Mathematical Biosciences and Engineering

http://www.aimspress.com/journal/MBE

MBE, 17(4): 3925–3952.
DOI: 10.3934/mbe.2020218
Received: 12 March 2020
Accepted: 20 May 2020
Published: 28 May 2020

Research article

Immunization strategies in directed networks

Junbo Jia, Wei Shi, Pan Yang and Xinchu Fu*

Department of Mathematics, Shanghai University, Shanghai 200444, China

* **Correspondence:** Email: xcfu@shu.edu.cn; Tel: +862166132664; Fax: +862166133292.

Abstract: Many complex systems can be modeled as directed networks, which can be regarded as a generalization of undirected networks. In this paper, epidemic dynamics and immunization strategies in directed networks are studied. First, a Susceptible-Infected-Susceptible (SIS) model on a directed network is established employing the mean-field method, and its dynamics and epidemic threshold of the network are studied. Then based on the continuous degree technique, namely, considering the degree of a node as a continuous variable, we propose a method to calculate the epidemic threshold of the immunized network. Besides, some immunization strategies, including optimal immunization, random immunization, combined targeted immunization, and combined acquaintance immunization, and three special networks are considered. Finally, through numerical analysis, all immunization strategies are simulated and compared on different types of networks. We find that the nodes with the largest product of in-degree and out-degree are the most worthy of being immunized.

Keywords: directed network; optimal immunization; random immunization; combined targeted immunization; combined acquaintance immunization

1. Introduction

Complex network is a powerful tool to describe and study complex systems [1]. Generally, the network is represented by a graph in mathematics, including nodes, which represent the individual component in the system, and edges between the nodes, which represent the relationship or interaction between individuals in the system. Many fields in the real world can be studied by using complex networks, such as computer networks [2], social networks [3–5], gene regulatory networks [6], etc.. An important reason why the application of complex networks is so widespread is that it can be used to characterize the topological features of these systems, such as small-world property [4,7], scale-free property [8], community structures [9–12], and hierarchical structures [13].

Disease or information transmission on the network is an important topic in the spreading dynamics on networks [1, 14–17]. It can be applied to study the spread of the epidemic in contact

networks, viruses in computer networks, and opinion in social networks. Depending on whether the edge has a direction, the underlying networks can be divided into undirected networks, directed networks [18, 19] and semi-directed networks [20, 21]. In this paper, directed networks are considered. There are two main reasons for this. First, the directed network can be regarded as a generalization of the undirected network and semi-directed network. For example, undirected networks can be seen as directed networks with two-way edges. Second, it is based on the fact that sometimes the edges are directional or the infection between two nodes is asymmetric. For example, some computer viruses can only be transmitted from server to client through downloading or visiting. The public opinion of social networks can only flow from followed to followers. Trade flows among manufacturers, wholesalers and retailers are well modeled by directed networks [22]. Besides, the health care workers who take protective measures are less likely be infected by patients, but in turn, the infected doctors in the clinic will likely to infect the patients, causing cross-infection, such as influenza, hepatitis B, and COVID-19 (or 2019-nCoV).

Also, designing and implementing efficient immunization strategies are of great significance for the prevention and control of disease or information outbreaks on various types of networks [23–27]. This is also an indispensable part of the study of spreading dynamics on networks. Here, by immunization we mean to make some nodes immune or resistant to a disease or pathogenic agent, especially by inoculation. After a node is immunized, it will neither be infected by the infected nodes nor spread the disease to the susceptible. Therefore, the immunized nodes can be regarded as the removed nodes, which change the topological structure of the network and protect the whole network from a disease outbreak. Many researchers have considered the immunization on different networks [24–26, 28, 29], including homogeneous networks and heterogeneous networks. The results indicate that the nodes with a large degree are worthy of being immunized because they are not only susceptible but once infected they can easily spread the disease. Besides, different immunization strategies have been developed as well [29], including random immunization, targeted immunization, and acquaintance immunization [26]. However, in front of the directed network, new problems arise. A node in a directed network has in-degree and out-degree, which makes the nodes with large in-degree not necessarily have large out-degree. This means that nodes that are easy to be infected are not necessarily easy to infect others. Similarly, nodes that are prone to infect others are not necessarily easy to be infected as well. Besides, when we consider immunizing the nodes with both large in-degree and out-degree, such nodes may not even exist, because in-degree and out-degree can obey different distributions.

In this paper, we study epidemic spread and immunization strategies in a directed network. First, using the mean-field method, the Susceptible-Infected-Susceptible (SIS) model on the directed network is established and the epidemic threshold is defined. By considering the immunized nodes as being removed from the network together with their corresponding in-edges and out-edges, the immunized network is regarded as a new network. Then, by continuous degree method [30], namely, considering the degree of a node as a continuous variable, the epidemic threshold of the immunized network is calculated. We also analyze four families of immunization strategies, including optimal immunization, random immunization, combined targeted immunization and its three variants, and combined acquaintance immunization and its two variants. Besides, immunization strategies for specific cases are considered as well, such as undirected networks, networks with an uncorrelated joint degree, and epidemic spread considering infectivity. Finally, the numerical analysis of all

immunization strategies for four different types of networks is given. We find that optimal immunization is the most efficient strategy, which is to immunize the nodes from top to bottom according to the product of in-degree and out-degree. Interestingly, the importance of in-degree and out-degree of the nodes are different. The strategies tending to immunize the nodes with large out-degree are slightly inferior to those of large in-degree in networks where only the in-degree is heterogeneous, but in the networks where the out-degree is heterogeneous, the former will be significantly better than the latter.

The rest of the paper is organized as follows. In section 2, the SIS model on the directed network is established, and the global stability analysis and expression of epidemic threshold are given. In section 3, based on the continuous degree, the epidemic threshold of the immunized network is calculated, and some immunization strategies and special cases are considered. In section 4, the numerical analysis is provided to complement the theoretical analysis. Finally, the conclusions and discussions are given in section 5.

2. The SIS model in directed network

In this section, we first introduce the directed network and build an epidemic spread model in it. Then, the global stability analysis of the model is given.

2.1. Directed network

First, let us introduce a directed network. Directed network means the edges in the network are directed. So, each node attached with two kinds of edges, in-edges, and out-edges. For a node in the network, we use in-degree and out-degree represent the numbers of in-edges and out-edges, respectively, and by joint degree (k, l) we mean that for a node its in-degree is k and its out-degree is l. Here, we use k_* and k^* to represent the minimum and maximum of the in-degree, respectively, and l_* and l^* correspond to out-degree. Denote by $N_{k,l}$ the number of nodes with a joint degree (k, l) and by N the size of the whole network, then

$$N = \sum_{k,l} N_{k,l}. \tag{2.1}$$

Denote the number of nodes with in-degree k by $N_{k,\cdot}$ and that of with out-degree l by $N_{\cdot,l}$, and we have

$$N_{k,\cdot} = \sum_{l} N_{k,l}, \tag{2.2}$$

and

$$N_{\cdot,l} = \sum_{k} N_{k,l}. \tag{2.3}$$

In an undirected network, the degree distribution $P(k)$ is defined to be the fraction of nodes with degree k. Thus if there are N nodes in total in a network and N_k of them have degree k, we have $P(k) = \frac{N_k}{N}$. However, in a directed network, degree distribution $P(k)$ will be replaced by joint degree distribution $P(k, l)$, which is defined to be the fraction of nodes with a joint degree (k, l) instead. And, we have

$$P(k, l) = \frac{N_{k,l}}{N}. \tag{2.4}$$

Besides, the in-degree distribution and out-degree distribution will be

$$P_{in}(k) = \frac{N_{k,\cdot}}{N} = \frac{\sum_l N_{k,l}}{N}, \tag{2.5}$$

and

$$P_{out}(l) = \frac{N_{\cdot,l}}{N} = \frac{\sum_k N_{k,l}}{N}, \tag{2.6}$$

respectively.

Accordingly, we have the relationship between the marginal degree distribution and degree distribution

$$P(k,\cdot) = \sum_l P(k,l) = \frac{\sum_l N_{k,l}}{N} = P_{in}(k), \tag{2.7}$$

and

$$P(\cdot,l) = \sum_k P(k,l) = \frac{\sum_k N_{k,l}}{N} = P_{out}(l). \tag{2.8}$$

And for each node in the network, if its joint degree is independent each other, we will have the relationship

$$P(k,l) = P_{in}(k)P_{out}(l). \tag{2.9}$$

Besides, by the joint degree distribution $P(k,l)$, we can obtain moment $\langle k^\phi l^\varphi \rangle$ about zero, which equals to

$$\langle k^\phi l^\varphi \rangle = \sum_{k,l} k^\phi l^\varphi P(k,l), \quad \phi,\varphi = 0,1,2,3,\cdots. \tag{2.10}$$

If $\phi = \varphi = 1$, $\langle k^\phi l^\varphi \rangle$ will be mixture moment $\langle kl \rangle$. If $\phi = 1, \varphi = 0$, or $\phi = 0, \varphi = 1$, it will be the mean in-degree or mean out-degree, which can be denoted by $\langle k \rangle$ and $\langle l \rangle$, respectively. Here, we should note that whole the in-degree must equal whole the out-degree because, for each directed edge, it must be an in-edge of certain node and an out-edge of another node at the same time. So we obtain

$$\langle k \rangle = \langle l \rangle. \tag{2.11}$$

2.2. Epidemic model in directed network

Based on the directed network described above, we now consider the epidemic spread model, including the assumptions and the modeling of the propagating mechanism.

First, let us introduce the health state of nodes into the network. All nodes are divided into two compartments: Susceptible denoted by S, and infected denoted by I. S node means that it can be infected and I node means it is infective. Meanwhile, we use $S(t)$ and $I(t)$ denote the amount of S nodes and the I nodes, respectively, at time t, and use $S_{k,l}(t)$ and $I_{k,l}(t)$ to represent the S nodes and I nodes with joint-degree (k,l). So, according to the health state of nodes and their joint degree, we have the following relationship

$$N = S(t) + I(t) = \sum_{k,l} S_{k,l}(t) + \sum_{k,l} I_{k,l}(t). \tag{2.12}$$

Besides, for the nodes with a joint degree (k,l), we have

$$N_{k,l} = S_{k,l}(t) + I_{k,l}(t). \tag{2.13}$$

If we let $s_{k,l}(t) = \frac{S_{k,l}(t)}{N_{k,l}}$ and $\rho_{k,l}(t) = \frac{I_{k,l}(t)}{N_{k,l}}$ denote the relative densities of S nodes and I nodes with respect to the nodes with joint degree (k, l), respectively, then from the last equation, we obtain

$$s_{k,l}(t) + \rho_{k,l}(t) = 1. \tag{2.14}$$

Next, we give the assumption about the propagating mechanism of the epidemic in the directed network described in the previous subsection. For each S node in the network, say A, it can be infected by a I node through the directed edge that is emitted by this I node and point to the node A, and all the directed edges like this can cause the infection of node A. The infection rate of each edge during a unit of time is set to β. On the other hand, for each I node, it can become an S node again due to recovery, and the recovery rate during a unit time is set to γ.

Based on this infection mechanism, the SIS model can be built via mean-field method, as below,

$$\frac{dI_{k,l}(t)}{dt} = \beta k S_{k,l}(t)\Theta_{k,l}(t) - \gamma I_{k,l}(t). \tag{2.15}$$

Where, the first item on the right hand represents the increment of $I_{k,l}(t)$ because of the infection of susceptible nodes $S_{k,l}(t)$, and the second item indicates the decrement of $I_{k,l}(t)$ caused by recovery. $\Theta_{k,l}(t)$ denote the possibility of an in-edge of S node with joint degree (k, l) is emitted by I nodes. For simplicity, we consider the relative densities and substitute (2.14) into (2.15) and assume the network does not have the degree-degree correlation between two end nodes of a directed edge, then the SIS model in directed network becomes

$$\frac{d\rho_{k,l}(t)}{dt} = \beta k(1 - \rho_{k,l}(t))\Theta(t) - \gamma\rho_{k,l}(t). \tag{2.16}$$

Where,

$$\Theta(t) = \frac{\sum_{k,l} lI_{k,l}(t)}{\sum_{k,l} lN_{k,l}} = \frac{1}{\langle l\rangle}\sum_{k,l} lP(k, l)\rho_{k,l}(t). \tag{2.17}$$

2.3. Epidemic threshold and global stability analysis

The epidemic threshold is a crucial parameter to determine whether the disease will outbreak or not. Generally, when the related parameter exceeds the threshold, the disease will break out, otherwise, disease tend to vanish. Here, we calculate the epidemic threshold.

First, let us consider the equilibria of the model, and let the right hand of (2.16) equals to zero, we get

$$\rho_{k,l}(t) = \frac{\beta k\Theta(t)}{\gamma + \beta k\Theta(t)}. \tag{2.18}$$

Substituting (2.18) into (2.17), we obtain a self-consistency equation with respect to $\Theta(t)$,

$$\Theta(t) = \frac{1}{\langle l\rangle}\sum_{k,l} lP(k, l)\frac{\beta k\Theta(t)}{\gamma + \beta k\Theta(t)}. \tag{2.19}$$

We define the function $f(\Theta)$ as the right hand of (2.19), and it has the following properties

$$f(0) = 0, \tag{2.20}$$

$$f(1) = \frac{1}{\langle l\rangle}\sum_{k,l} lP(k,l)\frac{\beta k}{\gamma+\beta k} < \frac{1}{\langle l\rangle}\sum_{k,l} lP(k,l) = 1, \tag{2.21}$$

$$\frac{df(\Theta)}{dt} = \frac{1}{\langle l\rangle}\sum_{k,l} lP(k,l)\frac{\beta k\gamma}{(\gamma+\beta k\Theta)^2} > 0, \tag{2.22}$$

$$\frac{d^2f(\Theta)}{dt^2} = \frac{1}{\langle l\rangle}\sum_{k,l} lP(k,l)\frac{-2\beta^2k^2\gamma}{(\gamma+\beta k\Theta)^3} < 0. \tag{2.23}$$

From these properties, we know that $f(\Theta)$ is a strictly monotone increasing and a strictly convex function. The value at origin is zero and is less than 1 at point 1. Therefore, the endemic equilibrium exists if and only if

$$\left.\frac{df(\Theta)}{dt}\right|_{\Theta=0} = \frac{\beta\langle kl\rangle}{\gamma\langle l\rangle} > 1. \tag{2.24}$$

So we define the epidemic threshold λ_c as

$$\lambda_c = \frac{\langle l\rangle}{\langle kl\rangle}. \tag{2.25}$$

When the joint-degree is independent, that is, when (2.9) holds, the epidemic threshold will be

$$\lambda_c^* = \frac{1}{\langle l\rangle}. \tag{2.26}$$

Next, we consider the global dynamics of the model (2.16).

Theorem 2.1. *For the model (2.16) and epidemic threshold λ_c defined in (2.25), if $\frac{\beta}{\gamma} \le \lambda_c$ then the disease-free equilibrium E_0 is globally asymptotically stable in Δ. While if $\frac{\beta}{\gamma} > \lambda_c$, model (2.16) admits a unique endemic equilibrium E_1, which is globally asymptotically stable in $\Delta - \{0\}$, where $\Delta = \{\rho_{k,l} \mid 0 \le \rho_{k,l} \le 1, \forall k_* \le k \le k^*, l_* \le l \le l^*\}$.*

Proof. Let $R_0 = \frac{\beta\langle kl\rangle}{\gamma\langle k\rangle}$, namely the basic reproduction number. Then we have $R_0 \le 1 \Longleftrightarrow \frac{\beta}{\gamma} \le \frac{\langle k\rangle}{\langle kl\rangle} = \lambda_c$ and $R_0 \ge 1 \Longleftrightarrow \frac{\beta}{\gamma} \ge \frac{\langle k\rangle}{\langle kl\rangle} = \lambda_c$. Besides, the model (2.16) is a special case of the model studied in [18] since the latter is generalized. By Theorem 1 in [18], we complete the proof. □

The Theorem 2.1 above shows that when the relative infection rate, $\frac{\beta}{\gamma}$, is less than or equal to the epidemic threshold, λ_c, the disease will gradually die out. Only when the relative infection rate is greater than the epidemic threshold, the disease will break out. Furthermore, Theorem 2.1 also implies that if we want to prevent the outbreak of the disease on the network, an effective way is to change the topological structure of the network such that the epidemic threshold λ_c of the network is greater than the relative infection rate of the disease.

Besides, it should be noted that the epidemic threshold defined in (2.25) is uniquely determined by the network topology and have nothing to do with the disease. However, the relative infection rate of disease, $\frac{\beta}{\lambda}$, is not only related to the disease itself, but also to the media, which refers to the directed network in this situation. Additionally, our model and conclusion are based on the assumptions that nodes are uniformly mixed up and there is no correlation between two ends of any directed edge, and the subsequent research on immunization strategies is also based on these assumptions.

3. Immunization strategies

In this section, we will study the immunization strategies in the directed network, which is an indispensable step for epidemic prevention. By immunization, we mean to make nodes immune to the epidemic or pathogenic agent, especially by inoculation. Generally, medical resources related to vaccination are high cost and often inadequate. Therefore, we need to design economical and efficient immunization strategies.

The immunizations we studied here are different from treatment and are carried out before the onset of the outbreak. In other words, before the emergence of the disease, some nodes in the network are vaccinated in advance, such that when an external infected node comes to the network or a certain node in the network suddenly becomes infected, the disease can tend to die out.

Let us briefly describe the principle behind the prevention of disease outbreaks through immunization. We call the immunized node the removed node, denoted by R, which is different from the S node and the I node. This is because the immunized nodes will not be infected and are not infectious as well. They do not participate in the spread of disease, just like they are removed from the network. Thus, immunization can be regarded as removing some nodes and the attached edges from the network, which changes the topological structure of the network and finally makes the epidemic threshold increase, thus reducing the prevalence or even preventing the outbreak of the disease.

3.1. Epidemic threshold of immunized network

We calculate the epidemic threshold of an immunized network by continuous degree method [30]. By continuous degree, we mean considering the degree of a node as the positive real number that varies continuously. We can do that mainly because when the size of the network is large enough, regarding the degree as a continuous variable is a good approximation method and is easy to deal with as well.

For the nodes with a joint degree (k, l), assume the rate of being immunized is $\alpha(k, l)$, where $0 \leq \alpha(k, l) \leq 1$. $\alpha(k, l) = 0$ means the nodes with a joint degree (k, l) are not been immunized, and $\alpha(k, l) = 1$ means the nodes are immunized completely. Alternatively, we can also think of $\alpha(k, l)$ as the probability that a node with joint degree (k, l) is immunized. Once the network is immunized, those immunized nodes are considered removed from the network and its total number are $\int_{l_*}^{l^*}\int_{k_*}^{k^*} \alpha(k, l)N_{k,l}dkdl$, where k_*, k^*, l_*, and l^* are as described in subsection 2.1. Accordingly, the attached $\int_{l_*}^{l^*}\int_{k_*}^{k^*} k\alpha(k, l)N_{k,l}dkdl$ in-edges and $\int_{l_*}^{l^*}\int_{k_*}^{k^*} l\alpha(k, l)N_{k,l}dkdl$ out-edges will be removed as well.

For the unremoved nodes with a joint degree (k, l), due to the removing of some edges their in-degree k and out-degree l will become k' and l', respectively, after immunization. Where,

$$\begin{aligned} k' &= k - \int_{l_*}^{l^*}\int_{k_*}^{k^*} l\alpha(k, l)N_{k,l}dkdl \frac{k}{\int_{l_*}^{l^*}\int_{k_*}^{k^*} kN_{k,l}dkdl} \\ &= k(1 - \tfrac{\langle l\alpha(k,l)\rangle}{\langle k\rangle}) \\ &= kA \end{aligned} \tag{3.1}$$

and

$$\begin{aligned} l' &= l - \int_{l_*}^{l^*}\int_{k_*}^{k^*} k\alpha(k, l)N_{k,l}dkdl \frac{l}{\int_{l_*}^{l^*}\int_{k_*}^{k^*} lN_{k,l}dkdl} \\ &= l(1 - \tfrac{\langle k\alpha(k,l)\rangle}{\langle l\rangle}) \\ &= lB. \end{aligned} \tag{3.2}$$

In (3.1), $\int_{l_*}^{l^*}\int_{k_*}^{k^*} l\alpha(k,l)N_{k,l}dkdl$ represents the total number of removed in-edges, which is from the perspective of surviving nodes, and the fraction part represents the ratio of in-edges of a node with in-degree k in the whole in-edges. So their product will be the reduction of in-degree for an node with in-degree k. And the items in (3.2) are similar. Besides, for simplicity, we also let A and B represent $1-\frac{\langle l\alpha(k,l)\rangle}{\langle k\rangle}$ and $1-\frac{\langle k\alpha(k,l)\rangle}{\langle l\rangle}$, respectively.

Conversely, for the node with a joint degree (k',l') in the immunized network, their original in-degree and out-degree in the original network are,

$$k=\frac{k'}{A}, \tag{3.3}$$

and

$$l=\frac{l'}{B}, \tag{3.4}$$

respectively.

For the original network, the in-degree k and out-degree l satisfy $k_*\le k\le k^*$ and $l_*\le l\le l^*$. Then for the immunized network, we have $k'_*\le k'\le k'^*$, $l'_*\le l'\le l'^*$, and the relationship

$$k'_*=k_*A,\quad k'^*=k^*A, \tag{3.5}$$

and

$$l'_*=l_*B,\quad l'^*=l^*B. \tag{3.6}$$

Next, the new joint degree distribution, or the new joint probability density, is

$$\begin{aligned} P'(k',l') &= \frac{N'_{k',l'}}{N'}\\ &= \frac{(1-\alpha(\frac{k'}{A},\frac{l'}{B}))N_{\frac{k'}{A},\frac{l'}{B}}}{\int_{l_*}^{l^*}\int_{k_*}^{k^*}(1-\alpha(k,l))N_{k,l}dkdl}\\ &= \frac{(1-\alpha(\frac{k'}{A},\frac{l'}{B}))P(\frac{k'}{A},\frac{l'}{B})}{\int_{l_*}^{l^*}\int_{k_*}^{k^*}(1-\alpha(k,l))P(k,l)dkdl}, \end{aligned} \tag{3.7}$$

which is completely determined by the original joint degree distribution $P(k,l)$ and the immune function $\alpha(k,l)$.

With the new joint degree distribution, we can obtain the new moments $\langle kl\rangle'$ and $\langle l\rangle'$, as follows,

$$\begin{aligned} \langle kl\rangle' &= \int_{l'_*}^{l'^*}\int_{k'_*}^{k'^*} k'l'P'(k',l')dk'dl',\\ &= \int_{l_*B}^{l^*B}\int_{k_*A}^{k^*A} AkBl\frac{(1-\alpha(\frac{k'}{A},\frac{l'}{B}))P(\frac{k'}{A},\frac{l'}{B})}{\int_{l_*}^{l^*}\int_{k_*}^{k^*}(1-\alpha(k,l))P(k,l)dkdl}dk'dl'\\ &= \int_{l_*}^{l^*}\int_{k_*}^{k^*} A^2B^2kl\frac{(1-\alpha(k,l))}{\int_{l_*}^{l^*}\int_{k_*}^{k^*}(1-\alpha(k,l))P(k,l)dkdl}P(k,l)dkdl\\ &= \frac{A^2B^2}{\int_{l_*}^{l^*}\int_{k_*}^{k^*}(1-\alpha(k,l))P(k,l)dkdl}\int_{l_*}^{l^*}\int_{k_*}^{k^*} kl(1-\alpha(k,l))P(k,l)dkdl, \end{aligned} \tag{3.8}$$

and

$$\begin{aligned} \langle l\rangle' &= \int_{l'_*}^{l'^*}\int_{k'_*}^{k'^*} l'P'(k',l')dk'dl'.\\ &= \int_{l_*B}^{l^*B}\int_{k_*A}^{k^*A} Bl\frac{(1-\alpha(\frac{k'}{A},\frac{l'}{B}))P(\frac{k'}{A},\frac{l'}{B})}{\int_{l_*}^{l^*}\int_{k_*}^{k^*}(1-\alpha(k,l))P(k,l)dkdl}dk'dl'\\ &= \int_{l_*}^{l^*}\int_{k_*}^{k^*} AB^2l\frac{(1-\alpha(k,l))}{\int_{l_*}^{l^*}\int_{k_*}^{k^*}(1-\alpha(k,l))P(k,l)dkdl}P(k,l)dkdl\\ &= \frac{AB^2}{\int_{l_*}^{l^*}\int_{k_*}^{k^*}(1-\alpha(k,l))P(k,l)dkdl}\int_{l_*}^{l^*}\int_{k_*}^{k^*} l(1-\alpha(k,l))P(k,l)dkdl. \end{aligned} \tag{3.9}$$

Thus we obtain the epidemic threshold of immunized network

$$\begin{aligned}\lambda_c' &= \frac{\langle l\rangle'}{\langle kl\rangle'} \\ &= \frac{\int_{l_*}^{l^*}\int_{k_*}^{k^*} l(1-\alpha(k,l))P(k,l)dkdl}{A\int_{l_*}^{l^*}\int_{k_*}^{k^*} kl(1-\alpha(k,l))P(k,l)dkdl} \\ &= \frac{\langle l\rangle}{\langle kl\rangle-\langle kl\alpha(k,l)\rangle}.\end{aligned} \tag{3.10}$$

Where, $\langle kl\alpha(k,l)\rangle = \int_{l_*}^{l^*}\int_{k_*}^{k^*} kl\alpha(k,l)P(k,l)dkdl$. If we substitute $\langle k\rangle$ with $\langle l\rangle$ due to (2.11), λ_c' can also be expressed as $\frac{\langle k\rangle}{\langle kl\rangle-\langle kl\alpha(k,l)\rangle}$.

For an immunization strategy, we can also define the immunization threshold, α_c, as the proportion of the total immunized nodes such that $\frac{\beta}{\gamma} = \lambda_c'$. If the proportion of total immunized nodes is less than α_c, the disease breakout. Otherwise, the disease will tend to die out. The smaller the immunization threshold α_c, the higher the efficiency, because a small α_c means fewer nodes needed to be immunized for preventing disease outbreaks.

3.2. Some examples of immunization strategies

Based on the calculation method proposed above, we now give some examples of immunization strategies in this subsection, including optimal immunization, random immunization, combined targeted immunization and combined acquaintance immunization. For convenience, we also respectively call these strategies *A*, *B*, *C* and *D*.

3.2.1. Strategy *A*: Optimal immunization

The first is the optimal immunization, which is to immunize the nodes in turn from the maximum value to the minimum value according to the *kl* value of the nodes, namely, the product of the out-degree and the in-degree. Here, the optimal means that the epidemic threshold λ_c' is the largest when the same number of nodes are immunized. In other words, its immunization threshold α_c defined at the end of the previous subsection is the smallest of all immunization strategies. This immunization strategy is inspired by the expression of the epidemic threshold in (3.10) because when the same number of nodes are being immunized, immunizing those nodes with the highest *kl* will maximize the epidemic threshold of the network.

The fact that strategy *A* is optimal is based on the global information of joint degree. Thus, one drawback is that we cannot give a specific immunization threshold α_c without the knowledge of the joint degree distribution of a specific network.

3.2.2. Strategy *B*: Random immunization

The second is random immunization, also known as uniform immunization [24], which is the simplest immunization strategy. That is to say, regardless of the degree of nodes, either in-degree, out-degree, or joint degree, each node is being immunized with the same probability, say constant α, where $0 \le \alpha \le 1$. So we have

$$\alpha(k,l) = \alpha. \tag{3.11}$$

Thus,

$$\langle kl\alpha(k,l)\rangle = \int_{l_*}^{l^*}\int_{k_*}^{k^*} kl\alpha P(k,l)dkdl = \alpha\langle kl\rangle. \tag{3.12}$$

So, substituting this into (3.10), we obtain the epidemic threshold

$$\lambda_c = \frac{\langle l \rangle}{(1-\alpha)\langle kl \rangle}. \tag{3.13}$$

Then let it equal the relative infection rate $\frac{\beta}{\gamma}$, we get the immunization threshold

$$\alpha_c = 1 - \frac{\gamma \langle l \rangle}{\beta \langle kl \rangle}. \tag{3.14}$$

Strategy B is effective because it can increase the epidemic threshold, and when the immunization proportion α is large enough, the epidemic threshold can be greater than the relative infection rate. One advantage of strategy B is that no information about node degree is required when the node is being immunized. Usually, this also makes it inefficient, and we often think of random strategy as a reference.

3.2.3. Strategy C: Combined targeted immunization

The third strategy is combined targeted immunization, which is a generalization of targeted immunization on undirected networks [24, 25]. In an undirected network, the nodes with a large degree are worthy of being immunized, because they are not only easy to be infected but also easy to spread disease. However, in a directed network, degree becomes joint degree, which contains two components, in-degree, and out-degree. The nodes with large in-degree do not necessarily have large out-degree, and the nodes with large out-degree do not necessarily have large in-degree as well. In other words, in the directed network, the nodes that are easy to be infected are not necessarily easy to spread the disease, and, similarly, the nodes that are easy to spread the disease may not be easy to be infected. Therefore, the nodes with both large in-degree and large out-degree are generally worthy of being immunized. So, here the immunization function $\alpha(k, l)$ is expressed as follows,

$$\alpha(k,l) = \begin{cases} 1, & k \geq k_c \text{ and } l \geq l_c, \\ 0, & \text{otherwise.} \end{cases} \tag{3.15}$$

Where, k_c and l_c represent the critical values of in-degree and out-degree, respectively. Thus, $\langle kl\alpha(k,l) \rangle$ will be

$$\begin{aligned} \langle kl\alpha(k,l) \rangle &= \int_{l_c}^{l^*} \int_{k_c}^{k^*} kl\alpha(k,l)P(k,l)dkdl \\ &= \int_{l_c}^{l^*} \int_{k_c}^{k^*} klP(k,l)dkdl \\ &= \omega_1. \end{aligned} \tag{3.16}$$

Where, ω_1 represents the integral of $klP(k,l)$ in area $\Omega_1 = \{(k,l) : k_c \leq k \leq k^*, l_c \leq l \leq l^*\}$. Then the epidemic threshold is

$$\lambda_c = \frac{\langle l \rangle}{\langle kl \rangle - \omega_1}. \tag{3.17}$$

Similar to the optimal immunization, we can not give a specific immunization threshold without specific k_c, l_c, and $P(k,l)$. And strategy C also requires global information about node degrees, so do strategy A. Besides, strategy C here is actually a series of immunizations, because different k_c and l_c can be considered as different immunization.

There are three variants associated with strategy C, as follows.

Variant C_1: Immunize the nodes either with large in-degree or with large out-degree. The immunization function is expressed as

$$\alpha(k,l) = \begin{cases} 1, & k \geq k_c \text{ or } l \geq l_c, \\ 0, & \text{otherwise.} \end{cases} \tag{3.18}$$

And, the $\langle kl\alpha(k,l)\rangle$ will be

$$\begin{aligned} \langle kl\alpha(k,l)\rangle &= \int_{l_*}^{l^*}\int_{k_*}^{k^*} kl\alpha(k,l)P(k,l)dkdl \\ &= (\int_{l_c}^{l^*}\int_{k_c}^{k^*} + \int_{l_c}^{l^*}\int_{k_*}^{k_c} + \int_{l_*}^{l_c}\int_{k_c}^{k^*})klP(k,l)dkdl \\ &= \omega_1 + \omega_2 + \omega_3. \end{aligned} \tag{3.19}$$

Where, ω_1 is the same as that in (3.16), and ω_2, ω_3 are the integrals of $klP(k,l)$ in areas $\Omega_2 = \{(k,l) : k_* \leq k \leq k_c, l_c \leq l \leq l^*\}$ and $\Omega_3 = \{(k,l) : k_c \leq k \leq k^*, l_* \leq l \leq l_c\}$, respectively. So the infection threshold is

$$\lambda_c = \frac{\langle l\rangle}{\langle kl\rangle - \omega_1 - \omega_2 - \omega_3}. \tag{3.20}$$

Variant C_2: The second variant is only immunizing the nodes with a large in-degree, regardless of the node's out-degree. And the immunization function is

$$\alpha(k,l) = \begin{cases} 1, & k \geq k_c, \\ 0, & \text{otherwise.} \end{cases} \tag{3.21}$$

Through the same calculation as above, we obtain the epidemic threshold

$$\lambda_c = \frac{\langle l\rangle}{\langle kl\rangle - \omega_1 - \omega_2}. \tag{3.22}$$

Where ω_1 and ω_2 are as described above.

Variant C_3: In contrast to the second variant, the third variant is only immunizing the nodes with a large out-degree. So the immunization function is

$$\alpha(k,l) = \begin{cases} 1, & l \geq l_c, \\ 0, & \text{otherwise.} \end{cases} \tag{3.23}$$

And the corresponding epidemic threshold is

$$\lambda_c = \frac{\langle l\rangle}{\langle kl\rangle - \omega_1 - \omega_3}. \tag{3.24}$$

Where ω_1 and ω_3 are as described above as well.

We also can not give the immunization threshold of these three variants, because these strategies require the global information about the joint degree and the specific values k_c and l_c, which is the same as the case in strategy C. For a particular network, when k_c and l_c are given, the epidemic threshold of C_1 is the maximum and that of C is the minimum among this family of immunization strategies. The

infection thresholds of C_2 and C_3 are undetermined, but both are between that of C and C_1. However, we can not say C_1 is more efficient than C due to the uncertainty of their immunization thresholds.

To maximize efficiency, strategy C and its three variants need to be implemented on different networks, although all of them can increase the epidemic threshold. Strategy C is suitable for the networks of nodes with heterogeneity in both in-degree and out-degree, and the networks where the nodes with large in-degree are usually the nodes with large out-degree. Variant C_1 is suitable for the networks that there is heterogeneity in both in-degree and out-degree but no correlation between in-degree and out-degree of a node. Variant C_2 is suitable for networks with heterogeneity in in-degree but homogeneity in out-degree. In contrast to variant C_2, variant C_3 is suitable for networks with heterogeneity in out-degree but homogeneity in in-degree.

3.2.4. Strategy D: Combined acquaintance immunization

The fourth is combined acquaintance immunization, which is a generalization of acquaintance immunization on the undirected network [26]. Its principle is that generally the nodes worthy of being immunized have a large degree, either in-degree or out-degree, and if we look for the targeted nodes in the opposite direction of the directed edges, these important nodes are more likely to be found. Therefore, we randomly immunize a neighbor of randomly selected a node each time, regardless of whether the neighbor is connected through the in-edges or the out-edges of the node. This strategy can also be called proportional immunization because the proportion (or the probability) of a node being immunized is proportional to the sum of the in-degree and out-degree of that node. And the immunization function is

$$\alpha(k,l) = \frac{a(k+l)}{\int_{l_*}^{l^*}\int_{k_*}^{k^*}(k+l)N_{k,l}dkdl} = \frac{a(k+l)}{2N\langle k\rangle}, \tag{3.25}$$

where a is the total number of nodes being immunized.

Calculating $\langle kl\alpha(k,l)\rangle$, then substituting it into (3.10), we can get

$$\lambda_c = \frac{\langle l\rangle}{\langle kl\rangle - \frac{a}{2N\langle k\rangle}(\langle k^2 l\rangle + \langle kl^2\rangle)}. \tag{3.26}$$

Let (3.26) equal relative infection rate $\frac{\beta}{\gamma}$, we have the immunization threshold

$$\alpha_c = \frac{a}{N} = \left(\langle kl\rangle - \frac{\gamma\langle l\rangle}{\beta}\right)\frac{2\langle k\rangle}{\langle k^2 l\rangle + \langle kl^2\rangle}. \tag{3.27}$$

Similar to strategy C, strategy D also has its variants, as follows.

Variant D_1: The first variant is also randomly immunizing a neighbor of a randomly selected node. However, the difference is that we look for the neighbor only along the in-edges of the node. In fact, this strategy can be considered as immunizing a node with the probability proportional to the out-degree of this node. This is because in this case, nodes with large out-degree are more likely to be immunized as neighbors. Here, the immunization function is

$$\alpha(k,l) = \frac{al}{\int_{l_*}^{l^*}\int_{k_*}^{k^*} lN_{k,l}dkdl} = \frac{al}{N\langle l\rangle}, \tag{3.28}$$

where, a is the number of immunized nodes, the same as above. By the same method as well, we obtain the epidemic threshold

$$\lambda_c = \frac{\langle l \rangle}{\langle kl \rangle - \frac{a}{N\langle l \rangle}\langle kl^2 \rangle}, \tag{3.29}$$

and the immunization threshold

$$\alpha_c = \frac{a}{N} = \left(\langle kl \rangle - \frac{\gamma \langle l \rangle}{\beta} \right) \frac{\langle l \rangle}{\langle kl^2 \rangle}. \tag{3.30}$$

Variant D_2: Unlike variant D_1, the second variant looks for the neighbors along the out-edges of the node. So variant D_2 can be regarded as immunizing node with the probability being immunized is proportional to its in-degree. Thus the immunization function is

$$\alpha(k, l) = \frac{ak}{\sum_{k,l} kN_{k,l}} = \frac{ak}{N\langle k \rangle}. \tag{3.31}$$

Then, the epidemic threshold λ_c and immunization threshold α_c will be

$$\lambda_c = \frac{\langle l \rangle}{\langle kl \rangle - \frac{a}{N\langle k \rangle}\langle k^2 l \rangle}, \tag{3.32}$$

and

$$\alpha_c = \frac{a}{N} = \left(\langle kl \rangle - \frac{\gamma \langle l \rangle}{\beta} \right) \frac{\langle k \rangle}{\langle k^2 l \rangle}, \tag{3.33}$$

respectively.

One of the advantages of strategy D and its two variants, D_1 and D_2, is that they require no knowledge of the node degree or any other global information [26], which makes them different from strategy A, C and variants C_1, C_2. Besides, only in the appropriate networks, the immunization strategies proposed in this subsection can maximize their performance. Strategy D is suitable for the networks with heterogeneity either in in-degree or out-degree. However, variant D_1 and D_2 are suitable for networks with heterogeneity only in in-degree and out-degree, respectively.

Finally, we should note that in the previous subsection we have obtained the expression of infection threshold of the immunized network by continuous degree method, so the integrals in this subsection can be replaced by the sum during the calculation for the specific network. Moreover, for convenience, all the related epidemic thresholds in this subsection are listed in Table 1, which also contains some epidemic thresholds in the next subsection.

3.3. Immunization for special networks

In the two subsections above, we have studied the calculation of the epidemic threshold of the immunized network and some immunization strategies. Next, in this subsection, we will consider three special cases of directed networks, which are undirected network, network with uncorrelated joint degree and epidemic spread with considering infectivity. This is because the design and implementation of immunization strategies are closely dependent on the topology of the network and the mechanism of disease spread.

Table 1. The related epidemic threshold λ_c[1].

	Directed networks	Case 1	Case 2
No immunization	$\frac{\langle l\rangle}{\langle kl\rangle}$	$\frac{\langle l\rangle}{\langle k^2\rangle}$	$\frac{1}{\langle k\rangle}$
Immunized	$\frac{\langle l\rangle}{\langle kl\rangle-\langle kl\alpha(k,l)\rangle}$	$\frac{\langle l\rangle}{\langle k^2\rangle-\langle k^2\alpha(k)\rangle}$	$\frac{\langle k\rangle}{\langle k\rangle^2-\langle kl\alpha(k,l)\rangle}$
Strategy A	/	/	/
Strategy B	$\frac{\langle l\rangle}{(1-\alpha)\langle kl\rangle}$	$\frac{\langle l\rangle}{(1-\alpha)\langle k^2\rangle}$	$\frac{1}{(1-\alpha)\langle k\rangle}$
Strategy C	$\frac{\langle l\rangle}{\langle kl\rangle-\omega_1}$	$\frac{\langle k\rangle}{\langle k^2\rangle-\omega_4}$	$\frac{\langle k\rangle}{\langle k\rangle\langle k\rangle-\omega_1}$
Variant C_1	$\frac{\langle l\rangle}{\langle kl\rangle-\omega_1-\omega_2-\omega_3}$	$\frac{\langle k\rangle}{\langle k^2\rangle-\omega_4}$	$\frac{\langle l\rangle}{\langle k\rangle l\rangle-\omega_1-\omega_2-\omega_3}$
Variant C_2	$\frac{\langle l\rangle}{\langle kl\rangle-\omega_1-\omega_2}$	$\frac{\langle k\rangle}{\langle k^2\rangle-\omega_4}$	$\frac{\langle l\rangle}{\langle k\rangle\langle l\rangle-\omega_1-\omega_2}$
Variant C_3	$\frac{\langle l\rangle}{\langle kl\rangle-\omega_1-\omega_3}$	$\frac{\langle k\rangle}{\langle k^2\rangle-\omega_4}$	$\frac{\langle l\rangle}{\langle k\rangle\langle l\rangle-\omega_1-\omega_3}$
Strategy D	$\frac{\langle l\rangle}{\langle kl\rangle-\frac{a}{2N\langle k\rangle}(\langle k^2l\rangle+\langle kl^2\rangle)}$	$\frac{\langle k\rangle}{\langle k^2\rangle-\frac{a\langle k^3\rangle}{N\langle k\rangle}}$	$\frac{\langle l\rangle}{\langle k\rangle\langle l\rangle-\frac{a}{2N}(\langle k^2\rangle+\langle l^2\rangle)}$
Variant D_1	$\frac{\langle l\rangle}{\langle kl\rangle-\frac{a}{N\langle l\rangle}\langle kl^2\rangle}$	$\frac{\langle k\rangle}{\langle k^2\rangle-\frac{a\langle k^3\rangle}{N\langle k\rangle}}$	$\frac{\langle l\rangle}{\langle k\rangle\langle l\rangle-\frac{a}{N}\langle l^2\rangle}$
Variant D_2	$\frac{\langle l\rangle}{\langle kl\rangle-\frac{a}{N\langle k\rangle}\langle k^2l\rangle}$	$\frac{\langle k\rangle}{\langle k^2\rangle-\frac{a\langle k^3\rangle}{N\langle k\rangle}}$	$\frac{\langle l\rangle}{\langle k\rangle\langle l\rangle-\frac{a}{N}\langle k^2\rangle}$

[1] Here, the case 1 and case 2 mean the undirected network in subsection 3.3.1 and the network with an independent joint degree in subsection 3.3.2, respectively. The symbol "/" for strategy A means we do not give the current epidemic threshold since it is dependent on both the joint degree distribution and the number of immunized nodes.

3.3.1. Case 1: Undirected network

The first special case is the undirected network. This is because an undirected network can be seen as a bi-directed network or two-way network. That is, for each undirected edge between two nodes, say A–B, we can see it as two directed edges, $A \to B$ and $A \leftarrow B$.

For the undirected network, its joint degree distribution follows that

$$P(k,l) = \begin{cases} P_{in}(k) = P_{out}(l), & k = l, \\ 0, & \text{otherwise.} \end{cases} \tag{3.34}$$

The in-degree distribution $P_{in}(k)$ and out-degree distribution $P_{out}(l)$ here are the identical distribution, and are actually the traditional degree distribution. So, for simplicity, let $P(k)$ represent the $P_{in}(k)$ or $P_{out}(l)$. We should note that, in general, in-degree and out-degree in a directed network follow different degree distributions.

The immunization strategies on undirected networks have been studied in references [24–26, 28], among which there are three main strategies: Random immunization, targeted immunization, and acquaintance immunization. If we regard the undirected network as a special directed network, the calculation method in subsection 3.1 is still works, and the immunization strategies in section 3.2 are still in accord with the existing results, which will be shown below.

If the underlying network we considered is undirected, the immunization function $\alpha(k,l)$, moments $\langle kl\rangle$, and $\langle kl\alpha(k,l)\rangle$ will become $\alpha(k)$, $\langle k^2\rangle$ and $\langle k^2\alpha(k)\rangle$, respectively. Thus, the epidemic thresholds of the original network and immunized network will be

$$\lambda_c = \frac{\langle k\rangle}{\langle k^2\rangle}, \tag{3.35}$$

and

$$\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k^2 \alpha(k) \rangle}, \tag{3.36}$$

respectively.

Therefore, for random immunization, if the probability of each node being immunized is still α, namely, $\alpha(k) = \alpha$ for all degree k, the corresponding epidemic threshold and immunization threshold will be

$$\lambda_c = \frac{\langle k \rangle}{(1-\alpha)\langle k^2 \rangle}, \tag{3.37}$$

and

$$\alpha_c = 1 - \frac{\gamma \langle k \rangle}{\beta \langle k^2 \rangle}, \tag{3.38}$$

respectively.

For the optimal immunization, combined targeted immunization, and its three variants, they are equivalent to the targeted immunization in the undirected network, in which

$$\alpha(k) = \begin{cases} 1, & k \geq k_c, \\ 0, & \text{otherwise.} \end{cases} \tag{3.39}$$

Therefore, the epidemic threshold is

$$\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle - \omega_4}, \tag{3.40}$$

where, ω_4 is the integral of $k^2 P(k)$ in the interval $[k_c, k^*]$. Besides, it should be noted that given the relative infection rate of the disease and the degree distribution of the network, immunization threshold and critical value k_c are uniquely determined, although it is difficult to obtain the analytical expressions of them.

Finally, the combined acquaintance immunization and its two variants are equivalent to the acquaintance immunization in the undirected network. If we plan to immunize a nodes, the immunization function will be

$$\alpha(k) = \frac{ak}{N\langle k \rangle}. \tag{3.41}$$

Accordingly, the epidemic threshold and immunization threshold will be

$$\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle - \frac{a\langle k^3 \rangle}{N\langle k \rangle}}, \tag{3.42}$$

and

$$\alpha_c = \frac{a}{N} = \left(\langle k^2 \rangle - \frac{\gamma \langle k \rangle}{\beta} \right) \frac{\langle k \rangle}{\langle k^3 \rangle}, \tag{3.43}$$

respectively.

Finally, for convenience, the related epidemic threshold obtained in this subsection are listed in Table 1 as well.

3.3.2. Case 2: Network with independent joint degree

The second special case is directed networks with independent joint degree. In this case, the joint degree distribution is equal to the product of the corresponding in-degree distribution and the out-degree distribution, that is, formula (2.9) is held. Here, the in-degree and out-degree in the joint degree are uncorrelated, and it is opposite of the undirected network studied in the first case, where the in-degree and out-degree have a linear relationship.

For a network with independent joint degree, the epidemic thresholds of the original network and immunized network are

$$\lambda_c = \frac{\langle l \rangle}{\langle k \rangle \langle l \rangle} = \frac{1}{\langle k \rangle}, \tag{3.44}$$

and

$$\lambda_c' = \frac{\langle k \rangle}{\langle k \rangle^2 - \langle kl\alpha(k,l) \rangle}, \tag{3.45}$$

respectively. Besides, the implementation of the four families of immunization strategies studied in subsection 3.2 remains unchanged, and to reduce redundancy, the corresponding infection thresholds are also listed in Table 1.

3.3.3. Case 3: Epidemic spread with considering infectivity

The last case is the epidemic spread in a directed network with considering infectivity, and its counterpart in the undirected network has been studied in [28, 31, 32]. Considering the infectivity is based on the observation that the spread of some disease requires to consume time for two individuals to make effective contacts, such as AIDS/HIV and severe acute respiratory syndrome (SARS), and the individual with large degree are usually impossible to make effective contact with all their neighbors during unit time. That is to say, for an infected node, its neighbors being infected during unit time have a saturation effect [31]. So, we assume that for the infected with a joint degree (k, l), the effective out-degree is $C(l)$, where $0 < C(l) \leq l$. Then the model (2.16) remains unchanged, but the $\Theta(t)$, the probability that an in-edge is emitted by infected nodes, will become

$$\Theta(t) = \frac{1}{\langle l \rangle} \sum_{k,l} C(l) P(k,l) \rho_{k,l}(t). \tag{3.46}$$

In fact, the model (2.16) in subsection 2.2 with (3.46) is more general than the model (2.16) with (2.17), because if the effective out-degree $C(l)$ equals out-degree l, the former model will become the latter. Besides, to characterize the saturation effect of infectivity, different analytical expressions of $C(l)$ with upper bound can be selected. For example, the constant function in [32], the piecewise function [31], and the continuous function with an upper limit in [28]. Here we will not give the specific analytical expression since the analysis is the same no matter which function is used, which will be shown as follows.

Similar to the process from (2.18) to (2.24) in subsection 2.3, we can also obtain the condition that the endemic equilibrium exists if and only if

$$\frac{\beta \langle kC(l) \rangle}{\gamma \langle l \rangle} > 1. \tag{3.47}$$

Similarly, in this case, we also define the epidemic threshold as

$$\lambda_c = \frac{\langle l \rangle}{\langle kC(l) \rangle}. \tag{3.48}$$

Then, by continuous degree method, the joint degree distribution of immunized network is the same as (3.7), and its moment $\langle kC(l)\rangle'$ will be

$$\begin{aligned}
\langle kC(l)\rangle' &= \int_{l'_*}^{l'^*} \int_{k'_*}^{k'^*} k'C(l')P'(k',l')dk'dl', \\
&= \int_{l_*B}^{l^*B} \int_{k_*A}^{k^*A} AkC(Bl) \frac{(1-\alpha(\frac{k'}{A},\frac{l'}{B}))P(\frac{k'}{A},\frac{l'}{B})}{\int_{l_*}^{l^*}\int_{k_*}^{k^*}(1-\alpha(k,l))P(k,l)dkdl} dk'dl' \\
&= \int_{l_*}^{l^*} \int_{k_*}^{k^*} A^2BkC(Bl) \frac{(1-\alpha(k,l))}{\int_{l_*}^{l^*}\int_{k_*}^{k^*}(1-\alpha(k,l))P(k,l)dkdl} P(k,l)dkdl \\
&= \frac{A^2B}{\int_{l_*}^{l^*}\int_{k_*}^{k^*}(1-\alpha(k,l))P(k,l)dkdl} \int_{l_*}^{l^*} \int_{k_*}^{k^*} kC(Bl)(1-\alpha(k,l))P(k,l)dkdl.
\end{aligned} \tag{3.49}$$

Where, symbols A and B refer not to the immunization strategies, but to the quantities $1 - \frac{\langle l\alpha(k,l)\rangle}{\langle k\rangle}$ and $1 - \frac{\langle k\alpha(k,l)\rangle}{\langle l\rangle}$ in subsection 3.1.

Then, substituting (3.49) and (3.9) into (3.48), we obtain the epidemic threshold for the immunized network

$$\begin{aligned}
\lambda'_c &= \frac{\langle l\rangle'}{\langle kC(l)\rangle'} \\
&= \frac{B\int_{l_*}^{l^*}\int_{k_*}^{k^*} l(1-\alpha(k,l))P(k,l)dkdl}{A\int_{l_*}^{l^*}\int_{k_*}^{k^*} kC(Bl)(1-\alpha(k,l))P(k,l)dkdl} \\
&= \frac{\langle k\rangle - \langle k\alpha(k,l)\rangle}{\langle kC(Bl)\rangle - \langle kC(Bl)\alpha(k,l)\rangle}.
\end{aligned} \tag{3.50}$$

The expression of epidemic threshold in (3.50) is general, and based on this, we can obtain the corresponding epidemic thresholds of various immunization strategies. But we will not go any further here, because their calculations are similar to those for the strategies in subsection 3.2.

4. Numerical analysis

In this section, we will present the numerical results of the epidemic transmission and immunization strategies on the directed network to complement the theoretical analysis above. And these results are mainly to illustrate and contrast the efficiency of different immunization strategies.

4.1. Four types of directed networks

First, based on the homogeneity or heterogeneity of the in-degree and out-degree, directed networks can be divided into the following four types:

Type 1: Both in-degree and out-degree are homogeneous;

Type 2: In-degree is homogeneous, but out-degree is heterogeneous;

Type 3: Out-degree is homogeneous, but in-degree is heterogeneous;

Type 4: Both in-degree and out-degree are heterogeneous.

This is different from the undirected network, which can only be divided into two types: Homogeneous and heterogeneous. To carry out the subsequent simulations, we construct four underlying networks, which correspond to the four types above. These networks, in turn, are called 'ER-ER', 'ER-SF', 'SF-ER', and 'SF-SF', because the generating mechanisms of homogeneous degree and heterogeneous degree are based on the ER random graph [7] and BA model [8], respectively. The sizes of all networks are 5000, and the corresponding in-degree distribution and out-degree distribution are shown in Figure 1. Besides, none of the four networks has degree-degree correlation, which is consistent with the previous model assumptions.

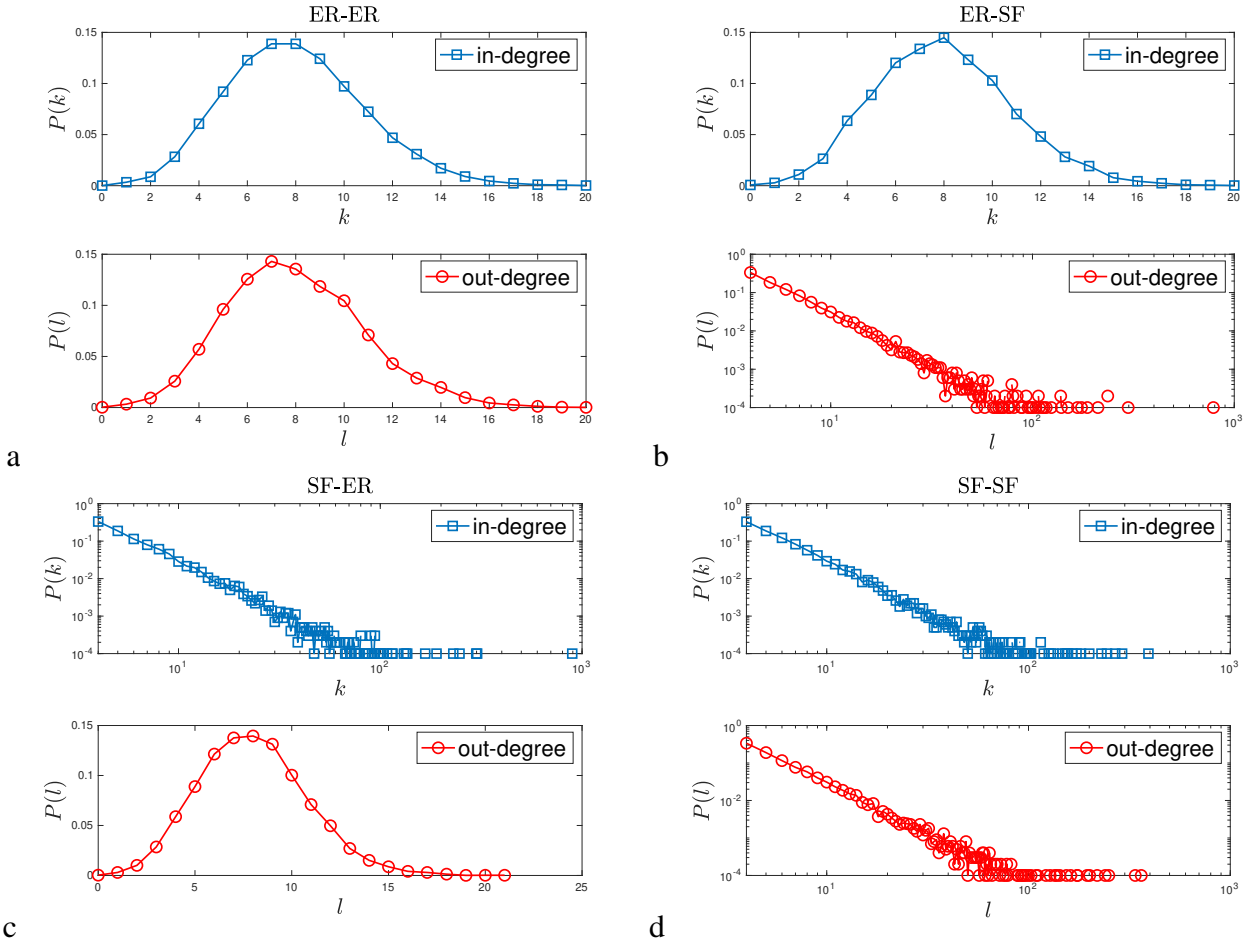

Figure 1. The in-degree distribution and out-degree distribution of the four underlying directed networks. $N = 5000$ for four networks. (a) is for 'ER-ER', where $\langle k\rangle = \langle l\rangle = 8$, $\langle kl\rangle = 64.0546$, and $\lambda_c \approx 0.1249$. (b) is for 'ER-SF', where $\langle k\rangle = \langle l\rangle = 8$, $\langle kl\rangle = 64.5226$, and $\lambda_c \approx 0.1240$. (c) is for 'SF-ER', where $\langle k\rangle = \langle l\rangle = 8$, $\langle kl\rangle = 63.9634$, and $\lambda_c \approx 0.1251$. (d) is for 'SF-SF', where $\langle k\rangle = \langle l\rangle = 7.9968$, $\langle kl\rangle = 65.0034$, and $\lambda_c \approx 0.1230$.

4.2. Infected densities over time

Based on the four networks constructed in the previous subsection, here we show the numerical results of the densities of the infected over time. In fact, there are two kinds of numerical results.

One is a numerical simulation. That is the numerical solution to the differential equation model (2.16) and (2.17), where the joint degree distribution $P(k, l)$ is calculated from the networks above. The other is stochastic simulation, that is, Monte Carlo simulation, which is performed by simulating the infection mechanism. Besides, we have selected two sets of parameters, which can be regarded as two diseases with different infection rates β and recovery rates γ. One is $\beta = 0.01$, $\gamma = 0.1$, and its relative infection rate is 0.1. Another is $\beta = 0.05$, $\gamma = 0.1$, and its relative infection rate is 0.5. But both the initial densities of infected are set to 0.1.

Figure 2 shows the changes in infected densities over time t for four networks under two sets of parameters. Each curve in the graph is the average of 50 stochastic simulations. We can see that for the same set of parameters, four different networks have the same changing trend, where disease tends to die out in Figure 2a and the disease tends to form endemic disease in Figure 2b. This is because the epidemic thresholds of four networks are larger than the relative infection rate in Figure 2a and smaller than that of in Figure 2b, respectively. However, for the same parameters, when the disease breaks out different networks can have different prevalence, namely, the infected densities at stationary state, because they can have different joint degree distributions. Interestingly, 'ER-ER' and 'ER-SF', as well as 'SF-ER' and 'SF-SF', have almost the same prevalence. This is probably because they have the same in-degree distribution, and the node is only infected through in-degree so that the probabilities both of nodes being infected and recovery are the same.

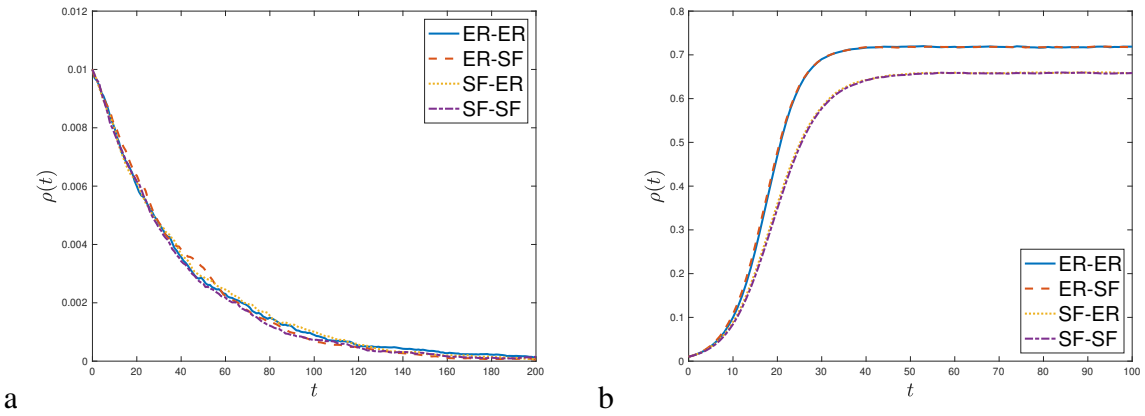

Figure 2. The change of infected densities over time for the four networks. Where, all the initial infected densities are set to 0.1, and each curve is the mean of 50 stochastic simulations. (a) is for the first set of parameters, where $\beta = 0.01$, $\gamma = 0.1$. (b) is for the second set of parameters, where $\beta = 0.05$, $\gamma = 0.1$.

Figure 3 is the comparison between the numerical simulations and the stochastic simulations of the infected densities over time. Here, the zigzag solid lines are the results of single stochastic simulations, and the dash-dot lines are the mean of 50 stochastic simulations, which are the same as that in Figure 2. The dashed lines are the corresponding numerical simulations. We can see that for these four networks, whether under the first set of parameters or the second set of parameters, the numerical simulations are in good agreement with the results of stochastic simulations, especially when the time is relatively large. In other words, in terms of the prevalence, the numerical simulation is consistent with the stochastic simulations. On the one hand, it shows that our model is reasonable. On the other hand,

it suggests that we can use numerical simulation instead of stochastic simulation to further study the model.

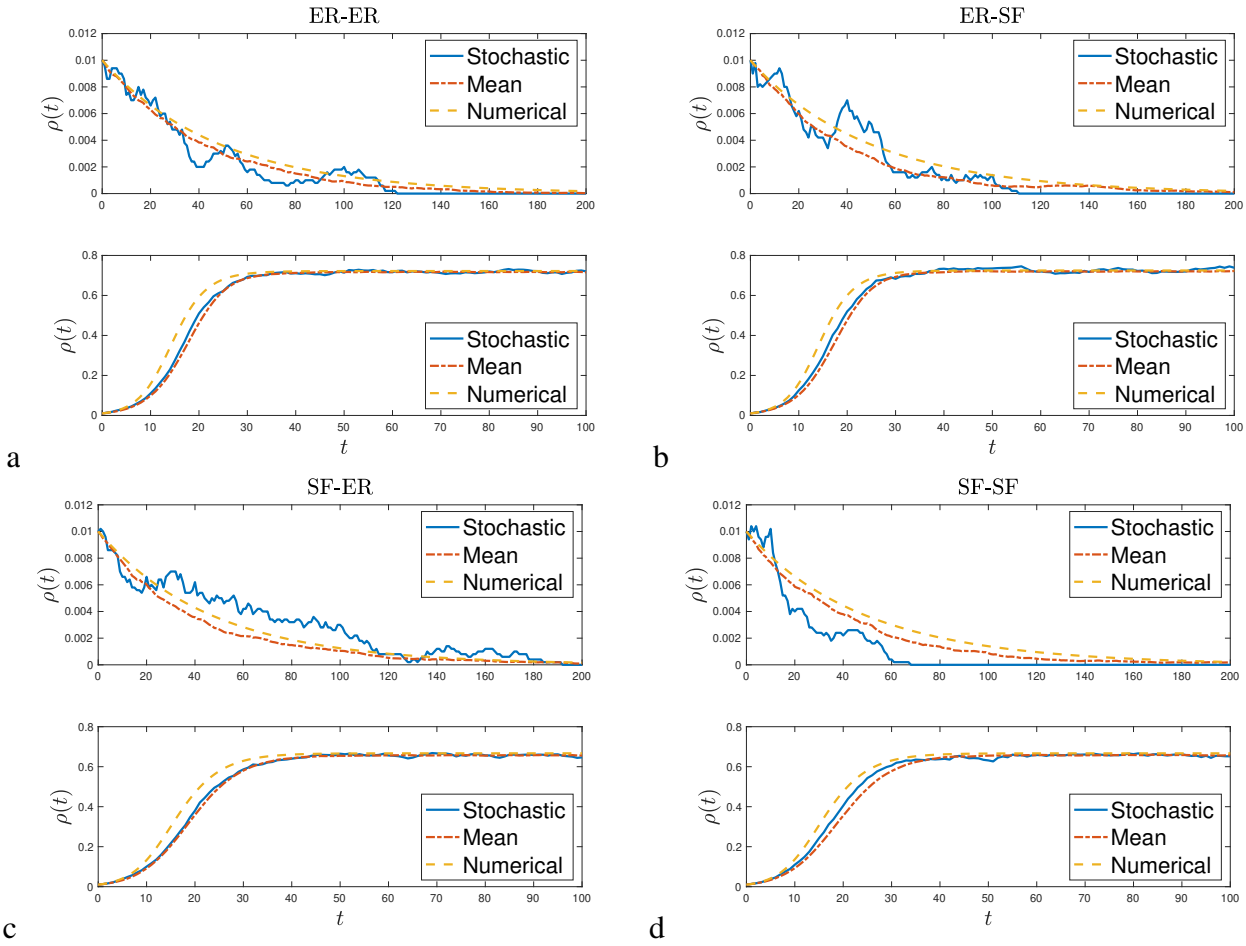

Figure 3. The change of infected densities over time for the four networks. Where, solid lines are the single stochastic simulations, and the dash-dot lines are the mean of 50 stochastic simulations. The dashed lines are the corresponding numerical simulations. All the initial infected densities are set to 0.1. (a), (b), (c) and (d) correspond to four networks, respectively, and each contains simulation results for both sets of parameters.

4.3. Comparison of immunization strategies

In this subsection, we will use numerical simulations to compare the efficiencies of different immunization strategies analyzed in subsection 3.2, including optimal immunization, random immunization, combined targeted immunization, combined acquaintance immunization and corresponding variants. For simplicity, we also call them strategy A, B, C, D, C_1, C_2, C_3, D_1, and D_2.

Figure 4 shows the change of infected densities over time on the four networks immunized by four strategies, A, B, C and D. Here, all the proportions of immunized nodes are 0.2, and the initial infected densities are set to 0.1. Besides, we use the second set of parameters for simulations, that is, $\beta = 0.05$, $\gamma = 0.1$. For convenience, let $k_c = l_c$ for combined targeted immunization. From Figure 4,

we can see that the four different immunization strategies are all effective because the prevalence after immunized are less than that of without immunization. Furthermore, no matter which network, the optimal immunization strategy is the best, followed by the combined targeted immunization, combined acquaintance immunization, and random immunization. For network 'SF-SF' in Figure 4d, where the in-degree and out-degree of nodes are heterogeneous, the effects of the four immunization strategies are obviously different. However, for 'ER-ER' in Figure 4a, where both the in-degree and out-degree are homogeneous, the differences in effects are not obvious, and the effects of strategies A and C, as well as B and D, are almost the same. And 'ER-SF' and 'SF-ER' are similar to 'SF-SF' and 'ER-ER', respectively.

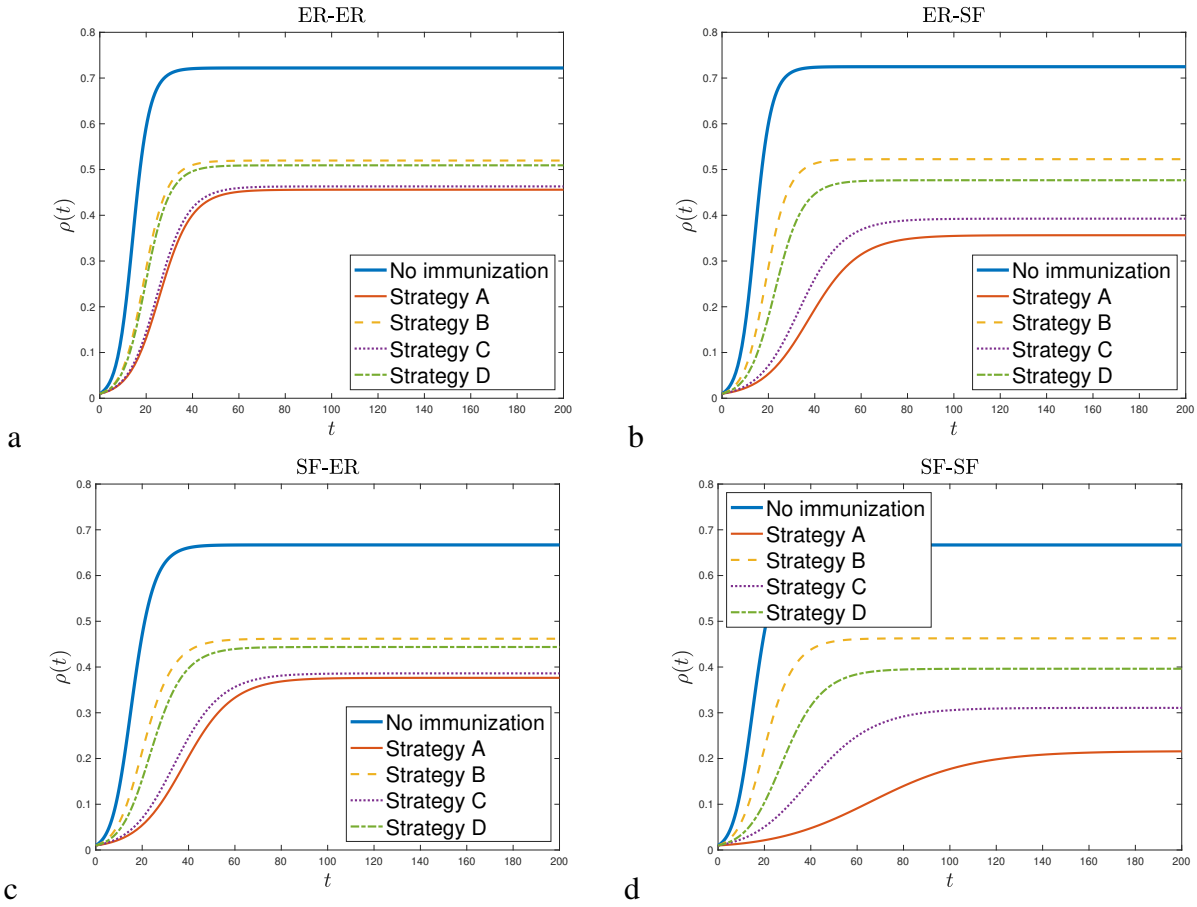

Figure 4. The change of infected densities over time for the four networks immunized by four strategies. Strategies A, B, C, and D are the optimal immunization, random immunization, combined targeted immunization, and combined acquaintance immunization, respectively. Here, all the proportions of immunized nodes are 0.2, and the initial infected densities are set to 0.1. $\beta = 0.05$, $\gamma = 0.1$. For combined targeted immunization, we let k_c equal l_c.

Figure 5 shows the relationship between the reduced prevalence, ρ_α/ρ_0, and the immunization proportion α. Here, ρ_α is the prevalence in the steady-state for the network with immunization proportion α. We can see that for each α and each network, the four strategies are A, C, D, and B in order of their immunization efficiency, which is consistent with the result in Figure 4.

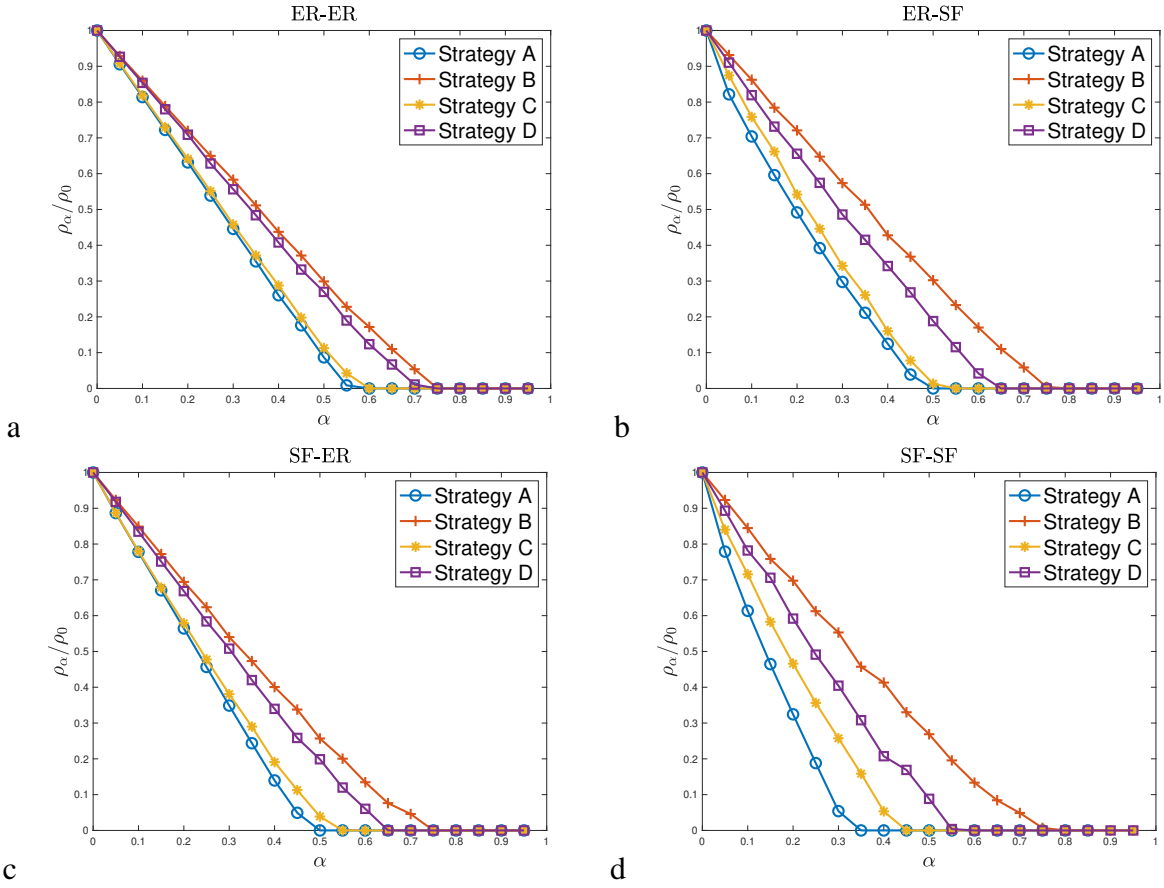

Figure 5. Reduced prevalence, ρ_α/ρ_0, with immunization proportion α. Where, ρ_α and ρ_0 represent the prevalences with immunization proportion α and without immunization, respectively. Strategies A, B, C, and D are the optimal immunization, random immunization, combined targeted immunization, and combined acquaintance immunization, respectively. Here, the initial infected densities are set to 0.1. $\beta = 0.05$, $\gamma = 0.1$. For combined targeted immunization, we let k_c equal l_c.

To compare the efficiency between the combined targeted immunization and its three variants, we also performed numerical simulations, as shown in Figure 6. Four solid lines, in top-to-bottom order, are the results of no immunization, random immunization, combined targeted immunization, and optimal immunization, which are shown as references. The dashed lines, dash-dot lines, and dotted lines are variant C_1, C_2, and C_3, respectively. We can see that strategy C and its variants, C_1, C_2, and C_3, do not always have a similar performance and efficiency ranking in four networks. In Figure 6a,c, there is no obvious difference between the four strategies. However, in Figure 6b,d, the efficiencies are different, and the variants C_1 and C_3 are better than C_2. Therefore, we can conclude that, generally, the nodes with large out-degree are more worthy of being immunized than the nodes with large in-degree. So, when both in-degree and out-degree are heterogeneous, strategy C_1 is the best, followed by C_3. And strategy C_3 is the best, followed by C_1 when only the out-degree is heterogeneous. In both cases, strategy C is also good, which is ranked behind C_1 and C_3, and strategy C_2 is the worst.

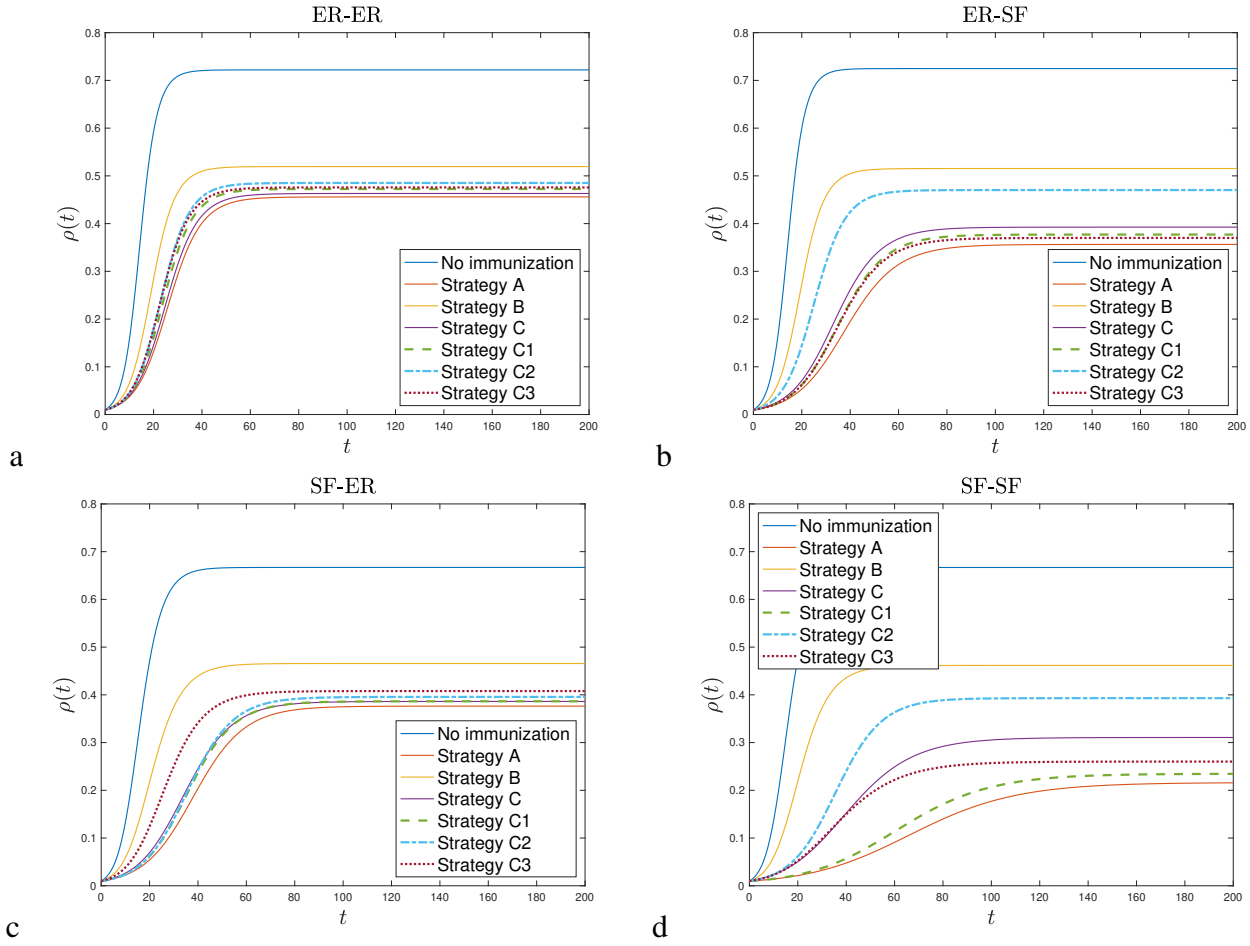

Figure 6. The prevalence of four networks immunized by combined targeted immunization, C and its three variants, C_1, C_2 and C_1. Here, all the proportions of immunized nodes are 0.2, and the initial infected densities are set to 0.1. $\beta = 0.05$, $\gamma = 0.1$. For these four strategies, we let $k_c = l_c$, as before.

Besides, we also compared the family of the combined acquaintance immunization, that is, D and its variants D_1 and D_2, as shown in Figure 7. The four solid lines in figures are for references, which are corresponding to no immunization, random immunization, combined acquaintance immunization, and optimal immunization from the top down. Dashed lines and dash-dot lines are corresponding to variant D_1 and D_2, respectively. In Figure 7, we can find a very similar phenomenon with that of in Figure 6, that is, in 'ER-ER' and 'SF-ER', the effects of D, D_1, and D_2 are similar, and their effects are different in 'ER-SF' and 'SF-SF'. Therefore, we can get the same conclusion as that in Figure 6, that is, generally, the nodes with large out-degree more important than the nodes with large in-degree. We should note that in strategy D_1, the probability of a node being immunized is proportional to its out-degree, although the node is looking for its neighbor though in-edges. Similarly, in D_2, the probability of a node being immunized is proportional to its in-degree. So, if the out-degree is heterogeneous, strategy D_1 will be better than D_2, as shown in Figure 7b,d. Otherwise, the effects of D_1 and D_2 will be similar, as shown in Figure 7a,c.

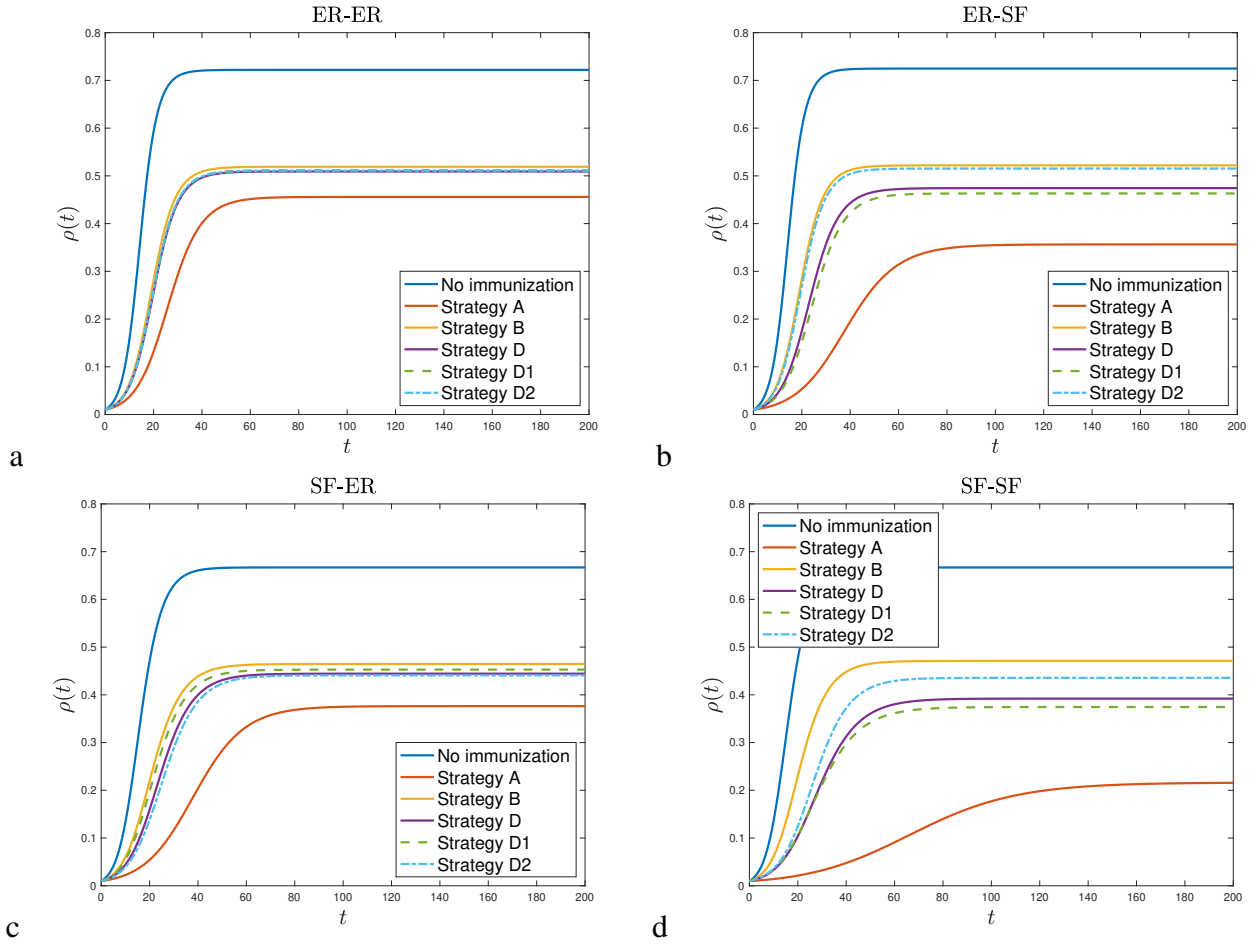

Figure 7. The prevalence of four networks immunized by combined acquaintance immunization, D and its two variants, D_1 and D_2. Here, all the proportions of immunized nodes are 0.2, and the initial infected densities are set to 0.1. $\beta = 0.05$, $\gamma = 0.1$.

Table 2. The reduced prevalence of all immunization strategies[1].

	ER-ER		ER-SF		SF-ER		SF-SF	
Ranking	ρ/ρ_0	Strategies	ρ/ρ_0	Strategies	ρ/ρ_0	Strategies	ρ/ρ_0	Strategies
	1	No	1	No	1	No	1	No
9	0.719	B	0.713	B	0.693	B	0.710	B
8	0.708	D_1	0.698	D_2	0.678	D_1	0.668	D_2
7	0.706	D	0.657	D	0.669	D	0.589	C_2
6	0.703	D_2	0.649	C_2	0.664	D_2	0.587	D_1
5	0.672	C_2	0.625	D_1	0.612	C_3	0.585	D
4	0.659	C_3	0.542	C	0.593	C_2	0.466	C
3	0.654	C_1	0.520	C_1	0.580	C_1	0.390	C_3
2	0.641	C	0.510	C_3	0.579	C	0.352	C_1
1	0.631	A	0.492	A	0.564	A	0.323	A

[1] Here, all the proportions of immunized nodes are 0.2, and the initial infected densities are set to 0.1. $\beta = 0.05$, $\gamma = 0.1$.

Finally, the prevalence ratios of the four networks are compared, including no immunization and nine immunization strategies, as shown in Table 2. For the sake of comparison, we let $k_c = l_c$. For the four networks, optimal immunization, A, and random immunization, B, are the best and the worst, respectively, and the rest of the immunization strategies are in between. If we ignore C_2, the family of combined targeted immunization is more efficient than the family of combined acquaintance immunization. These are easy to understand because the optimal immunization and the family of combined targeted immunization require the global information of nodes, and these strategies are deterministic. However, the family of combined acquaintance immunization only needs to know the local information, and even the random immunization does not need to know the information of the nodes, in which case they have certain randomness.

5. Conclusions and discussions

In this paper, the epidemic transmission and immunization strategies on directed networks are considered. There are two main reasons for the study of directed networks: First, a directed network can be regarded as a generalization of an undirected network. Second, some facts mentioned in the introduction show that some underlying networks of disease transmission are directed and the transmission between two nodes may be asymmetric. First of all, an SIS model on a directed network is established by using the mean-field method, and the conditions for disease outbreaks are obtained. That is, determine whether the epidemic threshold of a network, λ_c, is larger than the relative infection rate of the disease, $\frac{\beta}{\gamma}$. If $\lambda_c \geq \frac{\beta}{\gamma}$, the disease will die out. Otherwise, the disease will break out, leading to an endemic. Next, we provid the calculation method of the epidemic threshold of the immunized network by the continuous degree method by considering the degree of the node as a continuous variable. Then, four kinds of immunization strategies are studied, including optimal immunization, random immunization, the family of combined targeted immunization, and the family of combined acquaintance immunization. Besides, several special cases are also considered, including undirected networks, networks with independent joint degree, and epidemic spread with considered infectivity. Finally, numerical analysis is carried out to perform and complement the theoretical analysis.

Through the continuous degree method, we can obtain the epidemic threshold of a directed network immunized by various strategies. Generally, the epidemic threshold depends on the related immunization parameters and the joint degree distribution of the network. Besides, the immunization efficiencies of different strategies are compared through numerical simulation. We find that optimal immunization and random immunization are the best and the worst, which is independent of the degree distribution. If we do not consider the variant C_2, the family of combined targeted immunization is generally more efficient than that of combined acquaintance immunization. This is because the former needs to know the global information of the nodes and whether a node is immunized is determined. However, the latter only needs local information and whether a node is immunized is random. Finally, we also find that the node with the largest product value of in-degree and out-degree, kl, is the most worthy of being immunized. Secondly, relatively speaking, the nodes with the large out-degree is more worthy of being immunized than the nodes with large in-degree. That is why the optimal immunization is the best, and the strategies that tend to immunize nodes with large out-degree are usually better than the strategies that are related to in-degree.

Nevertheless, the design and implementation of efficient immunization strategies remain a

challenging topic [33]. The efficiency of strategy closely depends on the topological structure of the network, and the strategies studied in this paper are based on the assumption of mean-field where all nodes are uniformly mixed and there is no correlation between nodes. However, this may not be the case in the real system, which can have some special properties, such as assortativity and disassortativity [34], and some topological structures, such as star-shaped network, community structure [12, 35], hierarchical structure, and clusters, and even can be dynamic [17], that is, there are births and deaths of nodes and edges.

In addition, immunization strategies are also closely related to the mechanism of disease transmission. For convenience of study, we usually assume that the spread and infection of the disease are linear with out-degree and in-degree, respectively, which is sometimes inconsistent with reality. Although we have considered the infectivity of nodes, which can be used to characterize the saturation of the infection power of infected, the real infectivity function is still unknown for the specific disease. Therefore, both the particular structure of the network and the infectivity of the nodes are crucial in determining whether a node deserves to be immunized.

Acknowledgements

This work was jointly supported by the NSFC under grants 11572181 and 11331009. The authors would like to thank Dr. Mengfeng Sun for his helpful discussions and suggestions, and two anonymous Reviewers for their helpful comments on the previous version of the manuscript.

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. M. Newman, *Networks*, Oxford University Press, 2018.
2. M. Faloutsos, P. Faloutsos, C. Faloutsos, On power-law relationships of the internet topology, *Comput. Commun. Rev.*, **29** (1999), 251–262.
3. G. Kossinets, D. J. Watts, Empirical analysis of an evolving social network, *Science*, **311** (1999), 88–90.
4. D. J. Watts, S. H. Strogatz, Collective dynamics of small-world networks, *Nature*, **393** (1998), 440–442.
5. J. Scott, Social network analysis, *Sociology*, **22** (1988), 109–127.
6. E. H. Davidson, D. H. Erwin, Gene regulatory networks and the evolution of animal body plans, *Science*, **311** (2006), 796–800.
7. P. Erdős, A. Rényi, On the evolution of random graphs, *Publ. Math. Inst. Hung. Acad. Sci.*, **5** (1960), 17–60.
8. A. L. Barabási, R. Albert, Emergence of scaling in random networks, *Science*, **286** (1999), 509–512.

9. M. E. J. Newman, M. Girvan, Finding and evaluating community structure in networks, *Phys. Rev. E*, **69** (2004), 026113.
10. A. Clauset, M. E. J. Newman, C. Moore, Finding community structure in very large networks, *Phys. Rev. E*, **70** (2004), 066111.
11. S. Fortunato, Community detection in graphs, *Phys. Rep.*, **486** (2010), 75–174.
12. H. Cherifi, G. Palla, B. K. Szymanski, X. Lu, On community structure in complex networks: Challenges and opportunities, *Appl. Network Sci.*, **4** (2019), 1–35.
13. A. Clauset, C. Moore, M. E. J. Newman, Hierarchical structure and the prediction of missing links in networks, *Nature*, **453** (2008), 98–101.
14. R. Pastor-Satorras, A. Vespignani, Epidemic spreading in scale-free networks, *Phys. Rev. Lett.*, **86** (2001), 3200.
15. X. Fu, M. Small, G. Chen, *Propagation Dynamics on Complex Networks: Models, Methods and Stability Analysis*, John Wiley and Sons, 2013.
16. M. Jalili, M. Perc, Information cascades in complex networks, *J. Complex Networks*, **5** (2017), 665–693.
17. M. Nadini, K. Sun, E. Ubaldi, M. Starnini, A. Rizzo, N. Perra, Epidemic spreading in modular time-varying networks, *Sci. Rep.*, **8** (2018), 1–11.
18. J. Jia, Z. Jin, X. Fu, Epidemic spread in directed interconnected networks, *Commun. Nonlinear Sci. Numeri. Simul.*, **75** (2019), 1–13.
19. C. Pizzuti, A. Socievole, Epidemic spreading curing strategy over directed networks, in *International Conference on Numerical Computations: Theory and Algorithms*, Springer, Cham, (2020), 182–194.
20. L. Meyers, M. Newman, B. Pourbohloul, Predicting epidemics on directed contact networks, *J. Theor. Biol.*, **240** (2006), 400–418.
21. X. Zhang, G. Sun, Y. Zhu, J. Ma, Z. Jin, Epidemic dynamics on semi-directed complex networks, *Math. Biosci.*, **246** (2013), 242–251.
22. M. Moslonka-Lefebvre, T. Harwood, M. J. Jeger, M. Pautasso, SIS along a continuum (SISc) epidemiological modelling and control of diseases on directed trade networks, *Math. Biosci.*, **236** (2012), 44–52.
23. D. H. Zanette, M. Kuperman, Effects of immunization in small-world epidemics, *Phys. A: Stat. Mech. Appl.*, **309** (2002), 445–452.
24. R. Pastor-Satorras, A. Vespignani, Immunization of complex networks, *Phys. Rev. E*, **65** (2002), 036104.
25. N. Madar, T. Kalisky, R. Cohen, D. Ben-avraham, S. Havlin, Immunization and epidemic dynamics in complex networks, *Eur. Phys. J. B*, **38** (2004), 269–276.
26. R. Cohen, S. Havlin, D. Ben-Avraham, Efficient immunization strategies for computer networks and populations, *Phys. Rev. Lett.*, **91** (2003), 247901.

27. D. Chakraborty, A. Singh, H. Cherifi, Immunization strategies based on the overlapping nodes in networks with community structure, in *International Conference on Computational Social Networks*, Springer, Cham, (2016), 62–73.
28. 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** (2008), 036113.
29. Y. Yang, A. McKhann, S. Chen, G. Harling, J. P. Onnela, Efficient vaccination strategies for epidemic control using network information, *Epidemics*, **27** (2019), 115–122.
30. J. Jia, Z. Jin, L. Chang, X. Fu, Structural calculations and propagation modeling of growing networks based on continuous degree, *Math. Biosci. Eng.*, **14** (2017), 1215–1232.
31. J. Joo, J. L. Lebowitz, Behavior of susceptible-infected-susceptible epidemics on heterogeneous networks with saturation, *Phys. Rev. E*, **69** (2004), 066105.
32. T. Zhou, J. Liu, W. Bai, G. Chen, B. Wang, Behaviors of susceptible-infected epidemics on scale-free networks with identical infectivity, *Phys. Rev. E*, **74** (2006), 056109.
33. C. Nowzari, V. M. Preciado, G. J. Pappas, Analysis and control of epidemics: A survey of spreading processes on complex networks, *IEEE Control Syst. Mag.*, **36** (2016), 26–46.
34. R. Noldus, P. Van Mieghem, Assortativity in complex networks, *J. Complex Networks*, **3** (2015), 507–542.
35. N. Gupta, A. Singh, H. Cherifi, Community-based immunization strategies for epidemic control, in *2015 7th International Conference on Communication Systems and Networks (COMSNETS)*, IEEE, 2015.

AIMS Press

© 2020 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0)