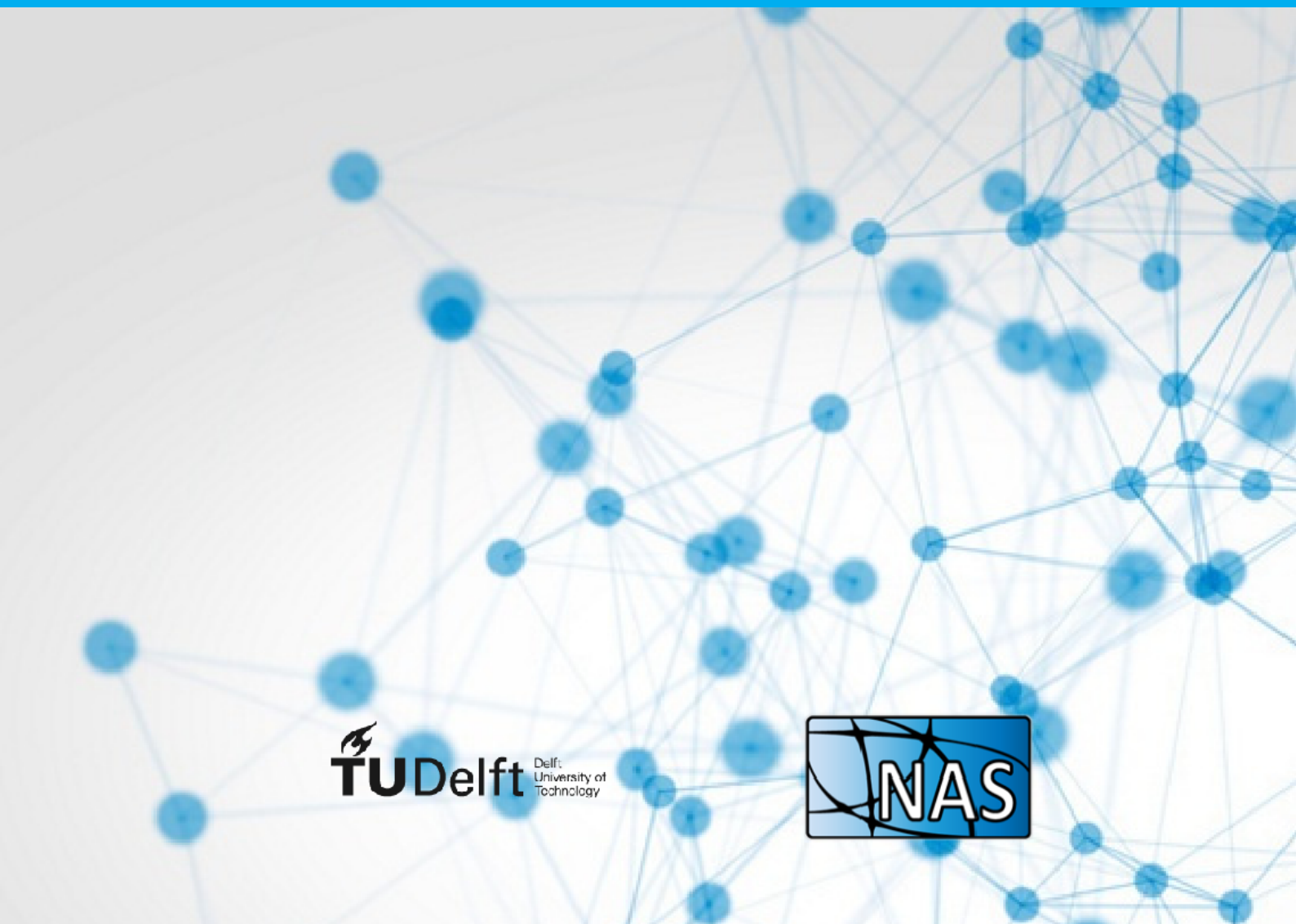
SIS Epidemics on Network with Non-Markovian Curing Process

X. Zhou

PVM2019-105



SIS Epidemics on Network with Non-Markovian Curing Process

by

X. Zhou

to obtain the degree of Master of Science
in Electrical Engineering
Track Telecommunications and Sensing Systems

at the Delft University of Technology,
to be defended publicly on Thursday December 19, 2019 at 9:00 AM.

Student number:	4646789	
Project duration:	February 25, 2019 – December 19, 2019	
Thesis committee:	Prof. dr. ir. Piet Van Mieghem,	TU Delft, supervisor
	Dr. Pasquale Cirillo,	TU Delft
	Dr. Alessandra Cipriani,	TU Delft
	Long Ma,	TU Delft, daily supervisor

Abstract

Susceptible-Infected-Susceptible (SIS) model is commonly used to describe the spreading of virus on networks. However, a real-life epidemic process is not necessarily Markovian. The spreading of diseases, behaviours and information in real systems are sometimes dependent on the characteristics and current status of individuals. Thus it is far from enough to just consider Markovian processes. We need to consider a more general model with non-Markovian processes. Although some recent works focus on the SIS model with a non-Markovian infection process, systematic research on the non-Markovian curing process is still lacking. Therefore, this thesis project is to study the influence of the non-Markovian curing process on the performance of SIS viral spreading on networks.

Through continuous-time SIS epidemics simulator, we find some dramatic effects of a non-exponential curing time (while still assuming an exponential infection time) on the prevalence and critical point of effective infection rate by considering Weibullean curing times with same mean, but different shape parameter α. For $\alpha \in [0.2, 10]$, the epidemic threshold satisfies $\tau_c = \frac{1}{\lambda_1}$, which is the same as the NIMFA conclusions of Markovian SIS process. Relatively, when α is too small, a large number of curing events synchronously happened at the beginning of the simulation, which will lead to collective deaths on finite network. The effect on initial condition of nodes further cause a decline on prevalence and an slow phase transition between healthy state and the metastable state. Furthermore, the heavy-tailed distribution of curing time leads to a small percent of nodes still surviving at the metastable state, even under a very low effective infection rate. The heavy-tailed distribution gives some nodes an extreme long curing time and thus can infect other nodes with a pretty small probability, thereby maintaining the virus' long-term spread in a small group of nodes. This spreading mode seems can explain some virus spreading phenomenon, like the spreading mode of hepatitis B virus (HBV). Additionally, when the shape parameter α of Weibull distribution is pretty large, the distribution of curing time is like a pulse or a Dirac delta function (δ function), thus a huge amount of nodes can get synchronously recovered. We find when we control the successful curing probability $p = 1 - 1/\mathrm{e} \approx 0.632$, the prevalence of pulse curing at the metastable state is equivalent to a Poisson curing process. Therefore, the pulse curing strategy can suppress the spreading of viruses and further save medical resources.

Acknowledgements

Before you lies the thesis "SIS Epidemics on Network with Non-Markovian Curing Process". This thesis has been written to obtain the degree Master of Science at the Delft University of Technology (TU Delft). The project was conducted at the Network Architectures and Services (NAS) group, within the faculty of Electrical Engineering, Mathematics and Computer Science at Delft University of Technology.

First of all, I would like to express my gratitude to Professor Piet Van Mieghem. Thank you for your help and giving me this precious opportunity to do my master thesis project at NAS group. I would also like to thank you for introducing me to take the extra project on brain networks at VU university medical center during last summer. I have benefited a lot from my studies and project researches. Moreover, my deepest gratitude goes first and foremost to Long Ma and Qiang Liu, my daily supervisors, for their constant encouragement and guidance. They taught me how to rigorously carry out project research and comprehend issues from various perspective. More importantly, I'm appreciated for their care and kind help to me, which makes my research more enjoyable. Besides, I would like to thank everyone at NAS group, for the discussions and joy we had during lunches and Friday afternoon drinks. Thanks Dr. Pasquale Cirillo and Dr. Alessandra Cipriani for being part of my thesis committee.

I am grateful to my friends I have met since the first day I came to delft, and thank them for always being with me! Last of all, I want to express my gratitude to my families, for all their support and belief in me!

Xiaoyu Zhou
Delft, December 2019

Contents

1 Introduction

1.1. Background

The propagation of diseases and information are rarely independent of each other in real-world networks [1]. In order to effectively explain the process of disease spreading, mathematical epidemic models have been widely investigated for a long period [2, 3]. The classical SIR epidemic model was proposed in 1927 [4]. And after that, there developed a huge number of publications on different epidemic dynamics. The Susceptible-Infected-Susceptible (SIS) model is one of the basis models in epidemiology [5], which is a first-order description of the spread of real epidemics. The range of applications is broad, including biological virus spread [6, 7], malware spread [8, 9], financial network contagion [10] and so on.

In general viral spreading epidemic models, various people are assumed to be homogeneously mixed, which means that susceptible people are infected with the same rate [11]. The Markovian SIS model is described using exponential waiting times for both infection and curing events [12]. However, a real-life epidemic process is not necessarily Markovian. Especially the current state of individuals can be affected by its characteristics and past experience. For instance, when a disease spreads among the people, recovered people often have strong resistance to the disease, so they are not easy to get re-infection or have the ability to get recovery more faster than others after re-infection. Moreover, different physical conditions of each individual also make differences in recovery. Thus it is far from enough to just consider Markovian processes. In order to model general

non-Markovian epidemic process, we consider the curing process as a renewal process and the waiting time interval following a more general distribution.

1.2. Motivation

Although some recent works focus on the SIS model with a non-Markovian infection process, systematic research on the non-Markovian curing process is still lacking. Therefore, this thesis project is to study the influence of the non-Markovian curing process on the performance of SIS viral spreading on networks. Generally, this project is to theoretically and numerically study the SIS epidemics with non-Markovian curing process. The motivation is to observe the influence of non-Markovian curing process on the performance of SIS epidemics on networks. The study focuses on the dependence of the epidemic threshold on the distribution of curing time and the applicability of N-Intertwined Mean-Field Approximation (NIMFA) results on various networks.

1.3. Contribution

In this project, we first simulate the SIS epidemics with the non-Markovian curing process on various networks. We assume curing time following a Weibull distribution or gamma distribution. Comparing the performance with previous theoretical conclusion, we find dramatic discrepancies between simulation results and theoretical conclusion.

Specifically, we first find that when the shape parameter of Weibull distribution is close to zero or larger than 10, the curing time have synchronicity, which means a high percentage of curing event trends to happen within a very short time. Especially when shape parameter is pretty large, the curing time follows a pulse-like distribution. That will cause a huge percentage of collective death and further affect the initial state of entire processes. Thus the discrepancies of prevalence are resulting from that reason. Inspired by the influence of pulse-like distribution of curing time, a kind of pulse curing strategy can be applied to suppress the spreading on networks to save cure resources in some extents. Furthermore, we can make use of the theoretical results of the effect of pulse strategy on the epidemic threshold to find the optimal curing plan for some existing viral spreading situations.

In addition to that, when shape parameter α is pretty small, the heavy-tailed distribution of curing time can causes a small percentage of infected nodes survive at the

metastable state even with a low effective infection rate. This special case describes the process in which some individuals obtained a long curing time and keep trying to infect neighbors within the infected period, resulting in the disease slowly and continuously spread between small group of nodes at the metastable state. The propagation pattern of hepatitis B virus (HBV) [13] is similar to this case.

1.4. Thesis Outlines

The Thesis will be described by 5 chapters and the remaining chapters of the thesis contain the following:

- Chapter 2 first provides the introduction to complex networks and SIS viral spreading model. This chapter next covers literature review on previous research of SIS epidemics both with Markovian process and further non-Markovian spreading process.
- Chapter 3 first gives an short introduction of the continuous-time SIS epidemics simulator and further describes the simulation procedures we designed to analyze the effects of non-Markovian curing processes on the prevalence and the critical point of phase transition. Eventually, we find some dramatic effects of non-Markovian processes when the distributions of curing time R are different. We present the influence of pulse-like distribution of curing time on the initial condition leading to some extents of collective deaths on finite network. That condition will finally affect the prevalence at the metastable state. In addition to that, we further discuss the effects of heavy-tailed distribution of curing time R and find a special spreading phase caused by this reason.
- Chapter 4 investigates the pulse curing strategy. The theoretical analysis of pulse curing effects are given through the theoretical deviation process. The impacts of pulse curing strategy are demonstrated through simulations. Besides, the suggestions of using pulse strategy will be shown at last.
- Chapter 5 gives a conclusion of the whole thesis project and some prospects for future research on non-Markovian SIS epidemics on networks.

2

The SIS Epidemic Spreading Model on Networks

In epidemiology, the Susceptible-Infected-Susceptible (SIS) is a typical model, which starts with infected items infecting their healthy neighbors. Compared with the discrete-time SIS model, the SIS model in continuous-time can describes the realistic epidemics process better [14], e.g. the flu spreading among people [15] and malware spreading process in computer networks [16]. One simple but practically meaningful method to study the effect of the network topology on the process that runs over the topology is applying the continuous-time SIS model epidemic process on graphs. Therefore, this chapter will first introduce two kinds of complex graphs based on their characteristic topology and a matrix that facilitates embodying topological properties. Furthermore, this chapter mainly focuses on the epidemic spreading process of continuous-time SIS process and the N-Intertwined Mean-Field Approximation (NIMFA), which is proposed by P. Van Mieghem [17]. However, the classical continuous-time SIS Markov model is a special case where both infection and cure process are exponentially distributed, which cannot describe real-life epidemics well. The second part of this chapter will concentrate on previous results on non-Markovian SIS process and point out the main objective of this thesis.

2.1. Complex Networks

Network science studies networks among different fields with various complexities and characteristics, such as telecommunication networks [18, 19], computer networks [20], social networks [21], biological networks [22] and so on. Graph can model pairwise relations between objects. In order to simulate the process of information spreading on networks and observe the performance of simulation. In this thesis, we consider information spread in an undirected graph G($\mathcal{N}$, $\mathcal{L}$), where $\mathcal{N}$ is the set with N nodes and $\mathcal{L}$ the set with L links. There is no limit to the direction of information spreading between random pair of nodes, but only one direction can be transmitted at a time. The topological characteristics of the graph G can be described by the adjacency matrix A.

2.1.1. Adjacency Matrix A

The adjacency matrix [23], which is a $N \times N$ matrix, describes adjacency between various nodes and contains the basic topological properties of the underlying network G. If there is a link between node i and node j, then the adjacency parameter a_{ij}= a_{ji}= 1. Conversely, there is no link between node i and node j, then the adjacency parameter a_{ij}= a_{ji}= 0. Therefore, adjacency matrix is a symmetric matrix $A^T = A$. It is assumed further in this paper that the graph G does not contain self-loops ($a_{ii} = 0$) nor multiple links between two nodes. The adjacency matrix A has N eigenvalues with $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq ... \geq \lambda_N$, and according to the Gerschgorin's theorem: $0 \leq \lambda_1 \leq d_{max}$. As another important graph metric, the degree d_i is the number of neighbors of node i, which is d_i=$\sum_{j=1}^{N} a_{ij}$.

2.1.2. Graph Models

Networks are mathematically described as graphs. According to the different characteristics of the network topology, the simple graphs can be divided into different classes. The following five kinds of network models are mainly involved in this thesis.

1. Deterministic Graph Models

Complete Graph: The complete graph K_N is the mother of all kinds of graphs with same number of nodes [24]. In the complete graph, each node is connected to all other nodes and the number of links $L = N(N-1)/2$.

Regular Graph: Each node in regular graph has the same degree $d_i = \overline{d} = k$ for i=1,2,...,N. The complete graph is a special case of regular graph.

Star Graph: A star graph S_k has $N = k + 1$ nodes and the number of links $L = k$. It has a center node connected with all the other nodes and degree is k. And other nodes only connect with the center node with one link for each.

2. Complex Graph Models

Erdős-Rényi Random Graph (ER): Random graphs are introduced by Erdős and Rényi [25] and they showed many monotone-increasing properties of random graphs. The Erdős-Rényi (ER) random graph $G_p(N)$ has N nodes in total and each node pair is connected independently with probability p. And any so generated graphs (i.e. any realizations) belongs to the class of ER random graph $G_p(N)$ with same N and p. Thus, a_{ij} is a Bernoulli random variable with mean p, with the probability of $a_{ij} = 1$ is p and the probability of $a_{ij} = 0$ is $1 - p$. The degree distribution of graph G is following the binomial distribution and the average number of links is $L = \frac{N(N-1)}{2}p$. The eigenvalues is determined by N and p. And If $p = 1$, then $G_1(N)$ is the complete graph K_N.

Scale-Free Graph of Barabási-Albert (BA) [26]: The scale-free network model is a network whose degree distribution follows a power law, at least asymptotically. It has a typical characteristics that most nodes in the network are connected to a few hubs, and those hubs are connected to a large number of nodes. Many realistic network have such scale-free features, such as the Internet, financial system networks, social networks, and so on [27].

2.2. The SIS Epidemic Model

In SIS (Susceptible-Infected-Susceptible) model, each node in this dynamical model can be in two states, infected and healthy. There are two state transition processes existing in this model: the infection process and the curing process. In classical continuous-time SIS Model, we assume the infection process per link is a Poisson process with infection rate β and the curing process per node i is a Poisson process with cure rate δ. The effective infection rate is $\tau = \frac{\beta}{\delta}$. Only when a node is infected, can it infects its direct healthy neighbors and recovered nodes can be infected again. For example, diseases such as gonorrhea or chlamydia fall into this group. All curing and infection Poisson processes are independent. Figure 2.1 shows that how individuals move through each state with SIS epidemic model in the reaction-diffusion process.

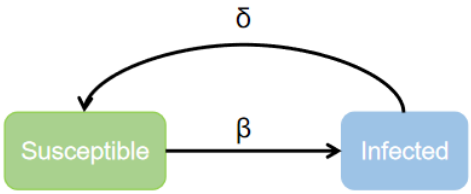


Figure 2.1: Diagrammatic representation of SIS model in terms of reaction-diffusion processes

2.2.1. The Continuous-time SIS Markov Process

We consider the virus spreading in a fixed graph G($\mathcal{N}$, $\mathcal{L}$), which does not change over time. The viral state of a node i at time t is specified by a Bernoulli random variable $X_i(t) \in 0, 1$: $X_i(t) = 0$ for a healthy but susceptible node and $X_i(t) = 1$ for an infected node. A node i at time t can be in one of two possible states: infected state with the probability $w_i(t) = Pr[X_i(t) = 1]$, and health state with probability $1 - w_i(t)$. Assuming the infection and cure process are Poisson process, as it satisfies that (1) $X(0) = 0$; (2) for all $t_0 = 0 < t_1 < \cdots < t_n$, the increments $X(t_1) - X(t_0), X(t_2) - X(t_1), \ldots, X(t_n) - X(t_{n-1})$ are independent random variables; (3) for the length of interval $t > 0$, time $\tau \geq 0$ and non-negative integers k, the increments have the Poisson distribution

$$P[(X(t+\tau) - X(\tau)) = k] = \frac{(\lambda t)^k e^{-\lambda t}}{k!} \quad k = 0, 1, \ldots \tag{2.1}$$

Thus the increments only depend on the length of interval t and not on the time τ. To conclude, the increments are stationary because $X(t+\tau) - X(\tau)$ possesses the same distribution for any time τ. And according to the properties of the Poisson distribution, the mean $E[X(t+\tau) - X(\tau)] = \lambda t$. Especially with $\tau = 0$ and $X(0) = 0$, the expected number of events in a interval with length t is

$$E[X(t)] = \lambda t, \tag{2.2}$$

which means that λ is the rate of the Poisson process. Furthermore, events occur one after another in the Poisson process and never happen at the same time. Therefore, the Poisson process describes the number of events occurred until time t when the interval from one event to the next is exponential and independent of all other intervals.

A stochastic process $\{X(t), t \in T\}$ is a Markov process if the future state of the Markov process only depends on the current state of the process and not on its past history. And formally, a stochastic process $\{X(t), t \in T\}$ is a continuous-time Markov process if

, for all $t_1 < t_2 < t_3 < ... < t_{n+1}$ of the index set T and for any set $x_0, x_1, ..., x_n + 1$ of the state space, it holds that, for the continuous-time Markov chain $\{X(t), t \geq 0\}$ with N states, the Markov property can be described as

$$\begin{aligned} Pr[X(t+\tau) = j \mid X(\tau) = i, X(u) = x(u), 0 \leq u < \tau] \\ = Pr[X(t+\tau) = j \mid X(\tau) = i], \end{aligned} \tag{2.3}$$

which reflects the fact that the future state at time $t+\tau$ only depends on the current state at time τ. Therefore, the Poisson process is a continuous time Markov process on the non-negative integers k where all transitions are a plus-one jump and the times between jumps are independent exponential random variables with the same rate parameter λ. This feature can be used to simplify the predictions about the future state of a node. And according to our assumption, the infection and curing processes are following the Poisson distribution with rate β and δ respectively.

2.2.2. The Governing Equations

The governing SIS equations is an intuitive approach that provides more insights into the spreading phenomena of the SIS model. Since $E[X_i]$=$Pr[X_i = 1]$, the exact SIS governing equation for node i equals

$$\frac{dE\left[X_i(t)\right]}{dt} = E\left[-\delta X_i(t) + (1 - X_i(t))\,\beta \sum_{k=1}^{N} a_{ki} X_k(t)\right]. \tag{2.4}$$

The SIS governing equation shows that the change over time of the probability of infection $E[X_i]$=$Pr[X_i = 1]$ of node i equals the average of two competing variables: (1) if node i is infected, then $\frac{dE[X_i]}{dt}$ decreases with rate equal to the curing rate δ and (2) if node i is healthy $(1 - X_i)$, it can be infected with infection rate β from each infected neighbor (its own self-infection is ignored). Since the network G does not change over time, the (2.4) can be simplified to

$$\frac{dE\left[X_i(t)\right]}{dt} = -\delta E\left[X_i(t)\right] + \beta \sum_{k=1}^{N} a_{ki} E\left[X_k(t)\right] - \beta \sum_{k=1}^{N} a_{ki} E\left[X_i(t) X_k(t)\right], \tag{2.5}$$

where the joint probabilities $E[X_i X_j] = Pr[X_i = 1, X_j = 1]$. When $i = j$, the governing equations are the (2.5) above. Because $\frac{dE[X_i X_j]}{dt} = E\left[X_j \frac{dX_i}{dt} + X_i \frac{dX_j}{dt}\right]$, the $\binom{N}{2}$ governing equations for $\frac{dE[X_i X_j]}{dt}$ can be derived for $i \neq j$ as follows (omitting the time-dependence

for brevity):

$$\begin{aligned}\frac{dE\left[X_iX_j\right]}{dt} =& E\left[X_j\left(\beta\left(1-X_i\right)\sum_{k=1}^{N}a_{ik}X_k-\delta X_i\right)+X_i\left(\beta\left(1-X_j\right)\sum_{k=1}^{N}a_{jk}X_k-\delta X_j\right)\right]\\ =& -2\delta E\left[X_iX_j\right]+\beta\sum_{k=1}^{N}a_{ik}E\left[X_jX_k\right]+\beta\sum_{k=1}^{N}a_{jk}E\left[X_iX_k\right]\\ &-\beta\sum_{k=1}^{N}\left(a_{ik}+a_{jk}\right)E\left[X_iX_jX_k\right]\end{aligned} \tag{2.6}$$

Since the equation (2.6) includes $E\left[X_iX_jX_k\right]$, which needs to be further determined. As mentioned above, by translating the SIS epidemic process directly in differential equations, we find that N equations for $E[X_i(t)]$ require the knowledge of the joint expectations $E\left[X_iX_j\right]$, whose $\binom{N}{2}$ differential equations require the knowledge of $E\left[X_iX_jX_k\right]$ and its $\binom{N}{3}$ differential equations require the joint fourth expectations and so on. Therefore, there are $\sum_{k=1}^{N}\binom{N}{k}=2^N-1$ equations and the conservation of probability equation $Pr[X_i=1]+Pr[X_i=0]=1$ are required in total. In conclusion, the defined infection process is a continuous-time Markov chain with 2^N states. Therefore, the exact SIS Markov process requires 2^N linear equations to solve, which is infeasible for large real-world networks. And then, finding a good approximate solution will be of vital importance for the further research.

2.2.3. The Steady State and Phase Transition

The SIS model assumes nodes can be infected over and over again, in a cycle $S\rightarrow I\rightarrow S$, thus the disease can be sustained forever under some conditions. In this case, the SIS spreading process has a steady state with the number of infections equal to the number of cure events.

We define the steady-state random variable as

$$X_i=\lim_{t\rightarrow\infty}X_i(t), \tag{2.7}$$

which obeys $\lim_{t\rightarrow\infty}\frac{dE[X_i(t)]}{dt}=0$. Thus the average fraction of infected nodes in the steady-state, also known as the prevalence, is represented by

$$y_\infty(\tau)=E[S_\infty]=\frac{1}{N}\sum_{i=1}^{N}X_i, \tag{2.8}$$

where τ is the effective infection rate. And in the steady-state, the equation (2.5) can be simplified as

$$\delta E\left[X_i\right] = \beta \sum_{k=1}^{N} a_{ki} E\left[X_k\right] - \beta \sum_{k=1}^{N} a_{ki} E\left[X_i X_k\right], \tag{2.9}$$

which also presents that the number of infections equal to the number of cure events in the steady state. The relationship between $y_\infty(\tau)$ and the effective infection rate τ depicts the physical aspects of the steady state, where the critical point is the epidemic threshold. The epidemic threshold τ_c is defined as the border between exponential die-out and a non-zero fraction of infected nodes in the metastable state. The epidemic threshold is analogous to the concept of phase transition in non-equilibrium systems [28]. The phase transition is defined as an abrupt change in the state of a system. Therefore, the determination of the critical point is significant in numerical simulations. To conclude, in any finite given sized network, the exact SIS epidemic threshold holds that

$$\tau_c \geq \tau_c^{(1)} = \frac{1}{\lambda_1}, \tag{2.10}$$

where λ_1 is the spectral radius of the adjacency matrix A. This theorem can be of great practical use: if the effective infection rate τ can be controlled such that $\tau \leq \tau_c^{(1)}$, then the network can be safeguarded from long-term and massive infection. While $\tau \geq \tau_c^{(1)}$, the spreading will enter the steady state. However, the steady-state of the exact SIS model is always reach the absorbing state after unrealistically long time, and the absorbing state is all-healthy state.

2.3. The Mean Field Analysis

In order to simplify the calculation progress and apply the SIS model to solve problems on real-world network with network size N is pretty large, we approximate the SIS epidemics by subsequent order expansions. The N-Intertwined Mean-Field Approximation (NIMFA) proposed by P. Van Mieghem [17], which apply the first-order mean-field expansion to approximates $E[X_i]$, can be applied to solve only with N linear equations.

The number of infected neighbors of node i, $\sum_{j=1}^{N} a_{ij} X_j(t)$, couples or intertwines' each of the N nodal infection states in the network and causes higher-order joint probabilities. In mean-field analysis, by assuming X_k be a sequence of independent random variables, where $E[X_i(t)X_j(t)] = E[X_i(t)]E[X_j(t)]$. Thus, equations for higher-order joint probabilities are not required. We denote $v_i(t) = Pr[X_i(t) = 1]$ under assumption. Thus, the

governing equation for a node i in NIMFA is

$$\frac{dv_i(t)}{dt} = -\delta v_i(t) + (1 - v_i(t))\,\beta \sum_{j=1}^{N} a_{ij} v_j(t). \tag{2.11}$$

By comparing the exact equation and NIMFA equation, it shows that the random variable X_k is replaced by its mean $v_k = E[X_k]$, irrespective of any correlation introduced by terms as $E[X_i X_j]$. Also, ignoring correlations leads to a replacement of the actual number of infected neighbors $\sum_{j=1}^{N} a_{ij} X_j$ by its mean $\sum_{j=1}^{N} a_{ij} v_j$. The mean $\sum_{j=1}^{N} a_{ij} v_j$ is an increasingly accurate estimator of the random variable $\sum_{j=1}^{N} a_{ij} X_j$ with N increasing.(p120)According to the Central limit Theorem, for each node i, NIMFA upper bounds $\frac{dE[X_i(t)]}{dt} \leq \frac{dv_i(t)}{dt}$ and $E[X_i] = w_{i\infty} \leq v_{i\infty}$. Thus, each node in the graph G obeys a differential equation as follows:

$$\begin{cases} \frac{dv_1(t)}{dt} = \beta \sum_{j=1}^{N} a_{1j} v_j(t) - v_1(t) \left(\beta \sum_{j=1}^{N} a_{1j} v_j(t) + \delta \right) \\ \frac{dv_2(t)}{dt} = \beta \sum_{j=1}^{N} a_{2j} v_j(t) - v_2(t) \left(\beta \sum_{j=1}^{N} a_{2j} v_j(t) + \delta \right) \\ \qquad\qquad \vdots \\ \frac{dv_N(t)}{dt} = \beta \sum_{j=1}^{N} a_{Nj} v_j(t) - v_N(t) \left(\beta \sum_{j=1}^{N} a_{Nj} v_j(t) + \delta \right) \end{cases}, \tag{2.12}$$

with $V(t) = \begin{bmatrix} v_1(t) & v_2(t) & \cdots & v_N(t) \end{bmatrix}^T$, and the matrix evolution equation of NIMFA is

$$\frac{dV(t)}{dt} = \beta A V(t) - \mathrm{diag}\left(v_i(t)\right)\left(\beta A V(t) + \delta u\right), \tag{2.13}$$

where u is the all-one vector and $diag(v_i(t))$ is the diagonal matrix with elements $v_1(t)$,$v_2(t)$,...,$v_N(t)$. Therefore, only N non-linear equations are required to solve in NIMFA at the expense of exactness. Besides, the NIMFA epidemic threshold is precisely $\tau_c^{(1)} = \frac{1}{\lambda_1(A)} < \tau_c$. With this threshold, there is a phase transition to the metastable state. To conclude, The NIMFA make the spreading process of SIS model analytically tractable even with pretty large number of nodes N in any graph and it provides a lower bound to the epidemic threshold.

Most studies on SIS epidemics in networks implicitly assume Markovian behaviour: the infection time T and the cure time R are exponentially distributed because the infection rate β and cure rate δ are following the Poisson distribution. And many efforts have been made to characterize the epidemic threshold on various networks. However, these exponential distributions in general do not describe real-life epidemics well [PR2013]. The classical continuous-time SIS Markov epidemic model is extended to incorporate infection and curing times characterized by a general distribution, which is called as the

generalized SIS (GSIS) model. The general distribution can be a gamma, Weibull and other lognormal distributions. And this extension is believed to be more applicable for the real-world epidemics.

With GSIS model, we assume all infection and cure processes are independent. If a node i gets infected at time t, we draw independently of everything else a recovery time $R_i(t)$. Given $R_i(t)$, we then draw independently, for each neighbor j of i, a random number $M_{ij}(t)$ of infecting times $T_{ij}^1(t) \leq \cdots \leq T_{ij}^{M_{ij}(t)}(t) \leq R_i(t)$, nodes i tries to infect node j at ties $t + T_{ij}^1(k)$ with $1 \leq k \leq M_{ij}(t)$. If node j is already infected at such a time, then nothing happens. Finally, node i recovers at time $t + R_i(t)$ and becomes healthy, but again susceptible to infection. Similarly, an exact analysis of the GSIS model on any network is very likely intractable, so that only an approximate a treatment seems possible.

2.3.1. Weibull Distribution and Gamma Distribution

To study the non-Markovian curing process, we consider gamma distribution and Weibull distribution as a more general distribution instead of the Poisson distribution in this project. And the description of gamma distribution and Weibull distribution are in the following:

1. Gamma Distribution
The gamma distribution [29] is as following,

$$f_{T_{\text{Gamma}}}(x;\xi) = \frac{\frac{1}{b_\Gamma}\left(\frac{x}{b_\Gamma}\right)^{\xi-1}}{\Gamma(\xi)} e^{-\frac{x}{b_\Gamma}}. \tag{2.14}$$

With mean $E[T_{Gamma}] = b_\Gamma \xi$ and variance $Var[T_{Gamma}] = \xi b_\Gamma^2$ and the corresponding pgf (probability-generating function):

$$\varphi_{T_{\text{Gamma}}}(z;\xi) = (1 + b_\Gamma z)^{-\xi}. \tag{2.15}$$

For ξ=1, the gamma distribution reduces to an exponential distribution. The key parameter for the gamma distribution is the shape parameter ξ, and b_Γ is the scale parameter. If $\xi = k \geq 1$ is an integer, then the gamma random variable equals the sum of k independent and identically distributed exponential random variables [30, p.45-46].

2. Weibull Distribution

The Weibull distribution [31] is as following:

$$f_T(x) = \frac{\alpha}{b}\left(\frac{x}{b}\right)^{\alpha-1} e^{-(x/b)^{\alpha}}, \tag{2.16}$$

for $x \geq 0$, with the expectation $E[T] = b\Gamma\left(1+\frac{1}{\alpha}\right)$, where α is a shape parameter. $\Gamma(x)$ is the gamma function. And $b = \left[\lambda\Gamma\left(1+\frac{1}{\alpha}\right)\right]^{-1}$ because, in order to compare Weibull distribution with the exponential distribution, the average curing time $E[T]$ is fixed to the inverse of the rate $1/\lambda$. Besides, the distribution function is

$$F_T(x) = \Pr[T \leqslant x] = 1 - e^{-(x/b)^{\alpha}}. \tag{2.17}$$

The Weibull distribution is heavy-tailed when $\alpha < 1$, the tail decreases slower, but the probability of small intervals increases as a power law, proportional to $x^{\alpha-1}$ [32]. The Weibull distribution reduces to the exponential distribution when $\alpha = 1$ and Gaussian-like when $\alpha > 1$.

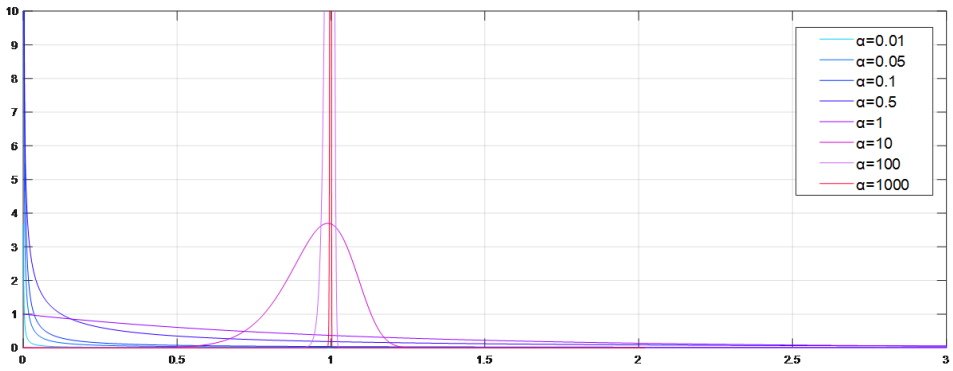


Figure 2.2: The Weibull distribution with various α

2.3.2. The Mean-field Approximation in GSIS

Assuming that a metastable state exists, thus the average recovery time $E[R_i]$ and the average number $E[M_{ij}]$ of infection event of node i are finite. Denoting v_i is the probability of node i getting infected in the metastable state. Besides, when we determine the effect of the neighbors on node i, we ignore the affects of node i to its neighbors.

Under these assumptions, we consider node j as a neighbor of node i. According to the elementary renewal theorem [30, p.165], the length of an infected period equals $E[R]$ and the number of node j got infected is asymptotically linear within a large time S.

Therefore, the number of infected periods is $v_j S/E[R]$. The average number of times that node j tries to infect node i is equal to $E[M]$. So that the total number number of node j attempting to infect node i is asymptotically equal to $v_j SE[M]/E[R]$. According to the mean-field approximation, the fraction of successful infections from node j to node i is equal to $1 - v_i$. Thus, the total number of successful infections that node i will receive in the time interval $[0, S]$ is asymptotically equal to

$$S \sum_{j \in U_i} \frac{E[M]}{E[R]} v_j \left(1 - v_i\right), \tag{2.18}$$

where we denote U_i as the index set of neighbors of node i and $U_i = \{j \mid a_{ij} = 1\}$. Therefore, in the metastable state, we have

$$S \sum_{j=1}^{N} a_{ij} \frac{E[M]}{E[R]} v_j \left(1 - v_i\right) = v_i \frac{S}{E[R]}. \tag{2.19}$$

The right-side of (2.19) is the number of infected periods of node i in interval $[0, S]$. And to simplify (2.19), we get

$$E[M] \left(1 - v_i\right) \sum_{j=1}^{N} a_{ij} v_j = v_i. \tag{2.20}$$

Besides, we consider the curing time R and infection time T are following the renewal processes, respectively. The renewal process is a counting process for which the inter-arrival times τ_n are i.i.d. random variables with distribution $F_\tau(t)$. Thus, through derivation, we finally arrive at the general expression of $E[M]$ is

$$E[M] = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{\varphi_T(z)\varphi_R(-z)}{1 - \varphi_T(z)} \frac{dz}{z}, \tag{2.21}$$

where $c > 0$ and $\varphi_R(z) = \int_0^\infty e^{-zu} f_R(u) du$ is the probability generating function (PGF) of the curing time R.

P. Van Mieghem [33] shows that if the infection process follows the Poisson distribution with rate β, the infection time T is exponentially distributed with mean $1/\beta$. And we only consider the curing time follows a general distribution with mean $1/\delta$, that $E[R] = 1/\delta$. Then the PGF is $\varphi_T(z) = \frac{\beta}{z+\beta}$, the (2.21) can be simplified to

$$E[M] = \frac{\beta}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{\varphi_R(-z)}{z^2} dz \quad (c > 0). \tag{2.22}$$

Thus when we close the contour over the negative $Re(z)$ plane, in which $\varphi_R(-z)$ is analytic and $|\varphi_R(-z)| \leqslant 1$. Thereby, by the Cauchy's integral theorem, we can get

$$E[M] = \beta \frac{d\varphi_R(-z)}{dz}\bigg|_{z=0} = \beta E[R] = \tau \tag{2.23}$$

Then we can know if the infection time T is exponentially distributed, then for any general distribution of the curing time R, NIMFA results are suitable to be applied.

2.4. Non-Markovian Infection Spread Epidemics

P. Van Mieghem [34] investigated the dramatic effect of a non-exponential infection time on the epidemic threshold by considering Weibullean infection times with the same mean, but different shape parameters α of Weibull distribution, on three basic classes of graphs. The average steady-state fraction of infected nodes is simulated from which the epidemic threshold is deduced. And the epidemic threshold increases with the shape parameters α. It mentioned that if the real epidemics are not infecting direct neighbors in an exponential time, a more deep-in research covers more types of graphs and other heavy-tailed distribution for both infection and curing is required.

Q. Liu [35] studied the Weibullean SIS process with the shape parameter α in two kinds of extreme situations with $\alpha \to 0$ and $\alpha \to \infty$. When $\alpha \to 0$, the Weibullean SIS epidemic threshold is zero, and for an arbitrary small α the mean-field epidemic threshold tends to zero. When $\alpha \to \infty$, the prevalence of the Weibullean SIS process in the long run, together with the NIMFA and the Markovian prevalence, the accuracy of the mean-field approximation is worst in the rectangular grid network with a minimum largest eigenvalue, and best in the scale-free network with a largest eigenvalue.

Even some recent works focus on the SIS model with a non-Markovian infection process, the systematic research on the SIS epidemics with non-Markovian curing process is still lacking. This thesis project extends curing process to a renewal process with inter-arrival intervals following a more general distribution, like Weibull distribution and gamma distribution. Combining the theoretical results from P. Van Mieghem [33] in 2013, this work demonstrated that if the infection time T is exponentially distributed with the mean $1/\beta$, then for any distribution of the recovery process, the NIMFA conclusion applies. Therefore, it holds that the non-Markovian SIS spreading process have the epidemic threshold that $\tau_c = \frac{1}{\lambda_1}$ under the NIMFA assumption. Therefore, this project will first examine that conclusion with simulations on exact networks. Through comparing

the discrepancies between simulation results and theoretical conclusions, we have a deep understanding of SIS epidemics with non-Markovian curing process on networks. A series of systematic simulations and results analysis will be presented in the next chapter.

3

Non-Markovian SIS Curing Epidemics

In previous research, the non-Markovian Susceptible-Infected-Susceptible (SIS) process with infection time interval following a more general distribution has been studied. However, systematic research on the non-Markovian curing process is still lacking. In order to explore the exact spreading process and the performance of non-Markovian curing process, various simulations on SIS epidemics with non-Markovian curing process will be implemented in this chapter.

3.1. Continuous-time SIS Epidemics Simulator

The simulator we used in this project is designed by Ruud van de Bovenkamp [36]. It is designed to model the graph dynamics on exact networks. The simulator provides a publicly available toolkit to extract graphs from data-sets or data streams and to analyse their properties. Everything that can be represented as a collection of entities that have a relation can be modeled as a network. And simulator provides various kinds of underlying networks, including Erdős-Rényi random graph, BA scale-free graph and so on. The simulator can simulate the SIS viral spreading process based on these networks.

The simulator allows to start simulation with different percentage of nodes in infected states. In this project, all nodes are infected in the initial state. The distribution of infection time T and cure time R can be designed by choosing different shape parameter

α and mean (infection rate and cure rate). During the simulation, the simulator records the number of infected nodes in each step. After 50 time units, the simulator can show us the prevalence of exact spreading process in meta-stable state.

During the simulation, the infected nodes will cure with curing rate δ irrespective of the state of the rest of the network. While the node is healthy and susceptible to disease, its infected neighbors will spread the infection with a combination rate. Therefore, the total infection rate that a healthy node i experiences is not constant over the healthy period of the node as infected neighbours can cure. Comparing with the assumptions of the the N-Intertwined Mean-Field Approximation (NIMFA) [17], the viral state of the nodes are not independent since clearly the rate of change from the healthy to the infected state of node i depends on the number of its infected neighbours, where $E[X_i(t)X_j(t)] \neq E[X_i(t)]E[X_j(t)]$. On the basis of that assumption in the simulator, we guess that this difference will cause discrepancies between the simulation results and the theoretical results.

3.2. Numerical Simulations

In order to have an overview of the non-Markovian SIS spreading phenomenon with the cure time R following a general distribution and to verify the theoretical results in previous researches [33, 35], we first simulate the spreading process and analyze the prevalence based on the simulator to observe how many nodes will get infected at the metastable state [37]. Besides, for finite systems, it is difficult to find the phase transition for numerical simulations. Therefore, we second focus on the critical point which separates the infected phase from the healthy phase.

3.2.1. Simulations on Prevalence

In order to make comparison with results of non-Markovian infection spread epidemics in previous researches [33, 35], we set the curing process to Weibull distribution with shape parameter $\alpha \in (0, \infty)$. We run the simulation on a ER random network with number of nodes $N = 50$ and the largest eigenvalue $\lambda_1 = 8.5232$. In order to ensure the spreading process arrive at the metastable state, the simulation runs for a long enough time with 50 time units with cure rate δ=1 and infection rate β=1. As a consequence, we accumulate the prevalence of infected nodes over time to observe the phenomenon of the non-Markov SIS spreading process. The prevalence is the average fraction of infected

nodes in the steady-state, which is denoted by $y_\infty(\tau) = \lim_{t\to\infty} \frac{1}{N} E\left[\sum_{j=1}^{N} X_j(t)\right]$.

3.2.2. Simulations on Phase Transition

In SIS epidemics, the epidemic threshold τ_c separates the infected phase from the healthy phase. However, because of random fluctuations, the absorption event can sometimes occur even in the active state well above the critical point [28]. In a nutshell, it is difficult to experimentally find the critical point of phase transition. Therefore, in order to get the epidemic threshold more accurately, we do simulation with 10^5 realizations and average the results. The epidemic threshold τ_c is defined as the border between exponential die-out and a non-zero fraction of infected nodes at the metastable state.

In order to figure out whether the Weibullean curing process have effects on the phase transition, the simulation aims to find the epidemic threshold with various distribution of curing process, determined by the shape parameter α. We first set the curing process to Weibull distribution with shape parameter α, ranging from 10^{-1} to 10^5. We run the simulation on a ER random network with number of nodes $N = 50$ and $\lambda_1 = 8.5232$. In order to observe whether the SIS epidemics with non-Markovian curing processes have obvious phase transition, we first did simulations with different Weibull distributions and focus on the curve of prevalence changing with the epidemic threshold $\tau = \beta/\delta$.

According to the analysis in [33], in the case with infection processes following the Poisson distribution and curing processes following any distribution, the theoretical NIMFA epidemic threshold is $1/\lambda_1$ as classical Poisson SIS process. Thus the epidemic threshold is only determined by the underlying network, which will not be influenced by the distribution of the curing process. However, the simulation results have conflict with that theory.

To further understand the impacts of different distributions of curing time R on the critical points of phase transition. The second step of the simulation focuses on the relationship between shape parameter α and the epidemic threshold at critical point. In general, the simulation runs for a long enough time with 50 time units with the cure rate δ =1. We consider the average of the maximum and minimum prevalence of the last complete time period as the main output value with specific infection rate β. And the prevalence is obtained by averaging over 10^5 realizations of simulation with all nodes infected initially to prevent the inaccuracy caused by the early die-out. Because the epidemic threshold τ_c is defined as the border between exponential die-out and a non-zero fraction of infected nodes in the metastable state. By fixing the value of shape

parameter α, we gradually reduce the value of infection rate β with the interval of 0.01 and we consider the process is die-out when the average prevalence is less than 0.1%. When the prevalence arrive at the range between [0.09%, 0.11%], we think the ratio of $\tau^\star = \beta/\delta$ is the epidemic threshold of this process. Thus, the first infection rate that causes the process to become a die-out process is the value at critical point. Then, we can get the distribution of $\tau^\star$ corresponding to various shape parameters α from 10^{-2} to 10^5.

Third, in order to rule out the impact of different network typologies to the epidemic threshold. we repeat the same procedure as the second step on various networks, including the BA scale-free network and regular graph. Besides, in order to explore the impact of network size and link density on epidemic threshold distribution, we run the simulations on regular graphs with various number of nodes and degrees.

Last but not least, since we only change the infection parameter α in previous simulations, there may exist large fluctuations and errors of simulation results when the infection rate is too small. In order to rule out the effects of simulation errors, we repeat the second simulation on the ER graph with fixed infection rate β =10 and reduce the curing rate to find the critical value at the epidemic threshold. Thereby, we can similarly get the distribution of the epidemic threshold at critical point, which is corresponding to various shape parameters α from 10^{-2} to 10^5.

3.3. Performance Analysis for Shape Parameter $\alpha \in [0.2, 10]$

According to the simulation settings in subsection 3.2, we analyze the data obtained from simulations in this section. Observing from the data, we find that the theoretical results are more practical for exponential-like distributions, whose shape parameter within a certain range $\alpha \in [0.2, 10]$. The detailed analysis will be presented in the following.

Figure 3.1 shows the prevalence $y_\infty(\tau)$ of infected nodes on an Erdős-Rényi (ER) Random Graph $G_p(N)$ with $N = 50$ nodes and number of links $L = 190$, versus the time for various α smaller than 10. The average steady state fraction increases with increasing α, which approaches 0.85. Figure 3.1 also illustrates that the average fraction $y_\infty(\tau)$ has more violent fluctuations at the beginning with α increasing. And when the number of infections is approximately equal to the number of cures, the average fraction $y_\infty(\tau)$ becomes steady, which is approximately equal to the final average fraction of Poisson process with α=1. Thus, when shape parameter α is larger than 0.2 and less than 10, the spreading process can arrive at the meta-stable state with same prevalence of classical

Markovian SIS process. In conclusion, when the shape parameter ranges from 0.2 to 10, the cure time R distribution is exponential-like. The NIMFA conclusions are applicable to these cases.

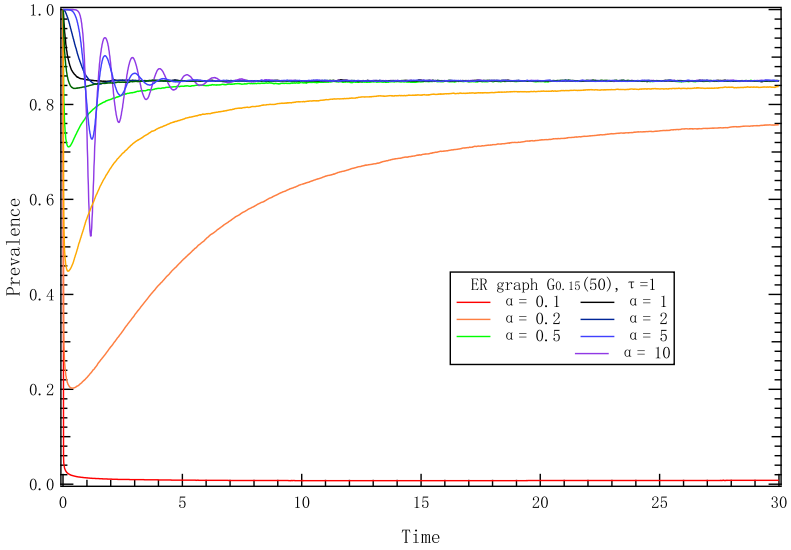


Figure 3.1: The prevalence over time on ER random graph with the network size $N = 50$ and number of links $L = 190$ and shape parameter of Weibull distribution belongs to $0.1 \leq \alpha \leq 10$

Figure 3.2 depicts that when shape parameter belongs to the range from 0.2 to 10, the curve shows a sharp phase transition from healthy state to infected state. Thus, these cases exist the epidemic threshold and the theoretical epidemic threshold $\tau_c = 1/\lambda_1 = 1/13.6229 \approx 0.0734$ provides a lower bound to the simulation results. To conclude, when the shape parameter of cure time R distribution ranges from 0.2 to 10, the theoretical conclusion are applicable to those cases.

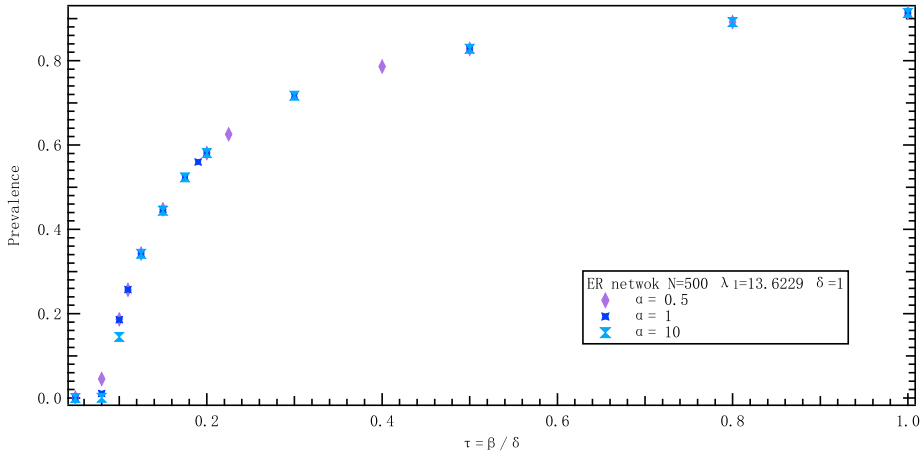


Figure 3.2: The prevalence over ratio of epidemic threshold $\tau^\star$ on ER random graph with the network size $N = 500$ and the largest eigenvalue $\lambda_1 = 13.6229$ and shape parameter of Weibull distribution belongs to $0.2 \leq \alpha \leq 10$

3.4. Performance Analysis for Shape Parameter $\alpha \in (0, 0.2]$

According to section 3.3, the red line (line close to time axis) illustrates that, when the shape parameter α is smaller than 0.2, the prevalence at the metastable state cannot arrive at the same value as Markovian SIS epidemic process. It is necessary to further understand how the phase transition was effected by different distributions of curing time R. The detailed analysis is presented in the following.

3.4.1. Effects on Phase Transition

Based on the simulations on phase transition, figure 3.3 indicates that, when the curing process follows Weibull distribution with shape parameter α is smaller than 0.2, the prevalence at critical point does not have a sharp phase transition between the healthy state and the metastable state. Comparing with cases close to Markovian SIS process, the prevalence with super small shape parameters grows pretty slowly with the effective infection rate τ.

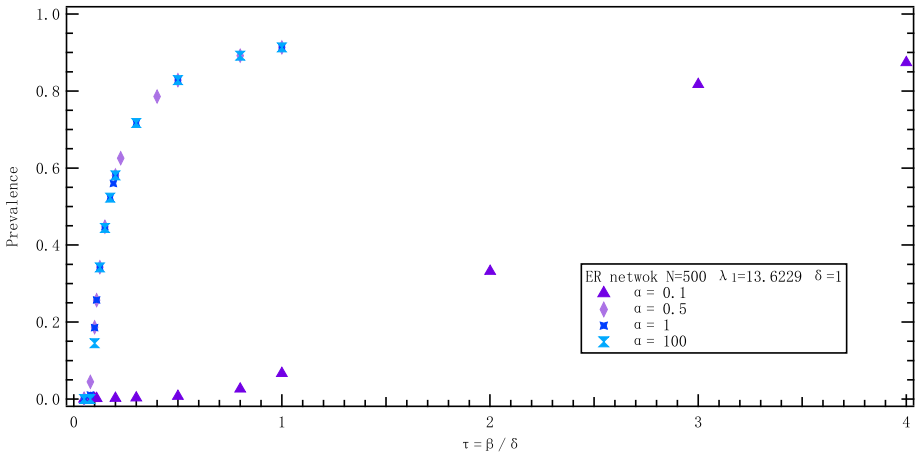


Figure 3.3: The prevalence over ratio of epidemic threshold $\tau^\star$ on ER random graph with the network size $N = 500$ and the largest eigenvalue $\lambda_1 = 13.6229$ and shape parameter of Weibull distribution belongs to $0.1 \leq \alpha \leq 10$

Therefore, in order to understand how the distribution of curing time R affect the phase transition, we need to further find out the relationship between the shape parameter α and the critical point τ_c of phase transition, which is shown in Figure 3.4. Here we define the critical point τ_c as the border between exponential die-out and a non-zero fraction of infected nodes in the metastable state.

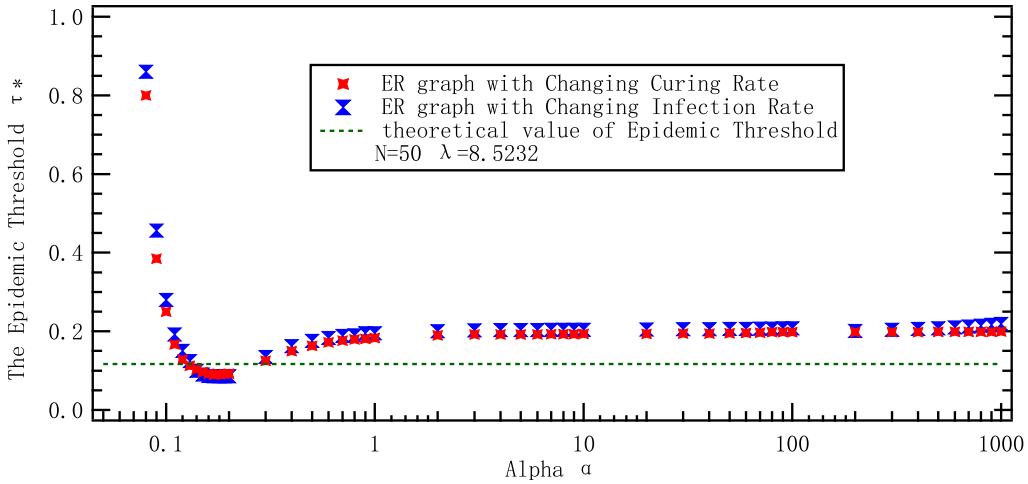


Figure 3.4: The epidemic threshold distribution with $0.08 \leq \alpha \leq 1000$. Comparison between changing infection rate and changing curing rate on ER random graph with $N = 50$ and $\lambda_1 = 8.5232$

Figure 3.4 shows that the distribution of epidemic threshold $\tau^\star = \beta/\delta$ over the shape parameter α within the range from 0.01 to 1000. The blue line shows a steep decline with α decreasing from 0.01 to approximately 0.18 and the curve approximately arrives at the lowest point with $\alpha = 0.18$, and after that, the curve gradually increases and eventually trend to be stable even with a large value of α. Comparing with the case that changing curing rate, both lines approximately show the same distribution and have the same lowest point.

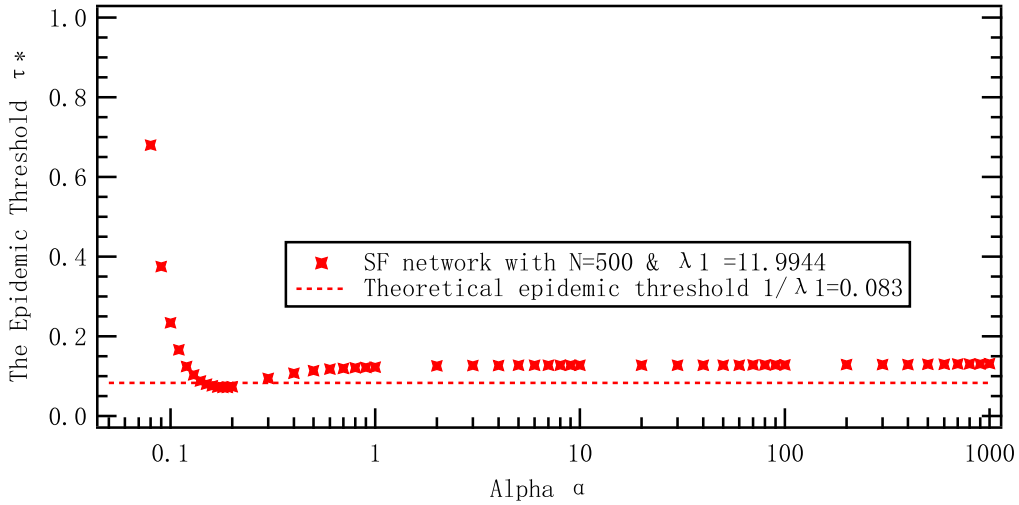


Figure 3.5: The epidemic threshold with $0.08 \leq \alpha \leq 1000$ on three kinds of networks: the BA scale-free network with the network size $N = 500$ and the largest eigenvalue $\lambda_1 = 11.9944$

Figure 3.4, 3.5 and 3.6 show that the epidemic threshold $\tau^\star = \beta/\delta$ distribution simulated on different networks. It depicts that all curves have approximately the same trends: the epidemic threshold decreases in the range of α from 0.01 to 0.18. And they all have

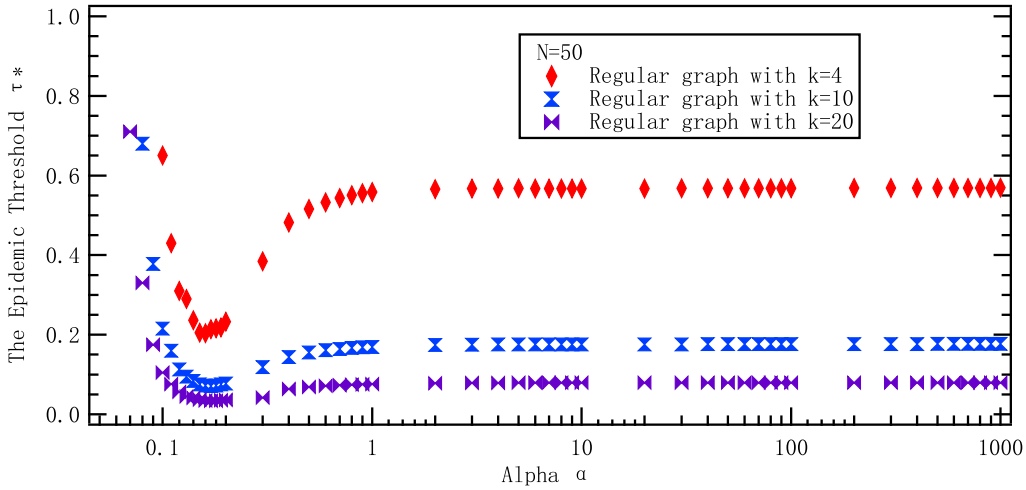


Figure 3.6: The epidemic threshold distribution with $0.08 \leq \alpha \leq 1000$ on regular network with various number of nodes N and various degree k

the lowest value at shape parameter $\alpha = 0.18$ and finally trend to be stable. Besides, Figure 3.6 depicts that even on the regular graph with different number of nodes and degrees, the epidemic threshold $\tau^\star = \beta/\delta$ has similar distribution. Three lines decrease first and arrive at the lowest point at approximately shape parameter $\alpha = 0.18$, and lines rise up slowly and trend to be finally stable.

To conclude the above results, when the shape parameter is smaller than 0.2, the distribution of curing time R will affect the prevalence of entire processes and the epidemic threshold at critical point. Moreover, the smaller the shape parameter it is, the larger epidemic threshold will be. But how exactly the distribution affect the prevalence and epidemic threshold? The details will be discussed in the following.

3.4.2. The Dramatic Effects of the High-peak and Heavy-tailed Distribution of Curing Time

By simulating with different shape parameter α, we find the number of infected nodes changes dramatically within a short period of time as shape parameter α changing and some processes could even die out at the beginning stage. Thereby, it will finally affect the prevalence at the metastable state.

Figure 3.7 depicts the percentage of die-out processes within all realizations. The curve is obtained from simulations on a ER network with the network size $N = 50$. Figure 3.7 illustrates that with shape parameter get smaller, the percentage of die-out processes get larger. Meanwhile, we can see that when the shape parameter α is larger than 500,

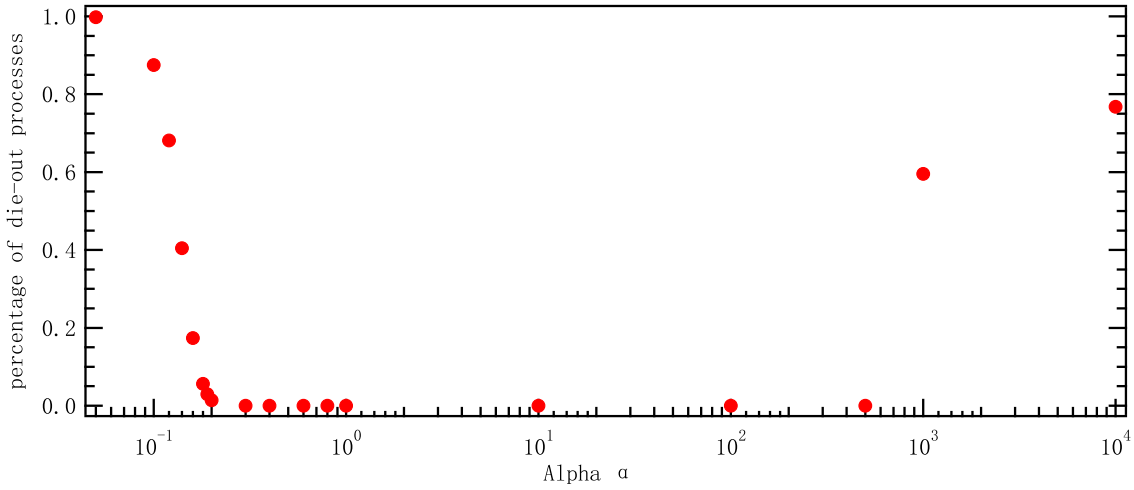


Figure 3.7: The percentage of die-out processes with 10^5 realizations over shape parameter α ranging from 0.05 to 10^4

the percentage of die-out processes becomes larger with the rising shape parameter α. This condition will be discussed in next section. In summary, a large amount of die-out processes lead to only a small percentage of infected nodes survive at the beginning and finally leads to a low prevalence at the metastable state. Thus, we can conclude that, when the shape parameter α is smaller than 0.2, a large number of infected nodes get cured immediately at the beginning of the simulation, which directly leads to the existence of a large proportion of die-out processes and ultimately affected the prevalence at the metastable state.

In Subsection 2.3.1, the Weibull distribution over various α shows that when shape parameter α is smaller than 0.2, it looks like an exponential distribution. The probability of small time intervals increases as a power law, which is proportional to $x^{\alpha-1}$ [32]. Thus, in order to clearly compare the probabilities of small intervals with different Weibull distribution, we accumulate the percentage of interval time which is smaller than 0.05. The differences among $\alpha = 0.1$, $\alpha = 0.05$ and $\alpha = 0.01$ are shown in Figure 3.8.

When shape parameter α is smaller than 0.2, the distribution of curing time R has a high peak at short curing interval and a heavy tail of long curing time. Moreover, with α decreases, the peak get higher and tail get heavier. We first choose an extreme small curing time 0.05. Figure 3.8 shows that the percentage of extreme short time drops with α increases from 0.01 to 0.1. In the case of $\alpha = 0.01$, the extreme short time can even reach 62.11%. To sum up, more than 50% curing events will happen within extreme short time.

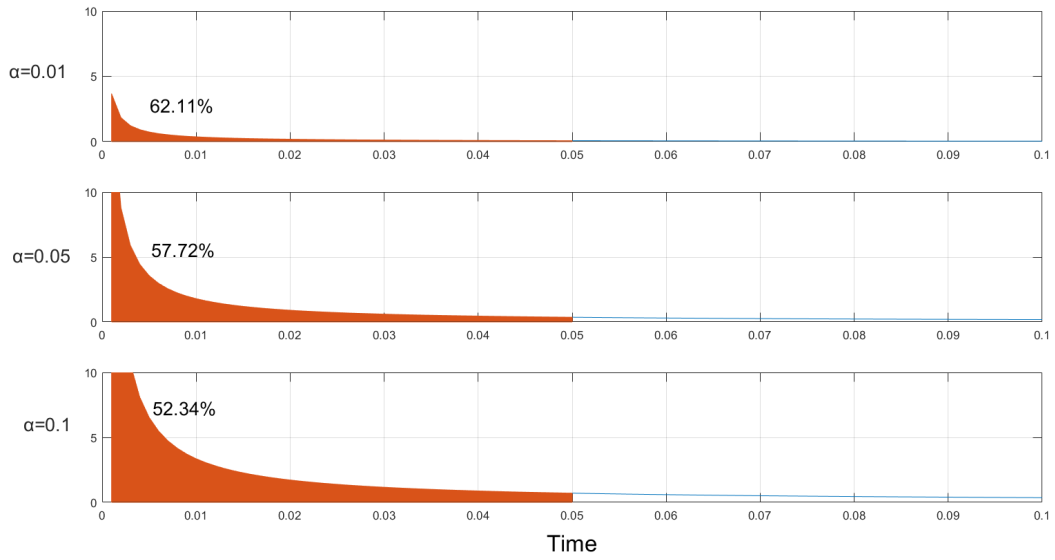


Figure 3.8: The percentage of curing events with interval time smaller than 0.05. Comparison among $\alpha = 0.1$, $\alpha = 0.05$ and $\alpha = 0.01$

In our simulations, all nodes start with infected states and get random curing time within a small interval. Therefore, a large percentage of nodes will get cured within extreme short time interval. This condition may cause all infected nodes getting cured in some realizations, so the spreading process will die out immediately at the beginning. Here we denoted this condition as initial condition. Thus, these initial conditions will influence the final simulation results.

However, does the non-Markovian curing process only affect the initial condition? Is there any other conditions are influenced by the distribution of curing time R? In order to figure out the answer of these questions, we run the simulations to get the prevalence of non-die-out processes. In this simulation, we only accumulate the prevalence of non die-out processes among 10^5 realizations over time. Similarly, we simulate on the ER random network with number of nodes $N = 50$ and $\lambda_1 = 8.5232$. In order to ensure the spreading process arrive at the metastable state, the simulation runs for a long enough time with 50 time units with curing rate δ=1 and infection rate β=1. In order to exclude the effect of initial die-out processes, we only accumulate the average prevalence of initial non die-out processes. Figure 3.9 and 3.10 show the comparison between prevalence of all processes and prevalence of non die-out processes.

Figure 3.9 shows that, when shape parameter is smaller than 0.1, the prevalence cannot arrive at the same value as Markovian SIS process. While Figure 3.10 depicts that the prevalence of non initial die-out processes can reach the same prevalence of case with

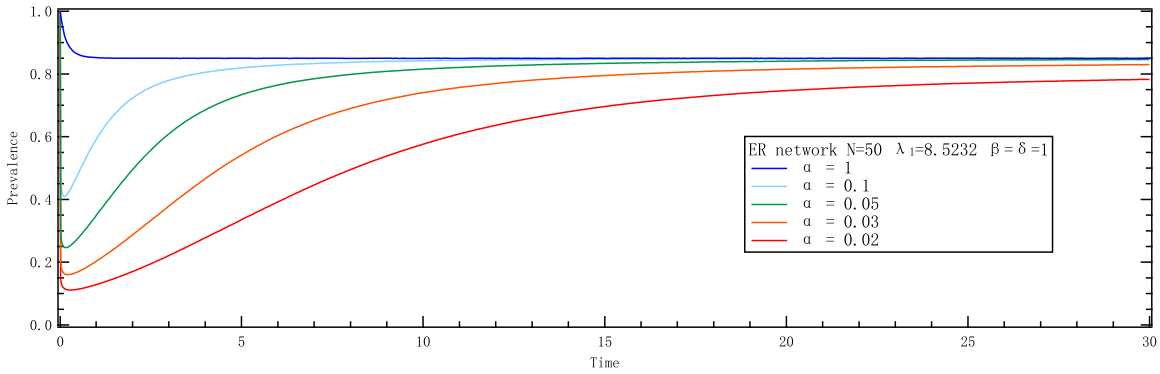


Figure 3.9: The prevalence of all processes over 10^5 realizations on ER network with N=50 and L=190

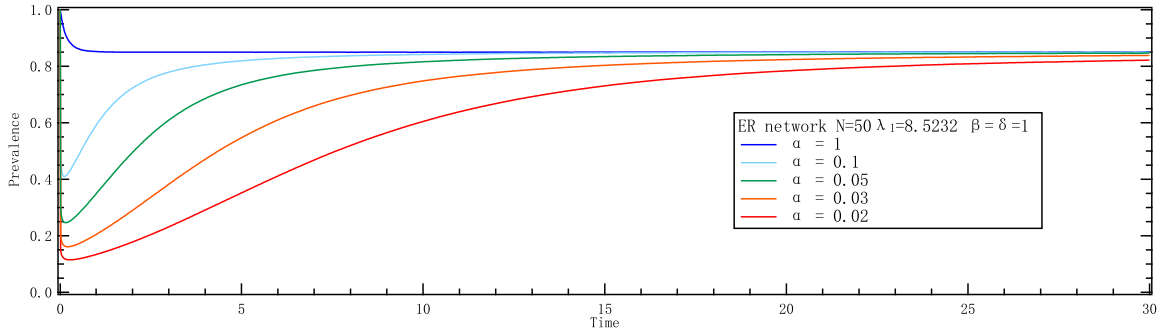


Figure 3.10: The prevalence of non initial die-out processes over 10^5 realizations on ER network with N=50 and L=190

$\alpha = 1$. Thus, ignoring the effect of the initial condition, the non-Markovian curing process does not influence the prevalence of non die-out processes.

Furthermore, we need to figure out whether the effect on initial condition will lead to changes on the critical point of phase transition. we run the simulation on phase transition with only considering non die-out processes. Figure 3.11 illustrates that the epidemic threshold of all processes and non die-out processes comparing with theoretical epidemic threshold. The curve of non die-out processes shows that, when shape parameter α is smaller than 0.2, the epidemic threshold is close to zero. That means even with a low infection rate, there still exists a small number of nodes in infected state at the metastable state. And the curve increases when α grows from 0.2 to 1. And when α is larger than 1, the curve of non die-out processes get closer to the curve of all prevalence. Therefore, the initial condition only has influence on the epidemic threshold of cases with shape parameter smaller than 0.2.

The epidemic threshold of case with $\alpha < 0.2$ is close to zero, which means there exists a small percentage of nodes surviving at the metastable state, even with a small effective infection rate. And that condition can be explained by heavy-tailed characteristics of

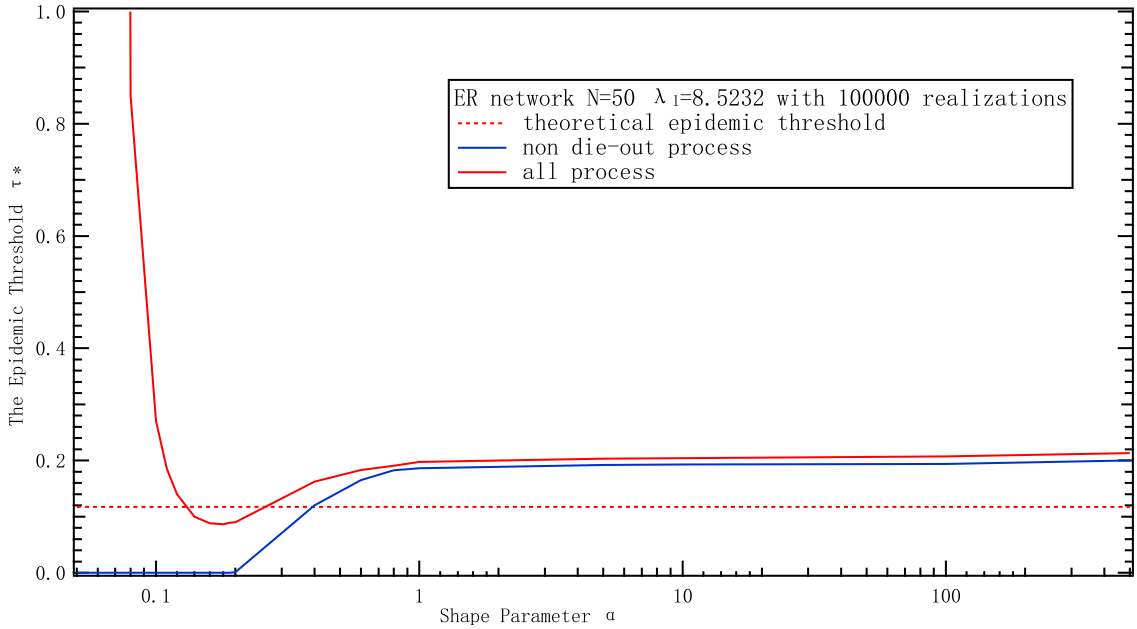


Figure 3.11: The epidemic threshold distribution with $0.1 \leq \alpha \leq 1000$. Comparison between all processes and non die-out processes on ER random graph with $N = 50$ and $\lambda_1 = 8.5232$

curing time distribution. Figure 3.12 shows that, the percentage of time interval longer than 10 is 35.94%, 32.56% and 28.39% with shape parameter $\alpha = 0.01$, $\alpha = 0.05$ and $\alpha = 0.1$ separately. Thus, with smaller the shape parameter, the percentage of extreme long time interval get larger. That means there always exist a small percentage of infected nodes have a enough long curing time to become healthy, so they can keep infected until the metastable state. Just like the hepatitis B virus (HBV) [13], which has a very long period to get recovered and a low probability to successfully infect others. but once their neighbors infected, they could also have an extreme long recovering period and slowly spread to others. In this way, virus can keep survive for a long time. And that's the reason why only a small group of nodes can keep infected at the metastable state.

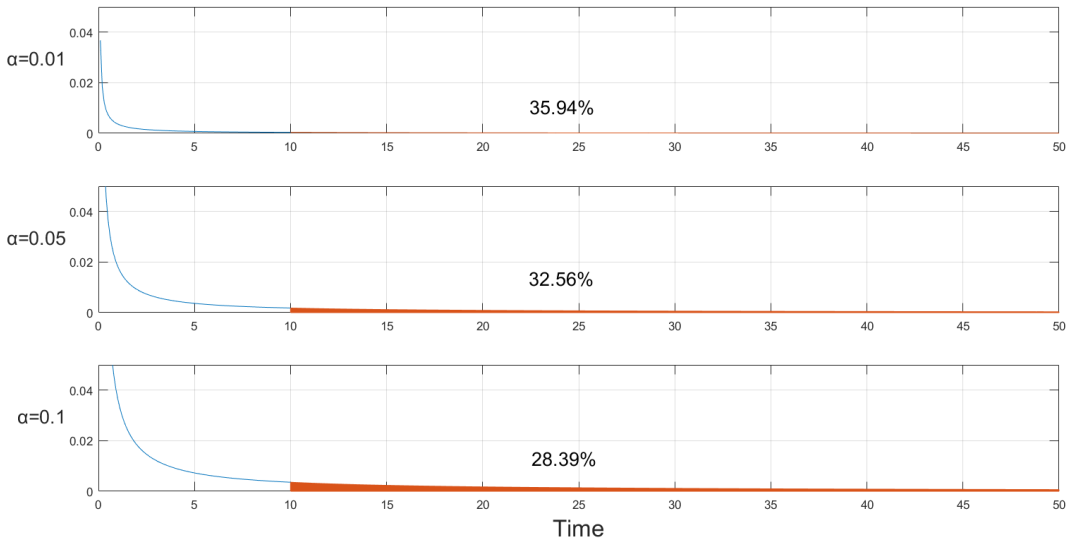


Figure 3.12: The percentage of interval time larger than 10. Comparison among $\alpha = 0.1$, $\alpha = 0.05$ and $\alpha = 0.01$

3.4.3. Summary of Dramatic Effects

To sum up, the impacts of non-Markovian curing processes mainly take up two aspects. On the one hand, the impact on the prevalence. The high peak of Weibull distribution causes a large percentage of nodes getting recovered at the beginning, which influence the initial condition of the entire spreading process and further affect the prevalence at the metastable state. On the other hand, the impact on the epidemic threshold. The heavy-tailed characteristics of Weibull distribution with shape parameter smaller than 0.2 causes a small group of nodes have an extreme long curing time and keep infected state for a long period. This spreading mode makes a small percentage of nodes survive at the metastable state.

3.5. Performance Analysis for Shape Parameter $\alpha \in [10, \infty)$

Considering the case that the shape parameter α is larger than 10, Figure 3.13 indicates that when $\alpha \gg 1$ (e.g. α=10 or 100 or 1000), the average fraction $y_\infty(\tau)$ all have a violent fluctuations at the beginning. And the steady-state average fraction decreases with the shape parameter α increases.

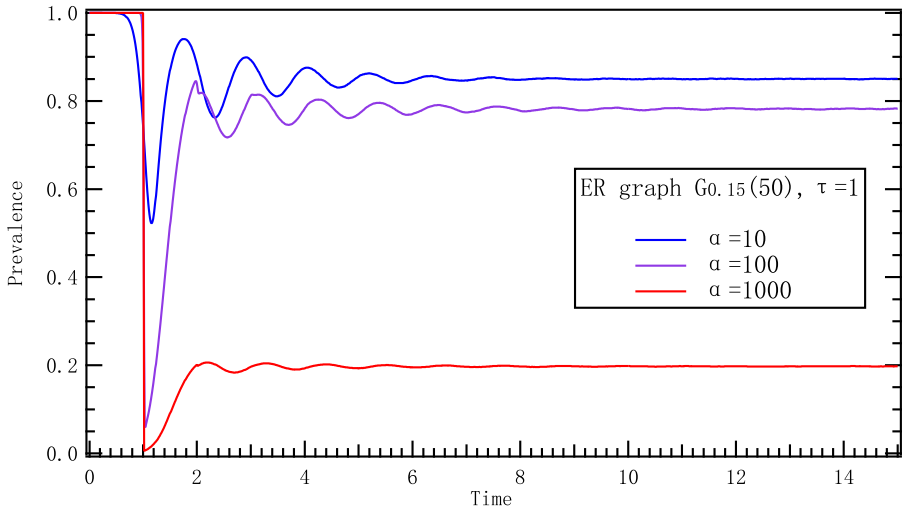


Figure 3.13: The prevalence over time on ER random graph with network size $N = 50$ and number of links $L = 190$ and shape parameter α=10, 100, 1000

That's because when $\alpha > 1$, the Weibull distribution is Gaussian-like. When α is large enough, a large percentage of cures concentrically occur at the same interval. Therefore, a large fraction of infected nodes are cured at the very beginning at the same time interval, and less infected nodes left when α increases, which resulting in a lower prevalence $y_\infty(\tau)$ at the steady state. And that condition is same with shape parameter smaller than 0.2, a large amount of infected nodes get recovered concentrate at the same time. Figure 3.14 depicts the percentage of curing time belongs to [0.99, 1.01] of case $\alpha = 10$, $\alpha = 100$ and $\alpha = 1000$. It shows that with $\alpha = 10$, 7.35% nodes can be seen as get recovered at the same time. And in the case of $\alpha = 1000$, almost 99.86% infected nodes get recovered at the same time. Thus, with shape parameter α increases, the percentage of infected nodes is growing up, thus leading to a smaller percentage of processes survive at the beginning. Figure 3.7 also shows that the percentage of non die-out processes increases with α increases from 500 to 10000. To conclude, when the shape parameter α is extremely large, the initial condition also affected by the focused curing at the first time unit.

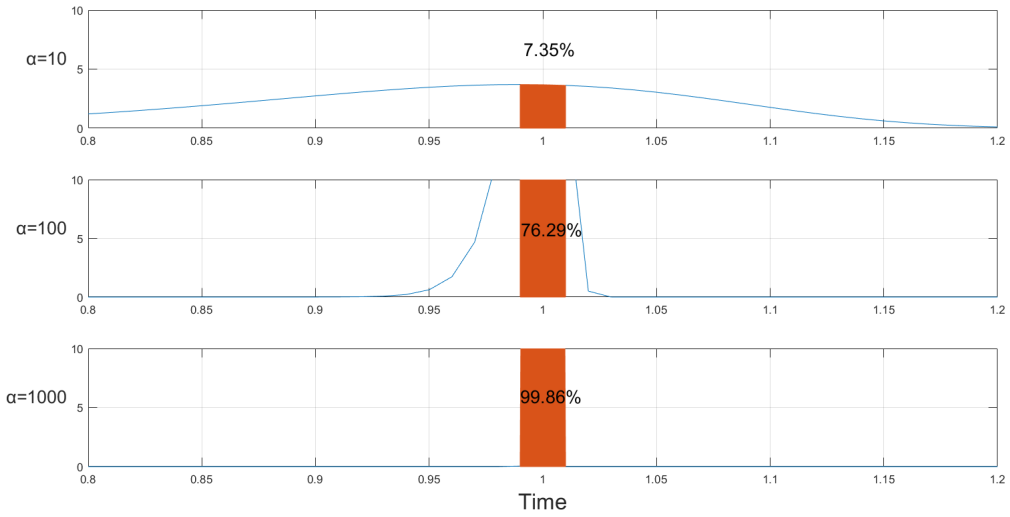


Figure 3.14: The percentage of interval time within [0.99, 1.01]. Comparison among $\alpha = 10$, $\alpha = 100$ and $\alpha = 1000$

Same as the experiments in 3.4.2, we ignore the influence of die-out processes on prevalence and collect data from simulation. Figure 3.15 and 3.16 show that, ignoring the effect on initial condition, the non-Markovian curing process does not influence the prevalence of non die-out processes.

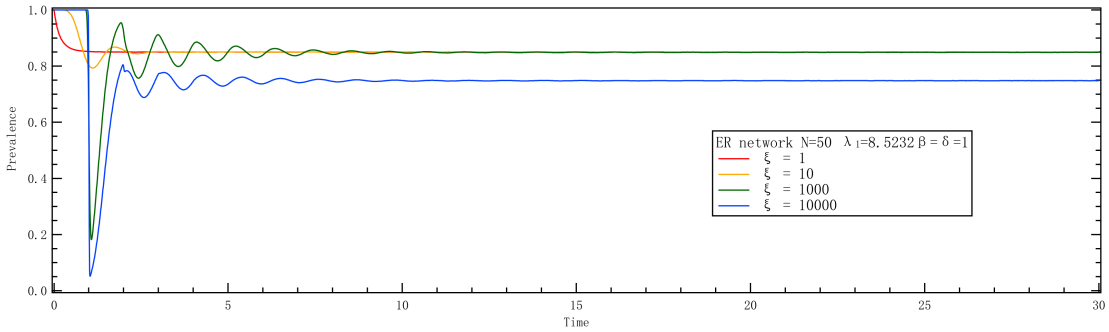


Figure 3.15: The prevalence of all processes over 10^5 realizations on ER network with N=50 and L=190

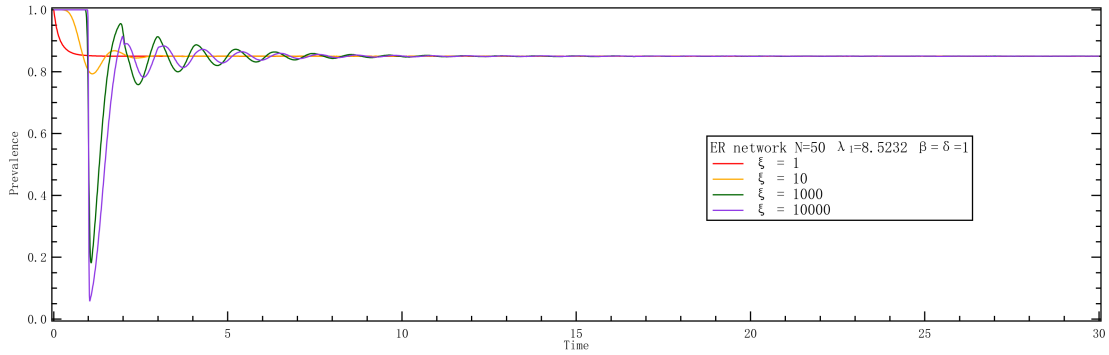


Figure 3.16: The prevalence of non initial die-out processes over 10^5 realizations on ER network with N=50 and L=190

Figure 3.17 depicts that all infected nodes will be cured at the same time. That's because when $\alpha \to \infty$, the Weibull distribution is approximates to a pulse, which means that all the nodes will get cured at the same interval. On the basis of special characteristic of pulse distribution, this strategy can be used to suppress the spreading on some extents, which will be introduced in the chapter 4.

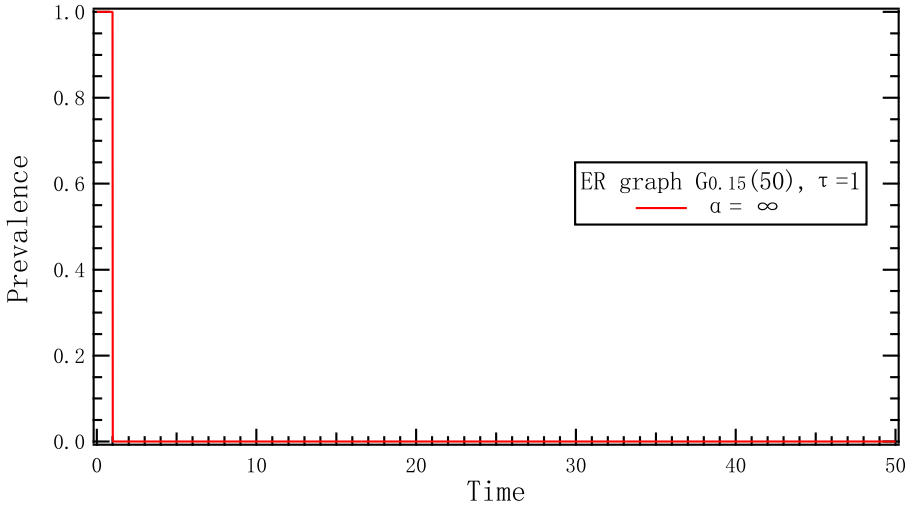


Figure 3.17: The prevalence over time on ER random graph with network size $N = 50$ and number of links $L = 190$ and shape parameter of Weibull distribution is infinite

3.6. Discussion

Based on above simulations and analysis, main effects of non-Markovian curing processes can be summarised into two aspects.

The first aspect is that synchronous curing can suppress spreading on networks. According to above sections, when the shape parameter α is extremely small (smaller than 0.2) or extremely large (larger than 10), there exists a high probability that nodes have similar curing time. In the simulation, all infected nodes synchronously start dynamic spreading processes, which causes a large percentage of nodes get covered at the same time or even all nodes become healthy, thereby first suppressing the spreading processes from the initial state. Similarly, in the process of node getting reinfected and re-cured, a large percentage of nodes get synchronous curing will also inhibit the spreading of disease. If we can control the successful curing probability in some extents, it could effectively suppress the spreading of virus and save medical resources. The detailed discussion will be presented in Chapter 4.

The second aspect is that when shape parameter α is pretty small, the heavy-tailed

characteristics of curing time distribution leads to a special spreading phase that infected nodes have pretty long curing time and low infection rate. Due to the long period of curing, infected nodes can try to infect their neighbors with a long time and in case some of its infected neighbors also have long curing time, the virus can keep surviving for pretty long time. This spreading mode could explain some real virus spreading phenomenon. Though the number of infected nodes is small, it can keep viruses alive for a long time. For example, the initial infection of hepatitis B virus (HBV) [13] rarely results in death and it can spread until the disease get recovered. In this way, hepatitis B virus keeps alive among a small percentage of people [38].

However, there exists some limitations on numerical simulations. First, due to simulations mainly on finite network with fixed topology, the simulation results will definitely have numerical difference. Figure 3.4 and 3.5 show that, when shape parameter α is larger than 1, the epidemic threshold $\tau^\star = \beta/\delta$ is slightly higher than theoretical value of $\tau_c = 1/\lambda_1$. And with the increasing of shape parameter α, the epidemic threshold $\tau^\star$ shows a slight increase and trend to be steady. Besides, although when shape parameter $\alpha \in [0.2, 10]$, the effect of non-Markovian curing processes on prevalence and epidemic threshold is slight or even can be ignored in this case. This range can be changed due to different sizes and typologies of underlying networks. But there is no doubt that, when shape parameter is pretty small and pretty large, the above two impacts of non-Markovian still exist.

4

Pulse Curing Strategy

As we discussed in Chapter 3, when the shape parameter α of Weibull distribution approaches to infinite, the SIS process can be seen as a SI process with periodically curing process. In this process, each node is spontaneously infected by a Poisson process with rate β. And in the pulse curing strategy, a synchronized curing process happened periodically with the successful probability $p < 1$. That means p percentage of infected nodes will be cured in each interval. Therefore, if $p = 1$, then the whole process converges to the all-healthy state.

In this chapter, we theoretically analyze the pulse curing effects under the mean-field theory to obtain the epidemic threshold. Furthermore, we point out that when $p = 1 - 1/\mathrm{e} \approx 0.632$, the pulse curing is equivalent to a Poisson curing process [39].

4.1. Theoretical Analysis of Pulse Curing Effects

Assuming that each node is cured with rate δ, the curing happens every $1/\delta$ time units. So the nodes can only been cured at k/δ with $k = 1,2,\ldots$. With the successful probability p, only a fraction p of nodes are randomly chosen to be cured within every $1/\lambda$ time units. And each infected nodes can infect its susceptible neighbors with rate β. Thus, the effective infection rate is determined by $\tau \triangleq \beta/\delta$. We present the random time $t = k/\delta + t^*$, where $t^* \in [0, 1/\delta)$. Considering the situation that only infection happens with $t^* \neq 0$, the mean-field equation of node i is

$$\frac{dv_i\left(k/\delta+t^*\right)}{dt^*}=\beta\left[1-v_i\left(k/\delta+t^*\right)\right]\sum_{j=1}^{N}a_{ij}v_j\left(k/\delta+t^*\right), \tag{4.1}$$

where $v_i\left(k/\delta+t^*\right)$ is the probability that node i is infected at time $t=k/\delta+t^*$ and a_{ij} is the element of adjacency matrix of the underlying network with N nodes. Since the correlation of the infection states between neighbors was disregarded as in the Markovian SIS process [40], Equation 4.1 is under the mean-field equation. And because the curing probability of each node at k/δ is p, the governing equation of the pulse curing process is

$$v_i\left(\frac{k+1}{\delta}\right)=(1-p)\lim_{t^*\to 1/\delta}v_i\left(\frac{k}{\delta}+t^*\right). \tag{4.2}$$

In the periodically pulse curing process, equation (4.1) does not have an explicit solution of the relationship between $v_i(k/\delta)$ and $\lim_{t^*\to 1/\delta}v_i\left(k/\delta+t^*\right)$ for general network. However, since we only care about the region where $v_i\left(k/\delta+t^*\right)\to 0$ to obtain the epidemic threshold, so we can linearize (4.1) at $v_i\left(k/\delta+t^*\right)=0$ for all i and obtain

$$\frac{d\mathbf{v}\left(k/\delta+t^*\right)}{dt^*}=\beta A\mathbf{v}\left(k/\delta+t^*\right), \tag{4.3}$$

where $\mathbf{v}\left(k/\delta+t^*\right)\triangleq\left[v_1\left(k/\delta+t^*\right),\ldots,v_N\left(k/\delta+t^*\right)\right]^T$. The general solution of (4.3) is $\mathbf{v}\left(k/\delta+t^*\right)=\mathrm{e}^{\beta At^*}C$. And $C=\mathbf{v}(k/\delta)$ is the initial value vector at $t^*=0$. Therefore, the solution of (4.3) evaluated at $t^*\to 1/\delta$ is

$$\lim_{t^*\to 1/\delta}\mathbf{v}\left(k/\delta+t^*\right)=\mathrm{e}^{\tau A}\mathbf{v}(k/\delta). \tag{4.4}$$

Substitute (4.4) into the governing equation (4.2), we obtain

$$\mathbf{v}\left(\frac{k+1}{\delta}\right)=(1-p)\mathrm{e}^{\tau A}\mathbf{v}\left(\frac{k}{\delta}\right). \tag{4.5}$$

The largest eigenvalue of $(1-p)\mathrm{e}^{\tau A}$ is $(1-p)e^{\tau\lambda_1}$. If $(1-p)e^{\tau\lambda_1}$ is smaller than 1, then the infection probability $\mathbf{v}\left(\frac{k}{\delta}\right)$ converges to 0 in the long run. Therefore, we can obtain

the epidemic threshold when $(1-p)e^{\tau\lambda_1} = 1$, which is

$$\tau_c^{(p)} \triangleq \frac{1}{\lambda_1} \ln \frac{1}{1-p}. \tag{4.6}$$

Thus, if $\tau < \tau_c^{(p)}$, the spreading of infection will disappear in the long run while the spreading can persist on the network when if $\tau > \tau_c^{(p)}$.

Comparing with the epidemic threshold of the classical SIS process with a uniform Poisson curing process, which is $\frac{1}{\lambda_1}$, if $\ln \frac{1}{1-p} = 1$, then the pulse curing process can be equivalent to a Poisson curing process in the traditional SIS model on any networks. In this situation, $p = 1 - 1/\mathrm{e} \approx 0.632$. Because the curing rates δ of the pulse curing process and the uniform Poisson curing process are the same, the pulse strategy can only consume 63.2% of the cost of the uniform strategy to suppress the spreading.

4.2. Simulation Verification

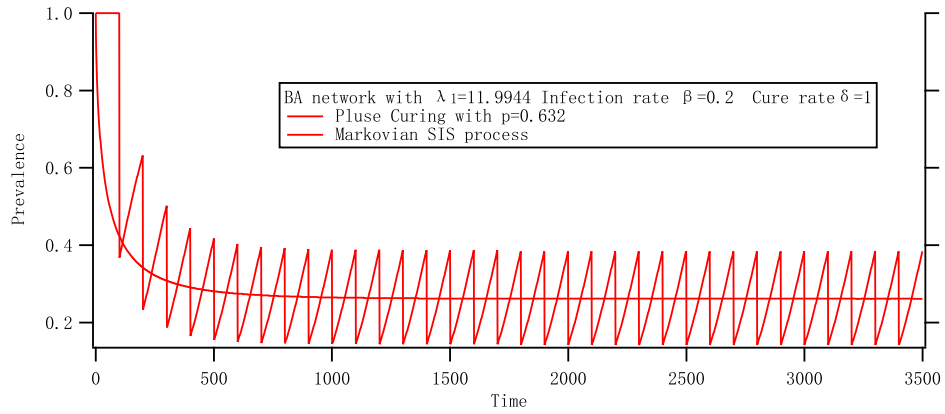


Figure 4.1: The prevalence of the Markovian SIS model and the pulse curing model with curing probability is $p = 0.632$. The simulation is implemented on a Barabási-Albert graph with 500 nodes and $\lambda_1 = 13.6229$. For both models, the infection rate is $\beta = 0.2$ and the curing rate is $\delta = 1$

Figure 4.1 illustrates that the steady prevalence of Markovian SIS process is exactly located in the middle of the curve of the prevalence of Weibullean curing process with curing probability $p = 0.632$. Therefore, the Figure 4.1 shows the conclusion that the pulse curing strategy with 63.2% of the cost required by the uniform strategy can achieve the same performance on the same underlying network structure. Furthermore, Figure 4.2 draws the prevalence of Markovian process and pulse curing process with curing probability $p = 0.632$, which could prove that the conclusion is invariant to the underlying network structure.

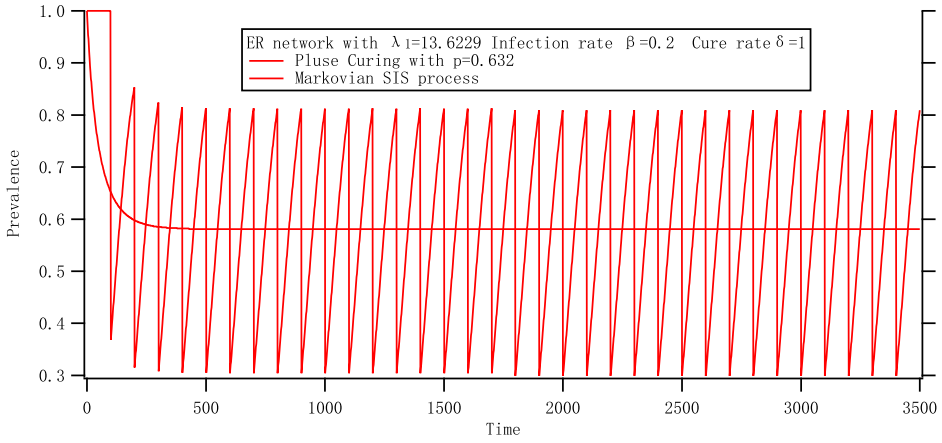


Figure 4.2: The prevalence of the Markovian SIS model and the pulse curing model with curing probability is $p = 0.632$. The simulation is implemented on a Erdős-Rényi Random model with 500 nodes and the largest eigenvalue $\lambda_1 = 11.9944$. For both models, the infection rate is $\beta = 0.2$ and the curing rate is $\delta = 1$

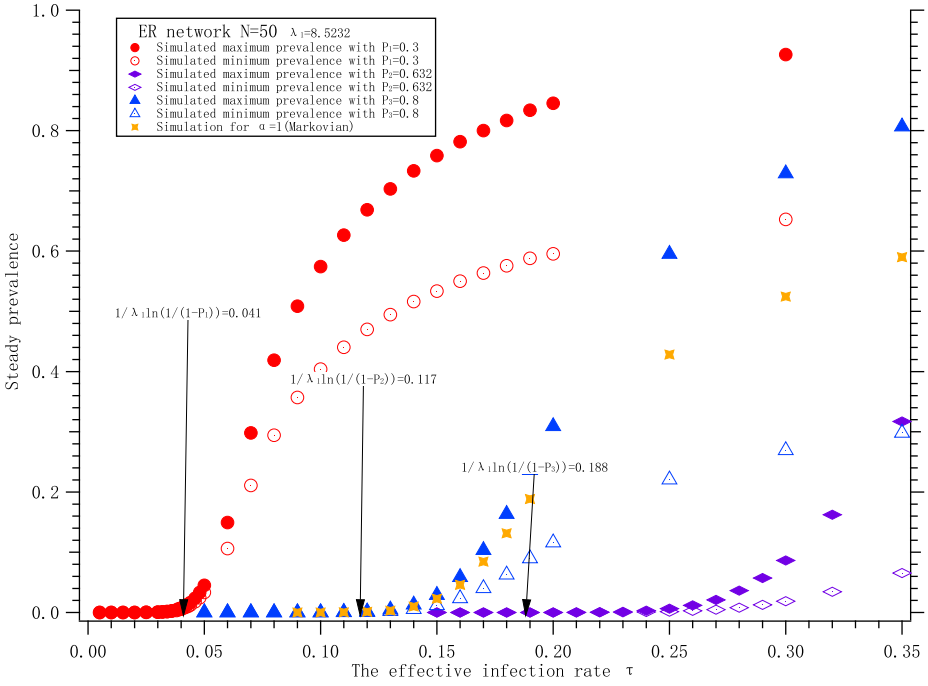


Figure 4.3: The steady prevalence of three pulse curing models with curing probability are $p = 0.3$, 0.632, 0.8. The simulation is implemented on a Erdős-Rényi Random model with 50 nodes and the largest eigenvalue $\lambda_1 = 8.5232$. For three models, the infection rate is $\beta = 1$

In order to observe the effect of the successful possibility p on the epidemic threshold of non-Markovian curing process, we pay attention to the relations between the effective infection rate δ with the prevalence at steady state. For a instance, we simulated the pulse curing process on a ER network with 50 nodes. The star markers presents the line of simple SIS spreading process with Markovian infection and curing process. Figure 4.1 illustrates that when the successful probability of pulse curing strategy is equal to

0.632, the average prevalence is same with the simple Markovian process. Therefore, it shows that the pulse curing strategy can efficiently save the cost of the uniform strategy to restrain the spreading.

While in the pulse curing strategy, the coverage of successful possibility $p < 1$, since if $p = 1$, then synchronous curing kills the spreading immediately. And for the reality perspective, due to the limited resource or other complications, the coverage of successful probability can not achieve full coverage with $p = 1$. Considering the effects on the performance of pulse curing process with $p \in [0, 1)$, Figure 4.3 plots the performance with successful probability selected on the both sides of $p = 0.632$. Both cases are simulated on the same Erdős-Rényi random graphs with same spreading process. With the successful probability $p = 0.3$, which is smaller than 0.632, the epidemic threshold is smaller than the Markovian process, while the steady prevalence is much larger than the general process. Conversely, with successful probability $p = 0.8$, the steady prevalence of pulse curing process is smaller than the Markovian process, but the epidemic threshold is much larger. That means more curing resources are consumed. Therefore, the pulse strategy is optimal when the successful probability is equal to 0.632.

4.3. Optimal Plan of Pulse Strategy

According to equation (4.6), we can get the relation between effective infection rate τ with the optimal successful curing probability p, which is shown in Figure 4.4 in red line. We consider the red line as the optimal plan of pulse strategy for suppressing the viral spreading. Figure 4.4 and 4.5 also present the changing process of prevalence at the metastable state starts with specific effective infection rate τ and specific curing probability p. Therefore, we can get similar optimal plan in case we know the characteristics of the underlying networks. If we know the effective infection rate of a viral spreading process, according to the red line, we can choose the optimal successful curing probability p to suppress the spreading. Furthermore, if we want to control the prevalence at the metastable state, it is also convenient to set up curing possibility p and even change effective infection rate τ to suppress disease spreading and save medical resources.

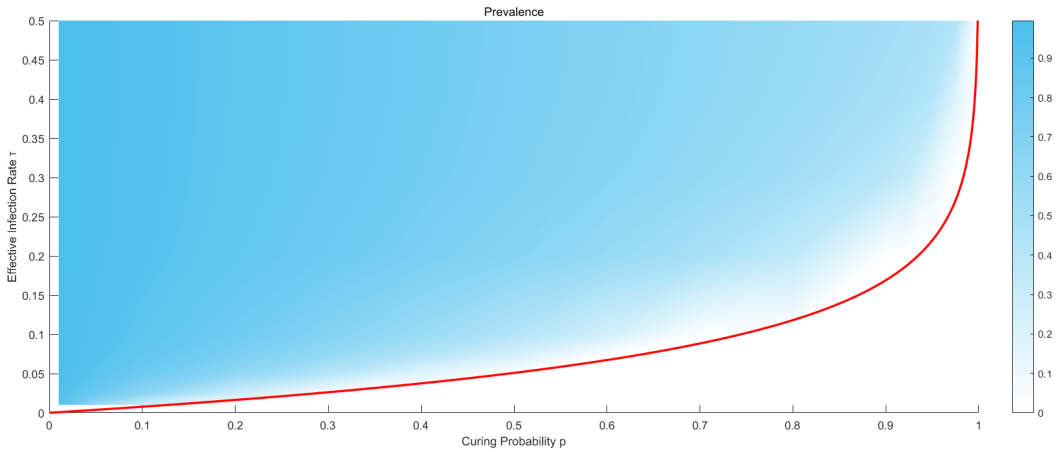


Figure 4.4: Heat map with optimal plan over time on ER random graph with $N = 50$ and the largest eigenvalue $\lambda_1 = 8.5232$

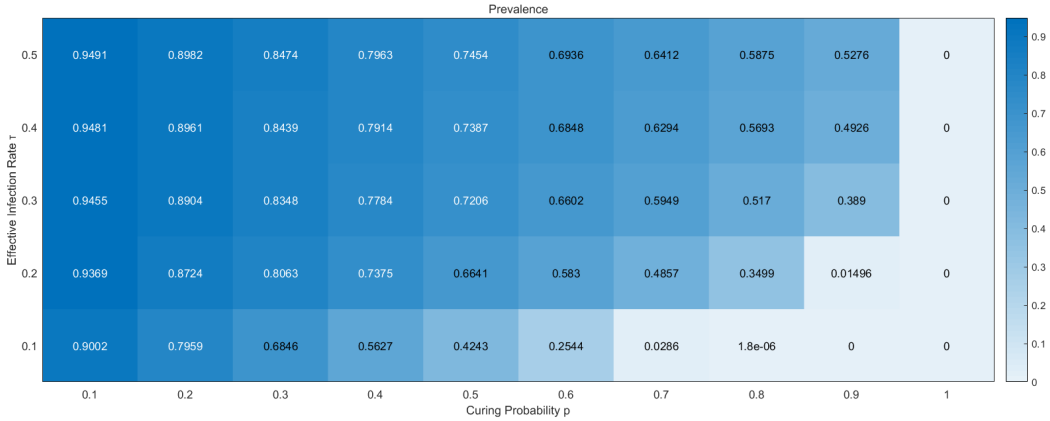


Figure 4.5: Heat map with prevalence at the metastable state over time on ER random graph with $N = 50$ and the largest eigenvalue $\lambda_1 = 8.5232$

5

Conclusions and Future Work

This chapter will summarize outcomes of this work and provide suggestions for further investigation. Section 5.1 presents conclusions based on our results and discussions, and Section 5.2 provides suggestions for future research.

5.1. Conclusions

In this thesis, we focus on the effects of non-Markovian curing processes on the spreading performance. Based on numerical simulations on finite networks, two dramatic impacts are shown in the research. The first impact is that synchronous curing can suppress viral spreading on networks and further optimal setting of curing probability can effectively save the medical resources. The second impact is that the heavy-tailed characteristics of curing time distribution leads to a spacial spreading phase that infected nodes have pretty long curing time and keep survive among a small group of individuals. This spreading of hepatitis B virus (HBV) is similar to this mode. Comparing with the theoretical conclusion that the epidemic threshold is $\tau_c = 1/\lambda_1$ if infection process following Poisson process with any distribution of curing time R, the non-Markovian curing processes have dramatic effects on the performance of viral spreading. With the shape parameter α getting pretty smaller or pretty larger, the effects are more obvious.

5.2. Recommendations for Future Work

This thesis is mainly a systematic study of the effects of non-Markov curing processes on SIS epidemics based on numerical simulations. Although we can find dramatic effects through the simulation results on prevalence and phase transition, the theoretical explanations are still required. Figure 3.4, 3.5 and 3.6 present that the epidemic threshold arrives at the lowest point when shape parameter α is approximately equal to 0.18. It is difficult to explain this performance only from the analysis, theoretical deviation and verification are still needed to support this result. Furthermore, if we can get the relationship between this lowest point of epidemic threshold value and other factors from a theoretical point of view, we can apply this conclusion in other cases. For example, when we know an underlying network, we can promote the spreading with applying the corresponding shape parameter at lowest epidemic threshold. Finally, a real-life epidemic may follow a more complex epidemic model. If both the infection process and curing process are following non-Markovian process, it could be quite interesting to know the spreading phenomenon on networks.

Appendices

A

Step Verification of Theoretical Conclusions

We reproduce the theoretical derivation results of each step in combination with the simulations to compare the simulation results with theoretical derivation results on the basis of the deviation process in [33], which is also described in details in 2.3.2.

Procedure :
For the simplicity and generality of the calculation, we apply the gamma distribution to distribute the cure time in this simulation. And ξ is the shape parameter of gamma distribution. According to (2.23), the average number of times that node j tries to infect node i is equal to $E[M]$, which should equal to the rate ratio of $\tau = \frac{\beta}{\delta}$. Based on the definition of $E[M]$, we first count the average number of times that an infected node infects one of its neighbors within a random cure interval, which is the simulation results of $E[M]$. On the basis of 10^5 realizations, we can verify the simulation results by comparing with the equation (2.23). At the same time, we run the simulation with three different infection rates $\beta = 0.5$, 1 or 2. Besides, in order to figure out the effect of various cure time distributions, we plot the relationship between $E[M]$ and shape parameter ξ, which is shown in Figure A.1. Since only when the parameters are too small, the simulation results are far from the theoretical results, so we only consider the case where the parameters are less than 0.1.

Second, in order to explore the impact of simulation time S on simulation results, according to the equation 2.21 and the probability generating function of gamma function, we can get the value of $E[M]$ based on calculation. Therefore, we assume a node is linked to an infected node and this node starts with the susceptible state. During the

simulation, we assume the node can be infected instantly and the cure time following the gamma distribution with different shape parameter $\xi = 1, 0.01, 0.001$ and cure rate $\delta = 1$. Set with different time period S, we can get the number of cure events occurred during S, which is shown in Figure A.2. By comparing the calculation results and simulation results, we can figure our the effect of simulation time S on results, which is shown in Figure A.3.

Third, In order to exclude the effect of infection process, we focus on the value of the number of infected periods. With setting cure rate $\delta = 1$, the number of infected periods is equal to $v_j S/E[R] = v_j S$. On one hand, we account the number of infected periods in simulator with different infection rate $\beta \in [0.1, 1]$. One the other hand, based on the simulation setting on Matlab in the second step, we fixed time S as a large time interval $S = 10^5$ and calculate the number of infection periods with infection rate from 0.1 to 1. Comparing two results from Matlab and simulator, we can know the effect of infection process on prevalence.

Performance :
The first step simulates the number of times that node j tried to infect node i during each infected period of node i (also the recovering period of node i). And based on results of 10^5 realizations, Figure A.1 illustrates the average number $E[M]$ of times over different shape parameters ξ ranging from 0.01 to 0.1. Simulating with cure rate $\delta = 1$, Figure A.1 shows that $E[M]$ is equal to the value of $\tau = \frac{\beta}{\lambda}$, and $E[M]$ does not change with different distribution of cure time R. Thus, we can prove that in the N-interwined mean-field approximation, the simulation results of $E[M]$ are consistent with the theoretical results.

The second step of simulation is to inspect the effect of simulation time S on the average number $E[M]$ of times. Since the simulator assumed when the spreading process runs for a long enough time with 50 time units, the whole process is in the meta-stable state. At the same time, the number of times of infection events calculated by equations (2.15) and (2.21) in a certain period of time S is compared with the simulation results.

Therefore, Figure A.2 plots that the process with a smaller shape parameter has a larger number of infection events occurred within short period of time S. And with the period of time S is larger than 10^3, three processes have the same number of infection events happened. Besides, the simulation results can be approved by the calculation results based on equation (2.15) and (2.21). To conclude, the simulation results is consistent with the theoretical results. What's more, the number of infection events with different cure-time distributions within a long enough time period S have the same number of

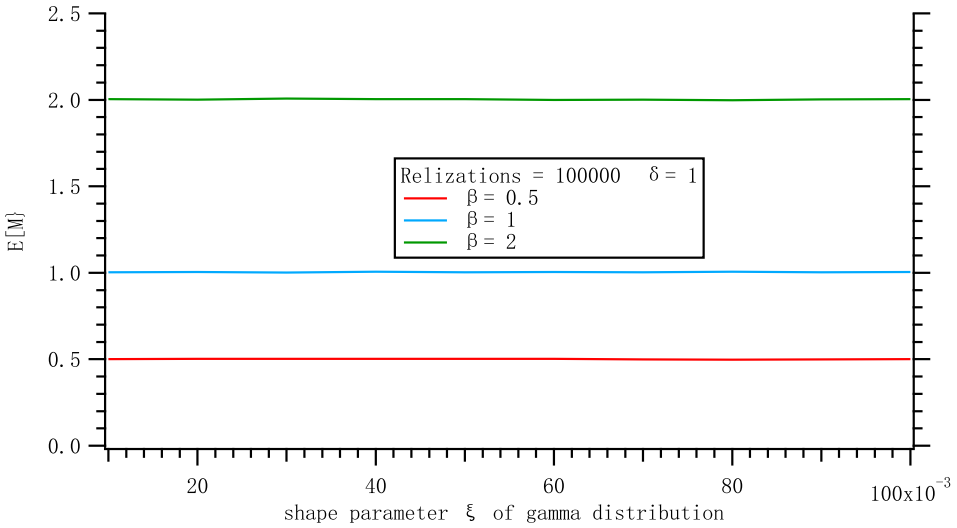


Figure A.1: The average number of $E[M]$ over various shape parameter $\xi \in [0.01, 0.1]$ based on 10^5 realizations with cure rate $\delta = 1$

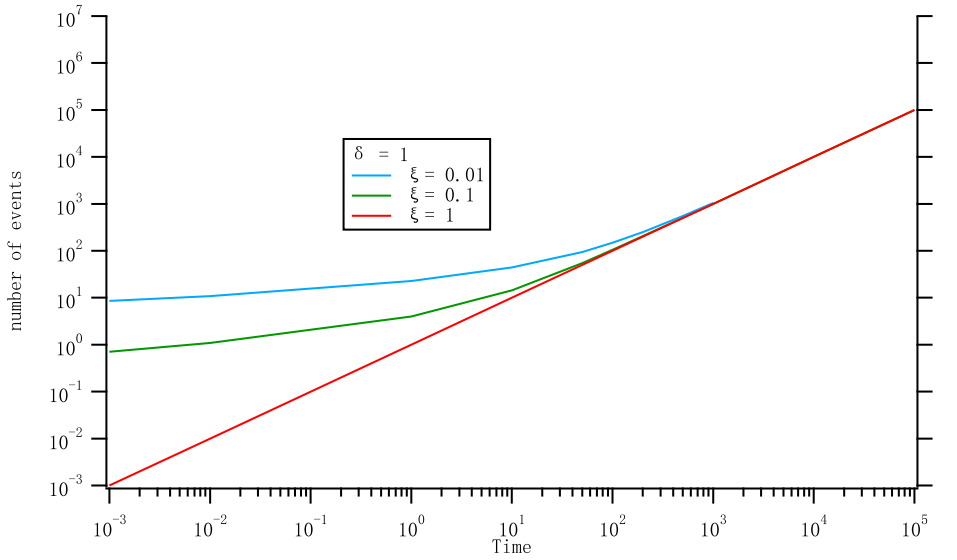


Figure A.2: The number of infection events over time

events happened. According to the Figure A.2, when the length of interval S is not long enough to make spreading process spreading into meta-stable state, S will slightly influence the prevalence. Since the simulator assumed the process will arrive at the meta-stable state after 50 time units. Thus, The result of process with small shape parameter α is not exactly the prevalence at the meta-stable state. To conclude, enough long time interval S is one of conditions to guarantee that the simulation results is consistent with theoretical results.

Based on the performance of the second-step simulation, we compare the number of infection periods from simulator and experiment on Matlab. Figure A.3 shows that

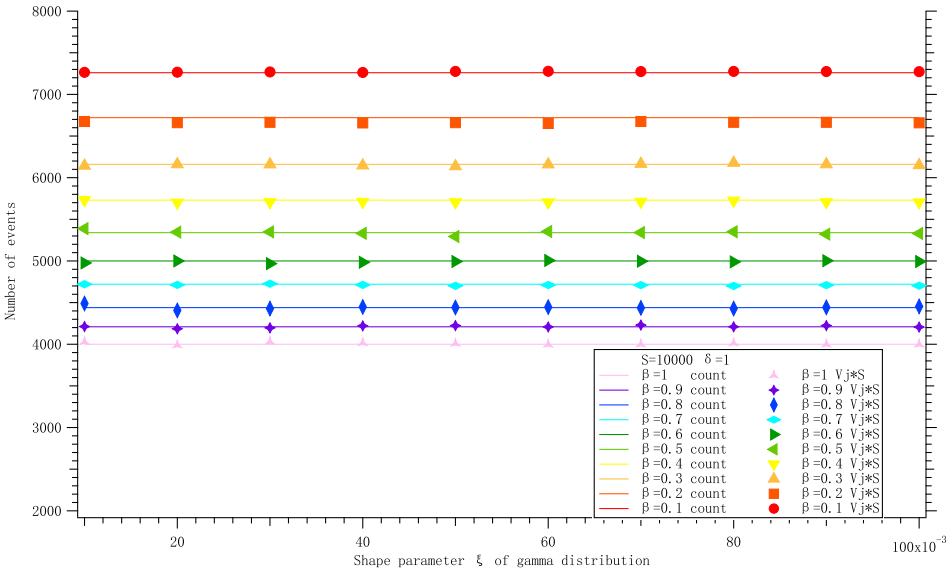


Figure A.3: The number of infection periods with different infection rate.

the result of experiment fluctuates around the account value from simulator. And both results are not influenced by the infection rate ranging from 0.1 to 1. Therefore, we can exclude the affect of infection process. And we set the infection rate $\beta = 1$ for the convenience in the following.

To conclude the above three experiments, ignoring the slight influence of time S, we can find the factor which makes the differences between the simulation results and the experimental results. And the simulation results is consistent with each deviation step.

Bibliography

[1] W. Wang, Q.-H. Liu, J. Liang, Y. Hu, and T. Zhou, "Coevolution spreading in complex networks," *Physics Reports*, 2019.

[2] N. B, *A Short History of Mathematical Population Dynamics*. Springer-Verlag London, Ltd, London, 2011.

[3] A. Vespignani, "Modelling dynamical processes in complex socio-technical systems," *Nature physics*, vol. 8, no. 1, p. 32, 2012.

[4] K. W. M. A.G., "A contribution to the mathematical theory of epidemics," *Proc. R. Soc. A*, vol. 115, no. 772, pp. 700–721, 1927.

[5] D.J.Daley and J.Gani, *Epidemic Modelling: An Introduction*. Cambridge University Press, Cambridge, UK, 1999.

[6] S. S. et al, "Identifying highly connected counties compensates for resource limitations when evaluating national spread of an invasive pathogen," *PLoS One 7*, vol. e37793, 2012.

[7] Y. Zhao and D. Jiang, "Dynamics of stochastically perturbed sis epidemic model with vaccination," in *Abstract and Applied Analysis*, vol. 2013. Hindawi, 2013.

[8] Z. Chen and C. Ji, "Spatial-temporal modeling of malware propagation in networks," *IEEE Transactions on Neural networks*, vol. 16, no. 5, pp. 1291–1303, 2005.

[9] R. Vida, J. Galeano, and S. Cuenda, "Vulnerability of state-interdependent networks under malware spreading," *Physica A: Statistical Mechanics and its Applications*, vol. 421, pp. 134–140, 2015.

[10] H. S. Rodrigues, "Application of sir epidemiological model: new trends," *arXiv preprint arXiv:1611.02565*, 2016.

[11] Z.-K. Zhang, C. Liu, X.-X. Zhan, X. Lu, C.-X. Zhang, and Y.-C. Zhang, "Dynamics of information diffusion and its applications on complex networks," *Physics Reports*, vol. 651, pp. 1–34, 2016.

[12] P. Van Mieghem, "Exact markovian sir and sis epidemics on networks and an upper bound for the epidemic threshold," *arXiv preprint arXiv:1402.1731*, 2014.

[13] R. Parveen, B. D. Gen, and L. H. Qing, "New awareness review on hepatitis b prevalence, transmission, and related complications."

[14] L. J. Allen, "Some discrete-time si, sir, and sis epidemic models," *Mathematical biosciences*, vol. 124, no. 1, pp. 83–105, 1994.

[15] D. Chakrabarti, Y. Wang, C. Wang, J. Leskovec, and C. Faloutsos, "Epidemic thresholds in real networks," *ACM Transactions on Information and System Security (TISSEC)*, vol. 10, no. 4, p. 1, 2008.

[16] S. Peng, S. Yu, and A. Yang, "Smartphone malware and its propagation modeling: A survey," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 2, pp. 925–941, 2013.

[17] J. S. O. Van Mieghem, P. and R. E. Kooij, "Virus spread in networks," *IEEE/ACM Transaction on NetworkingE*, vol. 17, no. 1, pp. 1–14, February. 2009.

[18] T. M. Tirpak, "Telecommunication network resource management based on social network characteristics," Nov. 11 2010, uS Patent App. 12/463,445.

[19] N. C. Valler, B. A. Prakash, H. Tong, M. Faloutsos, and C. Faloutsos, "Epidemic spread in mobile ad hoc networks: Determining the tipping point," in *International Conference on Research in Networking*. Springer, 2011, pp. 266–280.

[20] F. Harary and G. Gupta, "Dynamic graph models," *Mathematical and Computer Modelling*, vol. 25, no. 7, pp. 79–87, 1997.

[21] E. Otte and R. Rousseau, "Social network analysis: a powerful strategy, also for the information sciences," *Journal of information Science*, vol. 28, no. 6, pp. 441–453, 2002.

[22] Y. Iturria-Medina, R. C. Sotero, E. J. Canales-Rodríguez, Y. Alemán-Gómez, and L. Melie-García, "Studying the human brain anatomical network via diffusion-weighted mri and graph theory," *Neuroimage*, vol. 40, no. 3, pp. 1064–1076, 2008.

[23] F. Harary, "The determinant of the adjacency matrix of a graph," *Siam Review*, vol. 4, no. 3, pp. 202–210, 1962.

[24] P. Erdös, "Graph theory and probability," *Canadian Journal of Mathematics*, vol. 11, pp. 34–38, 1959.

[25] P. Erdős and A. Rényi, "On the evolution of random graphs," *Publ. Math. Inst. Hung. Acad. Sci*, vol. 5, no. 1, pp. 17–60, 1960.

[26] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Reviews of modern physics*, vol. 74, no. 1, p. 47, 2002.

[27] A.-L. Barabási, E. Ravasz, and T. Vicsek, "Deterministic scale-free networks," *Physica A: Statistical Mechanics and its Applications*, vol. 299, no. 3-4, pp. 559–564, 2001.

[28] C. C. P. V. M. Pastor-Satorras, R. and A. Vespignani, "Epidemic processes in complex networks," *Reviews of Modern Physics*, vol. 87, no. 3, pp. 925–979, 2015.

[29] M. Jambunathan, "Some properties of beta and gamma distributions," *The annals of mathematical statistics*, pp. 401–405, 1954.

[30] P. Van Mieghem, *Performance Analysis of Complex Networks and Systems*. Cambridge University Press, Cambridge, U.K., 2014.

[31] R. Jiang and D. Murthy, "A study of weibull shape parameter: Properties and significance," *Reliability Engineering & System Safety*, vol. 96, no. 12, pp. 1619–1626, 2011.

[32] R. A. Ibrahim, *Handbook of Structural Life Assessment*. Wiley Online Library, 2017.

[33] R. v. d. B. E. Cator and P. V. Mieghem, "Susceptible-infected-susceptible epidemics on networks with general infection and cure times," *Physical Review E*, vol. 87, no. 6, p. 062816, June. 2013.

[34] P. Van Mieghem and R. van de Bovenkamp, "Non-markovian infection spread dramatically alters the susceptible-infected-susceptible epidemic threshold in networks," *Physical Review Letters*, vol. 110, no. 10, p. 108701, Mar. 2013.

[35] Q. Liu and P. V. Mieghem, "Burst of virus infection and a possibly largest epidemic threshold of non-markovian susceptible-infected-susceptible processes on networks," *Physical Review E*, vol. 97, no. 2, p. 022309, Feb. 2018.

[36] R. van de Bovenkamp, "Epidemic processes on complex networks : Modelling, simulation and algorithms," Ph.D. dissertation, Delft University of Technology, 2015.

[37] R. V. D. BOVENKAMP, "Epidemic processes on complex networks: Modelling, simulation and algorithms," diploma thesis, Technische Universiteit Delft, 2015.

[38] Hepatitis b. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/hepatitis-b

[39] Q. Liu, X. Zhou, and P. Van Mieghem, "Pulse strategy for suppressing spreading on networks," *arXiv preprint arXiv:1904.10883*, 2019.

[40] E. Cator and P. Van Mieghem, "Nodal infection in markovian susceptible-infected-susceptible and susceptible-infected-removed epidemics on networks are non-negatively correlated," *Physical Review E*, vol. 89, no. 5, p. 052802, 2014.