



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Modeling and analysis of epidemic spreading on community networks with heterogeneity

Chanchan Li^{a,*}, Guo-ping Jiang^b, Yurong Song^b, Lingling Xia^a, Yinwei Li^a, Bo Song^{a,c}

^a School of Computer, Nanjing University of Posts and Telecommunications, Nanjing, 210003, China

^b School of Automation, Nanjing University of Posts and Telecommunications, Nanjing, 210003, China

^c Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, NSW 2007, Australia

HIGHLIGHTS

- A modified community network model with heterogeneity among communities is proposed.
- A mathematical epidemic model for each community is presented based on this network.
- Study the effect of the location of initial infection node on epidemic spreading.
- Study the impact of the heterogeneity among communities on epidemic spreading.

ARTICLE INFO

Article history:

Received 20 August 2017

Received in revised form 9 April 2018

Accepted 13 April 2018

Available online 27 April 2018

Keywords:

Community structure

Heterogeneity

Epidemic dynamics

Mean-field theory

ABSTRACT

A large number of real world networks exhibit community structure, and different communities may often possess heterogeneity. In this paper, considering the heterogeneity among communities, we construct a new community network model in which the communities show significant differences in average degree. Based on this heterogeneous community network, we propose a novel mathematical epidemic model for each community and study the epidemic dynamics in this network model. We find that the location of the initial infection node only affects the spreading velocity and barely influences the epidemic prevalence. And the epidemic threshold of entire network decreases with the increase of heterogeneity among communities. Moreover, the epidemic prevalence increases with the increase of heterogeneity around the epidemic threshold, while the converse situation holds when the infection rate is much greater than the epidemic threshold.

© 2018 Elsevier Inc. All rights reserved.

1. Introduction

With the development of network science, many common characteristics of real networks have been revealed, where one of the important characteristics is that a wide range of biological and social systems contains community structure [5,9,15,17,24–26]. Different individuals may form a community because of the same attributes, such as interest, age, profession and so on. As the large-scale outbreaks of epidemic diseases have occurred many times in different regions and different races in history, it can be said that the history of human history is a history of the struggle against different epidemic diseases. And with the progress of science and technology, the convenient transportation has also promoted the spread of the epidemic diseases to a great extent, for

example, the severe acute respiratory syndrome (SARS) in 2003, the H1N1 influenza A virus in 2009, the H7N9 avian influenza virus in 2013, the Ebola virus in 2014 and the Middle East respiratory syndrome coronavirus (MERS-CoV) in 2015. The problem of epidemic spreading has gained great attention over the years and people want to predict the epidemic spread trend and take effective public health measures with limited vaccine supply. As the topology properties of networks have a profound impact on the dynamics of epidemic spreading, it is necessary to consider the effect of community structure on epidemic spreading. So far, a lot of results on epidemic dynamics in community networks have been obtained [12,14,16,29,30].

To simulate the real network, many different kinds of community network models were constructed based on classical networks, and some individual behavior characteristics in real networks (such as random walk, long-range jump and awareness) were also taken into account. Liu and Hu studied the SIS (susceptible–infected–susceptible) dynamics on a random community network model with probability p (q) of intra- (inter-)

* Corresponding author.

E-mail addresses: 2011040135@njupt.edu.cn (C. Li), jianggp@njupt.edu.cn (G. Jiang), songyr@njupt.edu.cn (Y. Song), xialingling@jspi.edu.cn (L. Xia), 2016040232@njupt.edu.cn (Y. Li), bo.song@student.uts.edu.au (B. Song).

community links. They found that the existence of community makes the degree distribution get broader, and the epidemic threshold decreases with the increase of community strength (p/q) [11]. Huang and Li studied the SI dynamics on scale-free network with community structure, they found that the community structure makes the growth of the infection slow and the epidemic prevalence reduce [7]. Salathé and Jones investigated the epidemic spread in empirical and artificial networks with community structure, and got the similar conclusion: the community structure has inhibiting effect on epidemic spreading, and the peak value of epidemic prevalence decreases with the increase of community strength. Moreover, they found that immunizing bridge nodes is more effective than simply immunizing highly connected nodes in networks with strong community structure [20]. To understand the influences of community structure and clustering coefficient on SIR (susceptible–infected–removed) dynamical processes, Wu and Liu presented a network model with adjustable clustering coefficient and adjustable community strength, and found that for a fixed community strength, the efficiency of epidemic spreading will decrease with increase of the clustering coefficient [27]. Chu et al. investigated the epidemic spreading on weighted scale-free networks with community structure, and found that compared to the internal weighting exponent, the external weighting exponent has higher impact on slowing the epidemic spreading [3]. On this basis, Min et al. further studied the spread of disease on weighted scale-free community network with two mixing styles [13]: a dense-weak style and a sparse-strong style. In [2], the authors studied the epidemic dynamics on a network with overlapping community structure and found that the epidemic prevalence increases with the increase of overlapping region. Some scholars constructed a random network composed of a sparse community and a dense community, and they studied the phenomenon that the disease persists throughout the dense community, while the outbreak and extinction states occur periodically in the sparse community [18]. Shao et al. studied the epidemic dynamics in a small-world community structure network with consideration of a traffic-driven SIS epidemic model [21], and found that the community structure can accelerate the epidemic spreading in the traffic-driven model. Stegehuis et al. studied two random network models that create a network with similar community structure [4], and found that the community structure can both enforce as well as inhibit diffusion processes, and the exact internal structures of communities barely influence the behavior of spreading processes. Based on the SI (susceptible–infected) epidemic model, Shu et al. numerically studied how weak ties (connect pairs of nodes belonging to different communities) influence epidemic dynamics [23]. Some scholars studied the SEAIR (susceptible–exposed–asymptotically infected–symptomatically infected–recovered) [32] and the stochastic SIR [31] epidemic model on scale-free networks with community structure. Bonaccorsi et al. analyzed a continuous-time SIS model on a community structure network in which individuals that belong to different communities have different infecting probabilities [1]. Ren et al. investigated SIS epidemic spreading processes on a time-varying community structure network [19]. Xia et al. presented a new epidemic SIS model to investigate the spreading behavior on community structure networks with dynamical topology. They found that the epidemic threshold is only related with the population density within the community, and the long-range motion will make the original disease-free community become the endemic state [28].

In most existing community network models, the researchers often ignored the heterogeneity among different communities. However, in the real social networks, the physical contacts among individuals in some communities are more closer than in other communities, which reflects the heterogeneity among communities. For example, in real social networks, young students community interact with each other more frequently than elderly people

community. Furthermore, studies have suggested that heterogeneity in contact patterns among individuals in heterogeneity communities has an important effect on the epidemic spreading [10]. In addition, many scholars are concerned about finding ways to suppress the spread of the virus, such as looking for more effective immune strategies and raising awareness of the individuals. While some scholars hope to find the source of epidemic spreading and control the epidemic spreading in the early stage of epidemic spreading [6,8,22,33]. When there exists heterogeneity among the communities in the network, we are curious about whether the difference in the location of the source of infection will affect the spread of the virus, and want to study the transmission rule of epidemic on the community structure network with heterogeneity among communities, which is helpful for us to find an effective control strategy for epidemic spreading.

Mathematical modeling of the dynamics of infectious diseases, as a quantitative research method, has been widely used in the field of the epidemic spreading. In this paper, we firstly construct a new community structure network model with adjustable heterogeneity among communities. Then we establish a novel epidemic spreading model based on this network model and calculate the epidemic threshold. And we analyze the impacts of heterogeneity among different communities and modularity coefficient on epidemic spreading either. The rest of this paper is organized as follows. In Section 2, we introduce in detail our network model, and the network characteristics are discussed either. In Section 3, we define the epidemic model and introduce the mean field equations. Then we give some numerical and simulations which support the theoretical analysis in Section 4. Finally, this paper ends with concluding remarks in Section 5.

2. Community network model with heterogeneity

The possible mechanisms of constructing a community structure network have been revealed by many models [2–4,7,11,13,18–21,27,28], and these network models also take into account different actual features of the real networks, like clustering coefficient [27] and edge weight [3,13]. Most of the models ignore the heterogeneity among communities, and how it affects the spread of epidemic is a practical and meaningful problem. In [10], by observing the real populations in the populated city Hong Kong, the researchers find out that the heterogeneity of contact patterns of individuals within and between different age groups is an important impacting factor in the transmission of infectious diseases, so they take into account the age structure of a population and the different contact patterns among individuals in different age groups and propose an age-structured model. When the groups of different ages mentioned in Ref. [10] are regarded as different communities, the whole population in the city can be viewed as a community structure network with heterogeneity. In this section, we construct a new community structure network model with heterogeneity among communities based on random network, in which the average degree of each community may be different.

Assuming that there are m different communities, each community has n_i ($i = 1, 2, \dots, m$) individuals. The generating algorithm of the new community network with heterogeneity can be described as follows:

- (1) Each node is assigned to a single community in accordance with the communities' size.
- (2) At each time step, each pair of unconnected nodes in i -th community can be connected with probability p_i^{in} . It should be noted that the inter-community connection probability of each community (p_i^{in}) may be different.
- (3) At each time step, randomly chose two nodes which are not neighbors and belong to different communities, then connect them with probability p^{ex} to create an inter-community link where

$p^{\text{ex}} \ll \min(p_i^{\text{in}})$, which makes the community network has a strong community strength.

Through the above process, a community network with heterogeneity can be generated, in which the heterogeneity among communities is manifested in the difference of the average degree of each community $\langle k_i \rangle$. We can get community networks with different heterogeneity by adjusting p_i^{in} . When given a set of different networks, to measure the strength of heterogeneity among communities for each network, we set a heterogeneity strength function H and its expression is:

$$H = \frac{D(\langle k_i \rangle)}{\max[D(\langle k_i \rangle)]}. \quad (1)$$

Assume that these networks all have m communities. In Eq. (1), the numerator $D(\langle k_i \rangle)$ denotes the variance of the communities' average degree of a network, it can be expressed as:

$$D(\langle k_i \rangle) = \frac{1}{m} \sum_{i=1}^m (\langle k_i \rangle - \langle k \rangle)^2, \quad (2)$$

where the sub item $\langle k \rangle$ in Eq. (2) denotes the average degree of the entire network. According to Eq. (2), the heterogeneity strength of a network with the same characteristics among communities is 0. For a set of community structure networks, the greater the differences in the communities average degree the larger the variance of the communities average degree of this network $D(\langle k_i \rangle)$ is, namely, the range of the variance array $[D(\langle k_i \rangle)]$ is very wide. Thus we divide the maximum variance of communities average degree of these networks $\max[D(\langle k_i \rangle)]$ in Eq. (1).

Suppose that the total number of edges is E , we obtain:

$$E = \sum_{i=1}^m \frac{n_i(n_i - 1)p_i^{\text{in}}}{2} + \sum_{i \neq j} n_i n_j p^{\text{ex}}. \quad (3)$$

As an important measurement for the community strength of community networks, modularity coefficient [15] is defined as follows:

$$Q = \sum_i \left[e_{ii} - \left(\sum_j e_{ij} \right)^2 \right], \quad (4)$$

where e_{ij} denotes the proportion of edges between community i and j in the entire network. So e_{ii} and $\sum_j e_{ij}$ can be described as follows:

$$e_{ii} = \frac{n_i(n_i - 1)p_i^{\text{in}}}{2E}, \quad (5)$$

$$\sum_j e_{ij} = \frac{n_i(n_i - 1)p_i^{\text{in}}}{2E} + \sum_{j \neq i} \frac{n_i n_j p^{\text{ex}}}{E}, \quad (6)$$

where E represents the total edge number, which is the same as shown in Eq. (3). Thus, the modularity coefficient of our model is:

$$Q = \sum_{i=1}^m \left[\frac{n_i(n_i - 1)p_i^{\text{in}}}{2E} - \left(\frac{n_i(n_i - 1)p_i^{\text{in}}}{2E} + \sum_{j \neq i} \frac{n_i n_j p^{\text{ex}}}{E} \right)^2 \right]. \quad (7)$$

Therefore, from Eqs. (3) and (7), we can get the relationship between the modularity coefficient Q and other parameters, such as n_i , $\langle k_i \rangle$, p_i^{in} and p^{ex} . The values of p_i^{in} and p^{ex} can be adjusted to get community structure networks with given modularity Q , n_i and $\langle k_i \rangle$.

3. Epidemic model with heterogeneity among communities

The existence of heterogeneity among communities in the network may lead to a novel phenomenon on the epidemic spreading.

So in this section, different from the previous mathematical model, we will describe the spreading process by a novel SIR (susceptible–infected–recovered) mathematical model, which is based on each community and considers both of the inter- and intra-community epidemiological processes of all nodes in a community.

In the initial stage of epidemic transmission, there are only a small number of infected individuals in the network, and the rest of them are susceptible individuals. At each time step, an susceptible individual can be infected by one of its infected neighbors with probability λ , and an infected individual can be recovered with probability γ . The infection process contains two parts for each community in general: intra-community infection and inter-community infection. Let $\langle k_i \rangle$ denote the average degree of intra-community within communities i , and $\langle k_{ij} \rangle$ denote that of inter-community between community i and j . As p_i^{in} and p^{ex} are the inner-community and inter-community connection rates, thus $\langle k_i \rangle$ and $\langle k_{ij} \rangle$ can be expressed as follows:

$$\langle k_i \rangle = \frac{p_i^{\text{in}} n_i (n_i - 1)}{n_i} = p_i^{\text{in}} (n_i - 1), \quad (8)$$

and

$$\langle k_{ij} \rangle = p^{\text{ex}} n_i n_j / n_i = p^{\text{ex}} n_j. \quad (9)$$

Let parameters S , I and R respectively represent the density of susceptible, infected and recovered individuals in the entire network. Similarly, S_i , I_i and R_i represent the proportion of healthy, infected and recovered individuals of the i -th community. Then the new epidemic model for each community can be constructed as:

$$\begin{cases} \frac{dS_i}{dt} = -\lambda \langle k_i \rangle S_i I_i - \lambda \sum_{j \neq i} \langle k_{ij} \rangle \frac{1}{m-1} S_i I_j \\ \frac{dI_i}{dt} = \lambda \langle k_i \rangle S_i I_i + \lambda \sum_{j \neq i} \langle k_{ij} \rangle \frac{1}{m-1} S_i I_j - \gamma I_i \\ \frac{dR_i}{dt} = \gamma I_i. \end{cases} \quad (10)$$

As the communities of the network have different average degrees, each community has its own epidemic threshold which is called community threshold in this paper. The following question naturally arises. What is the relation between community thresholds and entire network threshold? In the following, we will derive community thresholds and entire network threshold.

Let n_i^S and n_i^I denote the numbers of susceptible and infected individuals within i -th community, so S_i and I_i in Eq. (10) equal $S_i = n_i^S / N$, $I_i = n_i^I / N$. Because of the complex structure of community network, theoretical analysis of the epidemic threshold is very difficult. For simplicity, we assume that each community has the same size n , namely, $n_i = n$ ($i = 1, 2, \dots, m$). We substitute Eqs. (8) and (9) into Eq. (10), then the simplified equations are as follows:

$$\begin{cases} \frac{dS_i}{dt} = -\lambda p_i^{\text{in}} (n-1) S_i I_i - \lambda \sum_{j \neq i} \left[\frac{p^{\text{ex}} n}{m-1} S_i I_j \right] \\ \frac{dI_i}{dt} = \lambda p_i^{\text{in}} (n-1) S_i I_i + \lambda \sum_{j \neq i} \left[\frac{p^{\text{ex}} n}{m-1} S_i I_j \right] - \gamma I_i \\ \frac{dR_i}{dt} = \gamma I_i. \end{cases} \quad (11)$$

Combining conditions $I = \sum_{i=1}^m I_i$, $\sum_{i=1}^m I_i + R_i + S_i = 1$ and Eq. (11), we can accurately estimate the threshold. For a community network, if the modularity coefficient Q is large enough, the impact of spread through the inter-community edges on the epidemic prevalence within a community almost can be ignored.

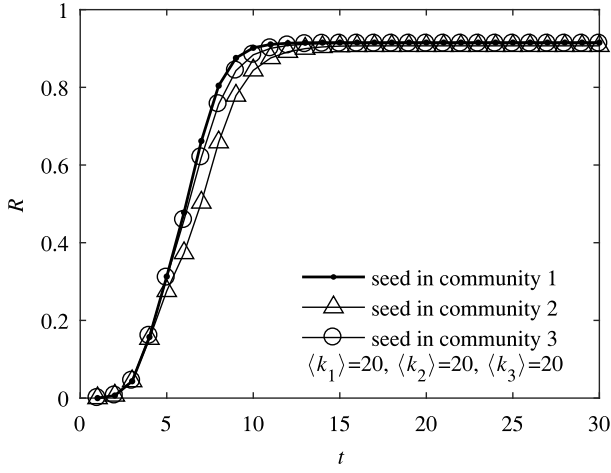


Fig. 1. Time evolution curves of epidemic prevalence for the network where the degree is 20 and the degrees of three communities are $\langle k_i \rangle = \{20, 20, 20\}$, $i = 1, 2, 3$.

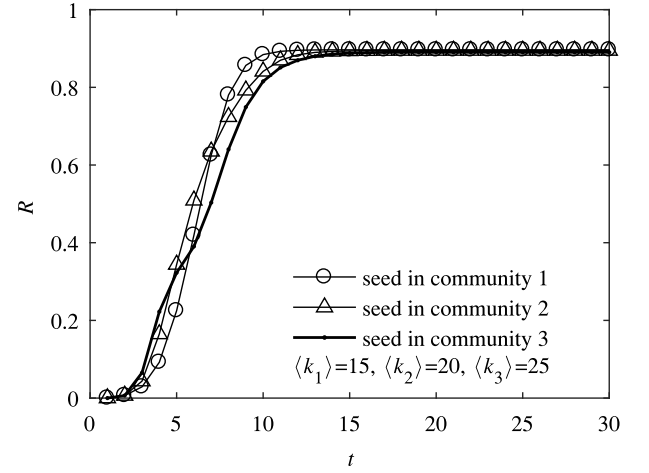


Fig. 2. Time evolution curves of epidemic prevalence for the network where the degree is 20 and the degrees of three communities are $\langle k_i \rangle = \{15, 20, 25\}$, $i = 1, 2, 3$.

Thus, we can approximately estimate the epidemic threshold of a community. The calculation process is as follows. Firstly, the first and third equations in Eq. (11) can be approximated:

$$\begin{cases} \frac{dS_i(t)}{dt} \approx -\lambda p_i^{\text{in}} n S_i(t) I_i(t) \\ \frac{dR_i(t)}{dt} = \gamma I_i(t), \end{cases} \quad (12)$$

where the initial condition $S_i(0) \approx 1/m$. From Eq. (12), we can eliminate the term of $I_i(t)$, namely,

$$\frac{1}{S_i(t)} \frac{dS_i(t)}{dt} \approx \frac{-\lambda n p_i^{\text{in}}}{r} \frac{dR_i(t)}{dt}. \quad (13)$$

Then compute the integral:

$$\int_0^\infty \frac{1}{S_i(t)} \frac{dS_i(t)}{dt} dt \approx \int_0^\infty \frac{-\lambda n p_i^{\text{in}}}{r} \frac{dR_i(t)}{dt} dt. \quad (14)$$

Thus, we get $S_i(\infty) \approx e^{-\lambda n (p_i^{\text{in}})^2 R_i(\infty)/r}$. When $t \rightarrow \infty$, there will be no infected nodes in the community network, namely, $S_i(\infty) + R_i(\infty) = 1/m$. Thus, we get:

$$R_i(\infty) = \frac{1}{m} - e^{-\lambda n p_i^{\text{in}} R_i(\infty)/r} \quad (15)$$

where $R_i(\infty) = 0$ is a trivial solution of Eq. (15), and this equation has other nonzero solutions when $d(1/m - e^{-\lambda n p_i^{\text{in}} R_i(\infty)/r})/dR_i(\infty)|_{R_i(\infty)=0} > 1$. Thus, the epidemic threshold of community i is

$$\lambda_i \approx r/(p_i^{\text{in}} n). \quad (16)$$

It is possible that part/one of the communities in the network are/is endemic, while the epidemic do not erupt throughout the network. So the theoretical threshold of the entire network λ_c is between the maximum and the minimum of the community thresholds.

4. Simulations and discussions

In this section, to discuss the influence of various factors on the dynamics of epidemic spreading, we give a set of simulations on different kinds of community networks with heterogeneity among communities and examine how the outbreak of epidemic is affected by both the heterogeneity among communities and

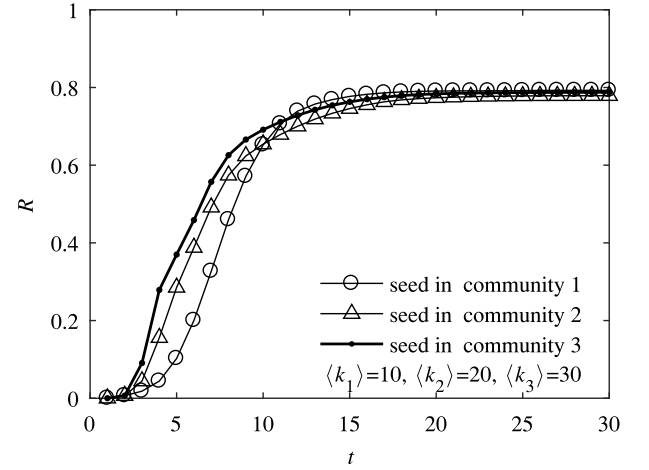


Fig. 3. Time evolution curves of epidemic prevalence for the network where the degree is 20 and the degrees of three communities are $\langle k_i \rangle = \{10, 20, 30\}$, $i = 1, 2, 3$.

modularity coefficient. In addition, we compare the effects of the infectious source locations on epidemic spreading either. In the following simulations, there is a small fraction of infected nodes in one of the communities initially, while the remaining individuals of the entire network are susceptible.

4.1. Based on homogeneous community network

Next, we make simulations and comparisons based on the homogeneous community structure network with heterogeneity among communities. It is well known that, in real networks, the outbreak of each epidemic is caused by a small number of infected individuals of a local area in the network. Since there are differences among communities, did the location of the initial infection node/nodes has effect on the epidemic spreading? To obtain a definitive answer, we do some simulations on three community structure networks, where each network is constructed based on ER random network model and has three communities with different heterogeneity.

Figs. 1–3 show the time evolution curves of epidemic size of three different networks, where the average degrees of three networks are the same, but with different average degrees in each

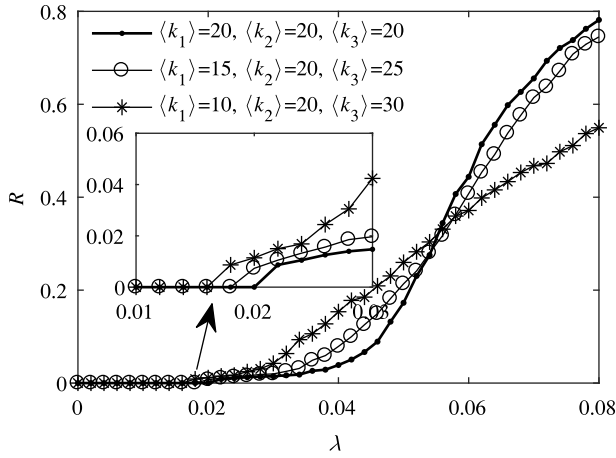


Fig. 4. The comparison curves of epidemic thresholds for three networks with different heterogeneity.

community. The average degrees for three communities are: $\langle k_i \rangle = \{10, 20, 30\}$, $\langle k_i \rangle = \{15, 20, 25\}$ and $\langle k_i \rangle = \{20, 20, 20\}$, where $i = 1, 2, 3$. In order to facilitate the comparison of the influence of the seed node locations, we assume that the network models have the same average degree of the entire network ($\langle k \rangle = 20$) and the community sizes are $n_1 = n_2 = n_3 = 500$. We find that, for each network, the location of the infection source affects only the propagation velocity rather than the epidemic prevalence. We also observe that the time duration to reach the steady state is inversely proportional to the heterogeneity among communities. In each simulation, an initial infected node is randomly chosen from a specified community. And for each network, 5000 simulations are taken with a randomly chosen seed node initially. The other simulation parameters are: the infection rate $\lambda = 0.1$ and the modularity coefficient $Q = 0.6$.

Fig. 4 shows the effects of heterogeneity among communities on epidemic threshold and prevalence. In the simulations, there are three communities within the network and the communities sizes are $n_1 = n_2 = n_3 = 500$. We keep the degree of entire network to be constant at $\langle k \rangle = 20$. The graph contains three sets of data, which compare the thresholds and prevalence of epidemic of these community networks. The lines with red dots, black circle and blue asterisk respectively correspond to the networks with intra-community degrees $\langle k_i \rangle = \{10, 20, 30\}$, $\langle k_i \rangle = \{15, 20, 25\}$ and $\langle k_i \rangle = \{20, 20, 20\}$. From Fig. 4, we find that the increase of heterogeneity among communities makes the infectious diseases easier to break out, namely the heterogeneity among communities has negative relation with the epidemic threshold. And the epidemic prevalence of a network with greater heterogeneity among communities is larger than others, whereas the converse situation holds when the infection rate is much greater than the epidemic threshold.

In Fig. 5, we study the effect of modularity coefficient Q on the epidemic threshold λ_c in three networks with different modularity coefficients (equal to 0.4, 0.5 and 0.6 respectively). In the simulations, there are three communities within each network and the average degrees of three communities are all $\langle k_i \rangle = \{15, 20, 25\}$, $i = 1, 2, 3$. The degrees of these networks are all $\langle k \rangle = 20$ and the communities sizes are $n_1 = n_2 = n_3 = 500$. It is found that the epidemic threshold decreases with the increase of the modularity coefficient of the network, and the epidemic size of the network with higher modularity coefficient is larger around the epidemic threshold, while the inverse situation occurs when the infection rate is far greater than the threshold. It is necessary to explain that although the threshold values obtained through

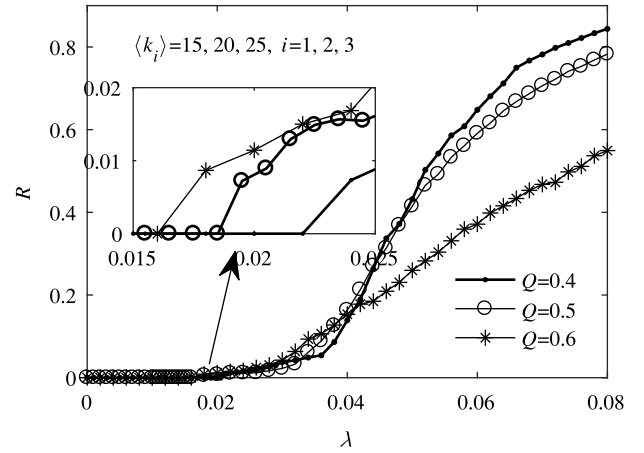


Fig. 5. The curves of epidemic thresholds for networks with different modularity and the average degrees of three communities are $\langle k_i \rangle = \{15, 20, 25\}$, $i = 1, 2, 3$.

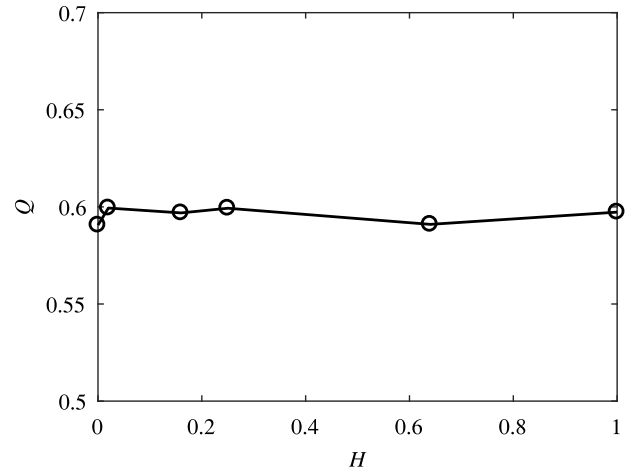


Fig. 6. The comparison curves of epidemic thresholds for three networks with different heterogeneity.

simulations cannot completely correspond to the threshold deduced theoretically, the relationship between the parameters of the network (especially the heterogeneity among communities H and the modularity coefficient Q) and the propagation threshold is in accordance with the theoretical deduction.

As the existence of heterogeneity among communities, the network models with same average degree of the entire network may show different modularity coefficients, and we cannot be sure whether the results of above simulations are only affected by the heterogeneity. So, we firstly investigate the relationship between heterogeneity and modularity of the community networks that was used before. In the following simulations, there are three communities in each network, and the community sizes are the same. Let the average degree of all networks be $\langle k \rangle = 20$. From Eq. (7), we can get the parameters (inter- and intra-community linking rate) that were used to construct the network model with given network configurations. As shown in Fig. 6, the vertical and horizontal axes in the graph stand for heterogeneity H (see Section 2) and modularity Q respectively. The curve is obtained with following network configurations: $\langle k_i \rangle = \{20, 20, 20\}$, $\langle k_i \rangle = \{18, 20, 22\}$, $\langle k_i \rangle = \{15, 20, 25\}$, $\langle k_i \rangle = \{14, 20, 26\}$, $\langle k_i \rangle = \{12, 20, 28\}$ and $\langle k_i \rangle = \{10, 20, 30\}$, and the modularity Q of corresponding networks model are 0.6. It shows that the changes of heterogeneity rarely affect modularity of community networks

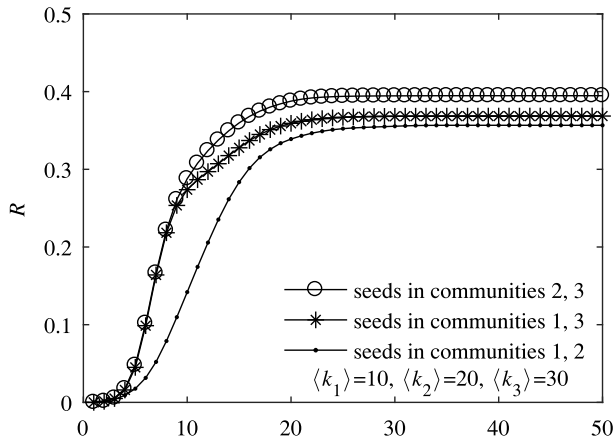


Fig. 7. Time evolution curves of epidemic prevalence for the network where the degree is 20 and the degrees of three communities are $\langle k_i \rangle = \{10, 20, 30\}$, $i = 1, 2, 3$.

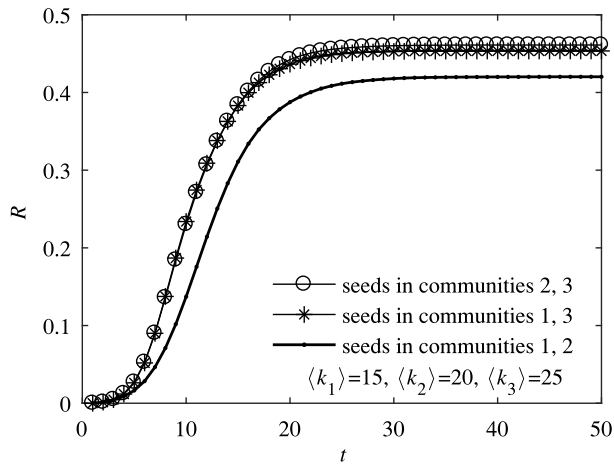


Fig. 8. Time evolution curves of epidemic prevalence for the network where the degree is 20 and the degrees of three communities are $\langle k_i \rangle = \{15, 20, 25\}$, $i = 1, 2, 3$.

with heterogeneity. So we can conclude that the results of above simulations in Figs. 1–5 are only affected by the heterogeneity.

The above simulations aim at the single source of infection. However, in real life people often move and contact to make virus appear in many regions at the same time, leading to multiple sources of infection. In the following, we will do some simulations on how the process of epidemic spreading is affected when multiple sources of infection are distributed among different communities in the community structure network with heterogeneity. In Figs. 7–9, we show the time evolution curves of epidemic size of three different networks, where the average degrees of three networks are the same, but with different average degrees in each community. We assume there are two sources of infection in the network, and that these two seed nodes are located in two communities. The parameters of the three networks used in the simulations are consistent with the previous single source of infection. The other simulation parameters are: the infection rate $\lambda = 0.05$ and the modularity coefficient $Q = 0.6$. For each network, 2000 simulations are taken with two randomly chosen seed nodes initially.

From Figs. 7–9, we find that the speed and prevalence of the epidemic spreading change with the locations of the sources of infection. Looking at each figure, we can see that the propagation speed is the fastest and the ultimate prevalence of epidemic

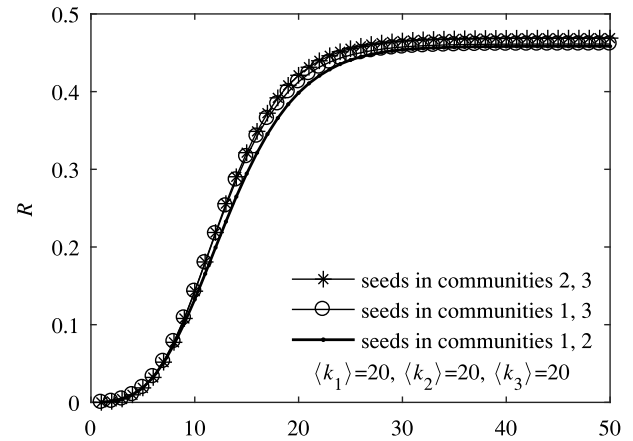


Fig. 9. Time evolution curves of epidemic prevalence for the network where the degree is 20 and the degrees of three communities are $\langle k_i \rangle = \{20, 20, 20\}$, $i = 1, 2, 3$.

spreading is the largest when two seed nodes are located in the two communities with largest average degree, and the opposite situation appears when the two seed nodes are located in the two communities with minimal average degree. And the larger the heterogeneity among communities, the greater the difference between curves. It should be noted that the difference between the three curves is not so large, which means that the locations of multiple sources of infection have little influence on the prevalence of epidemic spreading. Comparing these three figures, we find that the epidemic prevalence decreases with the increase of the heterogeneity among communities.

As the small world network is a kind of homogeneous network, the WS small world community structure network also belongs to the homogeneous networks homogeneous community networks, namely, the epidemic spreading mathematical model proposed before is still applicable to it. What is to be explained here is that we also do the same simulations in the WS small world community structure networks, and the simulation results are very similar to those obtained based on the ER random community structure networks. Therefore, we will no longer show these simulation curves here. In the last part of this section, we will compare and analyze how the epidemic spreading process is influenced by the topology of three kinds of community structure networks (the community networks respectively constructed based on the ER random network, the WS small world network and the BA scale-free network).

4.2. Based on heterogeneous community network

It is difficult to build an epidemic dynamic model based on the heterogeneous community networks with heterogeneity among communities. In the third section, some conclusions are obtained for homogeneous community network by analyzing the proposed epidemic dynamics model and verified by some Monte Carlo simulations. Next, we construct a heterogeneous community network model considering the heterogeneity among communities, and do Monte Carlo simulations to analyze the dynamics of epidemic spreading. In the following simulations, each network contains three communities, and the community sizes are $n_1 = n_2 = n_3 = 500$. The average degrees for three communities are: $\langle k_i \rangle = \{10, 20, 30\}$, $\langle k_i \rangle = \{14, 20, 26\}$ and $\langle k_i \rangle = \{20, 20, 20\}$, where $i = 1, 2, 3$, and the global average degree of each network $\langle k \rangle$ equals 20.

Figs. 10–12 show the time evolution curves of epidemic size of three different networks, where the average degrees of three

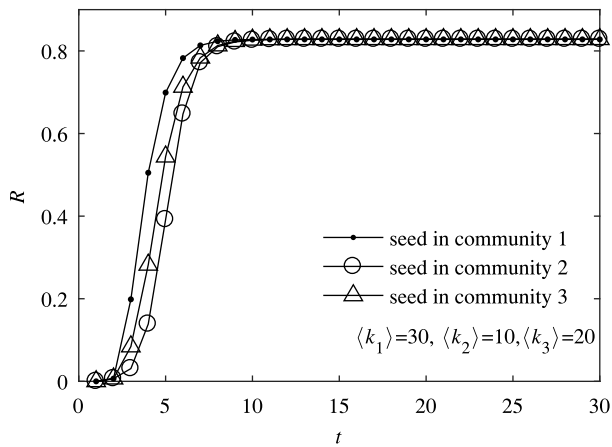


Fig. 10. Time evolution curves of epidemic prevalence with single source of infection for the scale free community structure network where the degree is 20 and the degrees of three communities are $\langle k_i \rangle = \{30, 10, 20\}$, $i = 1, 2, 3$.

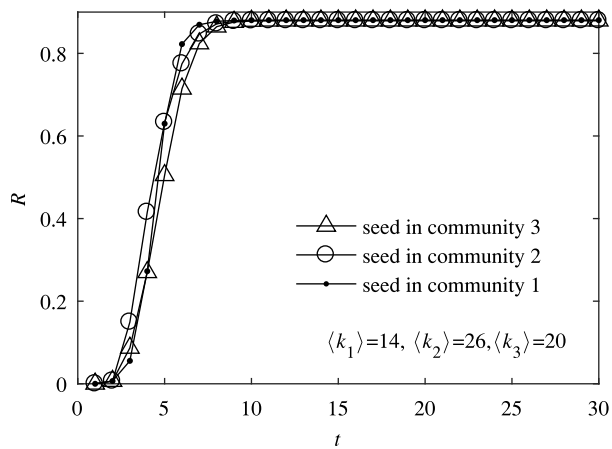


Fig. 11. Time evolution curves of epidemic prevalence with single source of infection for the scale free community structure network where the degree is 20 and the degrees of three communities are $\langle k_i \rangle = \{14, 26, 20\}$, $i = 1, 2, 3$.

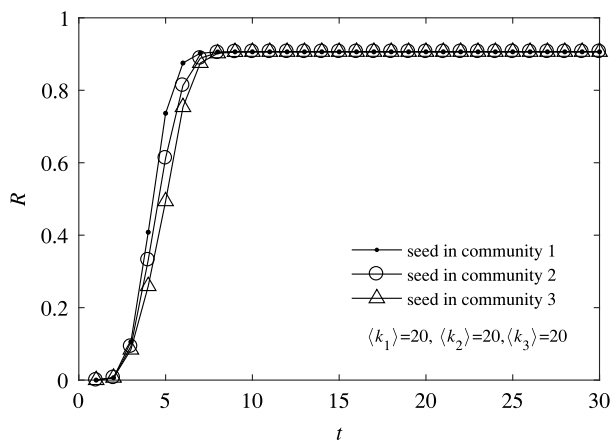


Fig. 12. Time evolution curves of epidemic prevalence for the network where the degree is 20 and the degrees of three communities are $\langle k_i \rangle = \{20, 20, 20\}$, $i = 1, 2, 3$.

networks are the same, but with different average degrees in each community. And for each network, 5000 simulations are taken with a randomly chosen seed node initially. The infection rate $\lambda =$

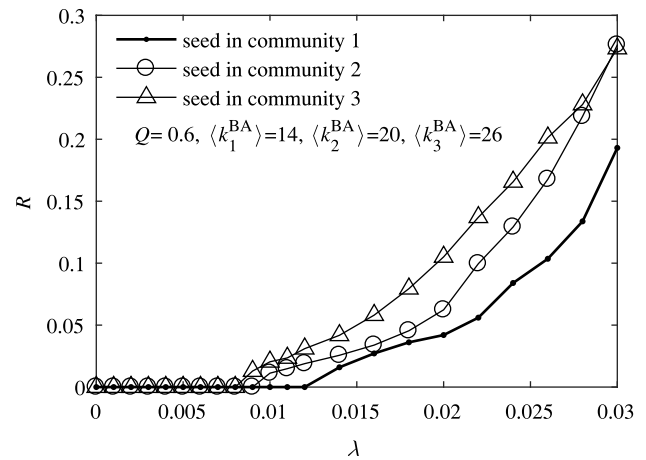


Fig. 13. The curves of epidemic threshold for each community.

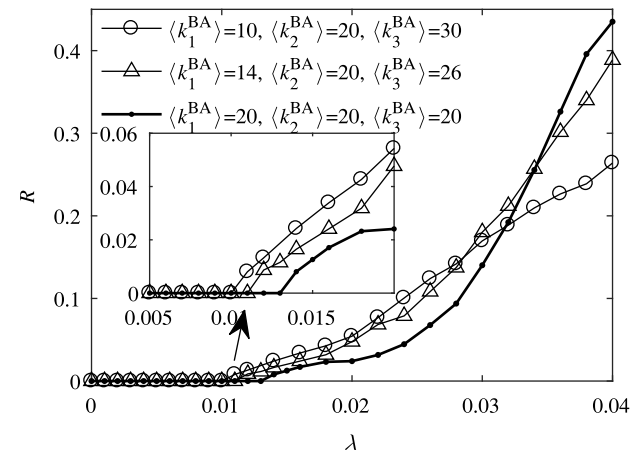


Fig. 14. The comparison curves of epidemic thresholds for three networks with different heterogeneity.

0.1 and the modularity coefficient $Q = 0.6$. In order to find out the impact of the location of infection source on the epidemic spreading process, we randomly chose an initial infected node from a specified community in each simulation. From these three figures, we find that the location of the infection source only affects the propagation velocity. The average degrees for three communities are shown in each figure. From Fig. 10 to Fig. 12, the heterogeneity among communities decreases, while the corresponding epidemic prevalence increases. This is because with the increase of the heterogeneity among communities, the links between nodes become closer in some communities, while the edges of the nodes in other communities are thinning which is analogous to the situation that deleting the inner edges of most communities and adding edges only in a small number of communities, so the epidemic will be suppressed when the virus spreads to the sparse community.

Fig. 13 shows the curves of epidemic threshold for each community in a heterogeneous community networks with heterogeneity among communities where the average degrees for three communities are $\langle k_i \rangle = \{14, 26, 20\}$, $i = 1, 2, 3$, and the community coefficient $Q = 0.6$. Communities with larger average degree of intra-community have smaller epidemic threshold. From Fig. 14, we can get the global epidemic threshold, and its value indeed between the maximum and the minimum of the communities' thresholds.

Fig. 14 shows the curves of epidemic thresholds for three networks with different heterogeneity. And Fig. 15 shows the curves

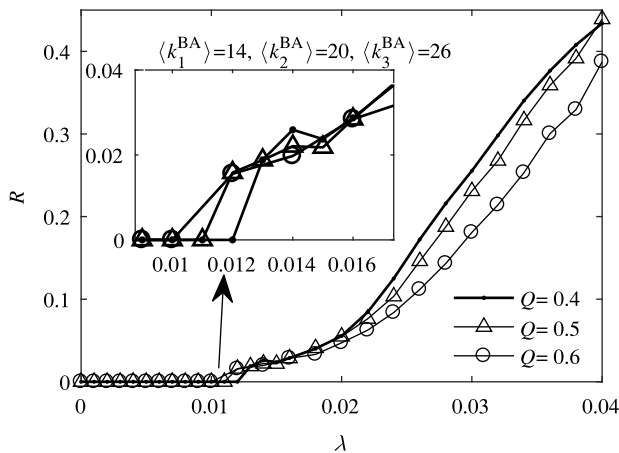


Fig. 15. The curves of epidemic thresholds for networks with different modularity and the average degrees of three communities are $\langle k_i \rangle = \{14, 26, 20\}$, $i = 1, 2, 3$.

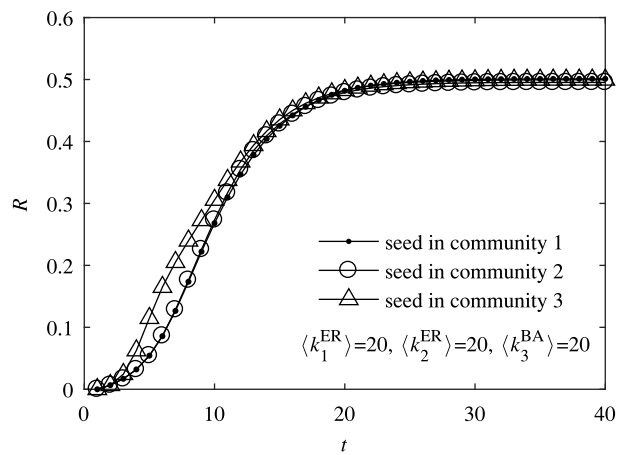


Fig. 16. Time evolution curves of epidemic prevalence with single source of infection for the mixing style network (communities 1, 2 are random networks and community 3 is a scale-free network) where the degree of entire network is 20 and the degrees of three communities are $\langle k_i \rangle = \{20, 20, 20\}$, $i = 1, 2, 3$.

of epidemic thresholds for networks with different modularity with a given heterogeneity ($\langle k_i \rangle = \{14, 26, 20\}$, $i = 1, 2, 3$) among communities. From Fig. 14, we find that the heterogeneity among communities has negative relation with the epidemic threshold, namely, the infectious diseases become easier to break out with the increase of heterogeneity among communities. And the epidemic prevalence of a network with greater heterogeneity among communities is larger than others, whereas the converse situation holds when the infection rate is much greater than the epidemic threshold. From Fig. 15, we find that the epidemic threshold decreases with the increase of the modularity coefficient of the network, and the epidemic size of the network with higher modularity coefficient is larger around the epidemic threshold, while the inverse situation occurs when the infection rate is far greater than the threshold. These conclusions are consistent with those obtained from homogeneous community networks with heterogeneity among communities.

4.3. Based on mixing style community network

The networks in real life is more complex, for example, the internal topology of some groups in the network tends to the scale-free network, and some groups are closer to the small world or

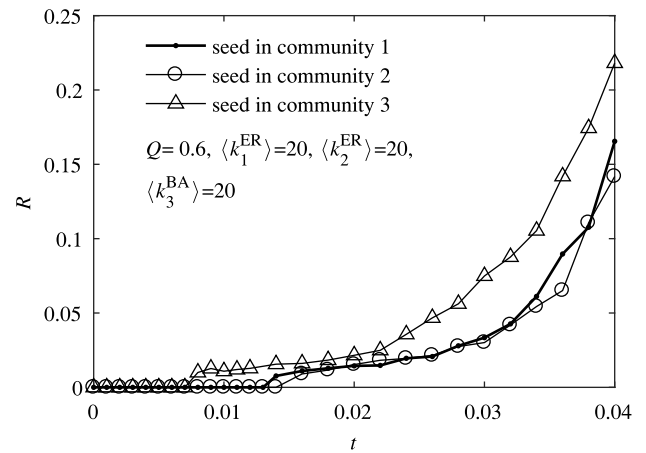


Fig. 17. The thresholds of each community of the mixing style network (communities 1, 2 are random networks and community 3 is a scale-free network) where the degree of entire network is 20 and the degrees of three communities are $\langle k_i \rangle = \{20, 20, 20\}$, $i = 1, 2, 3$.

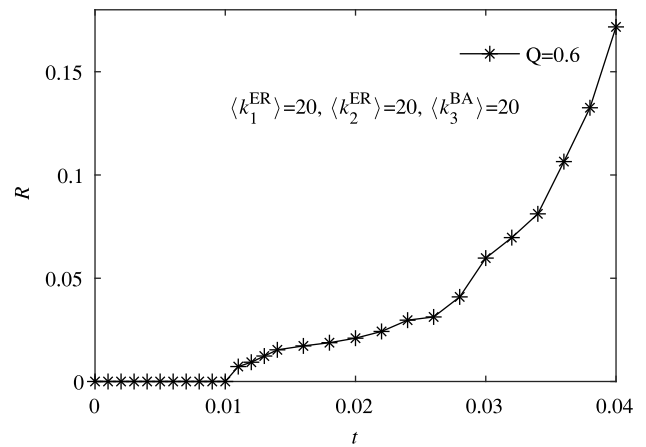


Fig. 18. The entire network threshold of the mixing style network (communities 1, 2 are random networks and community 3 is a scale-free network) where the degree of entire network is 20 and the degrees of three communities are $\langle k_i \rangle = \{20, 20, 20\}$, $i = 1, 2, 3$.

random networks, that is, several mixing style groups form a large network, we called it the mixing style community network. It is another kind of community network with heterogeneity among communities, which is more consistent with the actual network. However, its mathematical epidemic spreading dynamics model is more difficult to build. Next, we construct the mixing style community structure network based on scale-free network and random network, and analyze the epidemiology through Monte Carlo simulations.

We first try to see whether the location of the infection source of the mixed mode community structure network will have an impact on the propagation speed and the final transmission scale. Let the average degree in each community be the same, namely, $\langle k_i \rangle = \{14, 26, 20\}$, $i = 1, 2, 3$. From Fig. 16, we find that, only the speed of epidemic spreading is affected by the location of the infection source.

From Figs. 17 and 18, we find that the value of the global epidemic threshold is indeed between the maximum and the minimum of the communities thresholds. We mainly verify the formal conclusions, and more in-depth studies will be carried out later.

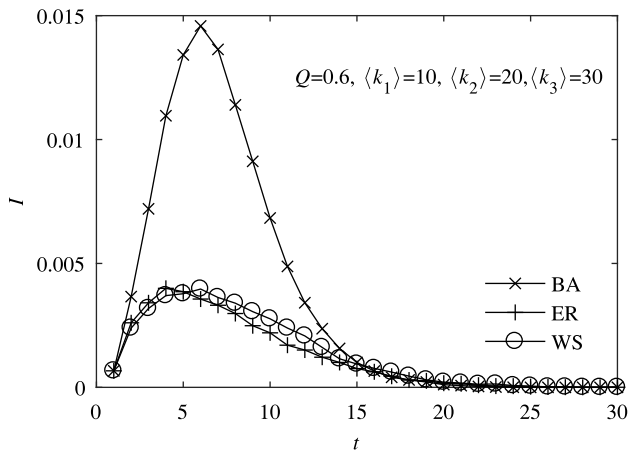


Fig. 19. Time evolution curves of epidemic prevalence for three kinds of community structure network with $\langle k_i \rangle = \{10, 20, 30\}$, $i = 1, 2, 3$.

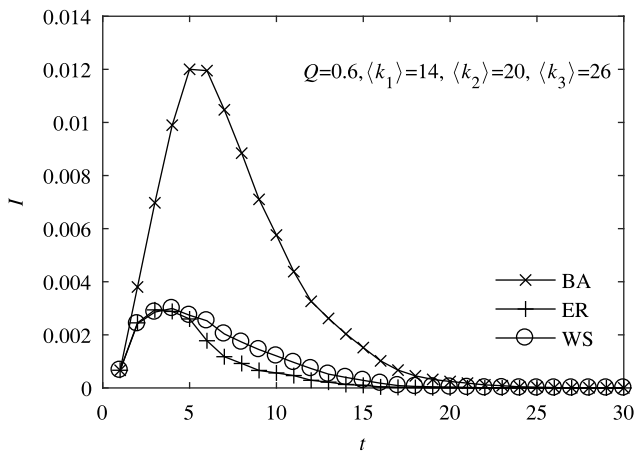


Fig. 20. Time evolution curves of epidemic prevalence for three kinds of community structure network with $\langle k_i \rangle = \{14, 20, 26\}$, $i = 1, 2, 3$.

4.4. Comparisons of epidemic dynamics between different networks

Next, according to ER random network model algorithm, WS Small World Network model algorithm and BA scale-free network model algorithm, we generate three kinds of community structure networks with different community heterogeneity among communities. Using the SIR propagation dynamics model, the epidemic spreading dynamics of infectious diseases in different network topologies are analyzed and compared by simulations. The network and related network parameters in simulation are as follows. There are three communities within each network and the communities sizes are $n_1 = n_2 = n_3 = 500$, the node degree of the entire network $\langle k \rangle$ is constant at 20. The graph contains three sets of data, which compare the thresholds and prevalence of epidemic of these community networks. Let the intra-community average degree $\langle k_i \rangle$ respectively be 10, 20 and 30 to generate community networks with larger heterogeneity among communities, and let $\langle k_i \rangle$ respectively be 14, 20 and 26 to generate community networks with smaller heterogeneity among communities.

Figs. 19 and 20 show the time evolution comparison curves of epidemic prevalence with single source of the infection in the BA scale-free community network, the ER random community network and WS small world community network, where the infection rate $\lambda = 0.025$. Here, the modularity coefficient Q is set to 0.6 for these three kinds of community networks, and

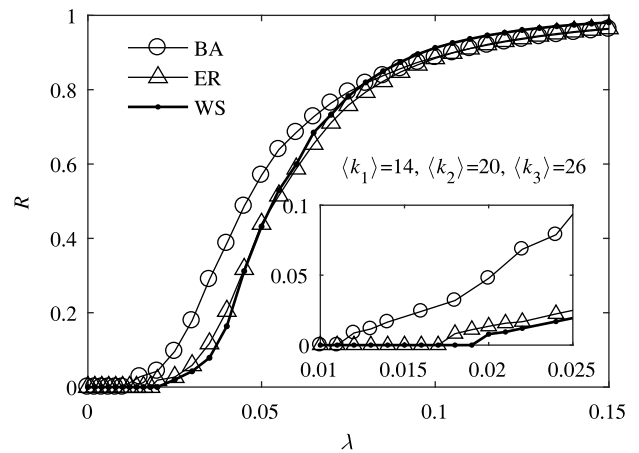


Fig. 21. The entire network threshold of the mixing style network for three kinds of community structure network with $\langle k_i \rangle = \{10, 20, 30\}$, $i = 1, 2, 3$ and $Q = 0.6$.

the intra-community average degree $\langle k_i \rangle = 10, 20, 30$ and 14, 20, 26, namely, Fig. 19/20 represents the simulation results in networks with larger/smaller heterogeneity among communities. For the propagation velocity of epidemic spreading, from these two figures, we can find that $BA > ER > WS$, which is consistent with the situation in the networks without community structure. And the peak value of the epidemic prevalence in the BA scale-free community network is obviously far greater than the other two, while the other two networks' peak values are basically the same.

From Fig. 21, we can find that, for these three community structure network models, the change trend of the epidemic threshold of the entire network is in order of $WS > ER > BA$, which is also consistent with the situation in the networks without community structure. In addition, when the infection rate is near the critical value, with the increase of transmission rate, the growth rate of epidemic prevalence of each network from large to small is $BA > ER > WS$, and when the rate of infection increases to a certain value, the situation reverses. Finally, the ultimate epidemic prevalence on the WS small world community network is the largest.

5. Conclusion

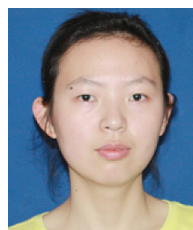
To summarize, we have proposed a modified community network model considering the heterogeneity among communities, which is more close to real networks. Based on this network with heterogeneity among communities, we have presented a mathematical epidemic model for each community and studied the epidemic dynamics in this network model. It has been found that community networks with different heterogeneities but the same average degree have different epidemic thresholds, and the epidemic threshold value of the entire network is between the maximum and minimum values of communities epidemic threshold. And when there is only one seed node in the network, the location of the initial infected individuals affects only the propagation velocity rather than the epidemic prevalence, while when there are multiple sources of infection in the network, the location of the sources of infection will have some influences on both the speed and prevalence of the epidemic spreading. In addition, the epidemic threshold is inversely proportional to the heterogeneity among communities. The study results are helpful to predict epidemic spreading trend more accurately and find useful strategies for controlling disease spreading.

Acknowledgments

We thank the National Natural Science Foundation of China (Grant Nos. 61373136, 61374180 and 61672298) for financial support.

References

- [1] S. Bonaccorsi, S. Ottaviano, P.F. De, A. Socievoli, M.P. Van, Epidemic outbreaks in two-scale community networks, *Phys. Rev. E* 90 (1) (2014) 012810.
- [2] J. Chen, H. Zhang, Z.H. Guan, T. Li, Epidemic spreading on networks with overlapping community structure, *Physica A* 391 (4) (2012) 1848–1854.
- [3] X. Chu, J. Guan, Z. Zhang, S. Zhou, Epidemic spreading in weighted scale-free networks with community structure, *J. Stat. Mech.* 2009 (07) (2009) P07043.
- [4] S. Clara, V.D.H. Remco, L.J.S.H. Van, Epidemic spreading on complex networks with community structures, *Sci. Rep.* 6 (2016) 29748.
- [5] M. Girvan, M.E.J. Newman, Community structure in social and biological networks, *Proc. Natl. Acad. Sci.* 99 (12) (2002) 7821–7826.
- [6] V. Hatzopoulos, M. Taylor, P.L. Simon, I.Z. Kiss, Multiple sources and routes of information transmission: implications for epidemic dynamics, *Math. Biosci.* 231 (2) (2011) 197–209.
- [7] W. Huang, C. Li, Epidemic spreading in scale-free networks with community structure, *J. Stat. Mech.* 2007 (01) (2007) P01014.
- [8] J. Jiang, S. Wen, S. Yu, Y. Xiang, W. Zhou, Identifying propagation sources in networks: State-of-the-art and comparative studies, *IEEE Commun. Surv. Tutor.* 19 (1) (2017) 465–481.
- [9] S. Kitchovitch, P. Liò, Community structure in social networks: applications for epidemiological modeling, *PLoS One* 6 (7) (2011) e22220.
- [10] K.O. Kwok, B. Davoudi, S. Riley, B. Pourbohloul, Early real-time estimation of the basic reproduction number of emerging or reemerging infectious diseases in a community with heterogeneous contact pattern: Using data from Hong Kong 2009 H1N1 Pandemic Influenza as an illustrative example, *PLoS One* 10 (9) (2015) e0137959.
- [11] Z. Liu, B. Hu, Epidemic spreading in community networks, *Europhys. Lett.* 72 (2) (2005) 315–321.
- [12] M.X. Liu, W. Wang, Y. Liu, M. Tang, S.M. Cai, H.F. Zhang, Social contagions on time-varying community networks, *Phys. Rev. E* 95 (5) (2017) 052306.
- [13] Y. Min, X. Jin, Y. Ge, J. Chang, The role of community mixing styles in shaping epidemic behaviors in weighted networks, *PLoS One* 8 (2) (2013) e57100.
- [14] A. Moinet, R. Pastor-Satorras, A. Barrat, Effect of risk perception on epidemic spreading in temporal networks, *Phys. Rev. E* 97 (1) (2018) 012313.
- [15] M.E.J. Newman, M. Girvan, Finding and evaluating community structure in networks, *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 69 (2) (2004) 026113.
- [16] W. Pan, G.Q. Sun, Z. Jin, How demography-driven evolving networks impact epidemic transmission between communities, *J. Theoret. Biol.* 382 (2015) 309–319.
- [17] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, A. Vespignani, Epidemic processes in complex networks, *Rev. Modern Phys.* 87 (3) (2015) 925.
- [18] X.L. Peng, M. Small, X.J. Xu, X. Fu, Temporal prediction of epidemic patterns in community networks, *New J. Phys.* 15 (11) (2013) 113033.
- [19] G. Ren, X. Wang, Epidemic spreading in time-varying community networks, *Chaos* 24 (2) (2014) 023116.
- [20] M. Salathé, J.H. Jones, Dynamics and control of diseases in networks with community structure, *PLoS Comput. Biol.* 6 (4) (2010) e1000736.
- [21] F. Shao, G.-P. Jiang, Traffic driven epidemic spreading in homogeneous networks with community structure, *J. Netw.* 7 (5) (2012) 850–855.
- [22] Z. Shen, W.X. Wang, Y. Fan, Z. Di, Y.C. Lai, Reconstructing propagation networks with natural diversity and identifying hidden sources, *Nature Commun.* 5 (2014) 4323.
- [23] P. Shu, M. Tang, K. Gong, Y. Liu, Effects of weak ties on epidemic predictability on community networks, *Chaos* 22 (4) (2012) 043124.
- [24] Y. Wang, S. Wen, Y. Xiang, W. Zhou, Modeling the propagation of worms in networks: a survey, *IEEE Commun. Surv. Tutor.* 16 (2) (2014) 942–960.
- [25] S. Wen, M.S. Haghighi, C. Chen, Y. Xiang, W. Zhou, W. Jia, A sword with two edges: Propagation studies on both positive and negative information in online social networks, *IEEE Trans. Comput.* 64 (3) (2015) 640–653.
- [26] S. Wen, W. Zhou, J. Zhang, Y. Xiang, W. Zhou, W. Jia, Modeling propagation dynamics of social network worms, *IEEE Trans. Parallel Distrib. Syst.* 24 (8) (2013) 1633–1643.
- [27] X. Wu, Z. Liu, How community structure influences epidemic spread in social networks, *Physica A* 387 (2) (2008) 623–630.
- [28] C. Xia, S. Sun, F. Rao, J. Sun, J. Wang, Z. Chen, SIS model of epidemic spreading on dynamical networks with community, *Front. Comput. Sci. China* 3 (3) (2009) 361–365.
- [29] Z. Xu, K. Li, M. Sun, X. Fu, Interaction between cluster synchronization and epidemic spread in community networks, *arXiv preprint*, 2017. mye-hostarXiv:1710.10539.
- [30] Y. Yu, G. Xiao, Infection spreading, detection and control in community networks, *J. Complex Netw.* 5 (4) (2017) 625–640.
- [31] H. Zhang, Z.H. Guan, T. Li, X.H. Zhang, D.X. Zhang, A stochastic SIR epidemic on scale-free network with community structure, *Physica A* 392 (4) (2013) 974–981.
- [32] J.-P. Zhang, Z. Jin, Epidemic spreading on complex networks with community structure, *Appl. Math. Comput.* 219 (6) (2012) 2829–2838.
- [33] Z. Zhang, H. Wang, X. Lin, H. Fang, D. Xuan, Effective epidemic control and source tracing through mobile social sensing over WBANs, in: *Proceedings of the 32nd IEEE International Conference on Computer Communications, INFOCOM*, Turin, Italy, 2013, pp. 300–304.



Chanchan Li was born in Shanxi, China, in 1989. She is a Ph.D. candidate in school of computer at Nanjing University of Posts and Telecommunications. Her current research interests include information security and complex dynamical networks.



Guo-ping Jiang, Professor and Ph.D., doctoral supervisor of Nanjing University of Posts and Telecommunications. His research interests include chaos synchronization and control, chaos-based communication, and complex dynamical networks.



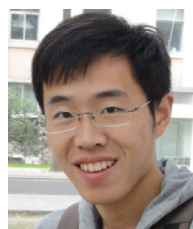
Yurong Song, Professor and Ph.D., doctoral supervisor of Nanjing University of Posts and Telecommunications. Her recent research interests include information dissemination and its control strategies in complex networks; modeling, simulation and intelligent optimization of adaptive networks.



Lingling Xia was born in Jiangsu, China, in 1988. She gets a Ph.D. in information security this year at Nanjing University of Posts and Telecommunications. Her research interests include information security and information propagation dynamics in complex networks.



Yinwei Li was born in Shandong, China, in 1981. He is a Ph.D. candidate in school of computer at Nanjing University of Posts and Telecommunications. His current research interests include information security and complex dynamical networks.



Bo Song was born in Shandong, China, in 1987. He is now a Ph.D. candidate both in faculty of engineering and information technology at University of Technology Sydney and in school of computer at Nanjing University of Posts and Telecommunications. His research interests include information security and modeling, simulation and intelligent optimization of adaptive networks.