Vol 14 No 11, November 2005
1009-1963/2005/14(11)/2153-05
Chinese Physics
© 2005 Chin. Phys. Soc.
and IOP Publishing Ltd

Dynamic properties of epidemic spreading on finite size complex networks*

Li Ying(李 昊)[a)b)†], Liu Yang(刘 旸)[a)], Shan Xiu-Ming(山秀明)[a)],
Ren Yong(任 勇)[a)], Jiao Jian(焦 健)[a)], and Qiu Ben(仇 贲)[a)]

[a)] *Department of Electronic Engineering, Tsinghua University, Beijing* 100084, *China*

[b)] *School of Business, Sun Yat-Sen University, Guangzhou* 510275, *China*

(Received 2 August 2004; revised manuscript received 17 May 2005)

The Internet presents a complex topological structure, on which computer viruses can easily spread. By using theoretical analysis and computer simulation methods, the dynamic process of disease spreading on finite size networks with complex topological structure is investigated. On the finite size networks, the spreading process of SIS (susceptible–infected–susceptible) model is a finite Markov chain with an absorbing state. Two parameters, the survival probability and the conditional infecting probability, are introduced to describe the dynamic properties of disease spreading on finite size networks. Our results can help understanding computer virus epidemics and other spreading phenomena on communication and social networks. Also, knowledge about the dynamic character of virus spreading is helpful for adopting immunity policy.

Keywords: epidemic spreading, complex networks, SIS model
PACC: 0250

1. Introduction

The epidemic spreading model was first established by biologists. Recent advances in complex network theory prompt development of the epidemic spreading model in complex networks and there appear to be some remarkable results in this field. Complex networks describe some systems in high technology and intelligence organism. For example, the cell can be presented as a complex network of chemicals connected by chemical reactions; the Internet is a complex network of routers and computers linked by various physical links. With the development of computer technology, now we can study the topology of many real systems which may contain millions of nodes and it is found that most of these networks have three main characters (small-world, scale-free, high-clustering).[1−5]

The highly influential network model of Barabási and Albert[2] is a kind of scale-free network which was initially used to simulate the growing process of World Wide Web. This model is also widely applied to analyse the epidemic spreading process. The network is constructed as follows: starting with a small number (m_0) of connected nodes, at each time step we add a new node with m ($m < m_0$) edges which link the new node to m different node that already exist in the network; the probability that the new node links to the node i only depends on the degree k_i of node i, usually proportional to the degree k_i. As the number of time steps is increasing, the network is growing to infinite size. If the network growing process stops at finite time steps, the network is a finite size network which contains finite nodes.

The study of the epidemic spreading model deals with an important process taking place on networks and greatly helps in the understanding of diseases spreading among organisms, information spreading over people and computer viruses pervading the Internet. This article analyses the dynamic properties of the SIS model spreading in finite size networks.

2.SIS model

The SIS model[3] is one of the most widely analysed epidemic spreading models. It is often studied as the model of tuberculosis and computer viruses. The

*Project supported by the National Nature Science Foundation of China (Grant Nos 90204004 and 90304005).
†E-mail: liying02@tsinghua.org.cn, liy@mail.ee.tsinghua.edu.cn

node can be cured, but without a permanent immunity, which means that the recovered nodes will lose their immunity gradually and may be infected again. In the SIS model, the node can exist in only two discrete states, susceptible or infected. At each time step, each susceptible node is infected with probability λ if it is connected to one or more infected nodes; at the same time, infected nodes are cured and become again susceptible with probability 1. Here λ is an effective spreading rate, and in general topology there is a nonzero epidemic threshold λ_c. If $\lambda \geq \lambda_c$, the infection spreads and becomes persistent in time; if $\lambda < \lambda_c$, the infection dies out fast. $\rho_k(t)$ denotes the relative density of infected nodes with given degree k. The dynamic process of epidemic spreading can be described by

$$\frac{\mathrm{d}\rho_k}{\mathrm{d}t} = -\rho_k(t) + \lambda k[1-\rho_k(t)]\Theta_k(t). \tag{1}$$

The creation of infected nodes is proportional to the spreading rate λ, the degree k, the density of susceptible nodes [1-$\rho_k(t)$], and the probability $\Theta_k(t)$ that any given link points to an infected node. The probability that any given link points to a node with degree k is $kP(k)/ < k >$. Thus the probability $\Theta_k(t)$ is irrelevant to the degree k,

$$\Theta(t) = \frac{1}{< k >}\sum_k kP(k)\rho_k(t). \tag{2}$$

When the epidemic spreading process reaches stationarity [$\mathrm{d}\rho_k/\mathrm{d}t = 0$], we obtain the stationary density

$$\rho_k = \frac{\lambda k\,\Theta}{1+\lambda k\,\Theta}. \tag{3}$$

By substituting Eq.(3) into Eq.(2), we obtain

$$\Theta = \frac{1}{< k >}\sum_k kP(k)\frac{\lambda k\,\Theta}{1+\lambda k\,\Theta}. \tag{4}$$

A nonzero epidemic threshold exists if and only if the right-hand side and the left-hand side of Eq.(4) cross in the internal $0 < \Theta \leq 1$:

$$\frac{\mathrm{d}}{\mathrm{d}\Theta}\left(\frac{1}{< k >}\sum_k kP(k)\frac{\lambda k\,\Theta}{1+\lambda k\,\Theta}\right)_{\Theta=0} \geq 1. \tag{5}$$

Thus we obtain

$$\lambda_c = \frac{< k >}{< k^2 >}. \tag{6}$$

In the network model of Barabási and Albert, giving the degree distribution $P(k)=2m^2k^{-3}$, we obtain

$$\Theta = \frac{1}{\lambda m(\mathrm{e}^{1/\lambda m}-1)}. \tag{7}$$

According to Eq.(7), by applying the relation

$$\rho = \sum_k P(k)\rho_k, \tag{8}$$

we can obtain

$$\rho \sim \mathrm{e}^{-1/\lambda m}. \tag{9}$$

In some complex networks, especially the scale-free networks, the value of $< k^2 >$ is infinite if the networks are infinite networks. Thus the value of epidemic threshold in these networks is 0. This result indicates that the disease at any given spreading rate can reach the stationarity and the density of infected nodes under this state will maintain a constant value. In finite size networks, people generally believe there is always a nonzero epidemic threshold as the result of the finite value of $< k^2 >$.[4]

In fact, Eq.(6) can only be applied to infinite size networks. The stationarity only exists in the infinite size networks; in finite size networks the disease will always die out. Here we will prove that the density of infected nodes has to decrease to zero when it satisfies the equation $\mathrm{d}\rho_k/\mathrm{d}t$=0 in the finite size networks. The number of nodes in finite size networks is denoted by N, and the state of node j at time t is $V_j(t)$ ($1 \leq j \leq N$) which can only take the value of 0 or 1. If node j is infected, $V_j(t)$=1, otherwise if node j is susceptible, $V_j(t)$=0. Thus, the vector $\boldsymbol{s}(t) = (V_1(t), V_2(t), \ldots, V_N(t))$ denotes the state of a network at time t. The collection of random variables $\boldsymbol{s} = \{\boldsymbol{s}(t), t > 0\}$ just constitutes a stochastic process. According to the infection mechanism of SIS model, the state at time t+1 of any given node in the network is determined by the state of this node and its neighbouring nodes at time t. $\{\boldsymbol{s}(t), t > 0\}$ is a time-homogeneous Markov chain with a finite state space. There exists one and only one absorbing state in the Markov chain. In the absorbing state, all nodes are susceptible. The state space is denoted by

$$\tilde{\boldsymbol{S}} = S \cup S_0, \tag{10}$$

where $S = \{\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3, \ldots, \boldsymbol{s}_d\}$ is the transient state set, and $S_0 = \{\boldsymbol{s}_0 = (0,0,\ldots,0)\}$ is the absorbing state set. The transition probability matrix $\tilde{\boldsymbol{P}}$ is given as

$$\tilde{\boldsymbol{P}} = \begin{pmatrix} P & P_0 \\ 0 & 1 \end{pmatrix}, \tag{11}$$

where P is the transition probability matrix between transient states, $P_0 = (I - P)\boldsymbol{e}$, $\boldsymbol{e} = (1, 1, \ldots, 1)^{\mathrm{T}}$ is a column vector of size $\boldsymbol{p}$. The epidemic spreading process can be described as a Discrete Phase-Type Distribution. Let τ be the time until absorption into state $\boldsymbol{s}_0$,

$$\tau = \inf\{t : t \geq 0, \quad \boldsymbol{s}(t) \in \boldsymbol{S}_0\}, \tag{12}$$

$$g_k = \mathrm{Pro}(\tau = k), \tag{13}$$

$$g_k(i) = \mathrm{Pro}(\tau = k | s(0) = s_i), \tag{14}$$

$$\boldsymbol{g}_k = (g_k(i), i = 1, ..., d)^{\mathrm{T}}. \tag{15}$$

Let the starting state probability vector be denoted by $\boldsymbol{\pi}(0) = (\alpha_0, \boldsymbol{\alpha})$, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d)$, $\alpha_k \geq 0$,

$$\sum_{s_k \in \tilde{S}} \alpha_k = 1. \tag{16}$$

We can obtain

$$g_0 = \alpha_0, \tag{17}$$

$$\forall k \geq 1, \quad g_k = \alpha P^{k-1} P_0. \tag{18}$$

P is the transition probability matrix between transient states, and so $\lim_{n \to \infty} = P^n = 0$. We obtain $\mathrm{Pro}(\tau = \infty)$=0, and thus the disease will die out in the network. As a result, there is not stationarity in finite size networks.

3. Stationary distribution

In order to describe the dynamical process of SIS model spread in finite size networks, here we define two parameters. Assuming that the disease starts spreading at time t=0, the definitions of the parameters are:

Survival probability $P_{\mathrm{s}}(t)$ is the probability that the disease is still spreading in the network at time t.

Conditional infecting probability $\rho_{\mathrm{s}}(j, t)$ is the probability that node j is infected at time t if the disease is still spreading at time t.

Mean conditional infecting probability $\rho_{\mathrm{s}}(t)$ is the average of $\rho_{\mathrm{s}}(j, t)$ over all nodes in the network, and it also equal to the mean density of infected nodes if the disease is still spreading at time t.

If we set S as the state space, in other words, that the states only transfer in the transient state space, we will find that the process is also a Markov chain. Let P' denote the transition probability matrix that is the conditional transition probability matrix between transient states. In the this Markov chain, there is no absorbing state, and if the stationary distribution of P' exists, the $\rho_{\mathrm{s}}(t)$ will remain a constant nonzero value in the stationary distribution. In computer simulation, we consider Barabási–Albert networks with m=3 and find that there indeed exists the stationary distribution of P' (see Fig.1).

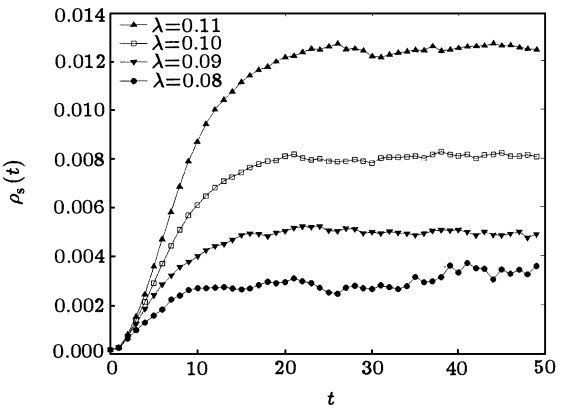


Fig.1. Mean conditional infecting probability as a function of time t Network size N=5000. The constructions of Barabási–Albert network are iterated 50 times and the SIS model spread processes are iterated 20 times in each network construction.

The probability that the state transfers from transient state set S to absorbing state set S_0 is denoted by $1 - p$ at each step, and p is the probability that the state remains in S at each step. If the disease spread process reaches the stationary distribution of P', p also remains a constant value. Under this condition, we can simplify the disease spread process into a Markov chain with only two states. Here we name one state the stationary transient state that is the stationary distribution of P' and the other the absorbing state. The transition probability matrix is

$$\tilde{\boldsymbol{P}} = \begin{bmatrix} p & 1-p \\ 0 & 1 \end{bmatrix}. \tag{19}$$

From Eq.(18) we obtain $\mathrm{Pro}(\tau = n) = p^{n-1}(1 - p)$. Assuming that the disease spread process starts with the stationary transient state, the survival probability $P_{\mathrm{s}}(t)$ is

$$\begin{aligned} P_{\mathrm{s}}(t) =& \mathrm{Pro}(\tau \geq t) = \sum_{n=t}^{\infty} \mathrm{Pro}(\tau = n) \\ =& \sum_{n=t}^{\infty} p^{n-1}(1-p) = p^{t-1}. \end{aligned} \tag{20}$$

If the disease spread process starts with some other transient state, and assuming that the process reaches the stationary transient state at time t_0, we obtain

$$P_{\mathrm{s}}(t) = P_{\mathrm{s}}(t_0) p^{t - t_0}. \tag{21}$$

According to the first equation, we find that $P_s(t)$ declines in an exponential form which is consistent with the simulation results (see Fig.2).

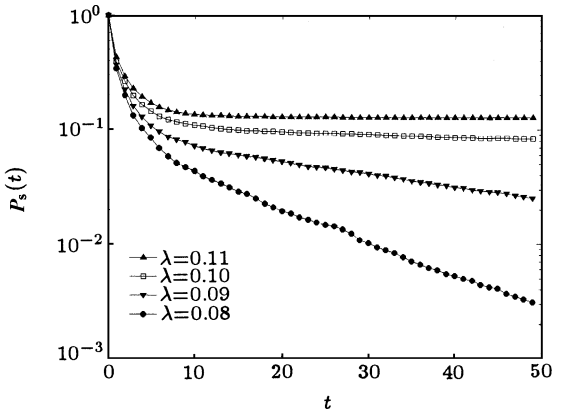


Fig.2. Survival probability as a function of time t, Network size N=5000. The constructions of Barabási–Albert network are iterated 50 times and the SIS model spread processes are iterated 20 times in each network construction.

The probability p that the process remains the transient state indicates how long the disease will spread in the network. This parameter is very helpful for us to understand and estimate the dynamic properties of epidemic spread. Here we call p the survival factor. λ is the effective spreading rate and M is the number of susceptible nodes which connect with the infected nodes. Thus the probability that all susceptible nodes remain susceptible is $(1-\lambda)^M$ and p satisfies the following equation:

$$p = 1 - (1 - \lambda)^M. \tag{22}$$

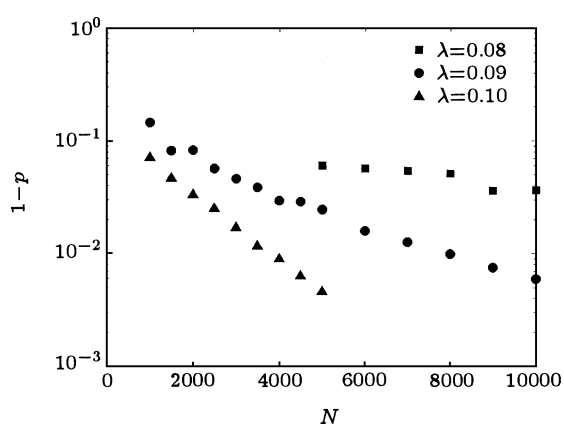


Fig.3. Survival factor as a function of network size N. The constructions of Barabási–Albert network are iterated 100 times and the SIS model spread processes are iterated 100 times in each network construction.

According to Fig.3, we find that the value of $\ln(1-p)$ is proportional to the network size N. Thus M is proportional to N. We can draw the conclusion that in the stationary transient state the constant value of $\rho_s(t)$ is independent of network size N. So the constant value of $\rho_s(t)$ equals to the density of infected nodes in the stationarity state of the infinite size networks. According to this the constant value of $\rho_s(t)$ in the stationary transient state will satisfy Eq.(9) (see Fig.4).

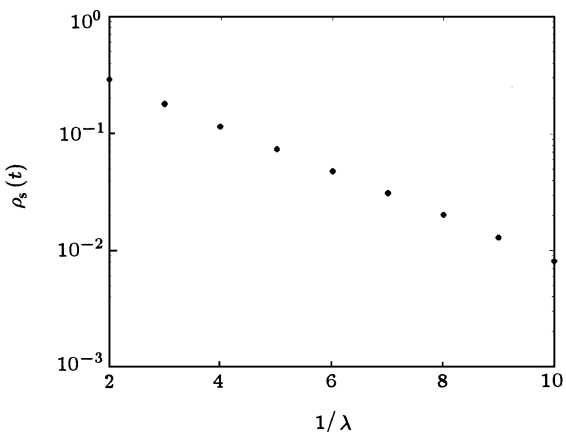


Fig.4. The constant value of $\rho_s(t)$ as a function of the effective spreading rate. The constructions of Barabási–Albert network are iterated 100 times and the SIS model spread processes are iterated 100 times in each network construction.

In homogeneous networks, if the density of infected nodes is ρ_s, the probability that a node which connects with infected nodes is susceptible is 1-ρ_s. Thus a best estimate of M is $\rho_s N < k > (1 - \rho_s)$ where $< k >$ is the average degree of the network,

$$p = 1 - (1 - \lambda)^{\rho_s N < k > (1-\rho_s)}. \tag{23}$$

4. Conclusion

In finite size networks the disease or virus spreading can reach a stationary transient state in which the mean conditional infecting probability $\rho_s(t)$ remains a constant value that is independent of network size N and the survival probability $P_s(t)$ declines in an exponential form. In finite size networks, the epidemic threshold λ_c indicates whether the disease spreading is able to reach the stationary transient state or not. In the stationary transient state, the mean conditional infecting probability $\rho_s(t)$ is equal to the mean density of infected nodes in infinite size networks. Thus, although $< k^2 >$ is finite in finite size networks, the epidemic threshold in such networks equals the value in infinite size networks of the same kind. The network size N only influences the value of survival factor

p which determines how long the disease will spread in finite size networks. To sum up, in order to get a nonzero epidemic threshold, we must change the topology structure of the network rather than depend on a finite size effect. Thus we must find proper immunization strategies to prevent the computer viruses spreading over the Internet.

References

[1] Albert R and Barabási A L 2002 *Rev. Mod. Phys.* **74** 47
[2] Barabási A L and Albert R 1999 *Science* **286** 509
[3] Pastor-Satorras R and Vespignani A 2001 *Phys. Rev.* E **63** 066117
[4] Pastor-Satorras R and Vespignani A 2002 *Phys. Rev.* E **65** 035108
[5] Li Y, Shan X M and Ren Y 2004 *Acta Phys. Sin.* **53** 3695 (in Chinese)