Front. Comput. Sci. China 2009, 3(3): 324–334
DOI 10.1007/s11704-009-0058-7

RESEARCH ARTICLE

An alternative approach to characterize the topology of complex networks and its application in epidemic spreading

Zonghua LIU[1] (✉), **Xiaoyan WU**[1], **Pak-Ming HUI**[2]

1 Institute of Theoretical Physics and Department of Physics, East China Normal University, Shanghai 200062, China
2 Department of Physics, The Chinese University of Hong Kong, Hong Kong, China

Abstract Based on the mean-field approach, epidemic spreading has been well studied. However, the mean-field approach cannot show the detailed contagion process, which is important in the control of epidemic. To fill this gap, we present a novel approach to study how the topological structure of complex network influences the concrete process of epidemic spreading. After transforming the network structure into hierarchical layers, we introduce a set of new parameters, i.e., the average fractions of degree for outgoing, ingoing, and remaining in the same layer, to describe the infection process. We find that this set of parameters are closely related to the degree distribution and the clustering coefficient but are more convenient than them in describing the process of epidemic spreading. Moreover, we find that the networks with exponential distribution have slower spreading speed than the networks with power-law degree distribution. Numerical simulations have confirmed the theoretical predictions.

Keywords complex networks, epidemic spreading, hierarchical layers, mean-field approach, fraction of degree for outgoing

1 Introduction

Epidemic spreading is nowadays a hot topic because of the hectic activity on complex networks [1, 2]. Compared to the previous works which focused on the epidemic threshold and the epidemic control, the present studies on epidemic spreading is mainly focusing on how the topological structure of complex network influences the epidemic spreading. It is found that the features of small world and scale-free networks can influence seriously the epidemic spreading on complex networks [3−20]. There are two typical models that describe epidemic spreading process by the contacts between infected and healthy individuals in traditional significance. They are the susceptible-infected-susceptible (SIS) model and the susceptible-infected-refractory (SIR) model. The SIS model is a two state model in which the individuals (nodes) can only exist in one of the states: susceptible or infected [3–5]. The SIS model has been studied in both small world [9] and scale-free networks [3]. The results show that, in the case of scale-free network, the infection can spread to the entire network even when the probability of the transmission is infinitely small. This result is in sharp contrast to the well-known threshold phenomenon in epidemiology [21]. While the SIR model is a three state model in which the individuals (nodes) can exist in one of the three states: susceptible, infected or refractory [6–8, 10, 11] where the refractory describes the immune state or the death of infected individuals. Compared with the SIS model, the SIR model incorporates the factor that the infected nodes may be self-recovery or become refractory. Once a node is in the status of refractory, it can be considered as being removed. In this paper, we choose the SIR model to study the spreading process of epidemic.

Epidemic can be gradually spread out from the infected seed in a specific way. For an infected node, which of its neighboring nodes will be infected at the next step is uncertain. However, epidemic spreading can not be simply treated as the random walk on complex network which is now attracting some attention [22–25]. An obvious difference between the random walk and the epidemic spreading is that

Received October 22, 2008; accepted June 1, 2009

E-mail: zhliu@phy.ecnu.edu.cn

for the former the number of moving agents is a constant and the individuals cannot duplicate themselves; while for the latter the infected number changes with time and it is possible for an infected node to make several neighboring nodes be infected at the next time step. Another difference is that the random walk can go back and forth, while the epidemic in SIR model can only go along the direction to the refractory. It is found that for a well mixed SIR model, the evolutions of S, I, and R nodes satisfy a set of mean-field equations [7, 8, 10, 21]. By these equations we can get a solution on how many nodes can be finally infected and how the network topology influences the final infected number. However, the mean-field approach can only show the infected number but cannot tell the detailed information of infection, such as how fast the epidemic is spread, how large fraction of the infected nodes is in the front of spreading, and which nodes will be infected in the next step etc. The purpose of this paper is to answer these questions.

The complex networks can be characterized by several structure parameters, such as the degree distribution $P(k)$, the clustering coefficient C, and the average degree $\langle k\rangle$ etc., which reflect the aspects of statistical average of network structure. The dynamics in complex networks, such as the synchronization, information propagation, and epidemic spreading etc., can be seriously influenced by the topological structure of networks or by the structure parameters [1, 2]. However, these structure parameters are not always convenient in describing the dynamical processes in networks. For example, in the case of epidemic spreading, these parameters can show how the network structure influences the total infected nodes but cannot answer the questions given in the last paragraph. Therefore, it is necessary to introduce some new characteristic parameters to describe the detailed spreading process. For implementing this purpose, in this paper we transform the topological structure of complex network into hierarchical layers and introduce a set of new parameters, i.e., the average fractions of degree for outgoing f^{+1}, ingoing f^{-1}, and remaining in the same layer f^0, to describe the infection process. Different from the mean-field method, this approach is of favor in characterizing the detailed process of infection. By this approach we find that f^{+1} is the key factor to determine the spreading speed and the front infected nodes while f^{-1} and f^0 determine the localized spreading in the complex networks. By checking two typical networks with exponential and power-law degree distribution, respectively, we find that the former has slower spreading speed than the latter. A heuristic argument is given to explain our findings.

The paper is organized as follows. In Section 2, we present a heuristic theory to analyze the detailed spreading process of epidemic in a complex network. Then in Section 3, we make numerical simulations to confirm the predictions given in Section 2. Finally, the discussion and conclusions are given in Section 4.

2 A novel approach to characterize epidemic spreading

We use the SIR model to study the epidemic spreading. Consider a network with N nodes. The infection can be only propagated from nodes to nodes through the links. That is, an infected node can only transmit the infection to its neighbors. At the beginning, we randomly choose one node as the infected seed and the remaining nodes are susceptible. We let the seed infect each of its neighbors with a probability λ. And then the infected nodes will continue to infect their neighbors. If the node i is susceptible and has k_i neighbors, of which k_{inf} are infected, then at the next time step the node i will become infected with probability $[1-(1-\lambda)^{k_{inf}}]$. We also assume a probability μ for each infected node to decay into the refractory class. Without loss of generality, we set $\mu = 1$ since it only affects the definition of the time scale of the virus propagation [21], i.e., the infected status can be kept only for one time step. We use $I(t)$ to denote the number of infected nodes at time t. With time going, the infected nodes will have larger and larger distances to the original seed and $I(t)$ will increase. At a finite time, $I(t)$ will begin to decrease until there is no longer any infected nodes in the network, then the process is over. The larger the probability λ is, the larger the number of final infected nodes. When $\lambda = 1$, the number reaches its maximum N and all the nodes of the network have been infected.

Take an infection process as an example. There are $I(t)$ infected nodes at time t. These nodes have different distances to the original seed. Suppose d denotes the shortest distance between the infected node i and the original seed. d can be calculated as follows: d equals 1 if the infected node i is one of the neighbors of the original seed and 2 if it is the neighbor's neighbor of the original seed, and so on. Generally, d will have a distribution of different values at time t for all the infected nodes except two situations. The first one is, for the case of $\lambda = 1$, the infection will be transmitted to all the nodes with the same distance, resulting $d = t$ for all the $I(t)$. The second one is for the tree-like network where the epidemic propagation will keep outgoing and all the distances d equal the time t. However, most of the realistic networks are not strictly tree-like but small world and scale-free [1,2]. There are local loops in the topological structure, such as the triangles. The existence of

loops will definitely influence the value of d. For example, the infection may be propagated to a node of a triangle and then it is possible to be propagated to the direction of backing to the original seed, resulting that some d decrease with time. Therefore, we here concern what is the distribution of d on different distances and how the distribution depends on the network structure. These questions cannot be answered by the mean-field method as the method does not distinguish the distances. To answer these questions, we transform the network into hierarchical layers with the same center and different radii where the center is the original seed and the distance between the neighboring layers is unity. The nodes on the first layer are the neighboring nodes of the original seed and have $d = 1$, and the nodes on the second layer are the neighbor's neighbors of the original seed and have $d = 2$, and so on. The existence of loops/triangles means that the infection can be transmitted from one node to another one on the same layer. Once it happens, this step does not make d increase 1. Further, the infection may go back to the nearest inner layer if the inner connected node is susceptible, see the schematic Fig. 1.

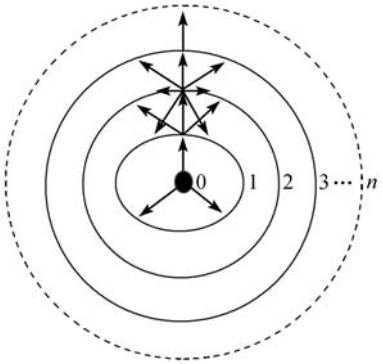

Fig. 1 Schematic illustration of epidemic spreading on a complex network where the center denotes the original seed, the numbers $0, 1, 2, \ldots$, denote the distance to the original seed, and the arrows denote the possible infection pathes.

An important problem in the control of epidemic is how fast the epidemic spread, i.e., the speed of spread. For the convenience of study, we consider here the distance as the topological distance but not the physical distance. Suppose the number of infected nodes with distance t is $I_1(t)$ and the total infected nodes is $I(t)$. Letting $\rho_1(t) = I_1(t)/I(t)$, $I_1(t)$ and $I(t)$ reflect the speed of spreading and $\rho_1(t)$ represents the fraction of front infected nodes in $I(t)$. Let's figure out how $I_1(t)$ and $\rho_1(t)$ depend on the structure of network.

Two typical quantities reflecting the network topology are the degree distribution $P(k)$ and the clustering coefficient C, where k denotes the degree of a node. The former represents how many nodes have the same links and the latter characterizes the propensity for two of one's neighbors to be neighbors also of each other [26]. The larger C means the larger density of triangles. These quantities are very useful in describing the dynamics on networks, such as the synchronization and information propagation etc. [27–30], but they are not the good candidates for describing the $I_1(t)$ and $\rho_1(t)$. Hence, we introduce some new parameters to obtain $I_1(t)$ and $\rho_1(t)$, which is related to $P(k)$ and C. Considering the fact that $I_1(t)$ and $\rho_1(t)$ are closely related to the outgoing links of a node, we use f_i^{+1} to denote the average fraction of outgoing part of a node on the layer i in Fig. 1 for a large number of realizations. Similarly, we use f_i^{-1} and f_i^0 to denote the average fractions of degree of ingoing and remaining in the same layer, respectively. As most of the realistic networks has the small world feature, the average distance between any two nodes, D, is not very large. For example, $D \approx \ln N/\ln\ln N$ for the scale-free network [1, 2]. It is very easy for the infection to reach the boundary of network. After that, the distance d can not be increased further. Therefore, the meaningful time for $I_1(t)$ and $\rho_1(t)$ is around D. We here limit the effective infection process to $t \sim 2D$. Noticing that the average number of nodes n_i on the layer i is different for different layers, the average degree on each layer should also be different. We use $\langle k_i\rangle$ to denote the average degree on the layer i. By these quantities, we figure out the average nodes on each layer is

$$
\begin{aligned}
n_1 &= \langle k\rangle \quad \text{for the layer 1,}\\
n_2 &= \frac{\langle k\rangle\langle k_1\rangle f_1^{+1}}{\langle k_2\rangle f_2^{-1}} \quad \text{for the layer 2,}\\
n_3 &= \frac{n_2\langle k_2\rangle f_2^{+1}}{\langle k_3\rangle f_3^{-1}} = \frac{\langle k\rangle\prod_{i=1}^{2}\langle k_i\rangle f_i^{+1}}{\prod_{i=2}^{3}\langle k_i\rangle f_i^{-1}} \quad \text{for the layer 3,}
\end{aligned}
$$

and so on. The infected nodes with distance t is

$$
\begin{aligned}
I_1(1) &= \lambda\langle k\rangle,\\
I_1(2) &= \frac{I_1(1)\langle k_1\rangle f_1^{+1}[1-(1-\lambda)^{\langle k_2\rangle f_2^{-1}}]}{\langle k_2\rangle f_2^{-1}},\\
&= \lambda n_2[1-(1-\lambda)^{\langle k_2\rangle f_2^{-1}}],\\
I_1(3) &= \frac{I(2)\langle k_2\rangle f_2^{+1}[1-(1-\lambda)^{\langle k_3\rangle f_3^{-1}}]}{\langle k_3\rangle f_3^{-1}},\\
&= \lambda n_3\prod_{i=2}^{3}[1-(1-\lambda)^{\langle k_i\rangle f_i^{-1}}],\\
I_1(t) &= \ldots.
\end{aligned}
\tag{1}
$$

Obviously, we always have $\rho_1(1) = 1$. Specifically, for the tree-like networks, we have $f_i^0 = 0$ and hence $\rho_1(t) = 1$. For the case of $\lambda = 1$, we also have $\rho_1(t) = 1$. For other situations, $\rho_1(t)$ can be determined by the set of parameters $f_i^{+1}, f_i^{-1}, f_i^0, \langle k_i\rangle$ and the contagion rate λ.

The parameters f_i^{+1}, f_i^{-1}, f_i^0, and $\langle k_i \rangle$ have close relation with the degree distribution $P(k)$ and the clustering coefficient C. Larger C means higher density of triangles in the local structure, implying a larger f_i^0. While $P(k)$ is related to $\langle k_i \rangle$. If the network has a heterogeneous structure, such as the Barabási and Albert (BA) model [1], the infection may be propagated prefer to the nodes with heavy links first and then to the nodes with smaller links, resulting a quick decrease of $\langle k_i \rangle$ with i. However, if the network has a homogeneous structure, such as the Watts and Strogatz (WS) model [31], the infection may prefer to be propagated in the local area and then gradually spreading out, resulting a constant $\langle k_i \rangle$ for $i \leqslant D$.

3 Numerical simulations

To confirm the above predictions of $I_1(t)$ and $\rho_1(t)$, in this section we use the WS model and Holme and Kim (HK) model [32] to make numerical simulations. The two models have the typical features of the complex networks, respectively. First, the clustering coefficient C can be easily changed in both the WS and the HK models. Second, the WS model is a small world network with an exponential distribution $P(k)$ while the HK model is a scale-free network with a power law distribution $P(k)$.

We first discuss the epidemic spread on the WS model. This model can be constructed as follows [31]:

(i) Start with a ring lattice with N nodes in which every node is connected to its first k neighbors ($k/2$ on either side);

(ii) Randomly rewire each edge of the lattice with probability p such that self-connections and duplicate edges are excluded.

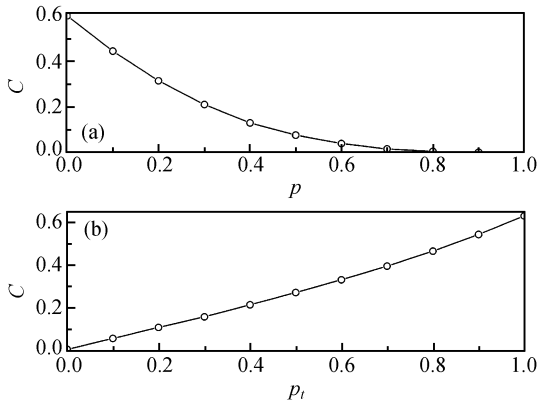

Fig. 2 **(a)** How the clustering coefficient C changes with the rewiring parameter p for the WS network with $N = 10^4$ and $k = 6$; **(b)** How the clustering coefficient C changes with the triad formation parameter p_t for the HK network with $N = 10^4$ and $k = 6$.

By adjusting the parameter p one can get different clustering coefficient C. It is a regular lattice when $p = 0$, small world network when $p \ll 1$, and random network when $p = 1$. In our numerical simulations, we fix $N = 10^4$ and $\langle k \rangle = 6$. Fig. 2(a) shows how C changes with p for $k = 6$. Obviously, C decreases with p.

According to Fig. 1, we first transfer the WS network to a layer structure and then calculate the structure parameters f_i^{+1}, f_i^{-1}, f_i^0, and $\langle k_i \rangle$, where the center node is randomly chosen. By a large number of realizations we get the average values of the parameters. Fig. 3 shows the results for different rewiring probabilities where (a) represents the relation for $\langle k_d \rangle$ versus d, (b) f_d^{+1} versus d, (c) f_d^0 versus d, and (d) f_d^{-1} versus d. From Fig. 3(a) it is easy to see that the average degrees on different layers $\langle k_d \rangle$ are approximate constants when $d \leqslant D$ and then keep constant for the case of $p = 0.1$ but decrease for the cases of $p = 0.4$ and 1 when $d > D$. The reason is that for the case of $p = 0.1$, only 10% of links are rewired and 90% of links keeps the constant degree $\langle k \rangle$, resulting the approximate constant $\langle k_d \rangle$ for different d. While in the case of $p = 1$, the network becomes a random network with an exponential degree distribution and diameter $D = \ln N / \ln\langle k \rangle \approx 5$ for $N = 10^4$ and $\langle K \rangle = 6$ [1, 2]. That is, most of them have the degree around $\langle k \rangle$. When $d \leqslant D$, there are a lot of nodes on each layer and their degree average should be around $\langle k \rangle$. When $d > D$, the nodes on each layer decreases very fast because of the boundary. Usually, these nodes are of smaller degree as the nodes with larger degree have a larger probability to connect to the seed or seed's neighbors and thus they do not stay at the layers with $d > D$. Therefore, the nodes on the layers with $d > D$ have smaller $\langle k_d \rangle$. After understanding these two limit cases, the case of $p = 0.4$ is just in between the cases of $p = 0.1$ and 1. Similarly, we can explain the behaviors of f_d^{+1} in Fig. 3(b). From the boundary effect it is also easy to understand the peak in Fig. 3(c) and the higher tail in Fig. 3(d). When d is slightly over D, the f_d^{+1} of boundary nodes is transferred to the f_d^0, resulting the increase of f_d^0. When d is further increased, the number of boundary nodes decreases and hence reduce the possibility for them to connect in the same layer and make the f_d^{+1} of boundary nodes be transferred to the f_d^{-1}, resulting the decrease of f_d^0 and the increase of f_d^{-1}. Finally, for those nodes on the layer with the largest d, they have only links to the inner layer, resulting f_d^{-1} close to unity.

From Fig. 3 we may also notice that $\langle k_d \rangle$ in (a) is slightly different for $p = 0.1, 0.4, 1$ when $d \leqslant D$, but the corresponding f_d^{+1} in (b) have larger differences. To understand it we show how $\langle k_d \rangle$ and $\langle k_d^{out} \rangle$ change with the rewiring probability p in Fig. 4(a) and (b), where $\langle k_d^{out} \rangle = \langle k_d \rangle f_d^{+1}$ de-

notes the average outgoing degree on the layer d and the lines with the "circles", "stars", "triangles", and "squares" represent the cases of $d = 1, 2, 3$, and 4, respectively. It is easy to see that $\langle k_d \rangle$ increases only about 8%, but $\langle k_d^{out} \rangle$ increases to 2 to 3 times when p changes from 0 to 1. The underlying reason comes from the big change in degree distribution $P(k)$. As we know, $P(k)$ has only one value for the regular

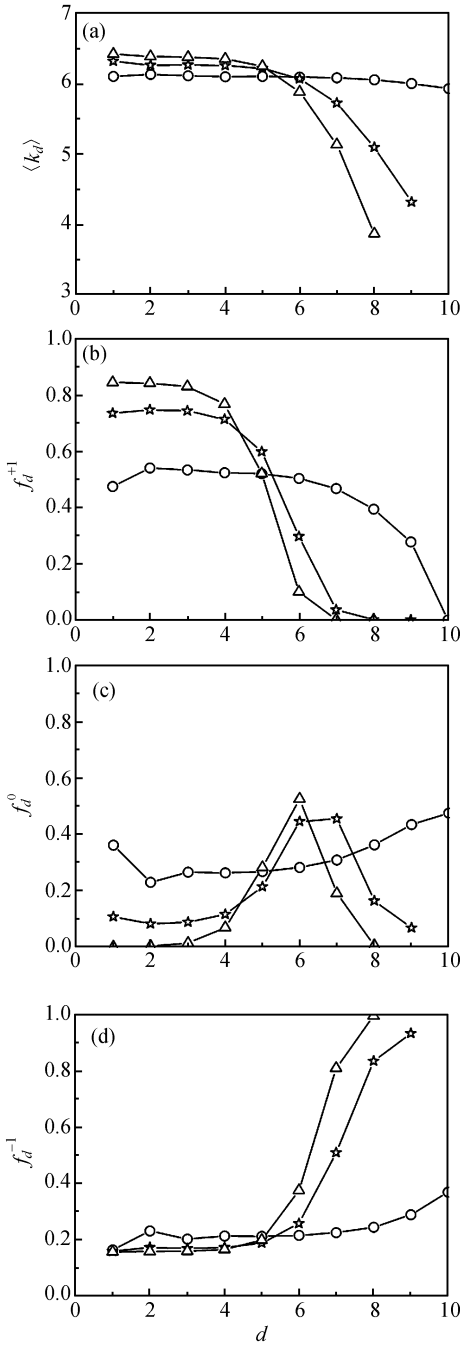

Fig. 3 How the structure parameters f_i^{+1}, f_i^{-1}, f_i^0, and $\langle k_i \rangle$ of the WS model change with the distance d where the lines with the "circles", "stars", and "triangles" represent the cases of $p = 0.1, 0.4$, and 1, respectively, and **(a)** $\langle k_d \rangle$ versus d, **(b)** f_d^{+1} versus d, **(c)** f_d^0 versus d, and **(d)** f_d^{-1} versus d.

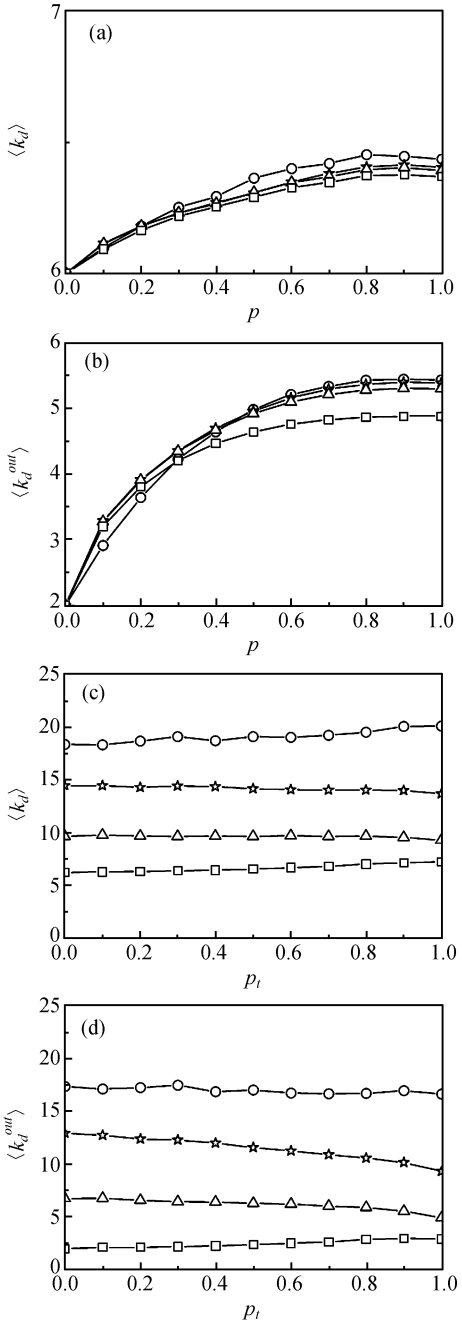

Fig. 4 **(a)** and **(b)** How the average degrees in the WS model change with the rewiring probability p where the lines with the "circles", "stars", "triangles", and "squares" represent the cases of $d = 1, 2, 3$, and 4, respectively, and **(a)** $\langle k_d \rangle$ versus p and **(b)** $\langle k_d^{out} \rangle$ versus p. **(c)** and **(d)** How the average degrees in the HK model change with the triad formation probability p_t where the lines with the "circles", "stars", "triangles", and "squares" represent the cases of $d = 1, 2, 3$, and 4, respectively, and **(c)** $\langle k_d \rangle$ versus p_t and **(d)** $\langle k_d^{out} \rangle$ versus p_t.

ring lattice of $p = 0$. In this case, there are equal possibility for outgoing, ingoing, and remaining in the same layer, hence $\langle k_d^{out} \rangle = \langle k \rangle / 3$. While for the case of $p = 1$, the random network has a very small clustering coefficient, hence it is a tree-like structure in the hierarchical layers, resulting

$\langle k_d^{out}\rangle \approx \langle k_d\rangle - 1$. Therefore, the big change of $P(k)$ results a big change of $\langle k_d^{out}\rangle$ when p changes from 0 to 1. Later, we will see that this phenomenon has a big influence to $\rho_1(t)$.

In the constructed WS network, we randomly choose a node as the infected seed and let it propagate on the network with contagion rate λ according to the rules given in Section 2. We find that both the infected nodes $I(t)$ and the infected front nodes $I_1(t)$ are depending on the parameters p and λ. Fig. 5 shows the results for different p where the lines with the "circles", "stars", "triangles", "squares", "asterisks", and "pluses" represent the cases of $p = 0, 0.1, 0.2, 0.4, 0.6$, and 1, respectively; (a) and (b) denote the case of $\lambda = 0.2$, and (c) and (d) the case of $\lambda = 0.4$. Obviously, both the $I(t)$ and $I_1(t)$ are small numbers in the case of $\lambda = 0.2$ and large numbers in the case of $\lambda = 0.4$, which can be understood from the nonlinear relationship between $I_1(t)$ and λ in Eq. (1).

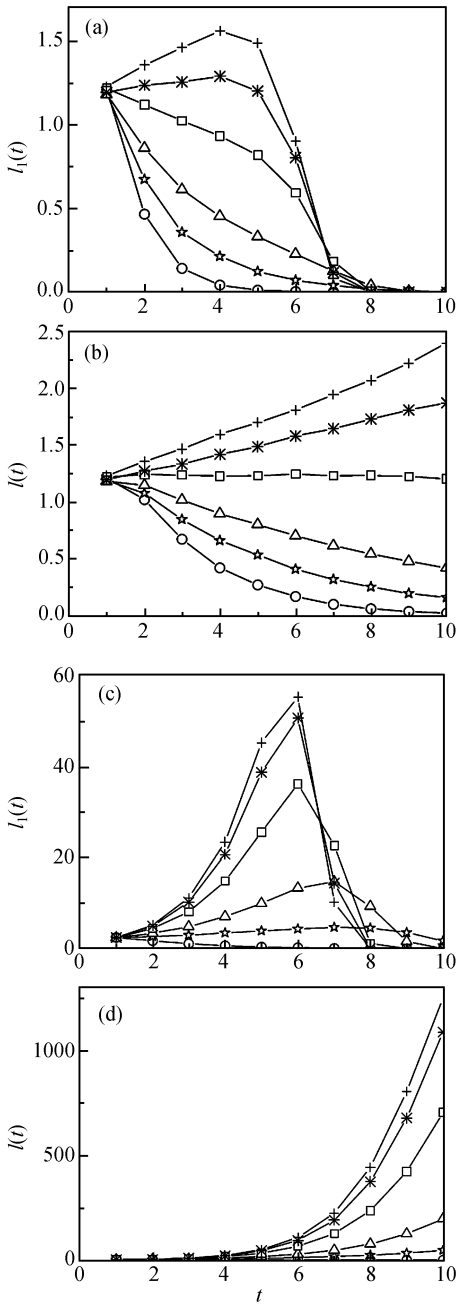

Fig. 5 Evolution of infected nodes with time in the WS model where the lines with the "circles", "stars", "triangles", "squares", "asterisks", and "pluses" represent the cases of $p = 0, 0.1, 0.2, 0.4, 0.6$, and 1, respectively; **(a)** and **(b)** denote the case of $\lambda = 0.2$, **(c)** and **(d)** $\lambda = 0.4$.

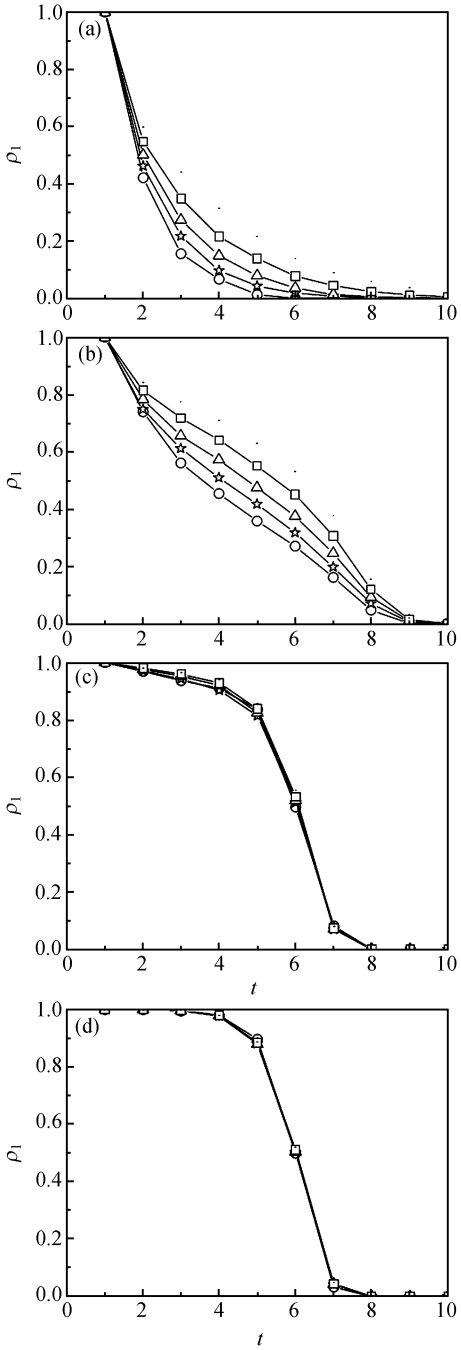

Fig. 6 Evolution of ρ_1 with time in the WS model for different λ where the lines with the "circles", "stars", "triangles", and "squares" represent the cases of $\lambda = 0.1, 0.2, 0.3$, and 0.4, respectively, and **(a)** denotes the case of $p = 0$, **(b)** $p = 0.2$, **(c)** $p = 0.6$, and **(d)** $p = 1$.

To understand how the structure parameters $f_i^{+1}, f_i^{-1}, f_i^0$, and $\langle k_i \rangle$ influence the $I(t)$ and $I_1(t)$, let's focus on the case of $\lambda = 0.2$. Fig. 5(b) shows that the total infected nodes $I(t)$ decrease with t for small p but increase with t for large p and Fig. 5(a) shows that $I_1(t)$ is a concave curve for small p but a convex curve for large p, implying that the random network is in favor of epidemic spreading than the small world network. This phenomenon can be understood by checking Fig. 3(b) or Fig. 4(b). Fig. 4(b) shows that $\langle k_d^{out} \rangle$ has a small value for small p but a large value for large p, i.e., small f_d^{+1} for small p and large f_d^{+1} for large p in Fig. 3(b). Therefore, the epidemic will spread out with a larger probability f_d^{+1} in the case with larger p and results a larger fraction of $I_1(t)$ in $I(t)$. Instituting f_d^{+1} into n_d and then both into Eq. (1) we can understand the behavior of $I_1(t)$. To see the fraction of $I_1(t)$ in $I(t)$ conveniently, we study the $\rho_1(t)$ versus t. We find that the variation of λ has larger influence for $\rho_1(t)$ for the cases of small p but small influence to the cases of larger p. Fig. 6 shows how $\rho_1(t)$ changes with λ for different p. Obviously, $\rho_1(t)$ curve has larger variation in Fig. 6(a) and (b) and small variation in (c) and (d). The reason is that for the radial spreading of large p, the larger λ will mainly increase $I_1(t)$ and results in a small deviation in $\rho_1(t)$; while for the local spreading with small p, the local part will lack of enough susceptible nodes to be infected and hence increase the fraction of $I_1(t)$ in $I(t)$, resulting in a large deviation in $\rho_1(t)$.

Now we turn to study the epidemic spreading in the HK model. The HK model can be constructed as follows [32,33]:

(i) Initially, the network consists of m nodes and no edges;

(ii) Then, we add a new node with m edges at each time step.

The first edge of the added node is attached to an existing node by the preferential attachment, i.e., with the probability proportional to its degree. Suppose the chosen node is i. The remaining $m-1$ edges of the added node are attached randomly to the neighbors of the node i with probability p_t and attached preferentially to the existing nodes with probability $1-p_t$. That is, the preferential links of a new node are $1+(m-1)(1-p_t)$ and the neighboring links are $(m-1)p_t$. This is called "triad formation" [32], which is based on our frequent everyday experience on how we are acquainted by newcomers: B becomes A's new friend since B is introduced by one of A's friends. This model has the same degree distribution with the BA model but with different C, and will reduce to the BA model when $p_t = 0$ [32, 33]. By changing p_t we can get different C. In our numerical simulations, we fix $N = 10^4$ and $m = 6$. Fig. 2(b) shows how C changes with p_t. Similar to the WS model, the HK model also has a small diameter $D \approx \ln N / \ln \ln N \approx 4$ for $N = 10^4$.

Doing the similar calculation as in the network of WS model, we get the evolution of structure parameter $f_i^{+1}, f_i^{-1}, f_i^0$, and $\langle k_i \rangle$ for the HK model as shown in Fig. 7. Comparing Fig. 7 with Fig. 3, one can see that their

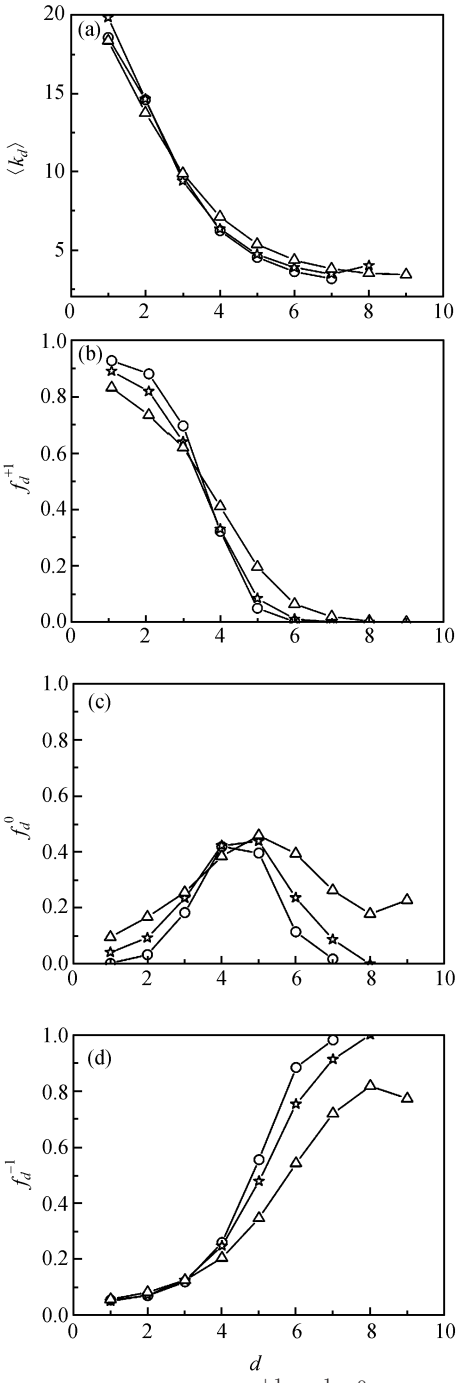

Fig. 7 How the structure parameters $f_i^{+1}, f_i^{-1}, f_i^0$, and $\langle k_i \rangle$ of the HK model change with the distance d where the lines with the "circles", "stars", and "triangles" represent the cases of $p_t = 0, 0.4$, and 0.8, respectively, and **(a)** $\langle k_d \rangle$ versus d, **(b)** f_d^{+1} versus d, **(c)** f_d^0 versus d, and **(d)** f_d^{-1} versus d.

common behaviors are that both the peaks in their panel (c) and both the increasing points in their panel (d) appear at about $d \approx D$. The reason is that both the WS and the HK models have the boundary restriction, whose consequence has been explained in the case of WS model. Their difference is that both $\langle k_d \rangle$ and f_d^{+1} for $d \leqslant D$ are approximate constants in the Fig. 3(a) and (b) but not in Fig. 7(a) and (b). This difference comes from their different degree distributions. Comparing with the WS model with short-range degree distribution, the HK model has a widely distributed $P(k)$ with power law. The hub nodes in power law distribution have much larger possibility to be connected by the seed than other nodes, i.e., the center in Fig. 1 will prefer to connect to the hub nodes. Therefore, the nodes with the largest links will most probably stay in the layer 1 and then for the same reason, the nodes with degree close to the hubs will most probably stay in the layer 2 and so on, indicating a decrease of $\langle k_d \rangle$ with d. In detail, we may figure out the value of $\langle k_d \rangle$ for a large number of realizations. Obviously, the average degree on the center is $\langle k_0 \rangle = \langle k \rangle$. Suppose the center has degree k for a specific realization, it will connect a node of degree k' on the layer 1 by the conditional probability $P(k'|k)$. The average degree of the connected node is $\sum k' P(k'|k)$, and the total degrees of all the k nodes on the layer 1 is $k \sum k' P(k'|k)$. As the possibility for the center to take degree k is $P(k)$, the total degrees of all the nodes on the layer 1 is $\sum P(k) k \sum k' P(k'|k)$. Hence, the average degree on the layer 1 is

$$\langle k_1 \rangle = \sum_k P(k) k \sum_{k'} k' P(k'|k) / \sum P(k) k. \quad (2)$$

For uncorrelated network, we have $P(k'|k) = k' P(k') / \langle k \rangle$ [34–36], so

$$\langle k_1 \rangle = \langle k^2 \rangle / \langle k \rangle. \quad (3)$$

Similarly, we may derive the formula of $\langle k_2 \rangle$ and so on. A direct consequence of the decrease of $\langle k_d \rangle$ with d is the decrease of the outgoing degree, resulting the decrease of f_d^{+1} in Fig. 7(b).

For the convenience of comparison, we also show the relation $\langle k_d \rangle$ versus p_t and $\langle k_d^{out} \rangle$ versus p_t in Fig. 4(c) and (d). Comparing Fig. 4(c) and (d) with Fig. 4(a) and (b), respectively, one can see that both $\langle k_d \rangle$ and $\langle k_d^{out} \rangle$ have significant change in the case of WS model but have only slight changes in the case of HK model. The reason is that $P(k)$ has a big change in the WS model when p changes from 0 to 1 while $P(k)$ does not change in the HK model when p_t changes from 0 to 1. Hence, $\langle k^2 \rangle$ has a big change for different p in the WS model but no change for different p_t in the HK model. By Eq. (3) we know that $\langle k_1 \rangle$ will change in the case of WS model but not change in the HK model, and the same reason for other $\langle k_d \rangle$ in Fig. 4(a) and (c).

Let us do the similar simulations of epidemic spreading as in the WS model. In the constructed HK network, we randomly choose a node as the infected seed and let it propagate on the network with contagion rate λ. Fig. 8 shows the results for different p_t where the lines with the "circles", "stars", "triangles", and "squares" represent the cases of $p_t = 0, 0.4, 0.8$, and 1, respectively, (a) and (b) denote the

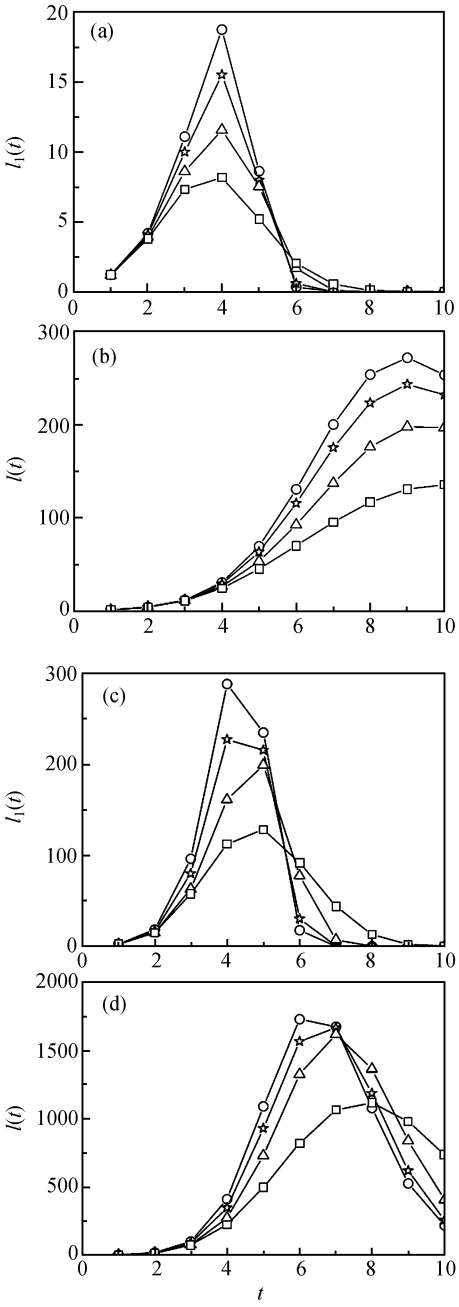

Fig. 8 Evolution of infected nodes with time where the lines with the "circles", "stars", "triangles", and "squares" represent the cases of $p_t = 0, 0.4, 0.8$, and 1, respectively, and **(a)** and **(b)** denote the case of $\lambda = 0.2$ and **(c)** and **(d)** $\lambda = 0.4$.

case of $\lambda = 0.2$ and (c)(d) $\lambda = 0.4$. Comparing Fig. 8 with Fig. 5 it is easy to see that both the $I(t)$ and $I_1(t)$ in the case of HK model are much larger than the corresponding values in WS model. This difference comes from the different $\langle k_d \rangle$ distribution, for example, the $\langle k_1 \rangle$ in Hk model is approximate three times of that in the WS model. Taking the values of $f_i^{+1}, f_i^{-1}, f_i^0$, and $\langle k_i \rangle$ into Eq. (1) we can understand this difference. Another point we notice here is that the peaks in Fig. 8(b) and (d) are much early than that in Fig. 5(b) and (d), indicating the evolution time of infection in HK model is shorter than that in WS model. This phenomenon is consistent with the value of $I(t)$ where the larger $I(t)$ makes the evolution process finish quick and hence half of the evolution time, i.e., the time for the peaks of $I(t)$, will come early.

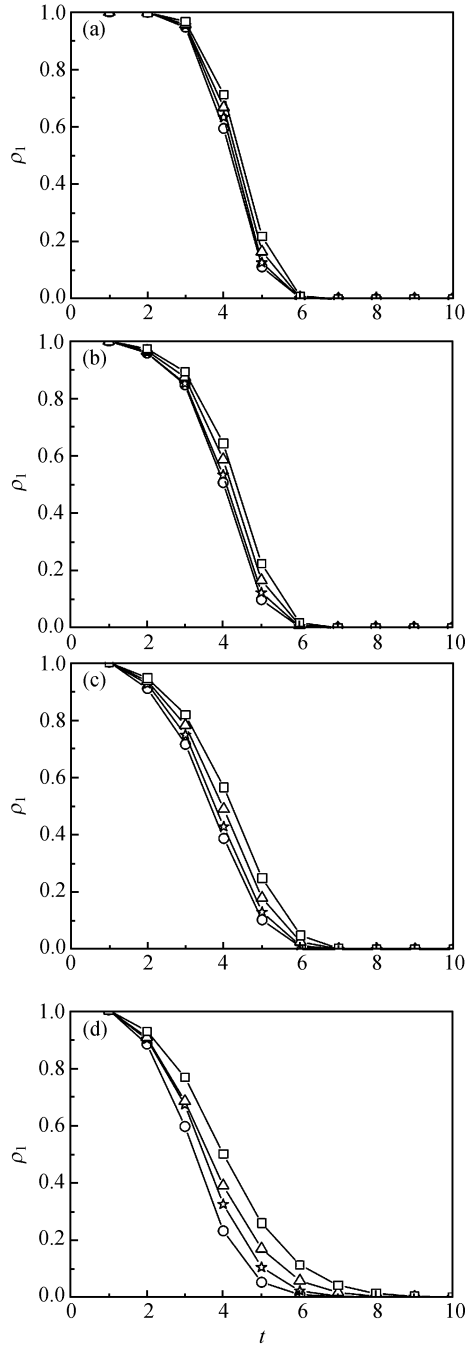

Fig. 9 Evolution of ρ_1 with time for different λ where the lines with the "circles", "stars", "triangles", and "squares" represent the cases of $\lambda = 0.1, 0.2, 0.3$, and 0.4, respectively, and **(a)** denotes the case of $p_t = 0$, **(b)** $p_t = 0.4$, **(c)** $p_t = 0.8$, and **(d)** $p_t = 1$.

Figure 9 shows how ρ_1 changes with t in HK model for different λ and p_t. Comparing Fig. 9 with Fig. 6 we see that ρ_1 are widely spread for different p in the WS model but relatively concentrated for different p_t in the HK model. This can be understood quantitative from the different structure parameters $f_i^{+1}, f_i^{-1}, f_i^0$, and $\langle k_i \rangle$. Namely, we may have the answer by instituting the corresponding parameters into Eq. (1). In qualitative, we may also explain it by the degree distribution $P(k)$ and the clustering coefficient C. As both $P(k)$ and C change with p in the WS model, thus ρ_1 are widely spread; while in HK model only C changes with p_t but $P(k)$ does not, thus ρ_1 are relatively concentrated. Moreover, the larger p_t corresponds to the larger C and hence reduce ρ_1 more.

To confirm the theoretic formula (1), we substitute the obtained values of $\langle k_d \rangle$, f_1^{+1}, and f_1^{-1} into Eq. (1) to calculate the predicted $I_1(t)$. We find that the theoretic $I_1(t)$ is in consistent with the results of numerical simulations. Fig. 10 shows the results where the dotted lines denote the theoretic results from Eq. (1): the solid lines denote the numerical simulations, (a) for the WS model; (b) for the HK model; in (a) the lines with "circles" are for $p = 0.1$ and the lines with "triangles" are for $p = 0.4$; in (b) the lines with "circles" are for $p_t = 0$ and the lines with "triangles" are for $p_t = 0.8$. From Fig. 10(a), we see that the theoretic result is

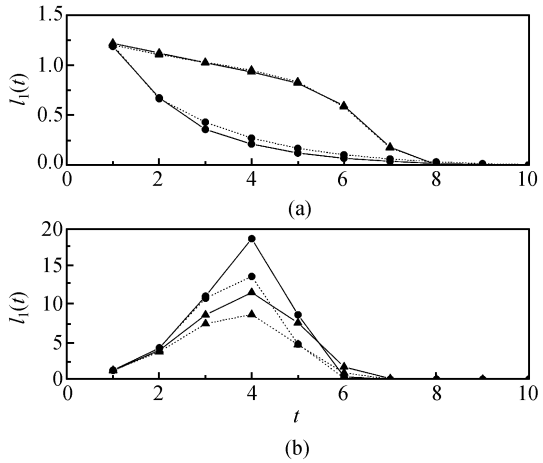

Fig. 10 Comparison of the theoretic results and the numerical simulations where the dotted lines denote the theoretic results from Eq. (1), the solid lines denote the numerical simulations, **(a)** for the WS model, **(b)** for the HK model; in **(a)** the lines with "circles" are for $p = 0.1$ and the lines with "triangles" are for $p = 0.4$, in **(b)** the lines with "circles" are for $p_t = 0$ and the lines with "triangles" are for $p_t = 0.8$.

almost overlapped with the numerical simulation in the case of approximate random network with $p = 0.4$ but they have a finite difference in the case of small world network with $p = 0.1$. From Fig. 10(b) we see that there is a finite difference between the theoretic and numerical results for both the BA model with $p_t = 0$ and the highly clustered network with $p_t = 0.8$. The reason is that Eq. (1) is derived based on a mean-field assumption. As both the small world network and the HK model are somehow away from the heterogeneous network, thus theoretic results have a finite difference from the numerical simulation. In sum, Eq. (1) grasps qualitatively the features of epidemic spreading in general.

4 Discussion and conclusions

Previous works of epidemic spreading focused on the spreading result, such as if the epidemic will break out and how many nodes can be infected, etc., but did not pay attention to the aspect of spreading process, such as the speed of spread and the fraction of front infected nodes. However, the speed of epidemic spread is very important. Its understanding may shed light on the control of epidemic spread, such as the SARS etc. As we know, the social network has high clustering coefficient [37]. How to give different efficient strategies to control the epidemic spreading in communities with different structures is significant. This work may stimulate some further research in this direction.

In conclusion, we have investigated the process of epidemic spread on complex networks by the SIR model. A set of new parameters are introduced to illustrate the speed of spread and the fraction of front infected nodes. We find that these parameters have close relation with the clustering coefficient and the degree distribution and can replace them in reflecting the influence of network structure on the epidemic spread. Numerical simulations have confirmed the theoretic predictions.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant Nos. 10775052, 10635040), and National Basic Research Program of China (973 Program) (2007CB814800).

References

1. Albert R, Barabasi A L. Statistical mechanics of complex networks. Reviews of Modern Physics, 2002, 74: 47–97
2. Boccaletti S, Latora V, Moreno Y, et al. Complex networks: structure and dynamics. Physics Reports, 2006, 424: 175–308
3. Pastor-Satorras R, Vespignani A. Epidemic spreading in scale-free netwoks. Physical Review Letters, 2001, 86: 3200–3203
4. Joo J, Lebowitz J L. Behavior of susceptible-infected-susceptible epidemics on heterogeneous networks with saturation. Physical Review E, 2004, 69: 066105
5. Barthelemy M, Barrat A, Pastor-Satorras R, et al. Dynamical patterns of epidemic outbreaks in complex heterogeneous networks. Journal of Theoretical Biology, 2005, 235: 275–288
6. Zheng D F, Hui P M, Trimper S, et al. Epidemics and dimensionality in hierarchical networks. Physica A, 2005, 352: 659–668
7. Liu Z, Lai Y C, Ye N. Propagation and immunization of infection on general networks with both homogeneous and heterogeneous components. Physical Review E, 2003, 67: 031911
8. Ben-Naim E, Krapivsky P L. Size of outbreaks near the epidemic threshold. Physical Review E, 2004, 69: 050901
9. Shao Z, Sang J, Zou X, et al. Blackmail propagation on small-world networks. Physica A, 2005, 351: 662–670
10. Newman M E J. Spread of epidemic disease on networks. Physical Review E, 2002, 66: 016128
11. Zhang H, Liu Z, Ma W. Epidemic propagation and microscopic structure of complex networks. Chinese Physics Letters, 2006, 23: 1050–1053
12. Eguiluz V M, Klemm K. Epidemic threshold in structured scale-free networks. Physical Review Letters, 2002, 89: 108701
13. Liu Z, Hu B. Epidemic spreading in community networks. Europhysics Letters, 2005, 72: 315–321
14. Serrano M A, Boguna M. Percolation and epidemic thresholds in clustered networks. Physical Review Letters, 2006, 97: 088701
15. Gross T, Dommar D'Lima C J, Blasius B. Epidemic dynamics on an adaptive network. Physical Review Letters, 2006, 96: 208701
16. Zhou J, Liu Z. Epidemic spreading in complex networks. Frontiers of Physics in China, 2008, 3: 331–348
17. Zhou J, Liu Z, Li B. Influence of network structure on rumor propagation. Physics Letters A, 2007, 368: 458–463
18. Zhou Y, Liu Z, Zhou J. Periodic wave of epidemic spreading in community networks. Chinese Physics Letters, 2007, 24: 581–584
19. Zhou J, Liu Z. Epidemic spreading in communities with mobile agents. Physica A, 2009, 388: 1228–1236
20. Tang M, Liu L, Liu Z. Influence of dynamical condensation on epidemic spreading in scale-free networks. Physical Review E, 2009, 79: 016108
21. Anderson R M, May R M. Infections Diseases in Humans. Oxford University Press, Oxford, 1992
22. Pandit S A, Amritkar R E. Random spread on the family of small-world networks. Physical Review E, 2001, 63: 041104
23. Adamic L A, Lukose R M, Puniyani A R, et al. Search in power-law networks. Physical Review E, 2001, 64: 046135
24. Noh J D, Rieger H. Random walks on complex networks. Physical Review Letters, 2004, 92: 118701
25. Parris P E, Kenkre V M. Traversal times for random walks on small-world networks. Physical Review E, 2005, 72: 056119
26. Watts D J. Small Worlds: the Dynamics of Networks Between Order and Randomness. Princeton University Press, Princeton, 1999
27. Yan G, Zhou T, Hu B, et al. Efficient routing on complex networks. Physical Review E, 2006, 73: 046108
28. Wang X G, Lai Y C, Lai C H. Oscillations of complex networks. Physical Review E, 2006, 74: 066104
29. Liu Z, Ma W, Zhang H, et al. An efficient approach of controlling traffic congestion in scale-free networks. Physica A, 2006, 370: 843–853

30. Zhang H, Liu Z, Tang M, et al. An adaptive routing strategy for packet delivery in complex networks. Physics Letters A, 2007, 364: 177–182
31. Watts D J, Strogatz S H. Collective dynamics of 'small-world' networks. Nature, 1998, 393: 440
32. Holme P, Kim B J. Growing scale-free networks with tunable clustering. Physical Review E, 2002, 65: 026107
33. Wu X, Liu Z. How community structure influences epidemic spread in social networks. Physica A, 2008, 387: 623–630
34. Newman M E J. Assortative mixing in networks. Physical Review Letters, 2002, 89: 208701
35. Catanzaro M, Boguna M, Pastor-Satorras R. Generation of uncorrelated random scale-free networks. Physical Review E, 2005, 71: 027103
36. Tang M, Liu Z, Zhou J. Condensation in a zero range process on weighted scale-free networks. Physical Review E, 2006, 74: 036101
37. Newman M E J. Clustering and preferential attachment in growing networks. Physical Review E, 2001, 64: 025102