Delft University of Technology

SIS epidemic spreading with correlated heterogeneous infection rates

Qu, Bo; Wang, Huijuan

SIS epidemic spreading with correlated heterogeneous infection rates

Bo Qu *, Huiijuan Wang

Delft University of Technology, Netherlands

HIGHLIGHTS

- We propose a correlated heterogeneous Susceptible–Infected–Susceptible model.
- The infection rates are correlated with nodal degrees.
- We investigate the influence of the correlation on the epidemic spreading.
- The negative correlation benefits the spreading with the small recovery rate.
- The positive correlation helps the spreading if the recovery rate is large enough.

ARTICLE INFO

ABSTRACT

The epidemic spreading has been widely studied in homogeneous cases, where each node may get infected by an infected neighbor with the same rate. However, the infection rate between a pair of nodes, which may depend on e.g. their interaction frequency, is usually heterogeneous and even correlated with their nodal degrees in the contact network. In this paper, we aim to understand how such correlated heterogeneous infection rates influence the epidemic spreading on different network topologies. Motivated by real-world datasets, we propose a Correlated heterogeneous Susceptible–Infected–Susceptible (CSIS) model which assumes that the infection rate $\beta_{ij}(= \beta_{ji})$ between node i and j is correlated with the degree of the two end nodes: $\beta_{ij} = c(d_i d_j)^{\alpha}$, where α indicates the strength of the correlation between the infection rates and nodal degrees, and c is selected so that the average infection rate is 1 in this work. The recovery rate is the same for all the nodes. In order to understand the effect of such correlation on epidemic spreading, we consider as well the corresponding uncorrected but still heterogeneous infection rate scenario as a reference, where the original correlated infection rates in our CSIS model are shuffled and reallocated to the links of the same network topology. We compare these two scenarios in the average fraction of infected nodes in the metastable state on Erdös–Rényi (ER) and scale-free (SF) networks with a similar average degree. Through the continuous-time simulations, we find that, when the recovery rate is small, the negative correlation is more likely to help the epidemic spread and the positive correlation prohibits the spreading; as the recovery rate increases to be larger than a critical value, the positive but not negative correlation tends to help the spreading. Our findings are further analytically proved in a wheel network (one central node connects with each of the nodes in a ring) and validated on real-world networks with correlated heterogeneous interaction frequencies.

© 2016 Elsevier B.V. All rights reserved.

* Corresponding author.
 E-mail addresses: b.qu@tudelft.nl (B. Qu), h.wang@tudelft.nl (H. Wang).

1. Introduction

Biological, social and communication systems can be represented as networks by considering the system components or individuals as nodes and the interactions or relations in between nodes as links. Viral spreading models have been used to model processes, e.g. the propagation of information and epidemics on such networks or complex systems [1–4]. The Susceptible–Infected–Susceptible (SIS) model is one of the most studied models. In the SIS model, the state of a node is a Bernoulli random variable at any time t, where $X_i(t) = 0$ represents that node i is susceptible and $X_i(t) = 1$ if it is infected. Each infected node infects each of its susceptible neighbors with an infection rate β. The infected node can be recovered to be susceptible again with a recovery rate δ. Both infection and recovery processes are independent Poisson processes. The ratio $\tau \triangleq \beta/\delta$ is called effective infection rate. When τ is larger than the epidemic threshold τ_c, the epidemic spreads out with a nonzero fraction of infected nodes in the metastable state. The average fraction of infected nodes y_∞ in the metastable state, ranging in [0, 1], indicates how severe the influence of the virus is: the larger the fraction y_∞ is, the more severely the network is infected.

In the classic SIS model, both the infection and recovery rates are assumed homogeneous, i.e. the infection rates are the same for all infected–susceptible node pairs and the recovery rates are the same for all infected nodes. However, the infection rates which can be reflected from, for example, the interaction frequencies, between nodes in real-world networks are usually heterogeneous and even dependent on the properties of the nodes. For examples, the number of flights between different pairs of airports in a month are different in the airline transportation network, and the number of collaborated papers between different pairs of authors in a year vary in a co-author network [5,6]. The interaction frequency is found to be correlated with the nodal degrees in e.g. airline transportation network and metabolic network [5,7,6]. Hence, we aim to understand the effect of correlated infection rate on the viral spreading in this work. We propose a correlated heterogeneous SIS (CSIS) model, in which the recovery rates are homogeneous but the infection rate $\beta_{ij}(=\beta_{ji})$ between node i and j is correlated with their degrees d_i and d_j in the way:

$$\beta_{ij} = c(d_i d_j)^\alpha \tag{1}$$

where α indicates the strength of the correlation and c is a constant to control the average infection rate to 1. This correlation between the interaction frequency, i.e. the infection rate, and the nodal degree is actually observed in various real-world systems, such as in the network of airports with the correlation strength $\alpha \approx 0.5$ (both in US [5,7] and China [6]) and in the metabolic network with the correlation strength $\alpha \approx 0.8$ [7]. We also find the relationship (1) in two real-world networks, i.e. the airline network and the co-author network, as shown in Section 5.

This paper aims to understand the effect of correlation between the infection rates and nodal degrees on viral spreading. Motivated by real-world networks, we consider the generic case where the heterogeneous infection rates are correlated with nodal degrees as described by (1) and the network topology is as well heterogeneous. For the heterogeneity of the network topology, we consider the different broadness of the degree distribution, i.e. Erdös–Rényi (ER) and scale-free (SF) networks. We consider also the corresponding uncorrelated heterogeneous infection rates scenario, where the correlated infection rates are shuffled and randomly assigned to all the links as a reference scenario. Note that the infection rates in the reference scenario are actually identically and independently distributed (i.i.d.). The study of the i.i.d. infection rates is an initial step to explore the influence of heterogeneous infection rates on epidemic spreading [8]. The real-world could be possibly correlated as mentioned, hence the study of the correlated infected rates is a further step to the reality, though more complex. In this work, we explore how the positive or negative correlation between the infection rates and nodal degrees influences the epidemic spread on the networks with different heterogeneities of the topology.

A few recent papers have taken into account the heterogeneous infection rates in the SIS model which is also addressed as the SIS model on weighted networks, where the weight represents the infection rate. The i.i.d. infection rates that follow, e.g. a Poisson or uniform distribution, which has a small variance tend to have negligible effect on epidemic threshold and prevalence [9]. We studied the SIS model with i.i.d. infection rates which may follow an arbitrary distribution that may have as well large variance and found that independent of the underlying network topology, though the heterogeneous infection rates may increase the probability that the epidemic spreads out when the recovery rate is large, the overall infection decreases as the variance of infection rates increases if the epidemic spreads out [8]. Buono et al. [10] considered a specific distribution of infection rates and observed slow epidemic extinction phenomenon. Preciado et al. [11] discussed how to choose the infection and recovery rates from given discrete sets to let the virus die out. Yang and Zhou [12] gave an edge-based mean-field solution of the epidemic threshold in regular networks (the degrees of all nodes are the same) with i.i.d. heterogeneous infection rates following a uniform or power-law distribution. Fu et al. [13] analytically studied the epidemic threshold when the infection rates are as a piecewise linear function of degrees by their mean-field method. Van Mieghem and Omic [14] developed the N-intertwined mean field approximation (NIMFA) [15] for heterogeneous SIS model, and discussed the bounds of epidemic threshold and the convexity of the steady-state infection probability of each node as function of the recovery rate. Heterogeneous infection rates have also been considered in interconnected networks, allowing component networks and their interconnections to have a different but constant infection rate [16, 17]. A Susceptible–Infected type model has been studied in a network where the weight or infection rate depends on the degree [18]. In this paper, we employ the continuous-time simulation of the exact SIS model instead of the mean-field methods and introduce a general correlation between the nodal degree and the infection rate which is motivated by real-world datasets to study the influence of the correlation on the epidemic spreading.
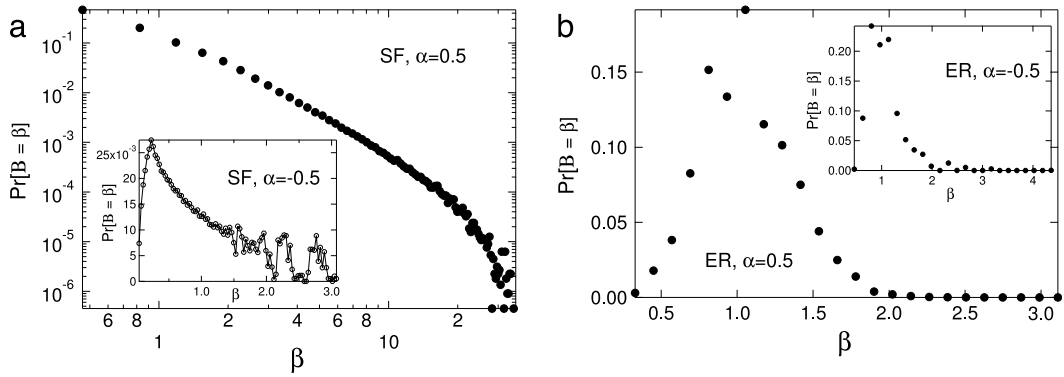
Fig. 1. The distribution of the heterogeneous infection rates for (a) SF networks and (b) ER networks, where the parameter $\alpha = 0.5$ in the main figures and $\alpha = -0.5$ in the insets.

2. Preliminary

In this section, we first introduce the construction of network models and the heterogeneous infection rates which are considered in our CSIS model. We then introduce the continuous-time simulation which is the main approach in this work. Finally, we introduce the mean-field approximation of the SIS and CSIS model which will be further used in our theoretical analysis of the CSIS model on the wheel network in Section 4.

2.1. Network models

The scale-free (SF) model has been used to capture the scale-free nature of degree distribution in real-world networks such as the Internet [19] and World Wide Web [20]: $\Pr[D = d] \sim d^{-\lambda}$, $d \in [d_{\min}, d_{\max}]$, where $d_{\min}$ is the smallest degree, $d_{\max}$ is the degree cutoff, and $\lambda > 0$ is the exponent characterizing the broadness of the distribution [21]. In real-word networks, the exponent λ is usually in the range [2, 3], thus we confine the exponent $\lambda = 2.5$ in this paper. We further employ the smallest degree $d_{\min} = 2$, the natural degree cutoff $d_{\max} = \lfloor N^{1/(\lambda-1)} \rfloor$ [22], and the size $N = 10^4$. Hence, the average degree is approximately 4.

The Erdös–Rényi (ER) model [23] has also been taken into account. In an ER random network with N nodes, each pair of nodes is connected with a probability p independent of the connection of any other pair. The distribution of the degree of a random node is binomial: $\Pr[D = d] = \binom{N-1}{d} p^d (1-p)^{N-1-d}$ and the average degree is $E[D] = (N-1)p$. For large N and constant $E[D]$, the degree distribution is Poissonian: $\Pr[D = d] = e^{-E[D]} E[D]^d / d!$. We consider ER networks with the size $N = 10^4$ and the average degree $E[D] = 4$.

2.2. The infection rates

Given the network topology, we build two heterogeneous infection-rate scenarios: (1) the correlated infection rates and (2) the uncorrelated or shuffled infection rates. In the scenario of the correlated infection rates, we assume that $\beta_{ij} = c \left(d_i d_j \right)^{\alpha}$ where α indicates the correlation strength. We select the constant c such that the average infection rate is 1, whereas we consider different values of homogeneous recovery rates. In this case, the infection rate of each link is determined by the given network topology and α. In the scenario of the uncorrelated infection rates, we shuffle the infection rates from all the links as generated in the first scenario and redistribute them randomly to all the links. In this way, we keep the distribution of infection rates but effectively remove the correlation between the infection rates and nodal degrees. The homogeneous infection rate is a special case of our heterogeneous infection rate construction where $\alpha = 0$. Clearly, in a homogeneous network where all the nodes have the same number of neighbors, the infection rates are homogeneous in both scenarios for any α.

As examples, we show the distribution of the heterogeneous infection rates when $\alpha = 0.5$ (the positive correlation) and $\alpha = -0.5$ (the negative correlation) for both ER and SF networks in Fig. 1. We find that the positive correlation between the infection rates and the nodal degrees leads to a heavy-tailed distribution of the infection rate in SF networks. However, the negative correlation in SF networks and both kinds of correlation in ER networks only generate the infection rates distributed in a small range.

A positive $\alpha > 0$ (or negative $\alpha < 0$), suggests a positive (or negative) correlation between the infection rates and nodal degrees. Too large or too small values of α could not be realistic. For example, [5,7,6] suggest that α is around 0.5 or 0.8 in their datasets, and we also find $\alpha \approx 0.14$ and -0.12 respectively in two real-world dataset as described in Section 5. Hence, we focus mainly on the range of $\alpha \in [-1, 1]$ in this paper, and discuss the extreme case when the absolute value of α is large in Section 3.2.

2.3. The simulations

In this paper, we perform the continuous-time simulations of the CSIS model on both ER networks and SF networks (the heterogeneity of the network topology is thus taken into account) with $N = 10\,000$ nodes. Given a network topology, a recovery rate δ and a value of α, we carry out 100 iterations. In each iteration, we construct the network as described in Section 2.1. We generate the heterogeneous infection rates as described in (1) for the scenario of the correlated infection rates and shuffle them for the scenario of the uncorrelated infection rates. Initially, 10% of the nodes are randomly infected. Then the infection and recovery processes of SIS model are simulated until the system reaches the metastable state where the fraction of infected nodes is unchanged for a long time. The average fraction y_∞ of infected nodes is obtained over 100 iterations for both scenarios of correlated and uncorrelated infection rates. Moreover, for simplicity, we use $y_{\infty,c}$ and $y_{\infty,u}$ to denote the average fraction of infected nodes in the scenarios of correlated and uncorrelated infection rates respectively.

A discrete-time simulation could well approximate a continuous process if a small time bin to sample the continuous process is selected so that within each time bin, no multiple events occur. However, a heterogeneous SIS model allows different as well large infection or recovery rates, which requires even smaller time bin size and challenges the precision of a discrete-time simulation. Hence, instead of performing discrete-time simulations, we further develop the continuous-time simulator for our CSIS model, based on the one firstly proposed by van de Bovenkamp and described in detail in [24] for homogeneous infection rates.

2.4. The N-Intertwined Mean-Field Approximation of the SIS model

The N-Intertwined Mean-Field Approximation (NIMFA) is an advanced mean-field approximation of the SIS model that takes the network topology into account. The governing equation for a node i in NIMFA for the classic SIS model with the homogeneous infection and recovery rates is

$$\frac{\mathrm{d}v_i(t)}{\mathrm{d}t} = -\delta v_i(t) + \beta(1 - v_i(t)) \sum_{j=1}^{N} a_{ij} v_j(t) \tag{2}$$

where $v_i(t)$ is the infection probability of node i at time t, and $a_{ij} = 1$ or 0 denotes if there is a link or not between node i and node j. In the steady state, defined by $\frac{\mathrm{d}v_i(t)}{\mathrm{d}t} = 0$ and $\lim_{t\to\infty} v_i(t) = v_{i\infty}$, we obtain the infection probability of node i:

$$v_{i\infty} = 1 - \frac{1}{1 + \tau \sum\limits_{j=1}^{N} a_{ij} v_{j\infty}}. \tag{3}$$

The trivial solution of (3) is $v_{i\infty} = 0$ for all nodes and indicates the absorbing state which will be reached after an unrealistically long time for a real network that is not small in size. Before reaching the absorbing state, there could exist the metastable state, indicated by the nonzero solution of (3).

The Laurent series of the steady-state infection probability is

$$v_{i\infty} = 1 + \sum_{m=1}^{\infty} \eta_m(i) \tau^{-m} \tag{4}$$

possesses the coefficients $\eta_1(i) = -\frac{1}{d_i}$ and

$$\eta_2(i) = \frac{1}{d_i^2} \left(1 - \sum_{j=1}^{N} \frac{a_{ij}}{d_j} \right)$$

and for $m \geq 2$, the coefficients obey the recursion

$$\eta_{m+1}(i) = -\frac{1}{d_i} \eta_m(i) \left(1 - \sum_{j=1}^{N} \frac{a_{ij}}{d_j} \right) - \frac{1}{d_i} \sum_{k=2}^{m} \eta_{m+1-k}(i) \sum_{j=1}^{N} a_{ij} \eta_k(j).$$

When the infection rates are heterogeneous, the governing equation becomes

$$\frac{\mathrm{d}v_i(t)}{\mathrm{d}t} = -\delta v_i(t) + \beta_{ij}(1 - v_i(t)) \sum_{j=1}^{N} a_{ij} v_j(t). \tag{5}$$

The infection probability of a node follows

$$v_{i\infty} = 1 - \frac{\delta}{\delta + \sum\limits_{j=1}^{N} \beta_{ij} a_{ij} v_{j\infty}} \tag{6}$$

3. Effect on the average fraction y_∞ of infected nodes

As mentioned, the average fraction y_∞ of infected nodes in the metastable state indicates how severe the network is infected. In this section, we explore how the average fraction y_∞ of the infected nodes depends on the parameter α in both scenarios: correlated and uncorrelated infection rates. We mainly consider the influence of the correlation between the infection rates and nodal degrees on epidemic spreading when the recovery rate varies and the absolute value of α is in the range $[-1, 1]$. The difference of the influence between ER and SF networks is also discussed. Then we briefly describe the influence of the correlation in an extreme case when the absolute value of α is much larger. In this case, the influence of the correlation is independent from the value of the recovery rate.

3.1. Realistic cases: $\alpha \in [-1, 1]$

In this subsection, we first intuitively explain that how the correlation influences the epidemic spreading as the recovery rate δ and the parameter α varies. To support our explanations, we then define an intermediate quantity and illustrate the effect of the correlation when the recovery rate is small and large respectively.

The average fraction y_∞ of infected nodes as a function of the scale parameter α in both ER and SF networks are shown in Fig. 2. We employ different values of the recovery rate δ to illustrate the influence of the correlation on the epidemic spreading over different range of recovery rates, i.e. different prevalence of the epidemic.

Our previous work [8] explored the SIS model with i.i.d. heterogeneous infection rates and showed that the average fraction y_∞ of infected nodes tends to decrease as the variance of the i.i.d. infection rates increases. In this paper, the uncorrelated infection rates can be considered as being independent. In this case, our previous results can be applied to derive the relationship between the average fraction y_∞ of infected nodes and the parameter α as shown in Fig. 2: as α increases when $\alpha > 0$ or as α decreases when $\alpha < 0$, the variance of the infection rates increases, hence y_∞ decreases in the scenario of the uncorrelated infection rates. In the scenario of the correlated infection rates, though the peaks of y_∞ are not at $\alpha = 0$ for both types of networks, y_∞ roughly decreases as the absolute value of α increases. This is because, as the absolute value of α increases, large infection rates are assigned to a small number of links, limiting the spread of the epidemic.

From Fig. 2, we find that (1) the negative correlation ($\alpha < 0$) between the infection rates and the degrees tends to enhance the epidemic spreading compared to the uncorrelated infection-rate scenario ($y_{\infty,c} > y_{\infty,u}$) when the recovery is small, but prohibit the spreading ($y_{\infty,c} < y_{\infty,u}$) when the recovery rate is large; (2) the positive correlation ($\alpha > 0$) tends to enhance the epidemic spreading ($y_{\infty,c} > y_{\infty,u}$) when the recovery is large, but prohibit the spreading ($y_{\infty,c} < y_{\infty,u}$) when the recovery rate is small.

In the scenario of the uncorrelated infection rates, the parameter α determines only the distribution of the i.i.d. infection rates in a network, and the infection rate between any pair of nodes is independent from their degrees. Compared to the scenario of the uncorrelated infection rates, a positive correlation between the infection rates and nodal degrees ensures that the infection rates between nodes with larger degrees are also larger, whereas a negative correlation suggests the other way around. Intuitively, we may think that if the nodes with larger degrees can be infected by larger infection rates, the infection probability of those nodes are higher and those nodes can more effectively infect their neighbors as well. Hence, the positive correlation between the infection rates and nodal degrees seems to contribute to the epidemic spreading. This is indeed the case when the recovery rate δ is large, i.e. the prevalence of the epidemic is low.

In contrast to this intuition, we observe that the positive correlation actually tends to prohibit the epidemic spreading when the recovery rate δ is small. This can be explained as follows: when the recovery rate δ is small, i.e. the prevalence is high, the infection probabilities of the large-degree nodes are already high, then the increment of the infection rates between the large-degree nodes may not significantly increase the infection probabilities of these nodes, and thus the infection probabilities of their neighbors may not be significantly increased. However, the negative correlation between the infection rates and nodal degrees leads to the higher infection rates between the small-degree nodes and effectively enhances the probabilities of the small-degree nodes compared to the scenario of the uncorrelated infection rates. Though the infection probabilities of large-degree nodes decrease in this case, the large amount of small-degree nodes ensures that the overall infection is on average enhanced.

To support our explanations, we define $y_\infty^{(d)}$ as the average infection probability of the nodes with degree d, and $y_{\infty,c}^{(d)}$ and $y_{\infty,u}^{(d)}$ as that for the scenario of the correlated infection rates and uncorrelated infection rates respectively. We show $y_\infty^{(d)}$ as a function of the degree d when the recovery rate is small ($\delta = 0.5$), and the correlation parameter $\alpha = -0.6$ in Fig. 3 (the main figures). We find that such a negative correlation ($\alpha = -0.6$) between the infection rates and nodal degrees indeed decreases the infection probabilities of large-degree nodes, but the infection probabilities of the small-degree nodes are also significantly lifted. Furthermore, the number of small-degree nodes is much larger than that of large-degree nodes in SF networks, and those are similar in ER networks. To illustrate the combined effect above of the two aspects, we define $\eta(d)$ as (7), the product of the probability that a node has the degree d and the difference between the average infection probability of the nodes with the degree d in the scenarios of correlated and uncorrelated infection rates:

$$\eta(d) = (y_{\infty,c}^{(d)} - y_{\infty,u}^{(d)}) \Pr[D = d] \tag{7}$$
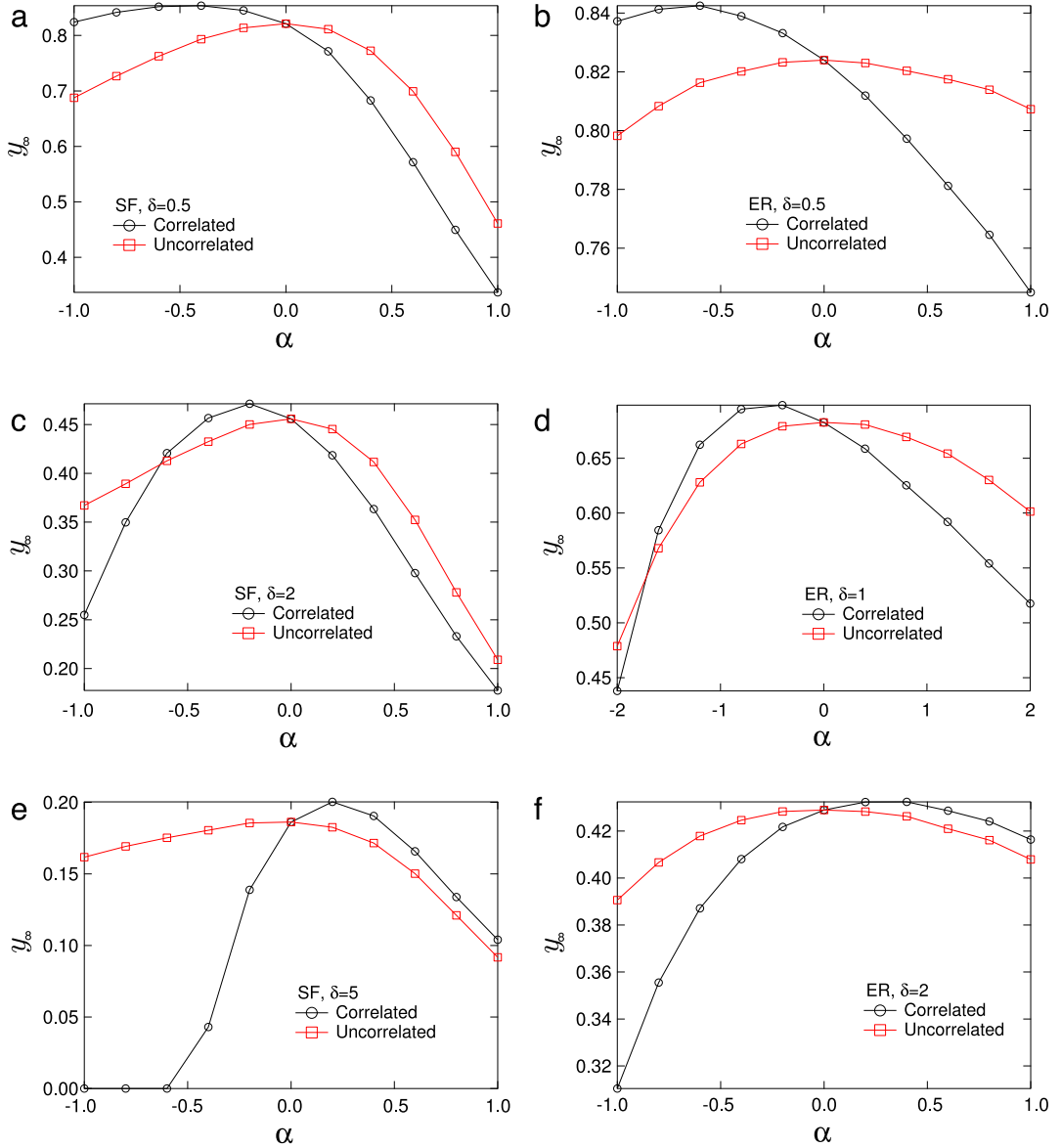
Fig. 2. The average fraction y_∞ as a function of α for (a) SF networks with the recovery rate $\delta = 0.5$, (b) ER networks with the recovery rate $\delta = 0.5$, (c) SF networks with the recovery rate $\delta = 2$, (d) ER networks with the recovery rate $\delta = 1$, (e) SF networks with the recovery rate $\delta = 5$ and (f) ER networks with the recovery rate $\delta = 2$ in both scenarios of correlated (○) and uncorrelated (□) infection rates.

and $y_{\infty,c} - y_{\infty,u} = \sum_{d=1}^{d=N-1} \eta(d)$. Note that a positive $\eta(d)$ always indicates that $y_{\infty,c} > y_{\infty,u}$, i.e. the correlation lifts the infection probability of nodes with the degree d compared to the scenario of the uncorrelated infection rates. As in the insets of Fig. 3, we plot $\eta(d)$ as a function of the degree d for both ER and SF networks. More plots[1] of η as a function of the degree d are shown in Fig. 4 where the cases with small degrees are shown in the main figures and those with relatively large degrees are shown in the insets. We find that in both ER and SF networks the value of η is significantly large for the small-degree nodes and contributes more to a higher prevalence of the epidemic when the recovery rate is small and the correlation is negative as shown in Figs. 3 and 4(c) and (d). The observation is consistent with our explanation about why the negative correlation tends to help the epidemic spreading when the recovery rate is small. In contrast, the observation, shown in Fig. 4(a) and (b), that the positive correlation does increase the infection probabilities of large-degree nodes but decreases

---

[1] We still select $\alpha = -0.6$ as the example of the negative correlation and $\alpha = 0.6$ as the comparison for ER network, but we select $\alpha = -0.2$ as the example of the negative correlation (and $\alpha = 0.2$ as the comparison) for SF networks since when $\delta = 5$ and $\alpha = -0.6$ the epidemic already dies out in the scenario of the correlated infection rates.
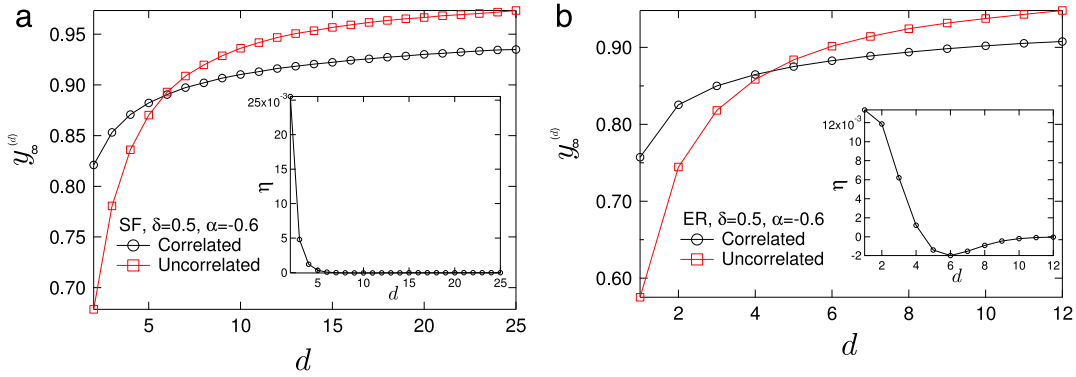
Fig. 3. In the main figures: the average infection probability $y_\infty^{(d)}$ of the nodes with degree d as a function of the degree d for (a) SF networks and (b) ER networks with the recovery rate $\delta = 2$ and $\alpha = -0.6$ for both scenarios of the correlated and uncorrelated infection rates. In the insets: η as a function of d with the same setting as the main figure.

those of small-degree nodes more when the recovery rate is small, also supports our explanation about how the positive correlation prohibits the spreading when recovery rate is small.

As the recovery rate becomes large thus the prevalence is low, the positive but not negative correlation, between the infection rates and nodal degrees may effectively enhance the infection probabilities of the large degree nodes comparing to the uncorrelated infection rates. When the recovery rate δ is large, for example, $\delta = 5$ for SF networks and $\delta = 2$ for ER networks in this paper, the positive correlation leads to the increment of the infection probabilities of large-degree nodes and could further lift the probabilities of their small-degree neighbors which are large in number. The infection probabilities of small-degree nodes are reduced, but the infection probabilities are already low and the small-degree nodes have few neighbors to infect. Hence, the overall infection increases on average. The explanations are supported by Fig. 4(a) and (b). The infection probabilities of the large-degree nodes increases for different recovery rates when α is positive, and the increment of the infection probabilities of large-degree nodes is also larger as the recovery rate increases. Though the infection probabilities of the small-degree nodes may decrease when the correlation is negative, the increment of the infection probabilities of the large-degree nodes also lifts the infection probabilities of the small-degree nodes. As a result, the infection probabilities of majority nodes are on average lifted when the recovery rate $\delta = 5$ and $\delta = 2$ for SF and ER networks respectively.

For both ER and SF networks, the positive correlation between the nodal degrees and the infection rates tends to enhance the spreading when the recovery rate is small, whereas the negative correlation tends to help when the recovery rate is large. As the recovery rate δ increases from 0 and if the absolute value of α is small, we expect that there is a critical value δ_c: when $\delta < \delta_c$ the negative correlation tends to enhance the spreading, otherwise ($\delta > \delta_c$) the positive correlation is likely to help the spreading. By comparing Fig. 2(c) and (f), we can observe that δ_c is larger in SF networks than ER networks. This difference is mainly caused by that the prevalence in SF networks tends to be higher than that in ER networks when the recovery rate and the parameter α are the same and the positive correlation tends to enhance the epidemic spreading when the prevalence is low as we discussed.

3.2. Extreme cases

We then discuss the influence of the correlation between the infection rates and the nodal degrees when the correlation is strong, i.e. the absolute value of α is large.

When the absolute value of α is large, the variance of the infection rates is large as well. In this case, most links possess small infection rates and few have large infection rates. A large proportion of the links have such a small infection rate that the infection processes driven by the small infection rate will hardly happen. The networks are actually filtered by the small infection rates. The other small proportion of the large infection rates then mainly determine the overall infection. In the scenario of the uncorrelated infection rates, the few large infection rates are randomly distributed, thus cannot form a connected cluster. Compared to the scenario of the uncorrelated infection rates, the positive correlation ensures that the large infection rates are distributed between the large-degree nodes which are more likely to connect with each other, forming a subgraph. A connected subgraph tends to help the spread of an epidemic. In the other way around, the negative correlation will almost surely stop the epidemic spreading since the few large infection rates distributed between the small-degree nodes could hardly form such a connected cluster. In summary, the positive correlation enhances the epidemic spreading whereas the negative one prohibits when the absolute value of α is large.
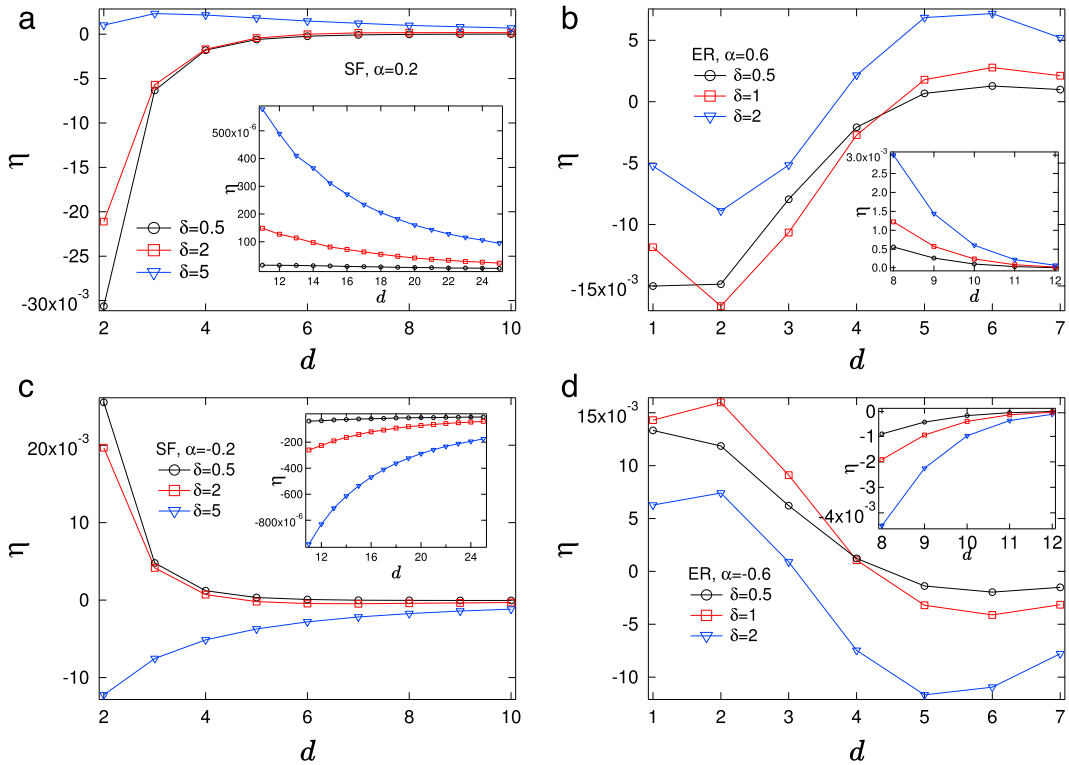
Fig. 4. The plot of η as a function of α for (a) SF networks and (b) ER networks with different recovery rates.

4. The wheel network

In this section, we further consider a special topology—the wheel network. We are going to prove that, compared to the uncorrelated heterogeneous infection rate, the negative correlation between the infection rates and the degrees tends to help the epidemic spreading in a large wheel network when the recovery rate is small, whereas the positive correlation tends to contribute to the epidemic spreading when the recovery rate is large.

In a wheel network, m side nodes compose a ring, i.e. node i connects with node $i + 1$ ($i = 1, 2, 3, \ldots, m - 1$) and node m connects with node 1, and all the m side nodes connect with one central node—node 0. In this section, we consider a large enough wheel network, and without loss of the generality, we still set both the homogeneous infection rate and the average of heterogeneous infection rates to be 1, and tune the recovery rate.

In the scenario of the uncorrelated infection rates, the infection rates are actually i.i.d. In our previous work, we found that, compared to the homogeneous infection rate, the i.i.d. heterogeneous infection rates always reduce the overall infection if the epidemic can spread out. We further verify this conclusion for the wheel network by simulations as shown in the Appendix, where we plot the average fraction of infection nodes as a function of α for the scenario of the uncorrelated infection rates and different values of the recovery rates. That is to say, the average fraction of infected nodes reaches the maximum when the infection rates are homogeneous, i.e. $\alpha = 0$ in the scenario of the uncorrelated infection rates. If the average fraction of infected nodes in the scenario of the correlated infection rates is larger than that when the infection rates are homogeneous, then the correlation enhances the epidemic spreading compared to the uncorrelated case.

We first consider the homogeneous infection rate in a wheel network, where the infection rates are homogeneous thus the same for all links. The infection probability $v_{0\infty}$ of the central node is

$$v_{0\infty} = 1 - \frac{1}{1 + m\tau v_{i\infty}} = 1 - \frac{\delta}{\delta + mv_{i\infty}} \tag{8}$$

where $v_{i\infty}$ is the infection probability of a side node, which is the same for all the side nodes. The infection probability of the side node is

$$v_{i\infty} = 1 - \frac{1}{1 + 2\tau v_{i\infty} + \tau v_{0\infty}} = 1 - \frac{\delta}{\delta + 2v_{i\infty} + v_{0\infty}}. \tag{9}$$

By solving (8) and (9), we obtain the positive solution

$$v_{0\infty} = \frac{m - \delta^2 + 2\delta + \frac{\delta\left((\delta-1)m+2\delta+\sqrt{(\delta^2-2\delta+9)m^2-(4\delta^2+12\delta)m+4\delta^2}\right)}{2m}}{\delta + m} \tag{10}$$

and

$$v_{i\infty} = \frac{(1-\delta)m - 2\delta + \sqrt{(\delta^2 - 2\delta + 9)m^2 - (4\delta^2 + 12\delta)m + 4\delta^2}}{4m}. \tag{11}$$

We consider a large m and a constant δ. In this case, $v_{0\infty} \approx 1$ and

$$v_{i\infty} \approx \frac{1 - \delta + \sqrt{\delta^2 - 2\delta + 9}}{4}. \tag{12}$$

The fraction of infected nodes is then

$$y_\infty = \frac{mv_{i\infty} + v_{0\infty}}{m + 1} \approx v_{i\infty}.$$

Now we consider the wheel network where the infection rates are correlated as defined before in (1). There are two kinds of infection rates in a wheel network: (1) the infection rate β_0 between the central node and a side node, i.e. $\beta_0 \sim (3m)^\alpha$; (2) the infection rate β_1 between a pair of connected side nodes, i.e. $\beta_1 \sim (3 * 3)^\alpha$. Since we consider a large m, we find $\beta_0 \gg \beta_1$ if $\alpha > 0$, whereas $\beta_0 \ll \beta_1$ if $\alpha < 0$. When the correlation between the infection rates and the degree is positive, the infection rates $\beta_0 \approx 2$ and $\beta_1 \approx 0$ and the network becomes a star network. In contrast, when the correlation is negative, the infection rates $\beta_1 \approx 2$ and $\beta_0 \approx 0$ and the network becomes a ring. By NIMFA, we can compute the average fraction of infected nodes[2] when $\alpha > 0$

$$y_\infty \approx v_{i\infty} = \frac{2}{2 + \delta}. \tag{13}$$

When $\alpha < 0$, the average fraction of infected nodes[3]

$$y_\infty = v_{i\infty} = 1 - \frac{\delta}{4}. \tag{14}$$

The parameter $\alpha = 0$ in the scenario of the correlated infection rates indicates the homogeneous infection rate which is the same as $\alpha = 0$ in the scenario of the uncorrelated infection rates. We have shown that the average fraction of infected nodes reaches the maximum when $\alpha = 0$ in the scenario of the uncorrelated infection rates, so we further compare the average fraction of infected nodes as shown in (12) when the infection rates are homogeneous i.e. $\alpha = 0$ with that as shown in (14) when $\alpha < 0$ or that as shown in (13) when $\alpha > 0$ to explore the influence of the correlation on the epidemic spreading. By comparing (12) and (14), we find that when the recovery rate $\delta < 2$, the average fraction y_∞ of infected nodes is higher if $\alpha < 0$ than if $\alpha = 0$. Hence, the negative correlation between the infection rates and nodal degrees helps the epidemic spreading if the recovery rate is small, i.e. $\delta < 2$. Similarly, if we compare (12) and (13), we find that when the recovery rate $\delta > 2$ the average fraction of infected nodes is higher if $\alpha > 0$ than if $\alpha = 0$, and conclude that the positive correlation between the infection rates and nodal degrees contributes to the epidemic spreading if the recovery rate is large. The theoretical results are consistent with our previous conclusions: when the recovery rate is small, the negative correlation tends to help the epidemic spreading, and as the recovery rate increases to be larger than a critical value, i.e. 2 in this case, the positive correlation enhances the spreading.

5. Real-world networks

As mentioned, the interaction frequency β_{ij} between node i and node j in a real-world network can be considered as the infection rate between them and it has been found that $\beta_{ij} \sim (d_i d_j)^\alpha$ in many networks. In this section, we choose two real-world networks as examples to illustrate how their heterogeneous infection rates affect the spread of SIS epidemics on these networks. We compare the average fraction y_∞ of infected nodes in the metastable state of the two networks in the two scenario: (1) the scenario of the correlated infection rates, where each network is equipped with its original heterogeneous infection rates (but normalized so that the average infection rate is 1) as given in the dataset; (2) the scenario of the uncorrelated infection rates, where each network is equipped with the infection rates (normalized as well) in the original dataset but the infection rates are shuffled and reassigned to each link. Our objective is to explore the relation

---

[2] Eq. (13) can be similarly derived by solving (3).

[3] Eq. (14) can be derived by applying the Laurent series as in (4).
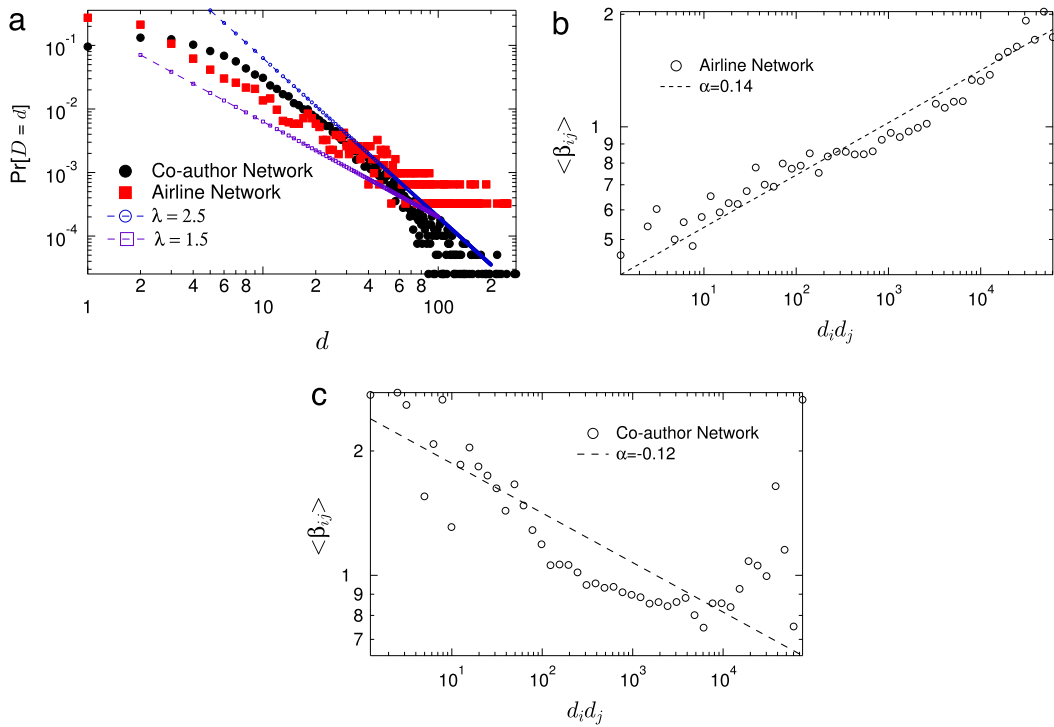
Fig. 5. (a) The degree distributions of two real-world networks: the airline and co-author networks. (b) The interaction frequency β_{ij} as a function of the nodal degrees $d_i d_j$ for the airline network. (c) The interaction frequency β_{ij} as a function of the nodal degrees $d_i d_j$ for the co-author network.

between the infection rates and average fraction of infected nodes in these 2 scenarios for both networks to verify our previous findings.

The first network is the airline network (with 3071 nodes and 15 358 links) where the nodes are the airports and the infection rate along a link is the number of flights between the two airports. We construct this network and its infection rates from the dataset of openFlights.[4] The other one is the co-author network (with 39 577 nodes and 175 692 links) where the nodes are the authors of papers, and the infection rate is the collaboration frequency depending on the number of collaborated papers and the number of authors in those papers [25].

As shown in Fig. 5(a), the degree distributions of the airline network and co-author network approximately follow a power law with the slope $\lambda = 1.5$ and 2.5 respectively. Moreover, we plot the average infection rates $\langle \beta_{ij} \rangle$ as a function of the product of the two nodal degrees $d_i d_j$ in Fig. 5(b) and (c) for the airline and co-author networks respectively. We find that, roughly $\beta_{ij} \sim (d_i d_j)^\alpha$ with $\alpha = 0.14$ and $\alpha = -0.12$ for the airline and co-author networks respectively.

We first discuss the case when the recovery rate is small, i.e. $\delta \in [0.5, 8]$ (as shown in Fig. 6), and then the case when δ is large, i.e. $\delta \geq 15$ (as shown in Fig. 7), since for both networks the recovery rate $\delta \in [0.5, 8]$ enables the high prevalence of the epidemic and $\delta \geq 15$ leads to the low prevalence. The average fraction y_∞ of infected nodes as a function of the recovery rate δ for both infection-rate scenarios is shown in Fig. 6 when the recovery rate δ is small. We find that the positive correlation ($\alpha = 0.14$) between the infection rates and nodal degrees in the airline network retards the spread of epidemics, whereas the negative correlation ($\alpha = -0.12$) in the co-author network contributes in the other way around. The observations are consistent with our previous conclusions that the positive correlation between the infection rates and nodal degrees probably retards the spread of epidemics whereas a negative correlation may help the spread when the recovery rate is small. The average fraction y_∞ of infected nodes as a function of the recovery rate δ when the recovery rate is large is shown in Fig. 7. We find that the positive correlation $\alpha = 0.14$ in the airline network (Fig. 7(a)) helps the epidemic spreading when the recovery rate is larger than 15, and the negative correlation $\alpha = -0.12$ in the co-author network suppresses the epidemic spreading only when the recovery rate is large, i.e. $\delta \leq 35$, that the overall infection is close to 0. The observations also agree with our conclusion that the positive correlation between the infection rates and nodal degrees tends to help the epidemic spreading but not the negative correlation when the recovery rate is large. Hence, the simulation results of the CSIS model on the real-world networks for both the small and large recovery rates agree with our previous conclusions.
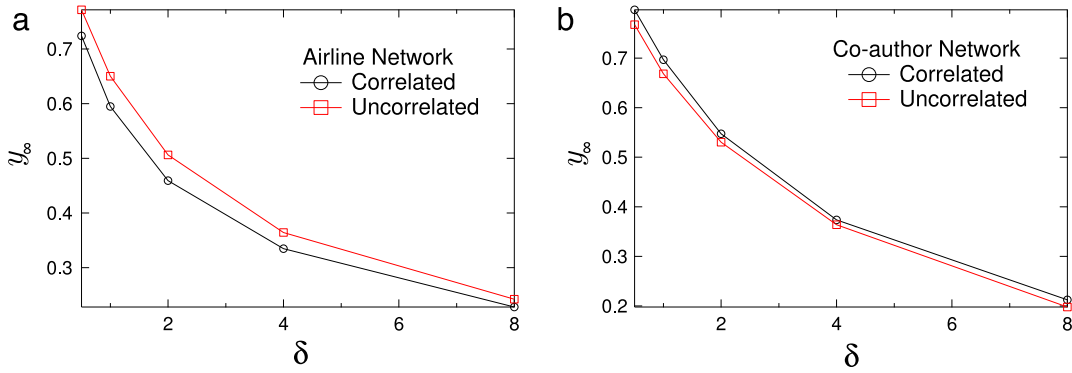
---

4 http://openflights.org/data.html.

Fig. 6. The average fraction y_{∞} as a function of the recovery rate δ for (a) the airline network and (b) the co-author network in both scenarios of correlated (○) and uncorrelated (□) infection rates. The recovery rate is in the range [0.5, 8].
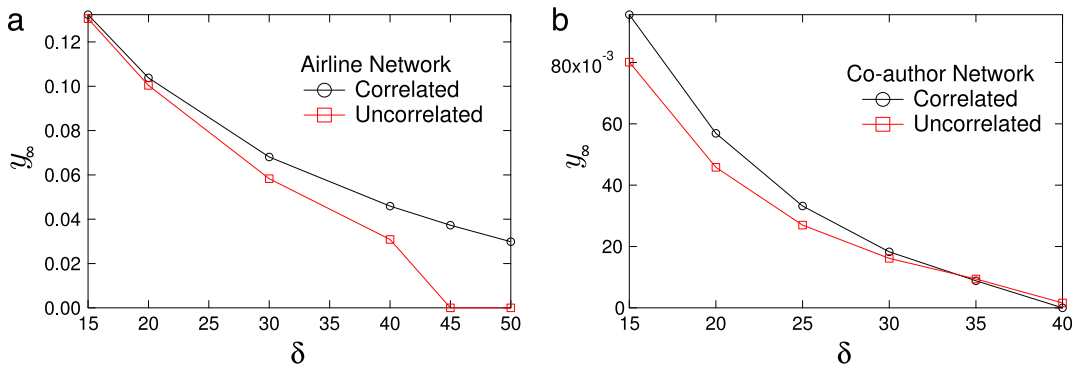

Fig. 7. The average fraction y_{∞} as a function of the recovery rate δ for (a) the airline network and (b) the co-author network in both scenarios of correlated (○) and uncorrelated (□) infection rates. The recovery rate is large.

6. Conclusion

In this paper, we study how the correlation between the infection rates and nodal degrees influences the epidemic spreading, compared to the uncorrelated case. By continuous-time simulations of our CSIS model in different infection-rate scenarios and networks with different topology heterogeneities, we find that, when the recovery rate is small, i.e. the prevalence of the epidemic is high, the negative correlation between the nodal degree and the infection rates tends to help the epidemic spreading. However, when the prevalence is high, the positive correlation is more likely to enhance the spreading. The validation on two real-world networks and the proof in the large wheel network agree with our conclusions. Our results shed light on that how the epidemic spreads in the real-world could be far away from the simple classic models. Not only the heterogeneity of infection rates but also the correlation between the heterogeneity of infection rates and network topologies could be various and complicated. Our work is the first step to study the correlated heterogeneous SIS model in heterogeneous networks.

Acknowledgment

Appendix. The wheel network with i.i.d. infection rates

As shown in Fig. A.8, we plot the average fraction y_{∞} of infected nodes as a function of the correlation strength parameter α for the wheel network with $m = 1000$ side nodes in the scenario of the uncorrelated heterogeneous infection rates. We employ the recovery rates $\delta = 0.5$, $\delta = 2$ and $\delta = 10$. We find that no matter what value the recovery rate is, the homogeneous infection rate ($\alpha = 0$) always leads to the highest overall infection.
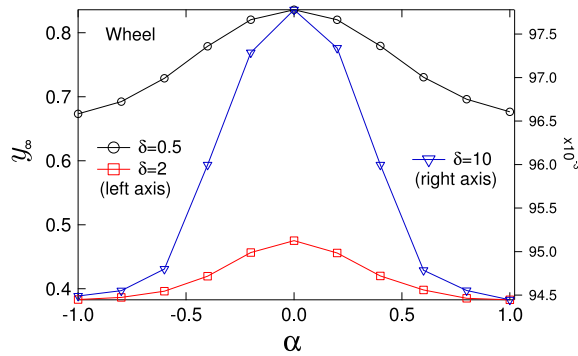
Fig. A.8. The average fraction y_∞ as a function of the correlation strength parameter α for a wheel network with $m = 1000$ side nodes in the scenario of the uncorrelated infection rates. The recovery rate is $\delta = 0.5$ ($\bigcirc$), $\delta = 2$ ($\square$) and $\delta = 10$ ($\triangledown$).

References

[1] D.J. Daley, J. Gani, J.M. Gani, Epidemic Modelling: An Introduction, Vol. 15, Cambridge University Press, 2001.
[2] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, A. Vespignani, Epidemic processes in complex networks, arXiv preprint arXiv:1408.2701.
[3] R. Pastor-Satorras, A. Vespignani, Epidemics and immunization in scale-free networks, in: Handbook of Graphs and Networks: from the Genome to the Internet, 2005, pp. 111–130.
[4] R. Pastor-Satorras, A. Vespignani, Epidemic dynamics and endemic states in complex networks, Phys. Rev. E 63 (6) (2001) 066117.
[5] A. Barrat, M. Barthelemy, R. Pastor-Satorras, A. Vespignani, The architecture of complex weighted networks, Proc. Natl. Acad. Sci. USA 101 (11) (2004) 3747–3752.
[6] W. Li, X. Cai, Statistical analysis of airport network of China, Phys. Rev. E 69 (4) (2004) 046106.
[7] P. Macdonald, E. Almaas, A.-L. Barabási, Minimum spanning trees of weighted scale-free networks, Europhys. Lett. 72 (2) (2005) 308.
[8] B. Qu, H. Wang, Sis epidemic spreading with heterogeneous infection rates, arXiv preprint arXiv:1506.07293.
[9] Y. Sun, C. Liu, C.-X. Zhang, Z.-K. Zhang, Epidemic spreading on weighted complex networks, Phys. Lett. A 378 (7) (2014) 635–640.
[10] C. Buono, F. Vazquez, P. Macri, L. Braunstein, Slow epidemic extinction in populations with heterogeneous infection rates, Phys. Rev. E 88 (2) (2013) 022813.
[11] V.M. Preciado, M. Zargham, C. Enyioha, A. Jadbabaie, G.J. Pappas, Optimal resource allocation for network protection against spreading processes, IEEE Trans. Control Syst. Technol. 1 (1) (2014) 99–108.
[12] Z. Yang, T. Zhou, Epidemic spreading in weighted networks: an edge-based mean-field solution, Phys. Rev. E 85 (5) (2012) 056106.
[13] X. Fu, M. Small, D.M. Walker, H. Zhang, Epidemic dynamics on scale-free networks with piecewise linear infectivity and immunization, Phys. Rev. E 77 (3) (2008) 036113.
[14] P. Van Mieghem, J. Omic, In-homogeneous virus spread in networks, arXiv preprint arXiv:1306.2588.
[15] P. Van Mieghem, The N-intertwined SIS epidemic network model, Computing 93 (2–4) (2011) 147–169.
[16] H. Wang, Q. Li, G. DAgostino, S. Havlin, H.E. Stanley, P. Van Mieghem, Effect of the interconnected network structure on the epidemic threshold, Phys. Rev. E 88 (2) (2013) 022801.
[17] F.D. Sahneh, C. Scoglio, P. Van Mieghem, Generalized epidemic mean-field model for spreading processes over multilayer complex networks, IEEE/ACM Trans. Netw. 21 (5) (2013) 1609–1620.
[18] W. Duan, R. Quax, M. Lees, X. Qiu, P.M. Sloot, Topology dependent epidemic spreading velocity in weighted networks, J. Stat. Mech. Theory Exp. 2014 (12) (2014) P12020.
[19] G. Caldarelli, R. Marchetti, L. Pietronero, The fractal properties of Internet, Europhys. Lett. 52 (4) (2000) 386.
[20] R. Albert, H. Jeong, A.-L. Barabási, Internet: Diameter of the world-wide web, Nature 401 (6749) (1999) 130–131.
[21] A.-L. Barabási, R. Albert, Emergence of scaling in random networks, Science 286 (5439) (1999) 509–512.
[22] R. Cohen, K. Erez, D. ben Avraham, S. Havlin, Resilience of the Internet to random breakdowns, Phys. Rev. Lett. 85 (2000) 4626–4628.
[23] P. Erdős, A. Rényi, On random graphs I, Publ. Math. Debrecen 6 (1959) 290–297.
[24] C. Li, R. van de Bovenkamp, P. Van Mieghem, Susceptible-infected-susceptible model: A comparison of n-intertwined and heterogeneous mean-field approximations, Phys. Rev. E 86 (2) (2012) 026116.
[25] M.E. Newman, The structure of scientific collaboration networks, Proc. Natl. Acad. Sci. 98 (2) (2001) 404–409.