Epidemic Thresholds in Networks: Impact of Heterogeneous Infection Rates and Recovery Rates

Zesheng Chen
Department of Computer Science
Indiana University - Purdue University Fort Wayne, Indiana 46805
Email: chenz@ipfw.edu

***Abstract*—The information diffusion and the influence propagation in an online social network behave similar to the epidemic process and have been studied through the classic epidemiological models. The epidemic threshold is a fundamental metric for the epidemic process to determine the condition when an epidemic can either survive or die out in a network. The previous works have been focused on finding the epidemic threshold when both infection rates and recovery rates are homogeneous or studying the transient behavior of an epidemic when the infection rates are heterogeneous. In this work, we explore the effect of the heterogeneity in both infection rates and recovery rates on the epidemic threshold through both analysis and simulations. We discover that heterogeneous infection rates and heterogeneous recovery rates have the opposite impact on the epidemic threshold. Specifically, the heterogeneity in infection rates leads to a larger epidemic threshold than in the homogeneous case. Moreover, as the degree of the heterogeneity of infection rates gets higher, the epidemic threshold increases. On the other hand, the heterogeneity in recovery rates generates a smaller epidemic threshold than in the homogeneous case. Furthermore, the epidemic threshold decreases as the degree of the heterogeneity of recovery rates gets higher.**

***Index Terms*—Epidemic thresholds, heterogeneous infection rates, heterogeneous recovery rates, Jensen's inequality.**

I. Introduction

In an online social network (OSN), a piece of information, either benign (*e.g.*, news or ideas) or malicious (*e.g.*, rumors or misinformation), can spread from friends to friends or from post to post, in a similar way to the epidemic process. As a result, the information diffusion or the influence propagation in OSNs has been modeled through the classic epidemiological models [7], [10]. For example, the spread of influence was modeled using the popular independent cascade model [11], which is essentially the classic *susceptible-infected-recovered* (SIR) model [15], [6]. Moreover, the propagation of rumors and anti-rumors has been described by the *susceptible-infected-cured* (SIC) model [20]. Furthermore, the behavior of blog cascading was characterized by the well-known *susceptible-infected-susceptible* (SIS) model [13]. Therefore, it is of great importance to understand the fundamentals of the epidemic process in order to facilitate the spread of benign information and defend OSNs against the propagation of malicious information.

A fundamental metric of the epidemic spread is the epidemic threshold, which provides a condition on which an epidemic can either die out or survive in a network [22], [3], [9], [14], [4], [15]. Specifically, in the classic SIS model, the infection rate (or the birth rate) has been used to denote the degree of intensity that an infected node can compromise one of its susceptible neighbors, whereas the recovery rate (or the death rate) has been applied to reflect the ability that an infected node can be recovered to be susceptible. If the ratio between the infection rate and the recovery rate is greater than the epidemic threshold, the epidemic survives and can exist in the network for a long time; otherwise, if the ratio is less than the threshold, the epidemic will die out eventually. The epidemic threshold is similarly defined in other epidemiological models [15]. As a result, the epidemic threshold is an important metric to evaluate when an epidemic can survive in a network for a long period of time.

The epidemic threshold has been widely studied, especially based on the SIS model [22], [3], [9], [14], [4]. Wang, Chakrabarti, and their colleagues pointed out that the epidemic threshold is exactly the inverse of the largest eigenvalue of the adjacency matrix of the network [22], [3]. Ganesh et al. and Mieghem et al. further provided the theoretical deviation to this result in [9], [14]. In our previous work, we found that the epidemic threshold also depends on the spatial correlation coefficient between neighboring nodes [4]. These previous works, however, assume homogeneous infections rates and recovery rates. That is, the infection rates are assumed to be the same for all links in a network, whereas the recovery rates are equal for all nodes.

The actual infection rates and recovery rates are heterogeneous, as pointed out by [2], [16], [5] and evidenced by SARS [21], the Witty worm [18], and Plasmodium falciparum infection [19]. Buono et al. showed that heterogeneous infection rates lead to the slow epidemic extinction [2]. Chen et al. discovered that the heterogeneity in scanning rates slows down worm propagation in the Internet [5]. Qu et al. considered different probability distributions for infection rates and found that the heterogeneity of infection rates on average retards the virus spreading [16]. These previous works on heterogeneous rates have been focused on the transient behavior of an epidemic. The impact of both heterogeneous infection rates and heterogeneous recovery rates on the epidemic threshold has not been studied systematically yet.

The goal of this work is to explore the effect of the heterogeneity in both infection rates and recovery rates on the epidemic threshold. Although we focus on the classic

SIS model, our analysis can be well applied to other epidemiological models. We use the probabilistic method and the Jensen's inequality to study the impact of heterogeneous rates. Specifically, we treat an infection rate or a recovery rate as a random variable and analyze the effect of randomness on the epidemic threshold. The Jensen's inequality helps compare the homogeneous case when all rates are the same and the heterogeneous case when rates can be different from each other. Furthermore, we apply the Taylor expansion to study the impact of the degree of the heterogeneity of rates on the epidemic threshold. Finally, we use simulations to verify our theoretical results, by finding the epidemic thresholds in three different networks (*e.g.*, a lattice, a power-law topology, and a coauthorship network).

We summarize our discoveries and contributions in the following:

- We show analytically and empirically that it is statistically easier for an epidemic to survive with heterogeneous recovery rates than with homogeneous recovery rates. That is, the epidemic threshold is smaller when recovery rates are heterogeneous than when they are homogeneous. Moreover, as the degree of the heterogeneity of recovery rates gets higher, the epidemic threshold becomes smaller.
- Through both analysis and simulations, we find that heterogeneous infection rates have the opposite effect on the epidemic threshold from heterogeneous recovery rates. That is, the epidemic threshold is larger when infection rate are heterogeneous than when they are homogeneous. Moreover, the epidemic threshold becomes larger when the degree of the heterogeneity of infection rates is higher.

The remainder of this paper is structured as follows. Section II introduces the system model. Section III uses the probabilistic method to analyze the impact of heterogeneous infection rates and recovery rates on the epidemic threshold, and Section IV applies simulations to verify our theoretical results. Finally, Section V concludes this paper.

II. System Model

We use the notation $G(V, E)$ to denote a network, where V is a set of nodes and E is a set of links. The number of nodes is denoted by $|V|$, and similarly the number of links is $|E|$. Generally, a network can be categorized into two types: regular graphs and irregular graphs. In a regular graph, each node has the same number of neighbors. Typical regular graphs include a ring and a lattice [12]. In an irregular graph, different nodes can have a different number of neighbors. Representative synthesized irregular graphs include an ER random graph [8] and a BA power-law topology [1]. Most real topologies are irregular graphs. In this work, we consider both regular and synthesized irregular graphs, as well as real topologies. A network can also be divided into two categories: directed graphs and undirected graphs. In an undirected graph, if node i connects to node j, then node j also connects to node i. A directed graph does not have such a constraint. In this work, we use the undirected graph in our simulations. However, our theoretical analysis can be applied to both directed and undirected graphs.

The epidemic threshold provides a condition on which an epidemic can either die out or survive in a network. Specifically, we consider the classic *susceptible-infected-susceptible* (SIS) model. A node in a network can have two states: susceptible or infected. If node i is susceptible at the current time, it can be infected by its infectious neighbor j with a probability called the infection rate β_{ji} ($0 < \beta_{ji} \leq 1$) at the next time step. Otherwise, node i is infected and can be recovered to the susceptible state with a probability called the recovery rate δ_i ($0 < \delta_i \leq 1$) at the next time step. The epidemic threshold τ is defined as the condition that if $(\sum_{(i,j)\in E} \beta_{ij}/|E|)/(\sum_{i\in V} \delta_i/|V|) > \tau$, the epidemic survives and infects a non-zero number of nodes; otherwise, $(\sum_{(i,j)\in E} \beta_{ij}/|E|)/(\sum_{i\in V} \delta_i/|V|) < \tau$, the epidemic dies out. In the previous works [22], [3], [9], [14], [4], it has been assumed that all nodes have the same recovery rate, *i.e.*, $\delta_i = \delta, \forall i \in V$, and all links have the same infection rate, *i.e.*,, $\beta_{ij} = \beta, \forall (i,j) \in E$. That is, the epidemic threshold is derived based on homogeneous infection rates and recovery rates. In this work, we assume that β_{ij}'s and δ_i's can be different in the network and study the impact of heterogeneous infection rates and recovery rates on the epidemic threshold.

III. Theoretical Analysis

In this section, we analytically show the impact of both heterogeneous recovery rates and heterogeneous infection rates on epidemic thresholds in networks.

A. Impact of Heterogeneous Recovery Rates

We assume that the recovery rate (or the death rate) δ_i ($\forall i \in V$) is a random variable with mean δ and variance σ_d^2 ($\sigma_d^2 \geq 0$). If $\sigma_d^2 = 0$, $\delta_i = \delta$, *i.e.*, all nodes have the same recovery rate, which represents the case of homogeneous recovery rates. Otherwise, $\sigma_d^2 > 0$, which means that different nodes can have the different recovery rate and stands for the case of heterogeneous recovery rates. Assuming node i is infected, we consider T_d as the time that it takes for node i to be recovered. Intuitively, if T_d is larger, an epidemic is easier to survive, and thus the epidemic threshold is smaller. In a discrete-time system, given δ_i, T_d follows a geometric distribution, *i.e.*,

$$\Pr(T_d = k|\delta_i) = \delta_i(1-\delta_i)^{k-1}, \quad k = 1, 2, \cdots \tag{1}$$

which leads to

$$E[T_d|\delta_i] = \frac{1}{\delta_i}. \tag{2}$$

In the homogeneous case,

$$E\left[T_d^H\right] = \frac{1}{\delta}, \tag{3}$$

where T_d^H denotes T_d when $\sigma_d^2 = 0$.

In the heterogeneous case,

$$E[T_d] = E[E[T_d|\delta_i]] = E\left[\frac{1}{\delta_i}\right]. \tag{4}$$

According to the Jensen's inequality [17], since function $f(x) = 1/x$ is a strictly convex function when $x > 0$ (*i.e.*, $f'(x) = -1/x^2 < 0$ and $f''(x) = 2/x^3 > 0$), we have

$$E[T_d] = E\left[\frac{1}{\delta_i}\right] \geq \frac{1}{E[\delta_i]} = \frac{1}{\delta} = E\left[T_d^H\right], \tag{5}$$

where the equality holds if and only if $\sigma_d^2 = 0$. Equation (5) shows that it takes a longer time for node i to be recovered in a heterogeneous case than in a homogeneous case, which leads to the following theorem.

Theorem 1: If the recovery rates δ_i's are i.i.d. random variables with mean δ and variance σ_d^2, then the epidemic threshold is smaller when $\sigma_d^2 > 0$ than when $\sigma_d^2 = 0$. That is, statistically the epidemic can be easier to survive with heterogeneous recovery rates than with homogeneous recovery rates.

Intuitively, in a network, when some nodes are infected and are with different recovery rates, the time for all these nodes to be recovered depends strongly on the node with the smallest recovery rate, which is less than the average of recovery rates. Therefore, in such a network, the epidemic is more difficult to die out.

Would the epidemic threshold be smaller when variance σ_d^2 is larger? To answer this question, we consider function $f(x) = 1/x$ and express it using the Taylor expansion at point a:

$$f(x) = \frac{1}{a} + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + H \tag{6}$$

$$= \frac{1}{a} - \frac{x-a}{a^2} + \frac{(x-a)^2}{a^3} + H, \tag{7}$$

where H contains the higher-order terms of $(x-a)$ and can be ignored. Set $x = \delta_i$ and $a = E[\delta_i] = \delta$. Then, we have

$$\frac{1}{\delta_i} \approx \frac{1}{\delta} - \frac{\delta_i - \delta}{\delta^2} + \frac{(\delta_i - \delta)^2}{\delta^3}. \tag{8}$$

Taking the expectation on both sides of the above equation, we find

$$E\left[\frac{1}{\delta_i}\right] \approx \frac{1}{\delta} + \frac{E[(\delta_i - \delta)^2]}{\delta^3} \tag{9}$$

$$= \frac{1}{\delta}\left(1 + \frac{\sigma_d^2}{\delta^2}\right). \tag{10}$$

Using $E[T_d^H]$ in Equation (3) and $E[T_d]$ in Equation (4), we have

$$E[T_d] \approx E\left[T_d^H\right]\left(1 + \frac{\sigma_d^2}{\delta^2}\right). \tag{11}$$

From the above equation, we can see that when variance σ_d^2 increases, $E[T_d]$ also increases, which leads to a smaller epidemic threshold. Therefore, we have the following conjecture.

Conjecture 1: When σ_d^2 is larger, the epidemic threshold becomes smaller. That is, it is statistically easier for the epidemic to survive with a higher degree of the heterogeneity in recovery rates.

B. Impact of Heterogeneous Infection Rates

Next, we consider the impact of heterogeneous infection rates on the epidemic threshold in a discrete-time system. We assume that the infection rate (or the birth rate) β_{ij} ($\forall (i,j) \in E$) is a random variable with mean β and variance σ_b^2 ($\sigma_b^2 \geq 0$). $\sigma_b^2 = 0$ represents the case of homogeneous infection rates, whereas $\sigma_b^2 > 0$ reflects the case of heterogeneous infection rates. We assume that node i is currently infected and will be recovered after t ($t \geq 2$) time steps. Note that $t = 1$ means that node i is recovered after only one time step and thus $\delta_i = 1$, so here we ignore such a corner case. Suppose that node i has m susceptible neighbors. We want to derive how many neighbors will be infected by node i during t time steps. Intuitively, if node i infects a fewer number of susceptible neighbors during t time steps, an epidemic is more difficult to survive, and thus the epidemic threshold is larger. We use the notation N_i to denote a set of susceptible neighbors of node i, and thus $|N_i| = m$.

Let X_j be an indicator as follows:

$$X_j = \begin{cases} 1, & \text{neighbor } j \text{ is infected by node } i \text{ during } t \\ & \text{time steps;} \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

Since for each time step node i can infect its susceptible neighbor j with probability β_{ij}, node i will *not* infect neighbor j during t time steps with the probability $(1-\beta_{ij})^t$. That is, given β_{ij}, neighbor j can be compromised by node i during t time steps with the probability

$$\Pr(X_j = 1 | \beta_{ij}) = E[X_j | \beta_{ij}] = 1 - (1-\beta_{ij})^t. \tag{13}$$

Let K denote the (random) number of neighbors of node i that are infected during t time steps, *i.e.*,

$$K = \sum_{j \in N_i} X_j. \tag{14}$$

Given β_{ij}'s, the average number of neighbors of node i that are infected during t time steps is

$$E[K | \beta_{ij}{}'s] = \sum_{j \in N_i} E[X_j | \beta_{ij}] = m - \sum_{j \in N_i} (1-\beta_{ij})^t. \tag{15}$$

In the homogeneous case,

$$E\left[K^H\right] = m - m(1-\beta)^t, \tag{16}$$

where K^H denotes K when $\sigma_b^2 = 0$.

In the heterogeneous case,

$$E[K] = E\left[E[K|\beta_{ij}{}'s]\right] = m - \sum_{j \in N_i} E\left[(1-\beta_{ij})^t\right]. \tag{17}$$

According to the Jensen's inequality [17], since function $f(x) = (1-x)^t$ is a strictly convex function when $t \geq 2$ and $0 < x < 1$ (*i.e.*, $f'(x) = -t(1-x)^{t-1} < 0$ and $f''(x) = t(t-1)(1-x)^{t-2} > 0$), we have

$$\sum_{j \in N_i} E\left[(1-\beta_{ij})^t\right] \geq \sum_{j \in N_i} (1 - E[\beta_{ij}])^t, \tag{18}$$

where the equality holds if and only if $\sigma_b^2 = 0$. That is,

$$E[K] \leq E[K^H]. \quad (19)$$

Equation (19) shows that on average an epidemic infects a fewer number of susceptible neighbors in a heterogeneous case than in a homogeneous case, which leads to the following theorem.

Theorem 2: If the infection rates β_{ij}'s are i.i.d. random variables with mean β and variance σ_b^2, then the epidemic threshold is larger when $\sigma_b^2 > 0$ than when $\sigma_b^2 = 0$. That is, statistically it is more difficult for an epidemic to survive with heterogeneous infection rates than with homogeneous infection rates.

To study the impact of the degree of the heterogeneity of infection rates on the epidemic threshold, we consider function $f(x) = (1-x)^t$ where $t \geq 2$ and $0 < x < 1$, and express it using the Taylor expansion at point a:

$$f(x) = (1-a)^t + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + H, \quad (20)$$

where H contains the higher-order terms of $(x-a)$ and can be ignored. Set $x = \beta_{ij}$ and $a = E[\beta_{ij}] = \beta$. Note that $f'(a) = -t(1-a)^{t-1}$ and $f''(a) = t(t-1)(1-a)^{t-2}$. Then, we have

$$\begin{aligned}(1-\beta_{ij})^t \quad &\approx \quad (1-\beta)^t - t(1-\beta)^{t-1}(\beta_{ij}-\beta) + \\ &\frac{t(t-1)}{2}(1-\beta)^{t-2}(\beta_{ij}-\beta)^2. \quad (21)\end{aligned}$$

Taking the expectation on the both sides of the above equation, we find

$$E\left[(1-\beta_{ij})^t\right] \approx (1-\beta)^t + \frac{t(t-1)}{2}(1-\beta)^{t-2}\sigma_b^2. \quad (22)$$

Using $E[K^H]$ in Equation (16) and $E[K]$ in Equation (17), we have

$$E[K] \approx E\left[K^H\right] - \frac{m}{2}t(t-1)(1-\beta)^{t-2}\sigma_b^2. \quad (23)$$

From the above equation, we can see that when variance σ_b^2 increases, $E[K]$ decreases, which leads to a larger epidemic threshold. Therefore, we have the following conjecture.

Conjecture 2: When σ_b^2 is larger, the epidemic threshold becomes larger. That is, it is statistically more difficult for an epidemic to survive with a higher degree of the heterogeneity in infection rates.

IV. Simulation Verification

In this section, we apply simulations to verify our theoretical analysis about the impact of heterogeneous infection rates and recovery rates on epidemic thresholds in networks.

A. Simulation Setup

We simulate the spread of an epidemic in a network by following the SIS model. Specifically, if node i is infected at time t ($t \geq 0$), it will recover with the probability δ_i or remain in the infected state with the probability $1-\delta_i$ at time $t+1$. Otherwise, if node i is susceptible at time t, it can be infected by its infectious neighbor j with the probability β_{ji} at time $t+1$. Here, the probabilities are obtained through a random number generator. For each scenario, we carry out 100 independent runs to find the mean and the variance of the epidemic threshold. At the beginning of each simulation run (*i.e.*, $t = 0$), we randomly select half of nodes to be infected, and assign an infection rate to a link (or a recovery rate to a node) according to the probability distribution of infection rates (or recovery rates). Moreover, we simulate each run long enough so that the epidemic reaches the steady state.

To generate the heterogeneous infection rates or the heterogeneous recovery rates, we consider a simple two-rate case. Taking the case of heterogeneous infection rates as an example, we assume that in a network 50% of links are associated with the infection rate β_1 and the other 50% of links have the infection rate β_2. Thus, the average of infection rates is $\beta = (\beta_1 + \beta_2)/2$, and the variance of infection rates is $\sigma_b^2 = (\beta_2 - \beta_1)^2/4$. In our simulations, we show how the epidemic threshold varies with the variance of infection rates (or recovery rates).

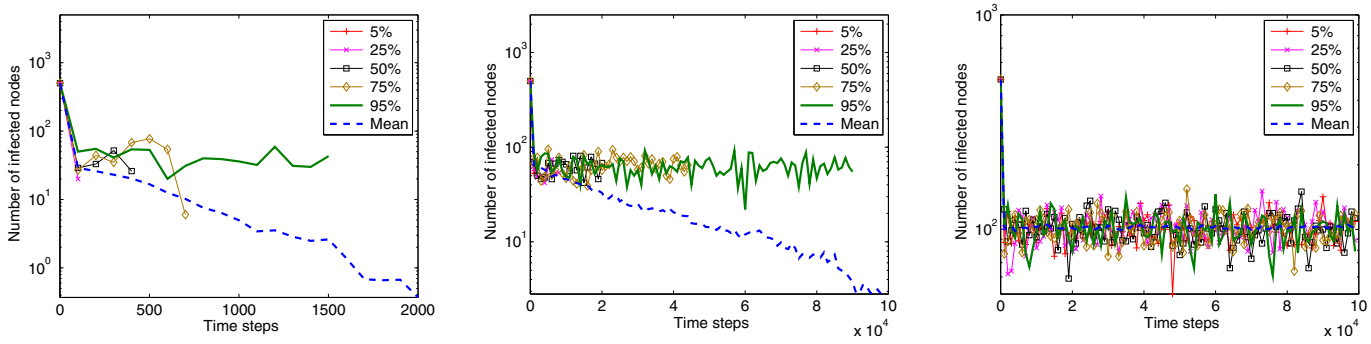
To find the epidemic threshold, we apply the binary search algorithm. Using the heterogeneous infection rates as an example, we show our binary search method for the epidemic threshold in a simulation run in Algorithm 1.

Algorithm 1 Finding the epidemic threshold τ for heterogeneous infection rates

```
Input: β1, β2
Output: τ
Set δ_low = 0 and δ_high = 1
while (δ_high − δ_low) > 0.01 ∗ δ_low do
  δ = (δ_high + δ_low)/2
  Simulate epidemic spread using β1, β2, and δ
  Calculate I_f: the average of the number of final infections over 100 runs
  if I_f > 0 then
    δ_low = δ
  else
    δ_high = δ
  end if
end while
τ = (β1 + β2)/(2 ∗ δ_high)
```

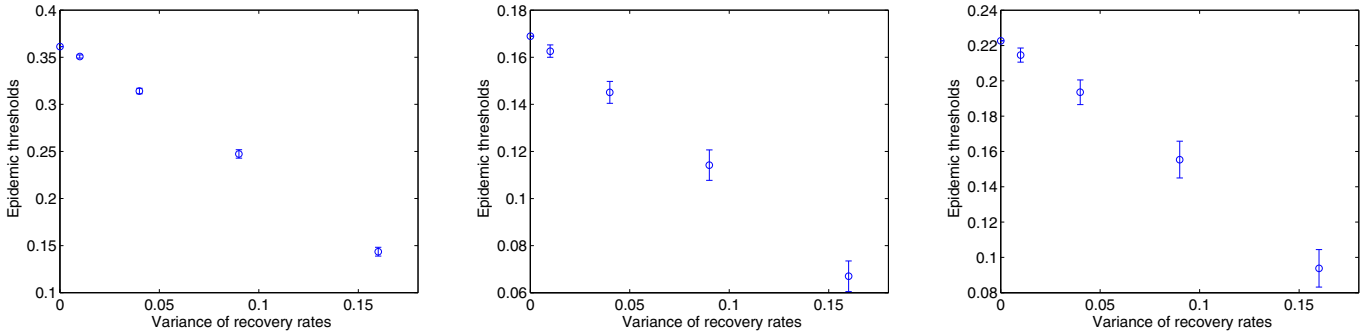
In our simulations, we consider three representative networks: a lattice [12] that is a regular graph, a BA power-law topology generated by BRITE [23] that is a synthesized irregular graph, and a coauthorship network [24] that is a real topology. The lattice contains 2,500 nodes, whereas the power-law topology has 1,000 nodes. The coauthorship network is not connected, so we only consider the giant component that contains 379 nodes and 914 links.

Figure 1 shows the sample runs of epidemic spread in a power-law topology when infection rates are heterogeneous (*i.e.*, $\beta_1 = 0.06$, $\beta_2 = 0.14$, and $\beta = 0.1$). That is, Figure 1 demonstrates the number of infected nodes over time when the recovery rate is assigned with different values (*i.e.*, $\delta = 0.7$, 0.6, or 0.5). Note that we use the log scale for the y



(a) Below the epidemic threshold ($\delta = 0.7$) (b) Around the epidemic threshold ($\delta = 0.6$) (c) Above the epidemic threshold ($\delta = 0.5$)

Fig. 1. Simple runs of epidemic spread in a power-law topology ($\beta_1 = 0.06$ and $\beta_2 = 0.14$).



(a) Lattice ($\delta = 0.5$) (b) Power-law topology ($\delta = 0.5$) (c) Coauthorship network ($\delta = 0.5$)

Fig. 2. Impact of heterogeneous recovery rates on epidemic thresholds.

axis. In each subfigure, the "mean" curve is the average over 100 simulation runs. The "5%" curve shows that the epidemic propagates no faster than this curve in 5 out of 100 runs. A similar definition is applied to other curves. It can be seen that when the recovery rate is large (*i.e.,* $\delta = 0.7$), the ratio between β and δ is below the epidemic threshold, and the epidemic dies out quickly in all 100 runs. When the ratio between β and δ is around the epidemic threshold (*i.e.,* $\delta = 0.6$), the epidemic still dies out, but with a much slower rate. When the ratio is above the threshold (*i.e.,* $\delta = 0.5$), we can see that the epidemic survives in all 100 simulation runs and the average number of final infections is around 100.

B. Impact of Heterogeneous Recovery Rates

We show in Figure 2 how the epidemic threshold (*i.e.*, τ) changes with the variance of recovery rates (*i.e.*, σ_d^2) in three networks. In each subfigure, we use the circle and the errorbar to denote the average and the standard deviation of the epidemic thresholds among 100 runs. In all three networks, we set the average of recovery rates to 0.5 (*i.e.*, $\delta = 0.5$). It can be seen that the epidemic threshold is smaller when $\sigma_d^2 > 0$ than when $\sigma_d^2 = 0$, which verifies Theorem 1. Moreover, the epidemic threshold decreases when σ_d^2 increases. That is, it is easier for an epidemic to survive when σ_d^2 becomes larger, which confirms Conjecture 1. Furthermore, the small value of the standard deviation of the epidemic threshold indicates that all 100 simulation runs converge to the similar results.

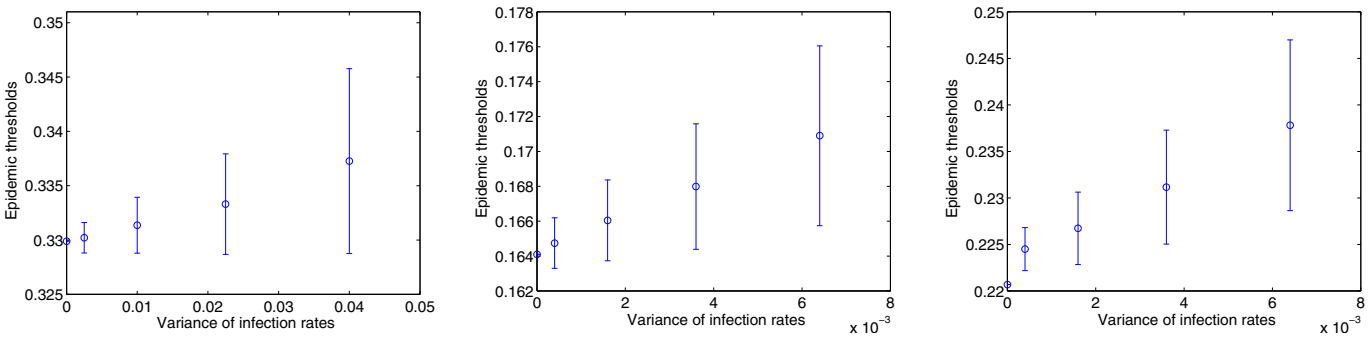
An interesting observation from Figure 2 is that the relationship between the epidemic threshold and the variance of recovery rates seems linear, and the absolute value of the slope of the line is close to $(1/\delta)^2$ times of the epidemic threshold for the case of $\sigma_d^2 = 0$. That is,

$$\tau \approx -\left(\frac{1}{\delta}\right)^2 \tau_0 \sigma_d^2 + \tau_0, \tag{24}$$

where τ_0 is the epidemic threshold for the case of homogenous recovery rates.

C. Impact of Heterogeneous Infection Rates

We plot simulation results in Figure 3 when infection rates are either homogeneous (*i.e.*, $\sigma_b^2 = 0$) or heterogeneous (*i.e.*, $\sigma_b^2 > 0$). According to [22], [3], [9], [14], [4], the epidemic threshold in the homogeneous case is about $1/\lambda_{max}(A)$, where $\lambda_{max}(A)$ is the largest eigenvalue of the adjacency matrix A of the network. To avoid the case that the infection rates are too large so that death rates are all 1, we choose the average of infection rates to be around $1/\lambda_{max}(A)$. Specifically, we use 0.25 as the average of infection rates for the lattice, whereas this value is set to 0.1 for both the power-law topology and the coauthorship network.



(a) Lattice ($\beta = 0.25$) (b) Power-law topology ($\beta = 0.1$) (c) Coauthorship network ($\beta = 0.1$)

Fig. 3. Impact of heterogeneous infection rates on epidemic thresholds.

It can be seen from Figure 3 that in all three networks, the epidemic threshold is larger when $\sigma_b^2 > 0$ than when $\sigma_b^2 = 0$, which verifies Theorem 2. Moreover, when σ_b^2 increases, the epidemic threshold also increases, which confirms Conjecture 2. Furthermore, compared with the case of heterogeneous recovery rates, heterogeneous infection rates lead to a much larger standard deviation of the epidemic threshold, especially when σ_b^2 becomes large. This indicates that when the variance of infection rates is large, 100 simulation runs may show different epidemic thresholds.

V. Conclusions

In this work, we have shown analytically the impact of heterogeneous infection rates and recovery rates on epidemic threshold. Specifically, we have found that the heterogeneous infection rates lead to a larger epidemic threshold, whereas the heterogeneous recovery rates result in a smaller epidemic threshold. Moreover, a higher degree of the heterogeneity in infection rates (or recovery rates) leads to the larger (or smaller) epidemic threshold. Our analytical results are verified by simulations.

As our on-going work, we plan to derive a closed-form expression for the epidemic threshold when infection rates and recovery rates are heterogeneous.

References

[1] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, Oct. 1999, pp. 509-512.

[2] C. Buono, F. Vazquez, P. A. Macri, and L. A. Braunstein, "Slow epidemic extinction in populations with heterogeneous infection rates," *Physics Reviews*, vol. 88, no. 2, 022813, Aug. 2013.

[3] D. Chakrabarti, Y. Wang, C. Wang, J. Leskovec, and C. Faloutsos, "Epidemic thresholds in real networks," *ACM Transactions on Information and System Security (TISSEC)*, vol. 10, no. 4, Jan. 2008.

[4] Z. Chen, "Toward understanding spatial dependence on epidemic thresholds in networks," in *Proc. of 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2016), Workshop on Social Influence (SI 2016)*, San Francisco, CA, Aug. 2016.

[5] Z. Chen and C. Chen, "Characterising heterogeneity in vulnerable hosts on worm propagation," *International Journal of Security and Networks*, vol. 11, no. 4, 2016, pp. 224-234.

[6] Z. Chen and K. Taylor, "Modeling the spread of influence for independent cascade diffusion process in social networks," in *9th International Workshop on Hot Topics in Planet-Scale Mobile Computing and Online Social Networking (HotPOST 2017)*, Atlanta, GA, June, 2017.

[7] D. J. Daley and J. Gani, *Epidemic Modeling: An Introduction*. Cambridge University Press, 2001.

[8] P. Erdös and A. Rényi, "On the evolution of random graphs," *Publ. Math. Inst. Hung. Acad. Sci.*, vol. 5, 1960, pp. 17-61.

[9] A. Ganesh, L. Massoulie, and D. Towsley, "The effect of network topology on the spread of epidemics," in *Proc. of IEEE INFOCOM'05*, vol. 2, Miami, FL, March 2005, pp. 1455-1466.

[10] A. Guille, H. Hacid, C. Favre, and D. A. Zighed, "Information diffusion in online social networks: a survey," *ACM SIGMOD Record*, vol. 42, no. 2, May 2013, pp. 17-28.

[11] D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the spread of influence through a social network," in *Proc. of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, 2003, pp. 137-146.

[12] J. O. Kephart and S. R. White, "Directed-graph epidemiological models of computer viruses," in *Proc. of the 1991 IEEE Computer Society Symposium on Research in Security and Privacy*, May 1991, pp. 343-359.

[13] J. Leskovec, M. McGlohon, C. Faloutsos, N. Glance, and M. Hurst, "Cascading behavior in large blog graphs: Patterns and a model," *SIAM International Conference on Data Mining (SDM)*, 2007.

[14] P. V. Mieghem, J. Omic, and R. Kooij, "Virus spread in networks," *IEEE/ACM Transaction on Networking*, vol. 17, no. 1, Feb. 2009, pp. 1-14.

[15] B. A. Prakash, D. Chakrabarti, M. Faloutsos, N. Valler, and C. Faloutsos, "Threshold conditions for arbitrary cascade models on arbitrary networks," in *Proc. of the 2011 IEEE 11th International Conference on Data Mining*, Vancouver, pp. 537-546.

[16] B. Qu and H. Wang, "SIS epidemic spreading with heterogeneous infection rates," *IEEE Transactions on Network Science and Engineering*, vol. 4, no. 3, 2017, pp. 177-186.

[17] S. M. Ross, *Stochastic Processes*, Second Edition. John Wiley & Sons, Inc., 1996.

[18] C. Shannon and D. Moore, "The spread of the Witty worm," *IEEE Security and Privacy*, vol. 2, no 4, Jul-Aug 2004, pp. 46-50.

[19] D. L. Smith, J. Dushoff, R. W. Snow, and S. I. Hay, "The entomological inoculation rate and Plasmodium falciparum infection in African children," *Nature*, vol. 438, no. 7067, Nov. 2005, pp. 492-495.

[20] J. Wang and W. Wang, "To live or to die: Encountering conflict information dissemination over simple networks," in *Proc. of the IEEE INFOCOM*, San Francisco, CA, Apr. 2016.

[21] W. B. Wang, Z. N. Wu, Z. M. Cao, and R. F. Hu, "Modelling the spreading rate of controlled communicable epidemics through an entropy-based thermodynamic model," *Science China Physics, Mechanics and Astronomy*, vol. 56, no. 11, Nov. 2013, pp 2143-2150.

[22] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos, "Epidemic spreading: An eigenvalue viewpoint," in *Proc. 2003 Symposium of Reliable and Distributed Systems*, Florence, Italy, Oct. 2003.

[23] BRITE, Power-Law Topology Generator [Online]. Available: http://www.cs.bu.edu/brite/ (January/2018 accessed).

[24] Collaboration Network in Science of Networks [Online]. Available: http://vlado.fmf.uni-lj.si/pub/networks/data/collab/netscience.htm (January/2018 accessed).