J Stat Phys (2010) 140: 1157–1166
DOI 10.1007/s10955-010-0034-5

Discrete Stochastic Modeling for Epidemics in Networks

Jia-Zeng Wang · Min Qian

Received: 28 October 2009 / Accepted: 22 July 2010 / Published online: 5 August 2010

Abstract In this paper, we construct a microscopic mechanism for the epidemic processes in heterogeneous populations, which contains two basic assumptions: all edges of the network are broken and reconnected at every time step among the epidemic process; the probability of randomly chosen two half-edges to make a pair is identical. We define the stochastic epidemic process and get the epidemic distributions numerically according to the transmission probability λ. Two different phases are observed, which means the onset of phase transition, and the threshold value is very small. Under the thermodynamic limit, the process can be approximated by a deterministic dynamical system. About this deterministic system, we get the analytical threshold value, which is consistent with the numerical results of the epidemic distributions.

Keywords Stochastic epidemic process · Heterogeneous network · Mass action principle

1 Introduction

Mathematical models of epidemic spreading are important tools in understanding the spreading dynamics (such as diseases, information, and states). Under the assumption that all individuals are homogeneous and well-mixed, the deterministic differential equations about the epidemic dynamics can be built. The effects of many features such as the age and social structure of the population, the contact pattern among individuals, and the stages of infection on epidemic evolution can be analyzed (Refs. [1, 2] and references therein).

This approach is called mean-field method in physical society. Some other works are essentially in this framework though they have different names in different literatures [1, 3, 4]. Based on the same assumption, stochastic epidemic models can also be built and analyzed (Chap. 4 in [5, 6]). In addition to the dynamics of mean epidemics, the outbreak distributions for susceptible-infected-removed(SIR) model are analyzed in Refs. [7, 8].

J.-Z. Wang (✉) · M. Qian
School of Mathematical Sciences, Peking University, Beijing 100871, China
e-mail: wangjiazen@yahoo.com.cn

Recently, the theories of complex networks have attracted a lot of interests within the physics community [9, 10]. With individuals or agents represented by nodes, and the interactions among them mimicked by links, many social, biological, and communication systems can be properly described by complex networks. It is discovered that many systems such as the sexual contact network, the World Wide Web, and the airport network show the property of power-law degree distribution [11–15]. So it is necessary to study epidemic spreading on networks. Using susceptible-infected-susceptible (SIS) [16] and susceptible-infected-removed (SIR) models [17, 18], it was shown that epidemic processes on many scale-free networks with infinite size do not exhibit the threshold phenomenon: computer viruses can spread and persist even when the probability of transmission is vanishingly small [16, 17]. In fact, the important role played by heterogeneity in population structure in determining properties of spreading processes has long appreciated by mathematical epidemiologists (see Chap. 11 in Ref. [1], and Ref. [18]).

Whereas models in all these works are built phenomenologically using deterministic differential equations. Similar to the homogeneous and well-mixed population case, the need of the stochastic epidemic modeling in heterogeneous populations is obvious: the spread of disease is naturally stochastic; one can get more detailed information about the epidemic process; and the deterministic model is just the limitation of certain stochastic model under the law of large number.

In this paper, we construct a microscopic mechanism of the epidemic process on networks, which contains two basic assumptions: one is that all edges in the network are broken and reconnected among every step of the epidemic process; another is that the probability of randomly chosen two half-edges to make a pair is identical, thus the probability of a given "susceptible" half-edges make pair with a "infectious" half-edge is proportional to the density of "infectious" half-edges. Based on these assumptions, a stochastic epidemic system is defined. We get the final epidemic distributions numerically according to different transmission probabilities, and observe two different phases: the single peak distribution which peaks at 0; the bimodal distributions which have two peaks, one at zero and a second at apart from zero. This phenomenon is qualitatively similar with which observed in case of homogeneous populations (Chap. 4 in [6, 8]).

Furthermore, we obtain the dynamics of the mean epidemics analytically from our stochastic system. As the number of individuals tends to infinity, we get limiting deterministic difference equations which can be used to approximate the original stochastic system. From this deterministic system, we discover the threshold phenomenon and get the critical value of the transmission probability λ.

The main body of our work is arranged in Sect. 2: we define the network ensemble and the probability measure on it, the microscopic mechanism, and the stochastic population process; get the final epidemic distributions numerically, the evolution equations of mean epidemics, and the deterministic difference equations. And the conclusions are made in Sect. 3.

2 The Stochastic Epidemic Process

2.1 The Network Ensemble

Since the individuals are heterogeneous, the classical mean-field practice is somewhat inappropriate. Noticing that all edges in the network are homogeneous, we build our model based on edges instead of on nodes. Here we assume that all edges in the network are broken into two half-edges at the beginning of every time interval, then these half-edges are

paired randomly and reconnected again, this process is accomplished immediately and it is independent of the epidemic process.

Since the edges are reconnected from time to time among the epidemic process, we need a network ensemble in our model rather than just a static network. Let us characterize the network ensemble and the probability measure on it first.

Set N to be the number of nodes for every network, and these nodes are ordered according to certain rules, i.e., the set of nodes is $\mathcal{N} = \{1, 2, \ldots, N\}$. The edge joining two nodes n_1 and n_2 is denoted by $\langle n_1, n_2\rangle$ and is assumed to be undirected. Let $\mathcal{L} = \{\langle n_1, n_2\rangle; n_1, n_2 \in \mathcal{N}\}$ denotes the set of edges, the network is just a graph $\mathcal{G} = \{\mathcal{N}, \mathcal{L}\}$. Since the node set of every network in the ensemble are identical, a network can be represented just by its edge set $\mathcal{L}$.

The most basic topological characterization of a graph $\mathcal{G}$ can be obtained in terms of the degree distribution $D(k), k = 1, 2, \ldots, K$, defined as the probability that a node chosen uniformly at random has degree k, here K denotes the maximum degree of the node in network. According to their degrees, N nodes can be classified into K subgroups: $\{SG(k); k = 1, 2, \ldots, K\}$. Let $N^{(k)}$ be the number of nodes in $SG(k)$, for simplicity we set the equation $D(k) = \frac{N^{(k)}}{N}$. The second restriction we impose on the network ensemble is that all networks in it have identical degree sequence $\{N^{(k)}; k = 1, 2, \ldots, K\}$.

For given parameters N and $D(k)$, our network ensemble contains all networks that satisfy the above mentioned two restrictions. The probability of every network in the ensemble is set to be identical.

2.2 The Microscopic Mechanism

In the standard compartment model, individuals are categorized according to their infection states: susceptible (S), infected (I), or recovered (R). There are two basic models: susceptible-infected-susceptible (SIS) and susceptible-infected-removed (SIR). In this paper, we study the SIR type epidemics.

The state of an individual node $n \in \mathcal{N}$ is one of the following three: "0", "1", or "2", which are interpreted as "0" = susceptible, "1" = infected, and "2" = removed.

Let X denotes the state space of the system, that is: $X = \{0, 1, 2\}^{\mathcal{N}}$. We define a time-discrete stochastic process $\eta = \{\eta_t; t = 0, 1, 2, \ldots\}$ which takes value on X, $\eta_t(n)$ means the state of node n at time t, $n = 1, 2, \ldots, N$. Its transition probabilities are modeled as follows.

At the beginning of the time interval t, the state of every node is determined. The edges are broken down and re-paired, whereas this process is accomplished immediately. So the neighbors of any given node are determined also. According to the SIR type epidemics, if the state of a node n is "2", it will never change; if it is "1", it will change to the state "2" deterministically at the end of the time interval; if it is "0", it may change to the state "1", the probability is determined by the number of infectious nodes among its neighbors. We can translate this statement into mathematical languages, i.e.:

$$\begin{array}{lll} \eta_t(n) = \text{“2”}, \quad \eta_{t+1}(n) = \text{“2”} & \text{at probability 1;} & \\ \eta_t(n) = \text{“1”}, \quad \eta_{t+1}(n) = \text{“2”} & \text{at probability 1;} & \\ \eta_t(n) = \text{“0”}, \quad \eta_{t+1}(n) = \text{“1”} & \text{at probability } 1 - (1 - \lambda)^{M(t,n)}. & \end{array} \tag{2.1}$$

Here $M(t, n)$ is the number of infectious nodes among the neighbors of node n; the parameter λ is the probability that a disease transmits through an edge, it can be explained as follows: suppose there is an edge $e = \langle n_1, n_2\rangle$, node n_1 is infectious and node n_2 is susceptible, we define the event "n_2 is infected by n_1 through e", then the probability that this event happens is λ.

The transmission of disease through every edge is assumed to be independent and identically distributed. Suppose a susceptible node has M edges connect to infectious neighbors, it will be infected if there is at least one edge transmits the disease—that is the reason of probability in the third formula of (2.1).

2.3 The Population Process

In statistical physics, an element in state space X is called a configuration. According to the transition probability in (2.1), it is easy to see that configurations contain components "0" and "2" but no "1" are absorbing states. Since the total number of nodes N is set to be finite, the process η will end at some absorbing states eventually, which means the stationary distribution of η for any given initial condition is exists. But it is difficult to get it directly. Here we present an indirect method, i.e., define an approximate process and get the properties of this process such as the threshold phenomenon and the outbreak distribution.

Through the process of edge breaking and reconnecting, the probability that every "susceptible" half-edges—which means it is stick to a susceptible node, pairing with an "infectious" half-edge should be considered to be identical. As a result, the probability of the transition of every susceptible node in subgroup $SG(k)$ should considered to be identical since all nodes have k half-edges.

Instead of studying the process η directly, we define random functions $S_t^{(k)}$, $I_t^{(k)}$, and $R_t^{(k)}$ based on it, which denote the number of susceptible nodes, the number of infected nodes, and the number of removed nodes in subgroup $SG(k)$ respectively. Of course we have $S_t^{(k)} + I_t^{(k)} + R_t^{(k)} = N^{(k)}$. Our purpose is to study the evolutions of the number of nodes with different states and degrees.

To pursue the transition rules of the process $\{(S_t^{(k)}, I_t^{(k)}, R_t^{(k)}); k = 1, 2, \ldots, K\}$, it is important to count the number of edges connecting nodes with different states. Here we adopt the mass action principle [1]. Suppose we have $\{(S_t^{(k)}, I_t^{(k)}, R_t^{(k)}) = (s^{(k)}, i^{(k)}, r^{(k)}); 1 \le k \le K\}$ at time t, there are $ks^{(k)}$ half-edges attached to the $s^{(k)}$ susceptible nodes at subgroup $SG(k)$, and all these half-edges are assumed to be independent with each other. The probability that one of these half-edge connected to a infective node is proportional to the density of the "infectious" half-edges, i.e., we have the parameter $u = \frac{\sum ki^{(k)}}{\sum kN^{(k)}}$. Let $L_t(s^{(k)})$ be the number of edges that connect the susceptible nodes in subgroup $SG(k)$ to the infected nodes in the whole network, by independent, it has binomial distribution:

$$P(L_t(s^{(k)}) = l) = C_{ks^{(k)}}^{l} u^l (1-u)^{ks^{(k)}-l}; \quad 1 \le l \le ks^{(k)}. \tag{2.2}$$

Now we have $\{(S_t^{(k)}, I_t^{(k)}, R_t^{(k)}) = (s^{(k)}, i^{(k)}, r^{(k)}); 1 \le k \le K\}\}$, and $L_t(s^{(k)}) = l$ at time t—which means there are totally l edges may transit disease independently, each one transits disease with probability λ. Thus we can get the distribution of $I_{t+1}^{(k)}$:

$$P(I_{t+1}^{(k)} = i'^{(k)}) = C_l^{i'^{(k)}} \lambda^{i'^{(k)}} (1-\lambda)^{l-i'^{(k)}}; \quad i'^{(k)} \le l, \text{ and } i'^{(k)} \le s^{(k)}. \tag{2.3}$$

And the other two equations are as follows:

$$R_{t+1}^{(k)} = R_t^{(k)} + I_t^{(k)}; \quad \text{and} \quad S_{t+1}^{(k)} = S_t^{(k)} - I_{t+1}^{(k)}. \tag{2.4}$$

Equations (2.2), (2.3) and (2.4) determine the transition probabilities of the epidemic process. Plus the initial distributions of $\{(S_0^{(k)}, I_0^{(k)}, R_0^{(k)}); 1 \le k \le K\}$, one can calculate the trajectory distribution of the process. It is easy to check that the states $\{I_t^{(k)} = 0; 1 \le k \le K\}$ are absorbing states. In order to get some intuitive images of the stationary distribution of the process, let us present our numerical results first.

2.4 Numerical Results

Our numerical method is somewhat a combination of the Gillespie algorithm [19] and the Monte Carlo method.

Firstly, we set total number of individuals $N = 5000$, the largest degree of the network $K = 100$, and a power-law degree distribution $D(k) \propto k^{-2.1}$ for the network. According to this degree distribution, we attach a randomly produced integer to every node as its degree. Thus we get a numerically realized degree distribution, its first order and second order moments are as follows: $\langle k \rangle = 2.7604$ and $\langle k^2 \rangle = 43.1096$, which means $\frac{\langle k \rangle}{\langle k^2 \rangle} = 0.064$.

At initial time, we set just one node with degree 3 to be infectious seed. At time t, as the values of $\{(s^{(k)}, i^{(k)}, r^{(k)}); 1 \le k \le K\}$ is known, the value of $u = \frac{\sum k i^{(k)}}{\sum k N^{(k)}}$ can be calculated. In subgroup $SG(k)$, there are $ks^{(k)}$ half-edges attached to suspectable nodes, so we do Bernoulli experiments $ks^{(k)}$ times with succeed probability u, the number of success $l^{(k)}$ means there are $l^{(k)}$ edges connect to infectious nodes.

As the $l^{(k)}$ is known, we do Bernoulli experiments $l^{(k)}$ times with succeed probability λ, the number of success is just the value of $I_{t+1}^{(k)}$. With other two equations in (2.4), we get a stochastic realization of the values $\{(S_{t+1}^{(k)}, I_{t+1}^{(k)}, R_{t+1}^{(k)}); k = 1, 2, \ldots, K\}$ based on the values of $\{(S_t^{(k)}, I_t^{(k)}, R_t^{(k)}); k = 1, 2, \ldots, K\}$.

Repeating these steps until all $I_t^{(k)} = 0, k = 1, 2, \ldots, K$, we get a trajectory of the process. Some typical trajectories of $R(t) = \sum_{k=1}^{K} r^{(k)}(t)$ are shown in Fig. 1, one trajectory represents a independent realization of the process, here $\lambda = 0.12$. One can see clearly that there are two kinds of trajectories, which means there are two possible outcomes of the system. In one outcome, a few individuals are infected and the epidemic is ended after one or two time intervals; in another outcome, a fairly large number of nodes are infected and time of durations in these trajectories are similar. It is interesting to note that there is no trajectory lies between these two outcomes.

In order to study the outbreak distribution more clearly, we generate 10^5 trajectories and calculate the distribution of the final prevalence $R(\infty) = \sum_{k=1}^{K} r^{(k)}(\infty)$. In Fig. 2, there are four pictures resulted from different λ. One can see from the first picture that the distribution has a single peak near zero when λ is very small, which means there is hardly any infection. It is called the first phase [6]. As λ increasing, there emerges gradually another peak besides the first one near zero, and the distributions are bimodal. There is hardly any probability measure between these two modal. Another interesting fact is that the transition from the single peak distribution to the bimodal distribution happens at a small λ: $0.06 \le \lambda_c \le 0.08$, and the threshold value got from the deterministic density models [1, 17] is $\frac{\langle k \rangle}{\langle k^2 \rangle} = 0.064$.

In case of homogeneous population, it is well known that the deterministic density system shows the threshold result of $\lambda_c = 1$. Besides the threshold value, the stochastic modeling in this case gives more detailed information about the out break scale. In [6], it is shown that the epidemic distribution is single peaked in the subcritical case, and it is bimodal in the supercritical case, here the initial seeds are finite. In [7, 8], the epidemic distribution at the threshold point has been got analytically.

While in heterogeneous population case, the deterministic density models illustrate that the threshold result is $\lambda_c \frac{\langle k^2 \rangle}{\langle k \rangle} = 1$. As for epidemic distributions, just as mentioned before, the distributions are qualitatively similar in two cases: single peaked at sub-threshold and bimodal at super-threshold. But we can not estimate the scales of the epidemic distribution in our model until now, so we can not make conclusions quantitatively.

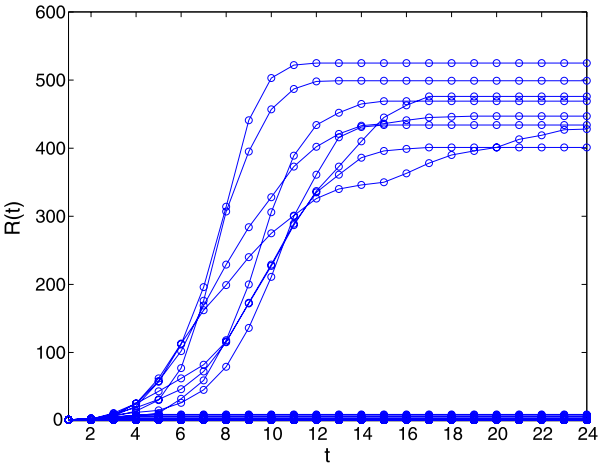

Fig. 1 Typical trajectories of $R(t)$, here $\lambda = 0.12$

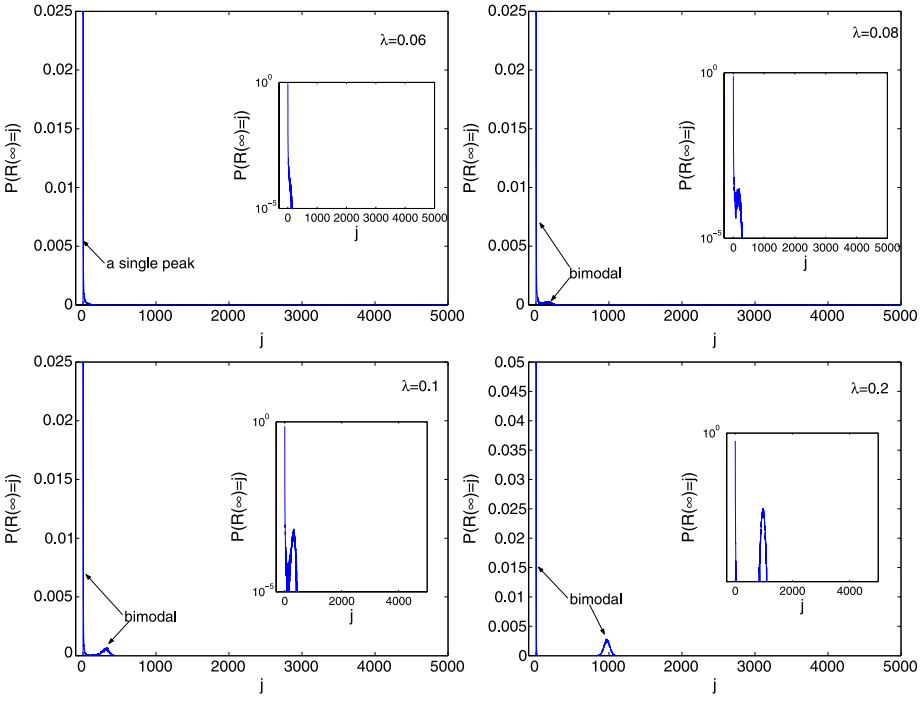

Fig. 2 The final epidemic distributions $R(\infty)$ according to different λ. To show the figures more intuitively, we present normal-log scale figures in the *inset*

2.5 Dynamics of the Mean Epidemics

Now let us investigate the evolution of the expectations of the random variables.

We have

$$E(I_{t+1}^{(k)}|(S_t^{(k')}, I_t^{(k')}, R_t^{(k')}); 1\le k'\le K) = E(I_{t+1}^{(k)}|L_t(s^{(k)}))E(L_t(s^{(k)})|(S_t^{(k')}, I_t^{(k')}, R_t^{(k')}); 1\le k'\le K). \tag{2.5}$$

Based on the conditional distributions in (2.2), we have

$$E(L_t(s^{(k)})|(S_t^{(k')}, I_t^{(k')}, R_t^{(k')}); 1\le k'\le K) = \sum_{k'=1}^{K} \frac{kk' S_t^{(k)} I_t^{(k')}}{\sum_{j=1}^{K} jN^{(j)}} \tag{2.6}$$

and

$$E(I_{t+1}^{(k)}|L_t(s^{(k)})) = \lambda L_t(s^{(k)}). \tag{2.7}$$

Combining (2.6) and (2.7), we have

$$E(I_{t+1}^{(k)}|(S_t^{(k')}, I_t^{(k')}, R_t^{(k')}); 1\le k'\le K) = \lambda \sum_{k'=1}^{K} \frac{kk' S_t^{(k)} I_t^{(k')}}{\sum_{j=1}^{K} jN^{(j)}}. \tag{2.8}$$

So we have

$$E(I_{t+1}^{(k)}) = E(E(I_{t+1}^{(k)}|(S_t^{(k')}, I_t^{(k')}, R_t^{(k')}); 1\le k'\le K)) = \lambda \sum_{k'=1}^{K} \frac{kk' E(S_t^{(k)} I_t^{(k')})}{\sum_{j=1}^{K} jN^{(j)}} \tag{2.9}$$

for $k = 1, 2, \ldots, K$. The other two classes of equations are as follows:

$$E(R_{t+1}^{(k)}) = E(R_t^{(k)}) + E(I_t^{(k)}), \tag{2.10}$$

and

$$E(S_{t+1}^{(k)}) = E(S_t^{(k)}) - E(I_{t+1}^{(k)}). \tag{2.11}$$

Equations (2.9), (2.10), and (2.11) describe the evolutions of the mean numbers of nodes with different states. It is widely agreed in physics and chemistry that the expectations can be used as macroscopic variables. But the existence of the terms $E(S_t^{(k)} I_t^{(k')})$ makes the solvation very difficult.

2.6 Thermodynamic Limit

Now suppose that $N\to\infty$, and $N^{(k)}\to\infty; k = 1, 2, \ldots, K$ accordingly so that the degree distribution $D(k)$ keeps invariant. For sufficiently large $\{N^{(k)}; 1\le k\le K\}$, we can substitute $I_t^{(k)}$, $R_t^{(k)}$, and $S_t^{(k)}$ to $\frac{I_t^{(k)}}{N^{(k)}}$, $\frac{R_t^{(k)}}{N^{(k)}}$, and $\frac{S_t^{(k)}}{N^{(k)}}$ and rewrite (2.9), (2.10), and (2.11) as follows:

$$E\left(\frac{I_{t+1}^{(k)}}{N^{(k)}}\right) = \lambda \sum_{k'=1}^{K} \frac{kk' N^{(k')} E(\frac{S_t^{(k)} I_t^{(k')}}{N^{(k)} N^{(k')}})}{\sum_{j=1}^{K} jN^{(j)}}, \tag{2.12}$$

$$E\left(\frac{R_{t+1}^{(k)}}{N^{(k)}}\right) = E\left(\frac{R_t^{(k)}}{N^{(k)}}\right) + E\left(\frac{I_t^{(k)}}{N^{(k)}}\right), \tag{2.13}$$

$$E\left(\frac{S_{t+1}^{(k)}}{N^{(k)}}\right) = E\left(\frac{S_t^{(k)}}{N^{(k)}}\right) - E\left(\frac{I_{t+1}^{(k)}}{N^{(k)}}\right). \tag{2.14}$$

The covariance of $\frac{I_t^{(k)}}{N^{(k)}}$ and $\frac{S_t^{(k)}}{N^{(k)}}$ is a higher order infinitesimal of $\frac{1}{N^{(k)}}$, so we can approximately set $E(\frac{S_t^{(k)}I_t^{(k')}}{N^{(k)}N^{(k')}}) = E(\frac{S_t^{(k)}}{N^{(k)}})E(\frac{I_t^{(k')}}{N^{(k')}})$. Let $i_t^{(k)}$, $r_t^{(k)}$, and $s_t^{(k)}$ be $E(\frac{I_t^{(k)}}{N^{(k)}})$, $E(\frac{R_t^{(k)}}{N^{(k)}})$, and $E(\frac{S_t^{(k)}}{N^{(k)}})$ respectively, we have macroscopic equations:

$$i_{t+1}^{(k)} = \lambda k s_t^{(k)} \sum_{k'=1}^{K} \frac{k' i_t^{(k')}}{\sum_{j=1}^{K} j N^{(j)}} = \lambda k s_t^{(k)} \sum_{k'=1}^{K} \frac{k' D(k') i_t^{(k')}}{\langle k \rangle}, \tag{2.15}$$

$$r_{t+1}^{(k)} = r_t^{(k)} + i_t^{(k)}, \tag{2.16}$$

$$s_{t+1}^{(k)} = s_t^{(k)} - i_{t+1}^{(k)}. \tag{2.17}$$

Here $\langle k \rangle = \frac{\sum_{j=1}^{K} j N^{(j)}}{N}$ is the average of degrees over the network.

Several time-continuous deterministic models with similar forms are built in Refs. [17, 18, 20]. In those models, the phenomenological equations are constructed directly without the microscopic mechanism. Furthermore, the physical means of the parameter λ are different, in our model λ is the probability of the disease transmission through an edge; while in ordinarily differential equation version it is explained as the rate of transmission.

In epidemiology, one uses the epidemic prevalence, which is the ratio of infected agent in total population, to measure the intensity of epidemics. In this paper, the epidemic prevalence is just $\sum_{k=1}^{K} D(k) r_\infty^{(k)}$, and its value depends on λ and the network structure. In the thermodynamic limit, statistical physicists interest in the onset of the phase transition as λ increasing and the threshold value if it is really exists. Now let us present the threshold phenomenon embodied in the deterministic system (2.15–2.17) and its physical explains

It is easy to see that all points with $\{i_t^{(k)} = 0; 1 \le k \le K\}$ are fixed points of this system. Now set initial conditions as $i_0^{(k)} = \epsilon = o(1)$, it can be treated as an infinitesimal disturbing at the point $\{(s^{(k)}, i^{(k)}, r^{(k)}) = (1, 0, 0); 1 \le k \le K\}$. Among the evolution of the process, if $r_t = \sum_{k=1}^{K} D(k) r_t^{(k)}$ keeps the order $o(1)$, we can always treat the $i_t^{(k)}$ as an infinitesimal disturbing. On the contrary, if r_∞ reaches a quantity with order $O(1)$, we must say that the point $\{(1, 0, 0); 1 \le k \le K\}$ is never be a stable steady state—thus happens the phase transition. Under this means, we can analyze the linear stability of (2.15) at the point and get the threshold of λ as follows:

$$\lambda_c = \frac{1}{\Delta_m} \quad \text{or} \quad \lambda_c \Delta_m = 1. \tag{2.18}$$

Here Δ_m is the largest eigenvalue of the matrix A:

$$A = \begin{pmatrix} \frac{D(1)}{\langle k \rangle} & \cdots & \frac{k' D(k')}{\langle k \rangle} & \cdots & \frac{K D(K)}{\langle k \rangle} \\ \frac{2D(1)}{\langle k \rangle} & \cdots & \frac{2k' D(k')}{\langle k \rangle} & \cdots & \frac{2K D(K)}{\langle k \rangle} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{kD(1)}{\langle k \rangle} & \cdots & \frac{kk' D(k')}{\langle k \rangle} & \cdots & \frac{kK D(K)}{\langle k \rangle} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{KD(1)}{\langle k \rangle} & \cdots & \frac{Kk' D(k')}{\langle k \rangle} & \cdots & \frac{KK D(K)}{\langle k \rangle} \end{pmatrix} \tag{2.19}$$

And the linearized system of equation (2.15) at the point $\{(1, 0, 0); 1 \le k \le K\}$ is as follows:

$$\overrightarrow{i_{t+1}} = \lambda A \overrightarrow{i_t}. \tag{2.20}$$

Here $\overrightarrow{i_t} = (i_t^{(1)}, i_t^{(2)}, \ldots, i_t^{(K)})^T$.

It is not difficult to check that $\Delta_m = \frac{\langle k^2 \rangle}{\langle k \rangle}$ is the simple eigenvalue of A and zero is the $K-1$ repeated eigenvalue.

Thus we get the threshold $\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle}$(or $\lambda_c \frac{\langle k^2 \rangle}{\langle k \rangle} = 1$). In [16, 17, 20], same threshold value is get by different method. But we can not explain why the threshold values gotten from time-discrete model and from time-continuous models are identical. Maybe it is because that essentially all treatments are doing some linearizing and the differences are wiped out. Anyway, all works point out that the structure parameter $\frac{\langle k \rangle}{\langle k^2 \rangle}$ of the heterogeneous populations is a key parameter to the dynamical processes in it, which can not be discovered if one just approximate the system to be well-mixed and homogeneous.

3 Conclusion

In this paper, we construct a microscopic mechanism of the epidemic process in heterogeneous populations. Since the individuals are not equal, we divide them into subgroups according to their degrees. Our work is based on two basic assumptions: edge repairing and edge reconnection at random. We consider that these assumptions are essentially the mass action principal for heterogeneous populations.

Based on the microscopic mechanism, we define the stochastic epidemic process. The numerical results show that there exist two kinds of phase, and phase transition happens at very small λ. Furthermore, we get the evolution dynamics of the mean epidemic. In the thermodynamic limit, the dynamical system of the means can be approximated by a deterministic system. From this deterministic system, we get the threshold $\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle}$, which is consistent with the numerical results.

This work revealed some properties of the outbreak distributions of the epidemic process in heterogeneous populations. Compared to the results gotten in homogeneous case, properties of stochastic models for heterogeneous case need to be revealed in detail by some analytical method. This would be the next task.

Acknowledgements The first author is supported by the China Postdoctoral Science Foundation funded project 20090450243. This work was supported partly by the National Basic Research Program of China(2006CB805900) and the National Natural Science Foundation of China(10625101,10531070).

References

1. Anderson, R.M., May, R.M.: Infectious Diseases of Humans: Dynamics and Control. Oxford Univ. Press, Oxford (1991)
2. Hethcote, H.W.: SIAM Rev. **42**, 599 (2000)
3. Durrett, R.: SIAM Rev. **41**, 677–718 (1999)
4. Boccara, N., Cheong, K.: J. Phys. A, Math. Gen. **25**, 2447–2461 (1992)
5. Bharucha-Reid, A.T.: Elements of the Theory of Markov Processes and Their Application. McGraw-Hill, New York (1960)
6. Anderson, H., Britton, T.: Stochastic Epidemic Models and Their Statistical Analysis. Springer, Berlin (2000)
7. Ben-Naim, E., Krapivsky, P.L.: Phys. Rev. E **69**, 050901(R) (2004)
8. Kessler, D.A., Shnerb, N.M.: Phys. Rev. E **76**, 010901 (2007)
9. Albert, R., Barabási, A.L.: Rev. Mod. Phys. **74**, 47–97 (2002)
10. Dorogovtsev, S.N., Mendes, J.F.F.: Adv. Phys. **51**, 1079 (2002)
11. Liljeros, F., Rdling, C.R., Amaral, L.A.N., Stanley, H.E., Abers, Y.: Nature **411**, 907 (2001)
12. Colizza, V., et al.: PNAS **103**, 2015–2020 (2006)

13. Hufnagel, L., Brockmann, D., Geisel, T.: PNAS **101**, 15124–15129 (2004)
14. Broder, A., et al.: Comput. Netw. **33**, 309–320 (2000)
15. Barabási, A.L., Albert, R., Jeong, H.: Physica A **281**, 69–77 (2000)
16. Pastor-Satorras, R., Vespignani, A.: Phys. Rev. Lett. **86**, 3200 (2001)
17. Moreno, Y., Pastor-Satorras, R., Vespignani, A.: Eur. Phys. J. B **26**, 521 (2002)
18. May, R.M., Lloyd, A.L.: Phys. Rev. E **64**, 066112 (2001)
19. Gillespie, D.T.: J. Phys. Chem. **81**, 2340–2361 (1977)
20. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.-U.: Phys. Rep. **424**, 175 (2006)