

# The Role of Community Mixing Styles in Shaping Epidemic Behaviors in Weighted Networks

Yong Min<sup>1,2</sup>, Xiaogang Jin<sup>3</sup>, Ying Ge<sup>2</sup>, Jie Chang<sup>2\*</sup>

**1** College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou, China, **2** College of Life Sciences, Zhejiang University, Hangzhou, China, **3** College of Computer Science, Zhejiang University, Hangzhou, China

## Abstract

The dynamics of infectious diseases that are spread through direct contact have been proven to depend on the strength of community structure or modularity within the underlying network. It has been recently shown that weighted networks with similar modularity values may exhibit different mixing styles regarding the number of connections among communities and their respective weights. However, the effect of mixing style on epidemic behavior was still unclear. In this paper, we simulate the spread of disease within networks with different mixing styles: a dense-weak style (i.e., many edges among the communities with small weights) and a sparse-strong style (i.e., a few edges among the communities with large weights). Simulation results show that, with the same modularity: 1) the mixing style significantly influences the epidemic size, speed, pattern and immunization strategy; 2) the increase of the number of communities amplifies the effect of the mixing style; 3) when the mixing style changes from sparse-strong to dense-weak, there is a 'saturation point', after which the epidemic size and pattern become stable. We also provide a mean-field solution of the epidemic threshold and size on weighted community networks with arbitrary external and internal degree distribution. The solution explains the effect of the second moment of the degree distribution, and a symmetric effect of internal and external connections (incl. degree distribution and weight). Our study has both potential significance for designing more accurate metrics for the community structure and exploring diffusion dynamics on metapopulation networks.

**Citation:** Min Y, Jin X, Ge Y, Chang J (2013) The Role of Community Mixing Styles in Shaping Epidemic Behaviors in Weighted Networks. *PLoS ONE* 8(2): e57100. doi:10.1371/journal.pone.0057100

**Editor:** Vittoria Colizza, INSERM & Universite Pierre et Marie Curie, France

**Received:** May 14, 2012; **Accepted:** January 21, 2013; **Published:** February 20, 2013

**Copyright:** © 2013 Min et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** Funding provided by China Postdoctoral Science Foundation (2012M510202), National Science Foundation of China (30970281, 31170305, and 60873033), and Natural Science Foundation of Zhejiang Province, China (R1090569, LZ12F02003, and LY12F02018), and the State Key Laboratory of Software Development Environment Open Fund (SKLSDE-2012KF-05). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: jchang@zju.edu.cn

## Introduction

Structural features (e.g. heterogeneous degree distribution, small-world property, clustering effect, etc.) are critical in explaining epidemic dynamics on complex networks [1–8]. Community structure (i.e. the groups of nodes with strong connectivity among their members and weak connectivity to outside nodes) is one of the most important structural features, and broadly exists in social networks, food webs, worldwide trade networks, etc. [9–12]. The existence of community structure can restrict the outbreak within a finite number of nodes in the network in the thermodynamic limit [13]. Additionally, the strength of community structure (i.e. the difference between the density of in-community connections and the expected value of that density in a random graph, and as measured by modularity [9]) can strongly affect epidemic dynamics [14]. Moreover, epidemic on metapopulation networks is also closely related to community structures [2,15]. Recently, studies demonstrated that a weighted network shows different mixing styles, i.e. the pair of connection density and weight distributions among communities [16,17]. For example, some social networks exhibit that only a few nodes can connect to other communities with large-weight external edges [12], but most of the nodes within trade networks have small-weight external edges connecting to other communities

[18]. Although the effect of mixing style on epidemic dynamics was implicitly observed in previous studies [16], the impact of various mixing styles was not explored systematically. Here, we try to isolate the effect of the mixing style from other structural features and describe how mixing style affects epidemic dynamics and the immunization strategies in weighted community networks.

In this study, we first develop a random weighted community network model, which generates networks with different mixing styles under the same modularity. Second, we simulate the epidemic dynamics within the generated networks using the standard Susceptible-Infected-Recovered (SIR) model [3], and compare average epidemic size, speed and pattern arising from different mixing styles. Third, we analyze the performance of a targeted immunization strategy with different mixing styles. Finally, we use branching processes to provide a mean-field solution of the epidemic threshold and size of weighted community networks with arbitrary external and internal degree distribution.

## Materials and Methods

### 2.1 The epidemic model

To investigate the spread of an infectious disease within a network, we use a common approach: nodes in a network represent individuals in a population, and the edges between the

nodes represent the contacts along which an infection can spread [3,19]. The network is abstracted by undirected, weighted edges (i.e. all contacts are reciprocal, and the weight of an edge indicates a transmission rate between the two nodes of the edge). Edges always link two distinct nodes (i.e. no self-loops), and there must be only one edge between any single pair of nodes (i.e., no parallel edges). According to the standard SIR model [3,19], each node has three possible states: (S) usceptible, (I) nfectd, or (R) ecovered. Initially, all nodes are susceptible except the one infected source. Then, during a number of discrete time steps, the initial infection can spread through the network, and the simulation is halted once there are no further infected individuals. At each time step, an infected node  $i$  infects a susceptible neighbor node  $j$  with a transmission rate  $\beta_{ij}$ , which is indicated by the weight of edge  $(i,j)$ . Infected nodes are toggled from infected to recovered with a recovery rate,  $\gamma$ , i.e. the probability of recovery of an infected node per time step.

Based on the SIR model and the structure of the community network, epidemic entropy [5] is introduced to discriminate the role of mixing styles on the spatiotemporal pattern of the epidemic process. Here, the spatiotemporal pattern indicates the distribution of all infected nodes in different communities at a time and the change of the distribution along with the epidemic process. Entropy is a quantity customarily used in information theory to quantify the level of disorder of a signal or system [5]. At each time step, a snapshot of the epidemic pattern is provided by the set of values of the prevalence,  $u_x = I_x/M_x$ , in each community  $x$ , where  $I_x$  is the number of infected nodes in the community and  $M_x$  is the total number of nodes in the community  $x$ . We define the normalized entropy function  $H$ :

$$H(t) = -\frac{1}{\log N} \sum_x \rho_x(t) \log \rho_x(t) \quad (1)$$

where  $\rho_x$  is the normalized  $u_x$ ,  $\rho_x = u_x / \sum u_x$ , and  $N$  is the number of communities in the network. If the epidemic is homogeneously affecting all communities (i.e., all prevalence are equal), the entropy attains its maximum value,  $H = 1$ . Starting from  $H(0) = 0$ , which represents one initial community with an infected node,  $H(t)$  increases as more communities becoming infected, thus reducing the level of epidemic heterogeneity.

After a simulation, we calculate the total number of infected cases (the epidemic size), the ratio of the total number of infected nodes to epidemic duration (epidemic speed) and the epidemic pattern over time (epidemic entropy). We use these parameters to describe the epidemic dynamics.

## 2.2 The network model

To generate weighted community networks with different mixing styles under the same modularity, we initialize a network by creating  $N$  communities consisting of  $M$  nodes. Each community is constructed using the configuration network model [20] with an arbitrary internal degree distribution  $p(k)$ , by which each node in a community is given the appropriate number of  $k_i$  of “stubs”, i.e. ends of edges emerging from the node. Pairs of these stubs are then chosen at random and connected to form edges (internal edges). The paring of stubs continues until none are left. The number of stubs must be even, otherwise we simply set  $k_i = k_i + 1$  for a random chosen node  $i$ . A node, therefore, has an average of  $d_I = \sum k_i p(k)$  internal edges linking to nodes within the same community. Then, we add another  $k'$  edges to each node with a distribution  $q(k')$  using the configuration model, making sure that the edges fell only between communities (external edges).

Each community has  $M * d_E$  ( $d_E = \sum k' q(k')$ ) edges connecting to other communities. Thus we create a network with  $N * M$  nodes and  $E = 0.5 * N * M * (d_I + d_E)$  undirected edges. Each edge from node  $i$  to  $j$  has a weight to represent the transmission rate ( $0 < \beta_{ij} \leq 1$ ) from an infected node to a susceptible one through the edge. In order to maintain the simplest model, the weights of external and internal edges are set equal to  $w$  and  $v$  respectively ( $0 < w \leq v \leq 1$ ). When fixing  $N$ ,  $M$ ,  $p(k)$  and  $v$ , we adjust  $q(k')$  and  $w$  to obtain community networks with different mixing styles among communities but having the same modularity.

The strength of community structure is generally measured as network modularity  $Q$  [9,21], which is defined as:

$$Q = \sum_x (e_{xx} - b_x^2) \text{ and } b_x = \sum_y e_{xy} \quad (2)$$

where  $e_{xy}$  is the fraction of all edges in the network that link nodes in community  $x$  to nodes in community  $y$ . According to our weighted community network model, in average:

$$e_{xy} = \begin{cases} \frac{v d_I}{v N d_I + w N d_E} & \text{if } x = y \\ \frac{2 w d_E}{(N - 1) \cdot (v N d_I + w N d_E)} & \text{if } x \neq y \end{cases} \quad (3)$$

If edges are to fall between nodes without any regard for the communities, we have a  $Q = 0$ , and  $Q \rightarrow 1$  indicates that networks display strong community structure [9,21].

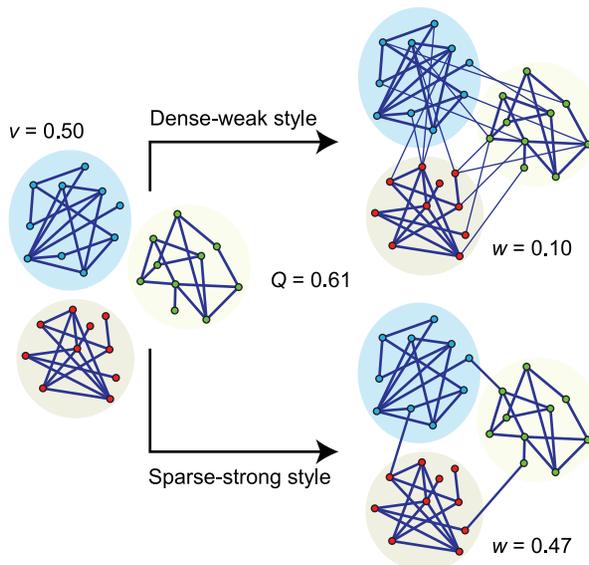
According to the definition of modularity in weighted networks, we define two mixing styles with the same  $Q$  value. The dense-weak style has a large  $d_E$  and a small  $w$  value. The style has many external edges but these edges has small weights. Contrarily, the sparse-strong style has a small  $d_E$  but a large value for  $w$ . The style is defined by having a few external edges with large weights. In order to compare the two styles, we set  $d_E = 2/M$  and 1 respectively, unless stated otherwise. A  $d_E = 2/M$  indicates that only two nodes in each community has an external edge on average, and this is the minimal value to maintain the connectivity of the generated networks. We retain all generated networks consisting of only a single connected component. A  $d_E = 1$  indicates that every node in a community has an external edge on average. An illustration of the difference between two mixing styles with the same  $Q$  is presented in figure 1.

In order to quantify the difference between two mixing styles, we generate 200 networks for each mixing style, and repeat 10 times of SIR simulation on each generated network with given recovery rate ( $\gamma$ ) and a randomly chosen initial infected node for each simulation. The average values of three epidemic parameters (epidemic size, speed and entropy) are calculated after 2000 simulations. Based on the comparison of mean values of three epidemic parameters between two styles, the difference from the two mixing styles can be illustrated.

## Results

### 3.1 Simulation results

We generate networks with different mixing styles maintaining the same modularity value. We set both the external and internal degree distributions as Poisson distributions, thus the average internal and external degree ( $d_E$  and  $d_I$ ) directly reflect the entire distribution and connection density. According to the simulating results, the average epidemic size and speed are strongly affected



**Figure 1. Schematic illustration of sparse-strong and dense-weak styles of weighted community networks.** The large circles denote the communities of the network, and the small circles denote the nodes. The lines denote the edges of the network, and its width represents the weight of the edges (i.e., the transmission rate between two nodes). The network parameters are:  $N=3$  (number of communities),  $M=10$  (number of nodes in each community),  $d_i=3$  (average internal degree),  $d_E=0.2$  or  $1$  (average external degree),  $v=0.5$  (weight of internal edges) and  $w=0.10$  or  $0.47$  (weight of external edges). The modularity  $Q=0.61$  for both styles.

doi:10.1371/journal.pone.0057100.g001

by the mixing styles (figure 2A and B). Although these networks have the same number of communities and modularity value, the dense-weak style networks result in a larger epidemic size (maximal 2.8-fold at  $\gamma=0.35$ ) and a faster epidemic speed than the sparse-strong style networks (figure 2A and B). The difference of each epidemic parameter between the two styles is obvious, therefore, the mixing style among communities has a significant effect on epidemic size and speed when the modularity values are equal.

In addition to the differences between the final status according to the two styles, epidemic entropy also demonstrates that the spatiotemporal pattern of epidemics depends upon the mixing style. The entropy and number of infected communities over time are shown in figure 2C and D. We find that diseases can rapidly reach more additional communities in dense-weak style networks, and synchronously spread during a long time window. Moreover, the appearance of maximal entropy in the dense-weak style is later than in the sparse-strong style, therefore, the later appearance and higher value of maximal entropy suggest that the end of an epidemic is rapid in the dense-weak style. In summary, the dense-weak style displays a relative “parallel” evolution ( $H$  is large) of the epidemics. Contrarily, we observe a different scenario for the sparse-strong style, where the evolution of epidemics is relative “sequential”. The maximum entropy of the sparse-strong style is significantly smaller than that of the dense-weak style, and a relatively slow decrease in the entropy curve indicates a long-lasting spreading effect in a relatively few communities (figure 2C and D). Therefore, the mixing style also affects the spatiotemporal pattern of an epidemic in networks with the same modularity.

In order to evaluate the effect of the number of communities on the epidemic profiles, we increase the number of communities

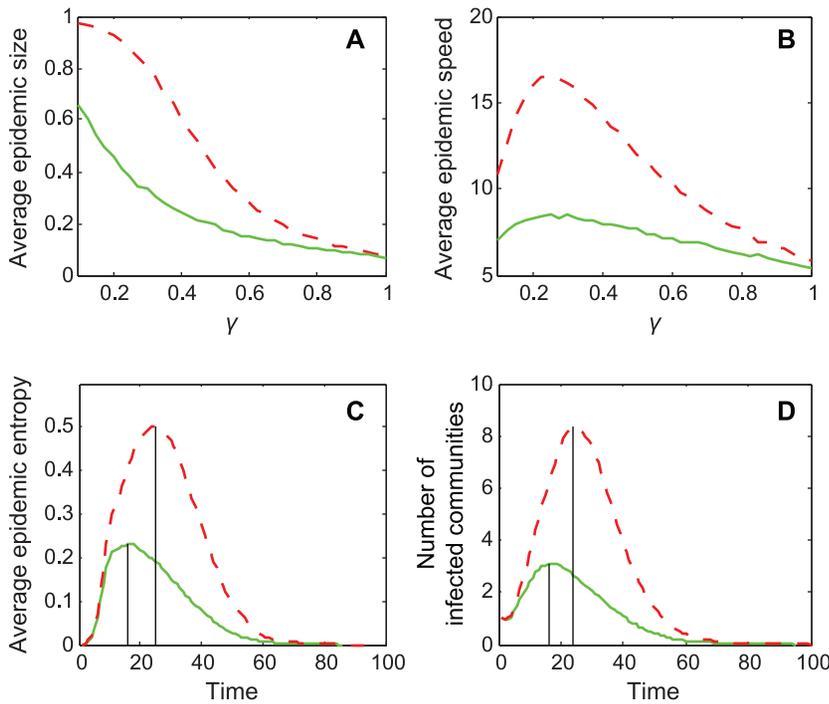
from  $N=2$  to 20 while keeping all the other parameters the same. We find that the increase in the number of communities actually enhances the distinction between epidemic behaviors on dense-weak and sparse-strong style networks (figure 3A and B). The result shows that each epidemic parameter is obvious different when  $N>4$  in the simulation. Although the effect of the mixing style is more obvious in the situation with the inclusion of more communities, the effect cannot be ignored when the number of communities was small.

In order to investigate the effects of the change of external connection density ( $d_E$ ) in detail, we simulate epidemic spreading for each  $M*d_E$  (from 2 to 25) with the same modularity, and calculate an average epidemic size and entropy. We find that the epidemic size and maximal entropy change nonlinearly with the increase of  $d_E$ , and exhibit a ‘saturation point’ (figure 3C and D). When  $M*d_E$  is below a point ( $\approx 10$  in the figure), two parameters change rapidly with the increase of external connection density, such that epidemic size and entropy are more than doubled. When  $M*d_E$  is above the point, however, the epidemic parameters become relatively stable. This observation is critical because the saturation point of the external connection density can be an important indicator of an epidemic spread within community networks. When the modularity value is the same, the mixing style could be a critical feature in controlling the spread of infectious diseases.

For halting or mitigating an epidemic, targeted immunization interventions or social distancing interventions focusing on changing the structure of the network of susceptible individuals can work in such a way as to make it more difficult for pathogens to spread [22–24]. Here, we rank the nodes of a network by their number of external edges as immunization targets. According to our network model setting, the rank reflects both the level of connection and the betweenness of a node in community networks [14,25,26]. The sparse-strong style experiences a rapid decrease in epidemic size when a few critical nodes are initially immune (figure 3E), due to the shortage of nodes with external edges in the sparse-strong style. However, the dense-weak style displays a smooth change in epidemic size with the increase of the proportion immunized (figure 3E); because high number of external edges between communities can effectively maintain the connectivity of entire networks. Therefore, the mixing style affects not only the epidemic behaviors on weighted community networks but also the application of immunization strategies.

### 3.2 Mean-field analysis

The simulation results described above demonstrate that the mixing style has special effects on epidemic size, speed and pattern, but another important epidemic parameter, global epidemic threshold, is still unclear. Colizza and Vespignani (2007) derived the expression of the epidemic threshold parameter for the case of SIR model in subpopulation networks where communities were coupled via Markovian mobility of individuals [15]. In this section, we use multiple-type branching processes to calculate the mean-field expression of the epidemic threshold and size on uncorrelated networks with arbitrary external and internal degree distribution [27]. According to the network model presented in section 3, we assume that the size of the network and the communities is approximately infinite ( $N, M \rightarrow \infty$ ) and the recovery rate  $\gamma=1$  in the theoretical study [20]. In this situation, the epidemic processes form a tree-like infectious component (figure 4): starting from an index case (the initial infected node), the pathogens can infect root nodes through external edges, and infect non-root nodes through internal edges. Here, ‘root’ indicates the node that was the first case infected by other communities. The distribution of potential



**Figure 2. Comparison of Epidemic properties in sparse-strong and dense-weak styles of weighted community networks.** (A) average epidemic size (i.e., density of infected nodes) as a function of recovery rate  $\gamma$ ; (B) average epidemic speed (i.e. the ratio of total infected nodes to total epidemic steps) as a function of recovery rate  $\gamma$ ; (C) average value of the epidemic entropy as a function of time steps; and (D) average number of total infected communities (it has at least one infected node at a time) as a function of time steps. The network parameters are:  $N=20$ ,  $M=50$ ,  $d_i=4$ ,  $d_E=0.04$  or  $1$ ,  $\nu=0.5$  and  $w=0.49$  or  $0.02$ . The modularity  $Q=0.94$  for both mixing styles. The recovery rate  $\gamma=0.35$  in (C) and (D). Solid curves represent sparse-strong style networks and dashed curves represent dense-weak style networks. The vertical lines indicate the appearing time of maximal value of average epidemic entropy and number of infected communities in C and D. doi:10.1371/journal.pone.0057100.g002

root and non-root nodes connected to an infected root node can be described by two generating functions:

$$F_{root}(x) = \frac{G'_{root}(x)}{G'_{root}(1)} = \frac{\sum_{k'=1}^{\infty} k'q(k')x^{k'-1}}{d_E} \quad (4)$$

$$G_{non-root}(x) = \sum_{k=0}^{\infty} p(k)x^k \quad (5)$$

where  $F_{root}(x)$  indicates the distribution of the number of external edges of a root node excluding the edge that the pathogen arrived along,  $G_{root}(x) = \sum q(k)x^k$  is the generating function for the external degree distribution and  $G_{non-root}(x)$  is the generating function for the internal degree distribution. By considering the different infection probabilities (the weights of edges) of external and internal edges, the distributions of actual numbers of infected cases from a root node follow the functions:

$$D_{root}(x) = F_{root}(1-w+wx) \quad (6)$$

$$D_{non-root}(x) = G_{non-root}(1-v+vx) \quad (7)$$

Similarly, the distributions of the actual numbers of infected cases from a non-root node follow the two generating functions:

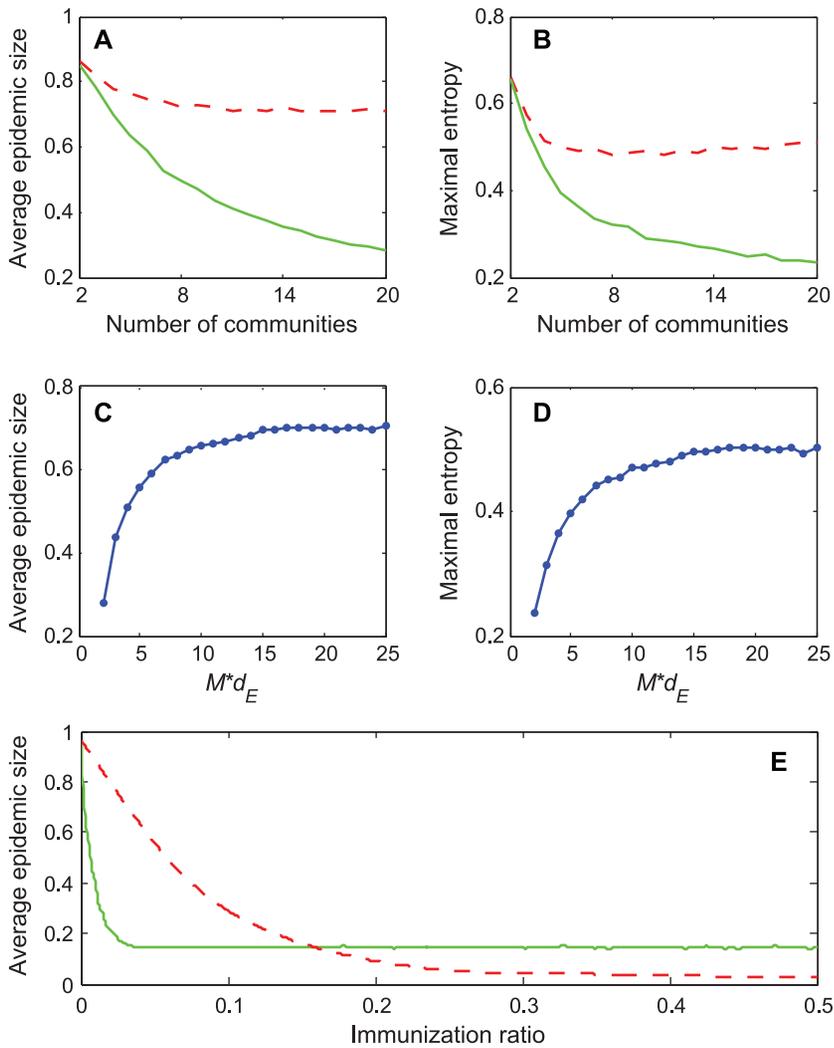
$$\begin{aligned} E_{root}(x) &= G_{root}(1-w+wx) \\ &= \sum_{k'=0}^{\infty} q(k')(1-w+wx)^{k'} \end{aligned} \quad (8)$$

$$\begin{aligned} E_{non-root}(x) &= F_{non-root}(1-v+vx) \\ &= \frac{G'_{non-root}(1-v+vx)}{G'_{non-root}(1)} = \frac{\sum_{k=1}^{\infty} kp(k)(1-v+vx)^{k-1}}{d_i} \end{aligned} \quad (9)$$

Therefore, a two-type branching process can describe the approximate tree-like infectious component. To define the behavior of the two-type process, we need two generating functions, each in two variables. The first generating function,  $f^{root}(x_1, x_2)$ , determines the distribution of the number of root and non-root nodes to be connected by a root node. Because the external and internal degree distribution is independent, we let:

$$\begin{aligned} f^{root}(x_1, x_2) &= \sum_{i,j \geq 0} P^{root}(i,j)x_1^i x_2^j \\ &= D_{root}(x_1)D_{non-root}(x_2) \end{aligned} \quad (10)$$

where  $0 \leq x_1, x_2 \leq 1$ , and  $P^{root}(i,j)$  is the probability that a root node connects to  $i$  root nodes and  $j$  non-root nodes. Similarly, the



**Figure 3. Detail effects of mixing style.** (A) and (B) average epidemic size and maximal epidemic entropy as a function of the number of communities. Network parameters are  $N = 2$  to  $20$ ,  $M = 50$ ,  $d_l = 4$ ,  $d_E = 0.04$  or  $1$ ,  $v = 0.5$  and  $w = 0.49$  or  $0.02$ . The modularity  $Q = 0.49$  to  $0.94$ , and is the same for both styles for each  $N$ ; (C) and (D) average epidemic size and maximal epidemic entropy as a function of the connection density. Network parameters are  $N = 20$ ,  $M = 50$ ,  $d_l = 4$ ,  $d_E = 0.04$  to  $0.5$ ,  $v = 0.5$  and  $w = 0.49$  or  $0.02$ . The modularity  $Q = 0.94$  for all networks with different connection densities; (E) average epidemic size (the ratio of infected nodes to total number of nodes in original networks) as a function of immunization ratio, which is the ratio of nodes with immune state (i.e. cannot be infected). The targeted immunization of nodes is according to the descending order of their external degree. The Network parameters are  $N = 20$ ,  $M = 50$ ,  $d_l = 4$ ,  $d_E = 0.04$  or  $1$ ,  $v = 0.5$  and  $w = 0.49$  or  $0.02$ . The modularity  $Q = 0.94$  for both mixing styles. Solid curves represent sparse-strong style networks and dashed curves represent dense-weak style networks. The recovery rate  $\gamma = 0.35$  in all SIR simulations.

doi:10.1371/journal.pone.0057100.g003

second generating function,  $f^{non-root}(x_1, x_2)$ , determines the distribution of the number of root and non-root nodes to be connected by a non-root node:

$$f^{non-root}(x_1, x_2) = E_{root}(x_1)E_{non-root}(x_2) \quad (11)$$

Let  $a_{s,t}$  be the expected number of type  $t$  offspring of a single type  $s$  infected node in one time step ( $s, t = \text{root or non-root}$ ), hence the mean matrix,  $A$ , of the two-type branching process is:

$$\begin{aligned} A &= \begin{bmatrix} a_{root,root} & a_{root,non-root} \\ a_{non-root,root} & a_{non-root,non-root} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial f^{root}}{\partial x_1}(1) & \frac{\partial f^{root}}{\partial x_2}(1) \\ \frac{\partial f^{non-root}}{\partial x_1}(1) & \frac{\partial f^{non-root}}{\partial x_2}(1) \end{bmatrix} \\ &= \begin{bmatrix} D'_{root}(1) & D'_{non-root}(1) \\ E'_{root}(1) & E'_{non-root}(1) \end{bmatrix} \end{aligned} \quad (12)$$

According to the basic theorem of branching processes, if the matrix  $A$  has a maximal eigenvalue  $ig_{max}(A) > 1$ , the extinction probability of the branching process is less than 1, in other words, the pathogen can occupy entire networks. Therefore, the condition of epidemic threshold is:

$$ig_{max}(A) = 1 \tag{13}$$

If all  $a_{s,t} > 0$ , the condition displayed as:

$$\frac{G'_{root}(1) - wG''_{root}(1)G'_{non-root}(1) - vG''_{non-root}(1)}{w(G'_{root}(1))^2} = 1 \tag{14}$$

In addition to epidemic threshold, multiple-type branching process also provides a mathematical solution of epidemic size for our model when  $ig_{max}(A) > 1$ . If  $ig_{max}(A) > 1$ , the extinction probability of the two-type branching process represents the probability that a node does not link to giant percolation component in the network. The extinction probability of the two-type process is the smallest nonnegative root  $u = (u_{root}, u_{non-root})$  of the questions of Eq.10 and 11:

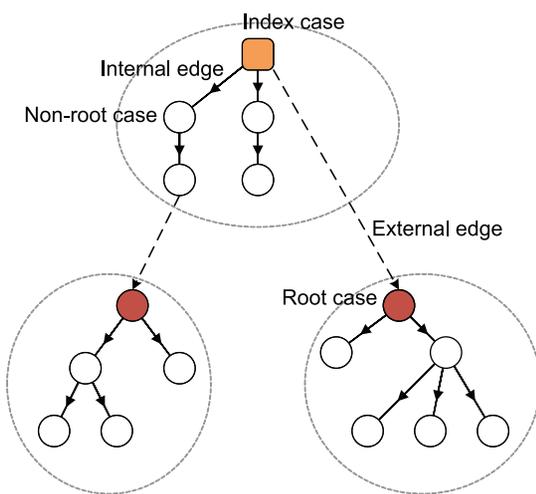
$$\begin{cases} r_{root} = D_{root}(r_{root})D_{non-root}(r_{non-root}) \\ r_{non-root} = E_{root}(r_{root})E_{non-root}(r_{non-root}) \end{cases} \tag{15}$$

where  $r_{root}$  is the extinction probability starting from a root node, and  $r_{non-root}$  is the extinction probability starting from a non-root node. By randomly choosing a node as index case, the average epidemic size equals to:

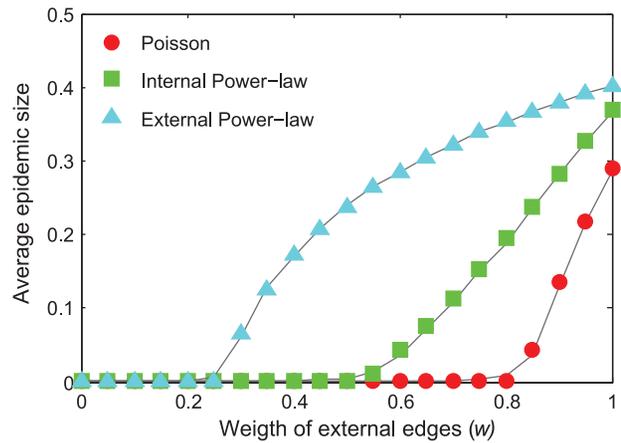
$$1 - G_{root}(1 - w + wr_{root})G_{non-root}(1 - v + vr_{non-root}) \tag{16}$$

The theoretical threshold and giant component size is identified by the simulation result (figure 5).

The Eq. 14–16 provide us with a clear description of the effect of mixing styles. First, the mean-field analysis extends the mixing style from some certain external and internal degree distribution



**Figure 4. Tree-like structure of infectious components in weighted community networks.** Due to the hypothesis regarding the infinite size of the network and community, the infecting processes will not cross the same node or community more than once. doi:10.1371/journal.pone.0057100.g004



**Figure 5. A comparison of the theoretical and the simulated epidemic size and threshold.** The curves depict the average epidemic size as a function of the external transmission rates in the simulations, and the points denote the theoretical values of the epidemic size in three situations (calculated using Eq. 16). The Network parameters are  $N=500, M=500, v=0.2$ . We use mean value and range to define a degree distribution. The mean value of internal Poisson

$(Pr(k) \sim \frac{\lambda^k e^{-\lambda}}{k!})$  and Power-law distribution  $(Pr(k) \sim k^{-\alpha})$  is 1.2, and the mean value of external Poisson and Power-law distribution is 1. The maximal internal degree does not exceed 10 in simulations. There are three combinations of internal and external degree distributions: internal Poisson with external Poisson (red circles), internal Power-law with external Poisson (green rectangles), and internal Poisson with external Power-law (cyan triangles). The recovery rate  $\gamma=1$  in the SIR simulations.

doi:10.1371/journal.pone.0057100.g005

(e.g. Poisson distribution in ref 13 and power-law distribution in ref 16) to the uncorrelated combination of arbitrary distributions. The equation shows that the second moment of the generating function of the external and internal degree distribution plays a critical role in determining the relation between connection density ( $d_I$  and  $d_E$ ) and transmission rate ( $v$  and  $w$ ) in the mixing style. Second, the equation demonstrates that the external and internal degree distribution does not simply sum up to determine the epidemic behaviors as the assumption of previous studies in unweighted community networks [13]. Third, the external and internal degree distribution plays a “symmetric” role in affecting epidemic threshold, which complements the previous simulation study [16].

## Discussion

A large number of human infectious diseases spread directly from one person to another, and early works on the spread of such diseases within a network were focused on the heterogeneous distribution of the node degree [3,14,19]. When considering additional features of the network structure within real populations, one of the most important findings was the demonstration that the modularity value can strongly affect epidemic dynamics [14]. Our results extend this finding and show that even in weighted networks with the same modularity, fundamentally different epidemic dynamics are observed due to different mixing styles among communities. Second, various mixing styles among communities have a wide existence in real networks [12,18], and this is the first time that the mixing style is clearly isolated from other structural features to determine nonlinear effects and the epidemic behaviors of weighted community networks. Third, our

study supports the idea that data on the strength of community structure will not be sufficient to predict the entire epidemic dynamics related to weighted networks [16]; the mixing style among communities complements the knowledge about community networks.

The impact of mixing style on epidemic behaviors suggests that it should be considered in the detection and measurement of community structure in complex networks. Current measures of community always emphasize the statistical distinction between the connection strengths inside and outside of the communities [11,21,28]. However, the effect of mixing style is still ignored in current metrics. Although the degree distribution is one of the most studied statistical features, mixing style is still not considered in the metrics of the community structure. The Eq. 14 clearly shows that external and internal degree distribution has obvious effects on diffusion dynamics in community networks. Therefore, the integration of modularity and mixing style can greatly improve our understanding of diffusion on weighted networks within the community structure.

The effect of mixing style also plays a critical role in the study of epidemic dynamic in networks with metapopulation structures [5,15,29–32]. A metapopulation consists of a group of spatially separated populations of the same species, which interact at some level. Colizza et al. (2006) showed that the heterogeneous connectivity patterns among subpopulations in airline transportation networks dynamically affected the behavior of the diffusion

processes occurring in the networks [5]. Our result regarding the effect of second moments of external or internal degree distributions supports the conclusion. Moreover, multiple-type branching process was used to study epidemic spreading in structured population [27]; our approach extends application of multiple-type branching process to the populations whose structure is described by explicit network structures. In the future, our approach has potential applications in metapopulation networks [5].

In this study, we initially consider the epidemic effect of mixing style among communities. In the future, other aspects of mixing style should also be considered, including the correlation between the internal and external structure of a community, and the dynamic contact structure among communities. Additionally, overlapping community structure is likely to play a critical role in infectious disease dynamics [33,34], because people are members of multiple, potentially overlapping communities. The exact effect of mixing style among communities on epidemic dynamics is still open for investigation.

## Author Contributions

Conceived and designed the experiments: YM XJ. Performed the experiments: YM. Analyzed the data: YM YG. Contributed reagents/materials/analysis tools: YM. Wrote the paper: YM YG JC.

## References

- Keeling MJ, Eames KT (2005) Networks and epidemic models. *J R Soc Interface* 2: 295–307.
- Vespignani A (2012) Modelling dynamical processes in complex socio-technical systems. *Nat Phys* 8: 32–39.
- Pastor-Satorras R, Vespignani A (2001) Epidemic Spreading in Scale-Free Networks. *Phys Rev Lett* 86: 3200.
- Newman MEJ (2002) Spread of epidemic disease on networks. *Phys Rev E* 66: 016128.
- Colizza V, Barrat A, Barthélemy M, Vespignani A (2006) The role of the airline transportation network in the prediction and predictability of global epidemics. *Proc Natl Acad Sci U S A* 103: 2015–2020.
- Balcan D, Colizza V, Gonçalves B, Hu H, Ramasco JJ, et al. (2009) Multiscale mobility networks and the spatial spreading of infectious diseases. *Proc Natl Acad Sci U S A* 106: 21484–21489.
- Rocha LEC, Liljeros F, Holme P (2010) Information dynamics shape the sexual networks of Internet-mediated prostitution. *Proc Natl Acad Sci U S A* 107: 5706–5711.
- Smieszek T, Fiebig L, Scholz RW (2009) Models of epidemics: when contact repetition and clustering should be included. *Theor Biol Med Model* 6: 11.
- Girvan M, Newman MEJ (2002) Community structure in social and biological networks. *Proc Natl Acad Sci U S A* 99: 7821–7826.
- Kumpula JM, Onnela J-P, Saramäki J, Kaski K, Kertész J (2007) Emergence of Communities in Weighted Networks. *Phys Rev Lett* 99: 228701.
- Santo F (2010) Community detection in graphs. *Phys Rep* 486: 75–174.
- Newman MEJ (2012) Communities, modules and large-scale structure in networks. *Nat Phys* 8: 25–31.
- Liu Z, Hu B (2005) Epidemic spreading in community networks. *Europhys Lett* 72: 315–321.
- Salathé M, Jones JH (2010) Dynamics and Control of Diseases in Networks with Community Structure. *PLoS Comput Biol* 6: e1000736.
- Colizza V, Vespignani A (2007) Invasion Threshold in Heterogeneous Metapopulation Networks. *Phys Rev Lett* 99: 148701.
- Chu X, Guan J, Zhang Z, Zhou S (2009) Epidemic spreading in weighted scale-free networks with community structure. *J Stat Mech* 7: P07043.
- Kiss IZ, Green DM, Kao RR (2008) The effect of network mixing patterns on epidemic dynamics and the efficacy of disease contact tracing. *J R Soc Interface* 5: 791–799.
- Min Y, Chang J, Jin X, Zhong Y, Ge Y (2011) The role of vegetables trade network in global epidemics. *arXiv* 11101724.
- Pastor-Satorras R, Vespignani A (2002) Epidemic dynamics in finite size scale-free networks. *Phys Rev E* 65: 035108.
- Newman MEJ (2003) The structure and function of complex networks. *SIAM Rev* 45: 167–256.
- Newman MEJ, Girvan M (2004) Finding and evaluating community structure in networks. *Phys Rev E* 69: 026113.
- Albert R, Jeong H, Barabasi A-L (2000) Error and attack tolerance of complex networks. *Nature* 406: 378–382.
- Restrepo JG, Ott E, Hunt BR (2006) Characterizing the Dynamical Importance of Network Nodes and Links. *Phys Rev Lett* 97: 094102.
- Rosvall M, Bergstrom CT (2008) Maps of random walks on complex networks reveal community structure. *Proc Natl Acad Sci U S A* 105: 1118–1123.
- Guimerà R, Mossa S, Turtleschi A, Amaral LAN (2005) The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles. *Proc Natl Acad Sci U S A* 102: 7794–7799.
- May R, McLean A (2007) *Theoretical Ecology: Principles and Applications*, 3rd edn. New York: Oxford University Press.
- Vazquez A (2007) Epidemic outbreaks on structured populations. *J Theor Biol* 245: 125–129.
- Clauset A (2005) Finding local community structure in networks. *Phys Rev E* 72: 026132.
- Vespignani A (2008) Reaction-diffusion processes and epidemic metapopulation models in complex networks. *Eur Phys J B* 64: 5.
- Masuda N (2010) Effects of diffusion rates on epidemic spreads in metapopulation networks. *New J Phys* 12: 093009.
- Bajardi P, Poletto C, Ramasco JJ, Tizzoni M, Colizza V, et al. (2011) Human Mobility Networks, Travel Restrictions, and the Global Spread of 2009 H1N1 Pandemic. *PLoS ONE* 6: e16591.
- Balcan D, Vespignani A (2011) Phase transitions in contagion processes mediated by recurrent mobility patterns. *Nat Phys* 7: 581–586.
- Palla G, Derenyi I, Farkas I, Vicsek T (2005) Uncovering the overlapping community structure of complex networks in nature and society. *Nature* 435: 814–818.
- Chen J, Zhang H, Guan Z-H, Li T (2012) Epidemic spreading on networks with overlapping community structure. *Physica A* 391: 1848–1854.