k=0}^{+\infty}kP(k)\int_0^{+\infty}\rho_{\rm l}\left(k,t'\right)\mathrm{d}t'.\end{aligned} \tag{21}$$

In the final system, we have $\mathrm{d}\phi(\infty)/\mathrm{d}t=0$. Thus, inserting Eq. (20), we gain the following equation from Eq. (21).

$$\begin{aligned}0 &= \frac{1}{\bar{k}}\sum_{k=0}^{+\infty}kP(k)\left(1-\mathrm{e}^{-\lambda k p_{\rm lat}\phi(\infty)}\right) - p_{\rm lat}\phi(\infty)\\ &\quad+\frac{p_{\rm inf}p_{\rm lat}\phi(\infty)}{\left(p_{\rm inf}+(1-p_{\rm inf})\,\delta\right)}.\end{aligned} \tag{22}$$

Around the threshold, $\phi(\infty)$ is small. We make the Taylor expansion of $\mathrm{e}^{-\lambda k p_{\rm lat}\phi(\infty)}$ in Eq. (22), and we find

$$\begin{aligned}0 &= \frac{1}{\bar{k}}\left(\lambda p_{\rm lat}\overline{k^2}\phi(\infty) - \lambda^2 p_{\rm lat}^2\overline{k^3}\phi^2(\infty) + o\left(\phi^2(\infty)\right)\right)\\ &\quad- p_{\rm lat}\phi(\infty) + \frac{p_{\rm inf}p_{\rm lat}\phi(\infty)}{\left(p_{\rm inf}+(1-p_{\rm inf})\,\delta\right)}.\end{aligned} \tag{23}$$

Then, we obtain the expression of $\phi(\infty)$ as

$$\phi(\infty)\approx\frac{\lambda p_{\rm lat}\overline{k^2}/\bar{k} - p_{\rm lat} + \dfrac{p_{\rm inf}p_{\rm lat}}{\left(p_{\rm inf}+(1-p_{\rm inf})\,\delta\right)}}{\lambda^2 p_{\rm lat}^2\overline{k^3}/\bar{k}}. \tag{24}$$

Information can propagate in the network if $\phi(\infty)>0$. Thus, we obtain the threshold as

$$\lambda\overline{k^2}/\bar{k} > 1 - \frac{p_{\rm inf}}{\left(p_{\rm inf}+(1-p_{\rm inf})\,\delta\right)}. \tag{25}$$

Since in the final system, $\rho_{\rm i}(k,\infty)=0$ and $\rho_{\rm l}(k,\infty)=0$ hold true, we have

$$\begin{aligned}\rho_{\rm r}(\infty) &= 1-\rho_{\rm s}(\infty) = \sum_{k=0}^{\infty}P(k)\left(1-\rho_{\rm s}(k,\infty)\right)\\ &= \sum_{k=0}^{\infty}P(k)\left(1-\mathrm{e}^{-\lambda k p_{\rm lat}\phi(\infty)}\right).\end{aligned} \tag{26}$$

Therefore, in heterogeneous networks, the threshold of spreading probability is also independent of $p_{\rm lat}$, but relies on δ and $p_{\rm inf}$, which also holds true in homogeneous networks.

4. Simulation results

We conduct Monte-Carlo simulations to find the role of latency in the process of information diffusion, and the spreading thresholds in different networks are also investigated. In the beginning, a node is selected at random and it is set to be in the infected state, while all the other nodes are susceptible. Simulations are implemented synchronously. Each time, all infected nodes make a decision to take actions and spread information to neighbors, and then they may enter into the latent state. All latent nodes may recover their activity or withdraw from interactions. The dynamics continues until no more updates take place. Homogeneous networks and Barabasi–Albert

scale-free networks mediate the diffusion process, and a real social network is also used as the underlying topology.

Figure 3 shows the densities of nodes in different states each as a function of time. The underlying networks are homogeneous networks. The density of latent nodes $\rho_{\rm l}(t)$ stays a little smaller than that of infected nodes $\rho_{\rm i}(t)$, and $\rho_{\rm i}(t)$ increases earlier. Latent nodes and infected nodes can interconvert into each other, and will become extinct together. Increasing λ raises the peaks of $\rho_{\rm l}(t)$, $\rho_{\rm i}(t)$, and $\rho_{\rm r}(\infty)$. For larger λ, $\rho_{\rm l}(t)$ and $\rho_{\rm i}(t)$ increase rapidly in the early stage and reach their peaks within less time, but then they drop relatively slowly. $\rho_{\rm s}(t)$ becomes stable more rapidly than the densities of nodes in other states. Although infected nodes still exist in the system, all of their neighboring nodes enter into the refractory state and $\rho_{\rm s}(t)$ is prevented from decreasing. Larger λ accelerates the evolution of $\rho_{\rm s}(t)$, but it does not reduce the relaxation times of $\rho_{\rm i}(t)$ and $\rho_{\rm r}(t)$.

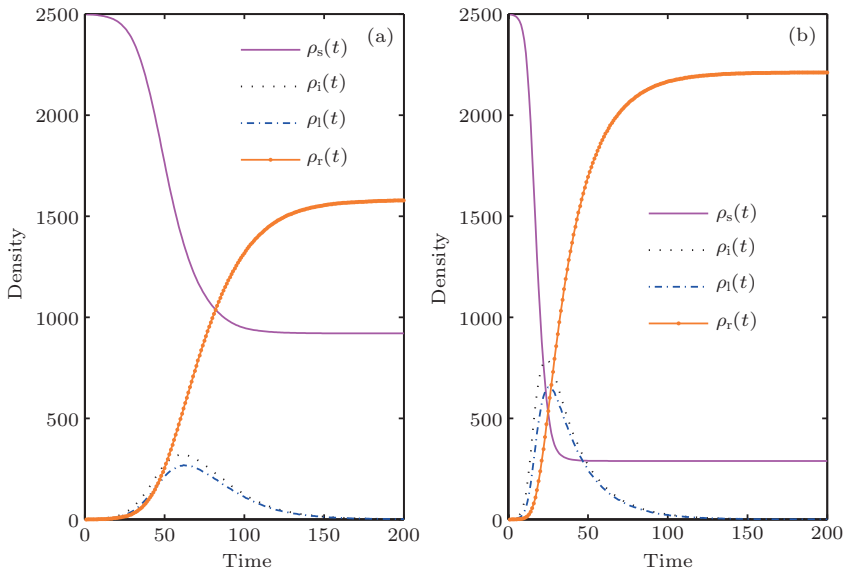


Fig. 3. (color online) Time evolutions of the density of susceptible, infected, latent, refractory nodes in homogeneous networks with the degree $\bar{k} = 4$. The system size is 2500, $\delta = 0.2$, $p_{\rm inf} = 0.5$, and $p_{\rm lat} = 0.5$. (a) $\lambda = 0.2$; (b) $\lambda = 0.5$.

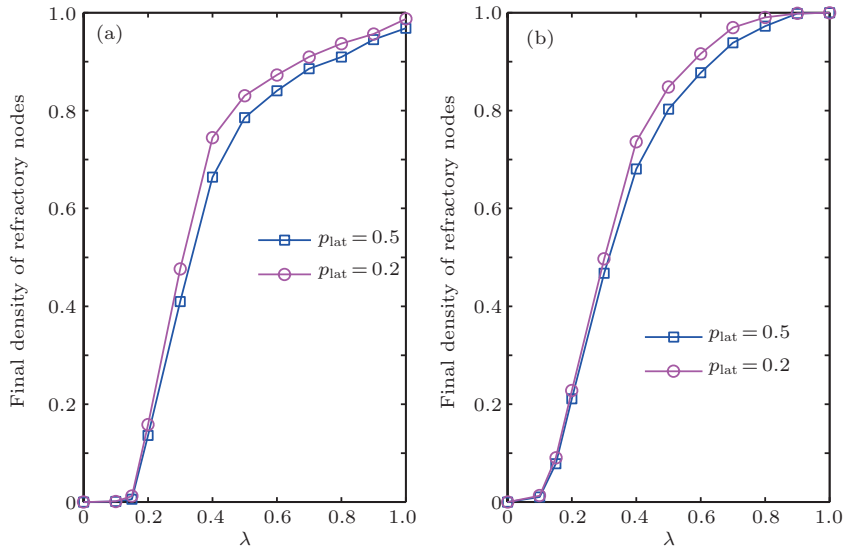


Fig. 4. (color online) Plots of $\rho_{\rm r}(\infty)$ versus λ, when $N = 2500$, $\delta = 0.5$, and $p_{\rm inf} = 0.3$. (a) The underlying network is a homogeneous network with $\bar{k} = 4$; (b) a scale-free network mediates the diffusion process. The results are averaged over 100 different realizations.

To demonstrate the threshold for the diffusion process, we calculate the final densities of refractory nodes with different spreading probabilities as shown in Fig. 4. Information diffusion witnesses a distinct threshold in both networks. In the left panel, the threshold in a homogeneous network is nearly the same as that in Fig. 2 with the equal values of parameters, and it accords with the analytical result. The parameter $p_{\rm lat}$ does not change the threshold, but it affects the final diffusion extent, which is in accordance with the mean-field analysis of Section 3. In a scale-free network, $\rho_{\rm r}(\infty)$ becomes slightly larger than in a homogeneous network due to the promotion of central nodes. In addition, the threshold in a scale-free net-

work is approximately half that in a homogeneous network. In heterogeneous networks, the threshold depends on $\bar{k}/\overline{k^2}$, as given in Eq. (25). However, in homogeneous networks, the threshold is proportional to $1/\bar{k}$. For a Barabasi-Albert scale-free network, $\overline{k^2} = \bar{k}^2 \log(N)/4$ holds true. Therefore, we have $\bar{k}/\overline{k^2} < 1/\bar{k}$.

Figure 5 shows the effects of latency probability $p_{\rm lat}$ on the diffusion extent. In the left panel, the underlying topology is the same scale-free network as that in Fig. 4. In the right panel, we use a real online network extracted from the Sina website that is a famous micro-blog in China. The real social network contains 5906 nodes with the average degree $\bar{k} = 4.5008$. The average shortest path length of the network is 4.99, and the clustering coefficient is 0.0772. From Fig. 5, it is dramatic that $\rho_{\rm r}(\infty)$ first decreases with the increase of $p_{\rm lat}$, and then it increases. Around $p_{\rm lat} = 0.5$, $\rho_{\rm r}(\infty)$ reaches the minimum. For small $p_{\rm lat}$, infected nodes become latent with a small chance. Since only latent nodes may enter into the refractory state and thoroughly stop their interactions, the diffusion process will last a long time. More infected nodes exist in the system with small $p_{\rm lat}$, and they have the ability to spread information, leading to a larger final diffusion. For large $p_{\rm lat}$, although infected nodes become latent more quickly, they may forward more messages to their neighbors before they pause their diffusion actions. After infected nodes forward a message, their susceptible neighbors can see the message and have the opportunity to be infected. Therefore, large $p_{\rm lat}$ speeds up the diffusion process, and also increases $\rho_{\rm r}(\infty)$. Obviously, the maximal density of infected nodes over time decreases with the increase of $p_{\rm lat}$, but the change for the density of latent nodes is opposite. We also calculate the maximum sum of densities of infected and latent nodes over time. When $\lambda = 0.2$, $\delta = 0.2$, and $p_{\rm inf} = 0.5$, the maximum sums in a scale-free network are 0.2792 for $p_{\rm lat} = 0.1$, 0.2573 for $p_{\rm lat} = 0.5$, 0.3071 for $p_{\rm lat} = 0.9$, respectively. We can see the bottom of the sum occurs around $p_{\rm lat} = 0.5$. Therefore, $p_{\rm lat}$ has a subtle role in the diffusion, and both large and small $p_{\rm lat}$ lead to a larger diffusion extent. Increasing $p_{\rm inf}$ makes latent nodes easier to retrieve their activity, and causes more spreaders, so $\rho_{\rm r}(\infty)$ with large $p_{\rm inf}$ increases markedly. In addition, more nodes are infected in a scale-free network than in the real social network, since the scale-free network has a smaller average shortest path length 4.37. Furthermore, $p_{\rm inf}$ in a scale-free network exerts a more obvious effect on the diffusion process.

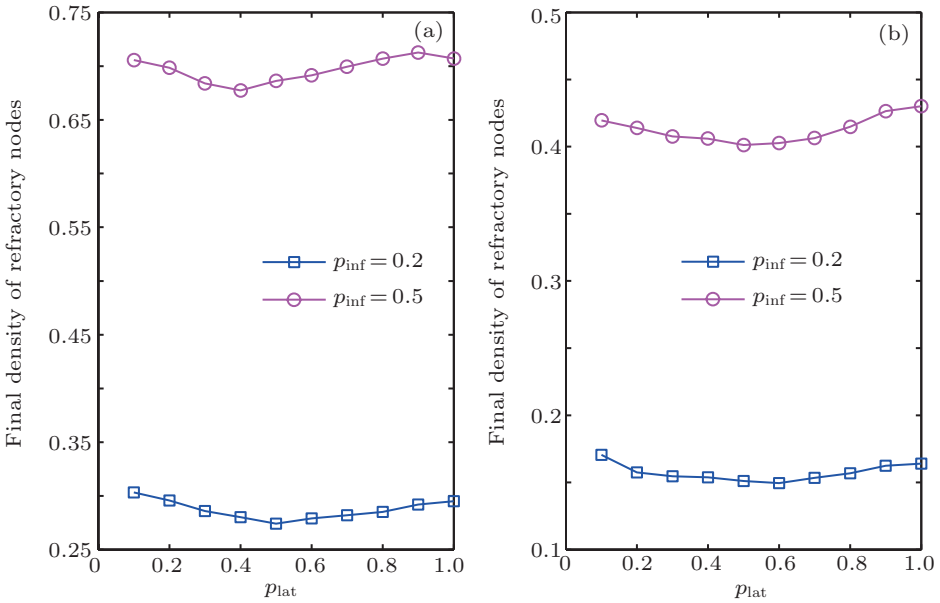


Fig. 5. (color online) Plots of $\rho_{\rm r}(\infty)$ versus $p_{\rm lat}$, when $\lambda = 0.2$ and $\delta = 0.2$. (a) A scale-free network is used; (b) a real social network is used. Each plot represents the results averaged over 100 different realizations.

Now, we concentrate on the dependence on the refractory probability δ. Intuitively, by increasing δ, more nodes thoroughly drop out of interactions, and information becomes less easy to propagate. As shown in Fig. 6, the final density of refractory nodes nearly decays with the increase of δ as a power law, that is, $\rho_{\rm r}(\infty) \propto \delta^{-\gamma}$. The power exponent is $\gamma = -0.6963 \pm 0.0718$ in a scale-free network, while it is $\gamma = -0.7307 \pm 0.0406$ in the real network. When δ is small, the information in a scale-free network almost occupies the absolute majority of the system, while in the real network, many susceptible nodes survive in the end.

We measure the influence of nodes in the diffusion process, and try to find the effective topological descriptor for influence maximization. We use the real social network as interaction topology. Then, we define $C(t) = \rho_{\rm i}(t) + \rho_{\rm l}(t) + \rho_{\rm r}(t)$, where $C(t)$ denotes the density of nodes that have ever been influenced by the information, and therefore, it can be used to measure the influence of initial spreaders at time t. We use

three topological descriptors to identify influential nodes, i.e., the degree, k-core index, and betweenness.

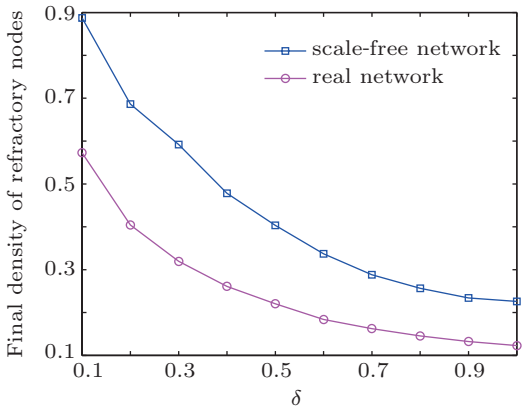


Fig. 6. (color online) Plots of $\rho_r(\infty)$ versus δ, when $\lambda = 0.2$, $p_{inf} = 0.5$, and $p_{lat} = 0.5$. Each plot represents the results averaged over 100 different realizations.

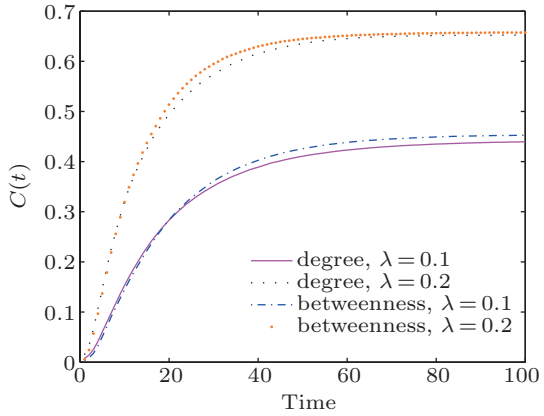


Fig. 7. (color online) Variations of $C(t)$ with time in the real network, when $\delta = 0.2$, $p_{inf} = 0.5$, and $p_{lat} = 0.5$. Initial infected nodes are assigned according to the degree and betweenness.

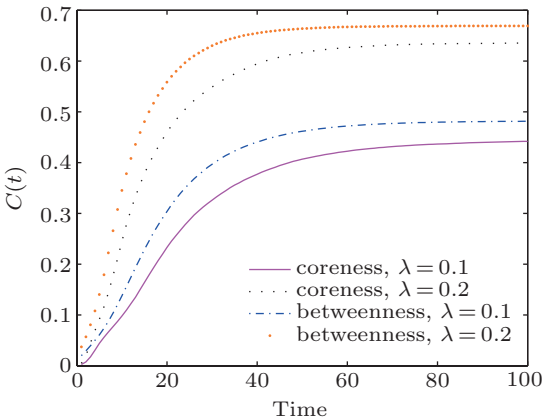


Fig. 8. (color online) Variations of $C(t)$ with time in the real network, when $\delta = 0.2$, $p_{inf} = 0.5$, and $p_{lat} = 0.5$. Initial infected nodes are assigned according to the coreness and betweenness.

We compare the performance of the degree with betweenness in Fig. 7, and compare the k-core index with betweenness in Fig. 8. In Fig. 7, we sort nodes according to the degree and betweenness, respectively, and select the top 10 nodes from each rank. To obtain a distinct result, we delete the common nodes occurring simultaneously in the top 10 of the degree and betweenness rank. For each topological descriptor, we retain 4 nodes as initial infected nodes. In Fig. 8, initial infected nodes are assigned similarly. From Fig. 7, the betweenness outperforms the degree a little in grasping the influential nodes in our diffusion model, despite the parameters. Moreover, the descriptor of betweenness has an obvious advantage compared with the k-core index from Fig. 8. Therefore, the betweenness is the best descriptor to mimic the node influence in information diffusion with latency.

5. Conclusions

In this paper, we study the effect of latency on the information diffusion process. We put forward a diffusion model that includes four node states, i.e., the susceptible, infected, latent and refractory state. After infected nodes take diffuse actions, they enter into the latent state and stop spreading information temporarily. Latent nodes may recover their activity, or they may become refractory and withdraw from interactions. We implement mean-field analysis and Monte-Carlo simulations to investigate the spreading threshold and final density of refractory nodes.

The results show that the threshold for information diffusion depends on the spreading probability, refractory probability, and reactivating probability, but the probability of becoming latent is unrelated to the threshold. The threshold in scale-free networks is nearly $4/\log(N)$ times as large as that in homogeneous networks. Small or large latency leads to a larger diffusion extent, but the essential causes for these two situations are different. Moreover, the final density of refractory nodes increases markedly with the reactivating probability increasing. We also detect influential nodes in the diffuse model. From the result, the spreading capabilities of nodes mainly correlate with their betweenness, while the coreness does not perform well for identifying node influence.

In online social networks, there are two cases of popular information. In the first case, users forward information rapidly, and the information propagates on a large scale within a short time, implying large latency probability in the model. In the second case, users forward information and become latent slowly, but the diffusion process lasts a long time. Many users are influenced by the information finally. From the results, if one hopes to restrain the spread of a rumor, the method of creating another topic to distract user attention may not be so effective in online social networks. Under the measure, active users enter into the latent state temporarily but they do not withdraw from the diffusion of the rumor. The rumor can propagate slowly and can cause a long-term impact gradually. A better way to restrain the diffusion may be to separate the rumor from neighbors and to clarify the truth at the nodes with large betweenness.

References

[1] Barthélemy M, Barrat A, Pastor-Satorras R and Vespignani A 2004 *Phys. Rev. Lett.* **92** 178701
[2] Nekoveea M, Morenob Y, Bianconic G and Marsili M 2007 *Physica A* **374** 457
[3] Lu Y, Jiang G and Song Y 2012 *Chin. Phys. B* **21** 100207
[4] Genois M, Vestergaard C L, Cattuto C and Barrat A 2015 *Nat. Commun.* **6** 8860
[5] Gong Y, Song Y and Jiang G 2012 *Chin. Phys. B* **21** 010205
[6] Reppas A I, De Decker Y and Siettos C I 2012 *J. Stat. Mech.* **2012** P08020
[7] Zhou T, Liu J, Bai W, Chen G and Wang B 2006 *Phys. Rev. E* **74** 056109
[8] Barthelemy M, Barrat A, Pastor-Satorras R and Vespignani A 2002 *Phys. Rev. Lett.* **92** 178701
[9] Parshani R, Carmi S and Havlin S 2010 *Phys. Rev. Lett.* **104** 258701
[10] Newman M E J 2002 *Phys. Rev. E* **66** 016128
[11] Silva R D and Fernandes H A 2015 *J. Stat. Mech.* **2015** P06011
[12] Wang Y, Cao J D, Alofi A, AL-Mazrooei A and Elaiw A 2015 *Physica A* **437** 75
[13] Castellano C, Fortunato S and Loreto V 2009 *Rev. Mod. Phys.* **81** 591
[14] Fotouhi B and Shirkoohi M K 2016 *Phys. Rev. E* **93** 012301
[15] Liu C, Xie J, Chen H, Zhang H and Tang M 2015 *Chaos* **25** 103111
[16] Li W, Tang S, Fang W, Guo Q, Zhang X and Zheng Z 2015 *Phys. Rev. E* **92** 042810
[17] Yagan O, Qian D, Zhang J and Cochran D 2013 *IEEE J. Sel. Area Comm.* **31** 1038
[18] Nematzadeh A, Ferrara E, Flammini A and Ahn Y Y 2014 *Phys. Rev. Lett.* **113** 088701
[19] Song B, Jiang G, Song Y and Xia L 2015 *Chin. Phys. B* **24** 100101
[20] Hu Q, Zhang Y, Xu X, Xing C, Chen C and Chen X 2015 *Acta Phys. Sin.* **64** 190101 (in Chinese)
[21] Kitsak M, Gallos L K, Havlin S, Liljeros F, Muchnik L, Stanley H E and Makse H A 2010 *Nat. Phys.* **6** 888
[22] Borge-Holthoefer J and Moreno Y 2012 *Phys. Rev. E* **85** 026116
[23] Lu Z, Wen Y, Zhang W, Zheng Q and Cao G 2015 *IEEE T. Mobile Comput.* **15** 1292
[24] Lu Z, Wen Y and Cao G 2014 *Proceedings of* 33*rd Annual IEEE International Conference on Computer Communications*, 2014, Toronto, Canada, p. 1932
[25] Li Y, Qian M, Jin D, Hui P and Vasilakos A V 2015 *Inform. Sci.* **293** 383
[26] Centola D 2010 *Science* **329** 1194
[27] Zhao J, Wu J and Xu K 2010 *Phys. Rev. E* **82** 016105
[28] Hu H, Wen Y, Chua T and Li X 2014 *IEEE Access Journal* **2** 652
[29] Wen S, Haghighi M S, Chen C, Xiang Y, Zhou W and Jia W 2015 *IEEE T. Comput.* **64** 640
[30] Wen S, Jiang J J, Xiang Y, Yu S, Zhou W L and Jia W J 2014 *IEEE T. Parall. Distr.* **25** 3306
[31] Jiang C, Chen Y and Liu K J 2014 *IEEE T. Signal Proces.* **62** 4573
[32] Lambiotte R, Saramaki J and Blondel V D 2009 *Phys. Rev. E* **79** 046107
[33] Goldenberg J, Libai B and Muller E 2001 *Market. Lett.* **12** 211