UNIVERSITY OF TECHNOLOGY SYDNEY

# [ Measurements and Evolution of Complex Networks with Propagation Dynamics ]

**by [ Bo Song ]**

Thesis submitted in fulfilment of the requirements for the degree of

**[Doctor of Philosophy]**

under the supervision of [ Prof. Y. Jay Guo, Prof. Ren Ping Liu ]

University of Technology Sydney
Faculty of [ Engineering and Information Technology ]

[ 09 2020 ]

# Certificate of Authorship/Originality

I, Bo Song declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the Faculty of Engineering and Information Technology at the University of Technology Sydney. This thesis is wholly my own work unless otherwise reference or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis. I certify that the work in this thesis has not been previously submitted for a degree nor has it been submitted as a part of the requirements for other degree except as fully acknowledged within the text. This thesis is the result of a research candidature jointly delivered with Nanjing University of Posts and Telecommunications as part of a Collaborative Doctoral Research Degree. This research is supported by the Australian Government Research Training Program.

Signature: Production Note:
Signature removed prior to publication.

Date: 21/9/2020

# Dedication

*This thesis is dedicated to my families.*
*This stands as a testimony for their endless support and love.*

*To my supervisors, for the academic guidance.*

*To my friends, for their encouragement.*

# Acknowledgements

My parents always told me to be grateful, so I would start by thanking my perfect families, my parents Yuzhen Song and Yanyun Zhao, my brother Tao Song. They support me to pursue research, allow me to be myself. Many Thanks!

Throughout the doctoral program, I have received a great deal of support and assistance. I would like to express my sincere gratitude to Prof. Y. Jay Guo for his invaluable support and immense knowledge. My deepest thanks also go to Prof. Ren Ping Liu for all the opportunities he offered and his patience and help. I would like to show my very profound gratitude to Dr. Wei Ni in CSIRO. My research would have been impossible without his support and supervision. I would also like to acknowledge Prof. Guoping Jiang and Prof. Yurong Song in NJUPT. I would also like to thank everyone in GBDTC and UTS who helped me so much.

I must express my sincere thanks to my co-author of papers, Lingling Xia, for the company throughout the whole Ph.D. period in China. A very special gratitude goes out to my mates Xu Wang, for all the help and unforgettable days with him in Sydney. Many thanks to my team partners in NJUPT and UTS, it is great having the awesome time with you guys. I am also grateful to my families and friends who have supported me along the way.

Bo Song

Nanjing, China, 2020

# List of Publications

**Published Journal Papers**

J-1. **Bo Song**, Xu Wang, Wei Ni, Yurong Song, Ren Ping Liu, Guo-Ping Jiang and Y. Jay Guo, Reliability Analysis of Large-Scale Adaptive Weighted Networks, IEEE Transactions on Information Forensics and Security. 2019, 15: 651-665.

J-2. **Bo Song**, Guo-Ping Jiang, Yurong Song, Ling-Ling Xia, Dynamic rewiring in adaptive weighted heterogeneous networks. International Journal of Modern Physics B, 2019, 33(9): 1950069.

J-3. **Bo Song**, Zhen-Hao Zhang, Yurong Song, Guo-Ping Jiang, Yin-Wei Li, Xiao-Ping Su, Preferential redistribution in cascading failure by considering local real-time information. Physica A: Statistical Mechanics and its Applications, 2019, 532(2019): 121729.

J-4. **Bo Song**, Guo-Ping Jiang, Yurong Song, Ling-Ling Xia, Rapid identifying high-influence nodes in complex networks. Chinese Physics B, 2015, 24(10): 100101.

J-5. Xu Wang, **Bo Song**, Wei Ni, Ren Ping Liu, Y. Jay Guo, Xinxin Niu, and Kangfeng Zheng Group-Based Susceptible-Infectious-Susceptible Model in Large-Scale Directed Networks. Security and Communication Networks, 2019: 1657164.

J-6. Ling-Ling Xia, **Bo Song**, Zheng-Jun Jing, Yurong Song, Liang Zhang, Guo-Ping Jiang, Dynamical interaction between information and disease spreading in populations of moving agents. CMC (Computers Materials & Continua), 2018: 123-144.

**Published Conference Papers**

C-1. **Bo Song**, Guo-Ping Jiang, Yurong Song, Epidemic dynamics in weighted adaptive networks. 3PGCIC 2014, pp: 69-73, January 2, Guangzhou, China, 2015.

C-2. **Bo Song**, Yurong Song, Guo-Ping Jiang, How clustering affects epidemics in complex networks. 2017 International Conference on Computing, Networking and Communications (ICNC). IEEE, Silicon valley, USA, 2017: 178-183.

**Accepted Papers**

S-1. **Bo Song**, Zhengjun Jing, Yingjie Jay Guo, Renping Liu, Qian Zhou. A novel measure to quantify the robustness of social network under the virus attacks, 6th International Symposium on Security and Privacy in Social Networks and Big Data, 2020.

# List of Figures

# List of Tables

# List of Abbreviations

| | |
|---|---|
| BA networks | Barabási-Albert networks |
| DDoS attack | Distributed denial-of-service attack |
| ER networks | Erdös-Rényi networks |
| HITS | Hyperlink-Induced Topic Search |
| NFV | Network function virtualization |
| PA networks | Preferential attachment networks |
| SDB | Spontaneous Defense Behavior |
| SEIR model | Susceptible-Exposed-Infected-Recovered model |
| SF network | Scale-free network |
| SI model | Susceptible-Infected model |
| SIS model | Susceptible-Infected-Susceptible model |
| SIR model | Susceptible-Infected-Recovered model |
| SQB | Spontaneous Quarantine Behavior |
| VANETs | Vehicular ad-hoc networks |
| VNFs | Virtual network functions |
| VMs | Virtual machines |
| WHO | World Health Organization |
| WS networks | Watts-Strogatz networks |
| WWW | World Wide Web |

# ABSTRACT

## Measurements and Evolution of Complex Networks with Propagation Dynamics

With the development of technology, we live in a world which is surrounded by complex networks, e.g., the power grid, transportation network, Internet, neural networks, social networks. Understanding the structure and dynamics of these extremely complex interactive networks has become one of the key research topics and challenges of life science in the 21st century. For example, the coronavirus disease 2019 (COVID-19) pandemic markedly changed human mobility patterns, hygienic habits and the communication methods. In order to control the virus spread, it is very necessary to analyze the network structures and epidemic dynamics, e.g., the importance of nodes in the networks, the influence of network structure measurements on propagation, the interaction between propagation dynamics and the structure measurements, and the construction of epidemiological models that can capture the effects of these changes in mobility on the spread of virus. Meanwhile, the results of these studies can also be used as a reference for the study of multiple propagation behaviors in other networks.

Complex network theory is to study the commonness of these seemingly different complex networks and the universal methods to deal with them. In 1998 and 1999, the finding of small world effects and scale-free property has attracted a great deal of attention of network structures and dynamics, which raises the science awareness for the real world. After the discovery of small world effects and scale-free property of networks, researchers gradually realize and study the complexity of networks. More network structure metrics are proposed, and more network characteristics are found with the development of complex network research. For example, many networks have community structures, e.g., the families, the schools, in which the internal connection of the community is much closer than the external connection. Meanwhile, studies on network structure related to the structure metrics are also in progress, such as the node influence identification, the community structure mining and the link prediction.

As one of the main subjects in the field of complex network theory, the study on dynamical behaviors in complex networks has assumed greater importance and attracted wider attention since the spreading phenomena on different type of real-world networks affect significantly human activities in social and economic environments. For example, the epidemic spread in the crowd, the cascading failures in the power grid and the information diffusion in online social networks. It is pointed out that the network structure measurements have an important impact on the propagation processes. For example, the epidemic threshold tends to zero in scale-free networks, which means that the virus is very easy to spread in scale-free networks because of a small minority of ‘super-spreader’. Compared with the scale-free network, the epidemic in the small world network is more difficult to break out due to the existence of the non-zero epidemic threshold.

While the network structure affects the propagation dynamics, the spreading process is also changing the network structure. For example, when a virus breaks out, people will selectively avoid symptomatic infected persons to protect themselves. The network structure has been changing dynamically due to the evasive behaviors, and the dynamic change in structure can also affect the spread of the virus in turn. In short, the network structure and the propagation dynamics in the network are co-evolving.

In the thesis, the influence of complex network structure measurements on the propagation processes and the dynamic relationship between network structures and the propagation processes are studied. Firstly, the influence of network structure measurement on the propagation process is studied and applied to the process of node influence identification, cascading failure and virus propagation. Based on the degree value of the nodes, a method to quickly identify the influence of the nodes, as well as a cascaded failure model considering the local real-time information priority redistribution strategy, is proposed, and a novel metric is proposed to measure the robustness in regard to virus attacks in social networks. Following on from this, the cooperative evolution of network structure and propagation process is studied, and the reliability of adaptive weighted networks is analyzed and discussed.

# CONTENTS

# Chapter 1

# Introduction

The rapid development of the Internet has brought our lives into the information network era. A large number of studies have shown that most networks and systems in nature and in human society, e.g., food chain networks, social networks, communication networks, transportation networks, etc., can be described as complex networks [1]-[5]. As a theory and tool for description and analysis, complex network provides a fresh view for the studies of the structures and dynamics of real-world complex systems.

In recent years, with the emergence and rapid development of computer technology, networks have brought unprecedented experiences to human life. With the deepening of the application of computer networks and the rapid expansion of the fields for those applications, both the concept and connotation of the idea of a "network" are constantly being updated. Conversely, the characteristics of networks, i.e., rapidity, efficiency, virtuality and openness, enable our production, life, communication and ways of thinking be extended across time and space. However, the gradual deepening of these networks is like a double-edged sword, which not only brings convenience to people's lives, but also brings many adverse effects and new challenges. For example, the rapid development of the transportation network makes it easier for biological viruses to spread around the world [6]-[9], computer viruses can quickly spread to all parts of the world through the Internet, causing adverse effects to network users [10]-[13], various rumors quickly spread through social networks [14]-[16], the personal and private details of users are invaded and sold on the internet unscrupulously [20]-[22], and large-scale cascading failures occur in power grids due to overloads [17]-[20].

Therefore, it is of great practical significance to study and understand these complex networks and the relationships between network structures and network dynamics. On the one hand, the research of complex network structure measurement and spreading behaviors can solve many problems in real world networks, such as establishing a corresponding propagation model for the spread of viruses through social networks, a practical cascading failure model for node loads and capacity in power grids, etc. On the other hand, the increasing degree of human social networking and the more complex network structures provide prospects for broad applications to complex networks.

Against the above backdrop, the thesis presents a study of the complex network dynamics process. It reveals the influence of network structure measurements on spread processes and the cooperative evolution and interaction between network structures and spread dynamics. Some effective network spread dynamics control strategies are proposed based on the studies. Furthermore, based on the similarity of spread processes, such as virus spread, information diffusion, and cascading failures, the thesis also provides some new concepts and methods for the study of multiple spread processes in complex networks.

## 1.1 Background

The earliest literature on network research is the famous Königsberg's seven bridge problem written by Leonhard Euler in 1736 [23], [24], where the mathematical descriptions of vertices and edges became the basis of graph theory, which further became a branch of mathematics that studies the properties of pairwise relations in complex network structures. Subsequently, graph theory continues to be developed and applied to different fields. In the 1950s, two Hungarian mathematicians, Erdös and Rényi, established the random graph theory [25], and created a systematic study of complex network theories in mathematics. In the 1990s, the study of complex networks entered a new era. In 1998, Duncan J Watts and Steven H Strogatz published a paper entitled “Collective dynamics of 'small-world' networks” [26] in the journal Nature, where the Small-World network model was proposed. In 1999, physicist Albert-László Barabási and his doctoral student Réka Albert presented a scale-free network model in a paper entitled “Emergence of scaling in random networks” [27]. These two kinds of network characteristics can accurately reflect the structures of real-world networks, deepening our understanding of real networks.

In the study of network theory, a complex network can be regard as a network structure composed of a large number of nodes and the complex relationships between those nodes. In the language of mathematics, it is a graph with complex topological structure features. Complex networks have characteristics that simple network structures, e.g., lattice networks and random graphs, do not have, and these characteristics often appear in real-world systems. Therefore, these highly complex systems are closely related to the research of complex networks. With the furthering and development of research, the academic field of complex networks, which studies real-world systems such as telecommunication networks, computer networks, biological networks and social networks, is called the field of network science. As an interdisciplinary subject, network science contains many scientific applications, e.g., the mathematical theories and methods of graph theory, statistical mechanics from physics, data mining and information visualization in computer science, statistical reasoning models in sociology and social structure. The National Research Council defines network science as “The study of network representations of physical, biological, and social phenomena leading to predictive models of these phenomena.” [28]. Therefore, network science has gradually become a stronger interdisciplinary subject and one of the most attractive fields of scientific research [29]-[31].

Based on different real-world network functions, the spread processes in network are diverse. As a typical transmission process in complex networks, epidemic spread has become an important topic in network science research in recent decades [6]-[11]. On one hand, many epidemics spread rapidly and easily cause social panic, which can lead to a series of (associated) social problems, and the process of this spread is closely related to the continuous development and evolution of human relationship network structures. On the other hand, for a specific epidemic, it is necessary to model it using a specific network, in order to reproduce the real-world dynamics of the epidemic and design methods to control or even eradicate the disease. In addition, based on epidemic spread models, many problems in complex networks have been solved, such as the measurement of node influences. Additionally, the established epidemic mathematical models can be extended to study other propagation behaviors on complex networks, such as cascading failures and information diffusions[14]-[18]. Because of the propagation characteristics of

different behaviors, new propagation models have been built and studied based on the mathematical models of epidemic spread.

## 1.2 Thesis Objectives and Arrangement

This thesis studies the relationship between the complex network structure measurements and the propagation processes, i.e., the influence of complex network structure measurements on the propagation processes and the coevolution between network structures and the propagation processes. The main research ideas are summarized in Figure 1.1. The influence of network structure measurements, e.g., degree, betweenness, closeness, on the propagation process is studied. Based on the degree value of the nodes, a method to quickly identify the influence of the nodes, as well as a cascaded failure model considering the local real-time information priority redistribution strategy is proposed, and a novel metric is proposed to measure the robustness in regard to virus attacks in social networks. Following on from this, the cooperative evolution of network structure and propagation process is studied, and the reliability of adaptive weighted networks is analyzed and discussed. In our research, the mean-field method [32], [33], differential dynamics [34] and other theoretical methods are used to model the network dynamics process, and nonlinear stability analysis [35] is used to study and discuss the critical characteristics and dynamics processes of propagation.

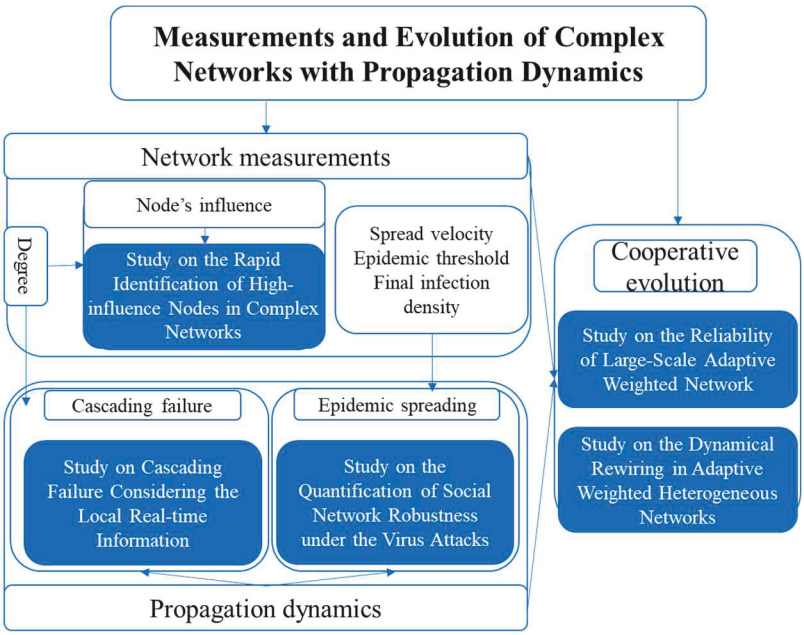



Fig 1.1 The framework of the thesis.

As shown in Fig 1.1, the most concerned network structure measurement in the thesis is the degree of nodes and the dynamical measurements are the spread velocity, the final infection scale and the epidemic threshold. A method in light of node degree and network structure to identify high-influence nodes rapidly in networks was first proposed. And then a novel cascading failure model was built and analyzed based on the degree of nodes. The impact of network structure measures on the propagation processes is reflected in the changes of network dynamical measures. A novel metric based on the network dynamical behavior measurements is proposed to quantify the network robustness in regard to virus attacks. Besides, the coevolution between network structures and the propagation processes are

analyzed to describe and understand the relationship between them.

The content of each chapter of the thesis is arranged as follows. The main work of the thesis is from Chapter 4 to Chapter 8. The first chapter introduces the significance of the research, the research background, research status, main work and the chapter arrangement of the thesis.

Chapter 2-3: The information of network structure measurement and cooperative evolution in communication dynamics is introduced, including complex network measurements, classical network models, typical propagation behaviors and virus propagation models. Finally, the interaction and co-evolution between network structure measurement and propagation behavior are discussed.

Chapter 4: Considering the dynamic change of networks, the privacy of network information and the big data of the network, this chapter focuses on how to quickly find the highly influential nodes in the network. A method based on the degree value of nodes to do this, which can quickly find a fraction of highly influential nodes in the network is proposed.

Chapter 5: Considering the real-time of network information and the rationality of load redistribution strategies, a new cascading failure model to analyze the robustness of a network due to node failure is proposed. Additionally, a network robustness and reallocation strategy are analyzed and discussed. The simulation results in artificial networks and real-world networks prove the effectiveness of the threshold analysis and reallocation strategy.

Chapter 6: Considering the network dynamical measurements, a novel robustness metric with respect to virus attacks in social networks is proposed. Simulation results show that the network becomes more vulnerable to the virus attacks as the average degree of the network grows in both homogeneous and heterogeneous networks.

Chapter 7: In this chapter, an adaptive weighted network model is proposed to describe the interactions between network structures and propagation processes. By using the mean-field method and Monte Carlo simulation, the influence of weight distribution and edge break reconnection probability on the propagation process is studied. The results show that the propagation threshold of the network increases with the decrease of the average network weight or the increase of the average edge break probability.

Chapter 8: Considering the diversity of human relationships, and the interaction between network topology and epidemic spread, this chapter proposes a new adaptive weighted heterogeneous network model based on the SIS spread process, which describes the interaction between node dynamic behavior and epidemic spread. Based on the dynamic epidemic propagation dynamics model, an effective edge disconnection and reconnection strategy is proposed to suppress epidemic propagation.

Chapter 9: Summary and prospects. This chapter summarizes the main contents and contributions of this thesis, and points out areas worthy of further study.

# Chapter 2

# Preliminaries

Network science, as a new interdisciplinary subject, covers mathematics, statistical physics, computation and various engineering technology sciences. As one of the main contents of network science research, the study on complex networks has attracted extensive interest in recent years, especially in the formation mechanism, evolution modes, structural measurements and dynamic behaviors on complex networks. Among them, network measurements have broadened people's understanding of complex networks, while network dynamics effectively depicts the essence of complex networks. The relationship between structural measurements and network dynamics reveals more profound and richer content of complex networks.

The history of research on complex networks can be traced back to 1736, when the famous mathematician Leonhard Euler studied and solved the famous Seven Bridge problem [24], thus turning a physical problem into a mathematical problem. One of Euler's contributions was to bring in a new branch of mathematics: Graph Theory. To study the measurement characteristics of different networks, effective analyzable tools are needed, among which graph theory is one of the best. A network can be regarded as a graph which contains nodes and the links which connect these nodes do so according to certain rules or formulae.

A network is described mathematically by a graph. According to the definition of a graph in graph theory, a network can be defined as the graph $G = (V, E)$ composed of nodes and their connecting links (called edges). $V$ and $E$ are respectively recorded as the nodes set and the edges set of $G$. Of course, in different fields of science, the definition of a network is also different. What kind of network can be called a complex network? The famous scientist, Xuesen Qian, believes that complexity is the dynamic characteristic of an open complex giant system [34]. The complexity of a complex network mainly lies in [1]:

- Complex structure: the number of nodes is large, and the network structure presents many different characteristics.
- Diversity of links: the weightings of links are different, and the links may have directionality.
- Diversity of nodes: nodes in a complex network can represent anything.
- Network evolution: the generation and disappearance of nodes or edges.
- Dynamics complexity: the dynamic change of a node’s state is a nonlinear dynamics system.
- Multiple complexity fusion: the interactions among all of the above complexity items, resulting in more unpredictable networks.

Complex systems are composed of many components that may interact with each other, such as social networks, biological systems, information networks, transportation systems, etc. While these

networks have different structural characteristics because of different functions, the communication behaviors on the network are also different, such as the spread of infectious diseases, information diffusion in social networks, traffic flow in traffic networks, the spread of computer viruses on the internet, etc. People are constantly studying real-world characteristics of these networks, trying to reveal the different propagation laws for each type of network, and trying to control the network measurements and network dynamics, which are the main contents of complex network research.

This chapter introduces the basic information related to those complex networks closely related to this thesis. It will be viewed from three aspects: (1) the introduction of the main complex network metrics, including the typical complex network structure metrics and typical network models, (2) the introduction of the network propagation processes related to the thesis, including the propagation model, measurements and the main results, and finally (3) the introduction of the co-evolution between the dynamic behaviors and the network structure.

## 2.1 Network Measurements and Network Models

### 2.1.1 Main Network Structure Measurements

**(1) Degree and degree distribution**

Degree is one of the simplest and most important concepts for describing the attributes of a single node. The number of edges connected to a node is its degree. Degree distribution is a general description of the degree values of nodes in a network. For a network, degree distribution refers to the probability distribution of the degree of each node in the graph.

For the degree of node $i$, $k_i$ refers to the number of nodes directly connected to $i$, which is the most basic static feature of a node. The average degree of all nodes in a network becomes the average degree of the network, which is recorded as $<k>$. Given the adjacency matrix $A = (a_{ij})_{N*N}$ of network $G$, we have

$$k_i = \sum_{j=1}^{N} a_{ij} = \sum_{j=1}^{N} a_{ji}, \tag{2.1}$$

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^{N} k_i, \tag{2.2}$$

where $N$ represents the number of nodes in $G$.

The degree distribution $P(k)$ of a network is the probability that the degree of a randomly selected node in the network is $k$. For example, the degree distribution of a random network and a small-world network model obey the approximate Poisson distribution, and the shape of the distribution graph is approximately bell shaped, and it decreases exponentially away from the peak value $<k>$. Therefore, these networks are also called homogeneous networks. However, it has been found that the degree distribution of many real-world networks, e.g., movie actor networks, power grids, etc., do not obey the Poisson distribution with uniform characteristics, but the power-law distribution as follows

$$P(k) \sim k^{-\gamma}, \tag{2.3}$$

where $\gamma$ is the power index, usually between 2 and 3. There are a large number of nodes having very low degrees and yet a few nodes having very high degree. The power-law distribution is more commonly called the scale-free distribution.

**(2) Clustering coefficient**

The degree of a network describes the number of neighbors a single node has. In fact, the relationship between the neighbors of nodes is very close. For example, in social networks, it is very likely that the two of your friends know each other. The clustering coefficient of a network, a parameter used to describe the clustering degree of nodes, can quantitatively describe the probability that those two friends are also friends with each other. The clustering coefficient $C_i$ of a node $i$ with a degree of $k_i$ in the network is defined as

$$C_i = \frac{E_i}{(k_i(k_i\text{-}1))/2} = \frac{2E_i}{k_i(k_i\text{-}1)}, \tag{2.4}$$

where $E_i$ is the number of existed edges in the network, and so obviously $0 \le C_i \le 1$. $C_i = 0$, if and only if any two neighbors of node $i$ are not neighbors or node $i$ has only one neighbor. Here, $E_i$ can be regarded as the number of triangles made with node $i$ as one of the vertices. The number of connected triangles with node $i$ as the center is actually the maximum possible number of triangles including node $i$, i.e., $(k_i(k_i\text{-}1))/2$. Therefore, from a geometric point of view, the definition of the clustering coefficient of node $i$ is as follows:

$$C_i = \frac{\text{the number of triangles containing node } i}{\text{the number of connected triples centered on node } i}. \tag{2.5}$$

The clustering coefficient $C$ of a network is defined as the average of the clustering coefficients of all nodes, namely

$$C = \frac{1}{N}\sum_{i=1}^{N} C_i. \tag{2.6}$$

**(3) Average path length**

Furthermore, not only are the neighbors of each node in a network closely related, but also the relationship among the nodes is very close. Despite the number of nodes in a real-world network being huge, the path between any two nodes in a network is short. For example, the average distance between two users on Facebook is only 4.74 steps. This is the so-called small-world phenomenon.

The shortest path between the two nodes $i$ and $j$ in a network is the path with the least number of edges directly connecting the two nodes. The distance $d_{ij}$ is defined as the number of edges on the shortest path connecting the two nodes $i$ and $j$. In this way, the average path length $L$ of the network is defined as the average value of the distances between any two nodes,

$$L = \frac{1}{\frac{1}{2}N(N-1)} \sum_{i \geq j} d_{ij} \cdot \tag{2.7}$$

**(4) Weights**

In many real-world networks, the relationships between nodes are different. For example, in the scientific research cooperation network, the number of papers that each researcher cooperates with is different; in the social network, people's intimacy with family members, friends or colleagues is different; in the internet or communication network, the traffic of each page/site is very different; in the traffic network, the traffic of each city is significantly different, and the traffic of big cities is different as the number of traffic lines is significantly greater than for those between small cities. Therefore, it is more suitable to describe these networks by weighted network models.

Let $w_{ij}$ represent the weight of the link between nodes $i$ and $j$. A weighted network can use the link weight matrix $(w_{ij})_{n \times n}$, $i, j=1, 2, \ldots, n$. For an undirected network, the weight matrix is symmetric, that is, $w_{ij} = w_{ji}$. The strength of node $i$ in the weighted network is defined as:

$$s_i = \sum_{j \in \Gamma(i)} w_{ij}, \tag{2.8}$$

where $\Gamma(i)$ represents the neighbor nodes set of node $i$.

**(5) Centrality measurements of networks**

The concept of node centrality [46] represents the importance of nodes in the network, which is an important concept in social network analysis. Based on different indicators that impact the importance of nodes, many different definitions of centrality have been proposed. The simplest index of centrality is node-degree. The larger the degree a node has, the greater its influence/centrality.

For a network with $N$ nodes, the betweenness centrality of node $i$, denoted by $BC_i$ is

$$BC_i = \sum_{s \neq i \neq t} \frac{n_{st}^i}{g_{st}}, \tag{2.9}$$

where $g_{st}$ is the number of shortest paths between nodes $s$ and $t$, and $n_{st}^i$ denotes the number of shortest paths between $s$ and $t$ which pass through node $i$.

Closeness of node $i$ is defined as the reciprocal of the average geodesic distances to all other nodes of $i$

$$CC_i = \frac{1}{d_i} = \frac{N}{\sum_{j=1}^{N} d_{ij}}, \tag{2.10}$$

where $d_{ij}$ is the geodesic distance between $i$ and $j$.

The nodes with the largest betweenness have the greatest control ability over the flow of information in the network, while the nodes with the largest closeness have the best observation for the flow of information. Therefore, the node with the largest betweenness and the largest closeness may not be the same node.

Another widely used centrality measure is eigenvector centrality. The importance of a node depends on the number of neighbor nodes and the importance of each neighbor node. The eigenvector centrality of node *i* is defined as being proportional to the sum of the eigenvector centers of the nodes to which node *i* is connected. That is

$$EC_i = \frac{1}{\rho} \sum_{j \in \Gamma(i)} EC_j = \frac{1}{\rho} \sum_{j=1}^{N} A_{ij} EC_j , \tag{2.11}$$

where $\rho$ is a constant and *A* is the adjacency matrix of the network.

### 2.1.2 Network Models

The previous section describes some basic indicators and statistical characteristics of network structures, which have been used to describe and evaluate real-world networks. The discovery of these different network structures in turn has prompted the theoretical research of network generation models. The usefulness of network models in the current context lies in the fact that they can be used as generators of synthetic networks to generate topological properties with particular requirements. In these network models, the dynamic behaviors, such as epidemics, can be studied in detail. In this section, some simple and general network models are introduced, which are used to explore the propagation processes from the perspectives of research motivation, construction process and important attributes.

#### (1) Random networks

The systematic study of random graphs [25], [94] was initiated in 1959 by two Hungarian mathematicians, Erdős and Rényi. The term "random graph" refers to the lack of order of connection arrangement between different nodes. Erdős and Rényi proposed a random graph model with *N* nodes and *M* connections, which is called the ER random network $G_{N,M}^{ER}$. Starting from *n* disconnected nodes, an ER random graph is generated by connecting randomly selected nodes with probability *p*. Multiple connections are forbidden between the same pair of nodes until the number of edges equals *M*.

Although it cannot reproduce most of the attributes in real-world networks, random graphs are the most widely studied model of graph models. The structural characteristics of ER random graphs change with the change of *p*. At that time, when $N \to \infty$ and $p \geq \ln(N)/N$, almost all networks generated by using the probability *p* become connected. For graph $G_{N,M}^{ER}$, the probability of any node with degree *k* is $p^k (1-p)^{N-1-k}$, so the degree distribution of a random network obeys binomial distribution

$$P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k} . \tag{2.12}$$

When $N \to \infty$ and $p$ is small, the degree distribution follows Poisson distribution

$$P(k) \approx \frac{(Np)^k \mathrm{e}^{-Np}}{k!} = \frac{\langle k \rangle^k}{k!} e^{-\langle k \rangle}. \tag{2.13}$$

**(1) Small-world networks**

The first small-world network model [26] was proposed by D. Watts and S. Strogatz in the paper "Collective dynamic of 'small-world' networks" published in the journal Nature in 1998. It was found that the clustering coefficient of a regular network is high, but the average path length of a network is also large, while the average path length of a random network is short, and its clustering coefficient is low. A real-world network is neither completely regular nor completely random. Therefore, based on a regular network, the authors randomly reconnected each edge with the probability $p$ (Self-connection and reclosing connection were not allowed), which introduced both randomness and long-range connections (or shortcuts) into the regular network. These small-amount and random long-range connections greatly reduced the average distance(s) between nodes in the network while maintain the high clustering characteristics of the original network, so that the network possessed a small-world characteristic. Adjusting the value of $p$ realized the transition from a completely regular network ($p$=0) to a completely random network ($p$=1).

The average path length $L$ and the average clustering coefficient $C$ are introduced to quantitatively analyze the small-world characteristics of the network. In Fig 2.1, two values are normalized, i.e. $C(p)/C(0)$ and $L(p)/L(0)$. It can be seen from Fig 2.1 that when $p$ increases from 0, the clustering coefficient of the network after a random reconnection decreases slowly, but the average path length decreases very quickly, that is, when $0 < p << 1$, $C(p) \sim C(0)$, $L(p) << L(0)$. This implies that after the rewiring, the resulting network shows small-world characteristics with a short average path length and a large clustering coefficient.

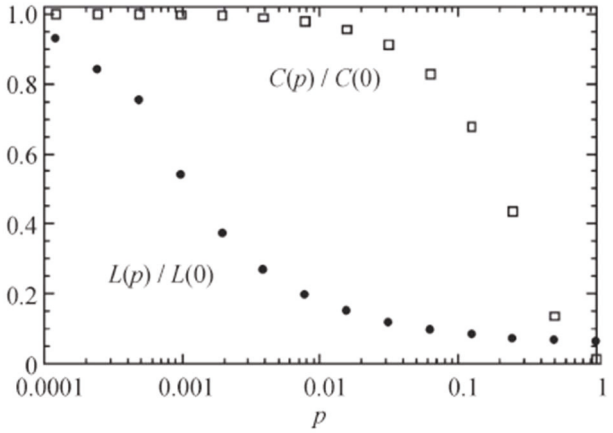



Fig 2.1 The relationship between the average path length and clustering coefficient of WS small-world network model with the probability of reconnection $p$. (from Ref. [1])

M. Newman and D. Watts improved the WS small-world network model [1], [3]. Without changing the connection edge between nodes in the original regular network, they chose two nodes at random, and allowed the probability $p$ to decide whether to add an edge between the two nodes or not. A random

connection edge was added between the two nodes. This kind of operation also introduces randomness and long-range connections in regular networks, which makes the network have small-world characteristics.

**(3) Scale-free networks**

In 1998, Albert-László Barabási and Réka Albert cooperated in a study of describing the World Wide Web (WWW). They found that the WWW network, which is composed of hyperlinks, web pages and files, did not have the same degree distribution as a general random network. It was found that there were a small number of highly connected pages with high-degree in the WWW network. The vast majority (more than 80%) of web pages have no more than 4 hyperlinks, but a few (less than 1/10000 of the total pages) have many links, more than 1000, and some pages even connect with more than 2 million other pages. As an analogy using the example of people's height, most of the nodes are "short men", but there are a few "giants" with a height of 100 feet. In 1999, Albert-László Barabási and Réka Albert proposed a scale-free network model called the BA scale-free network [27]. This model was based on two assumptions:

Growth mode: Many real-world networks are growing, such as the birth of new web pages on the internet, the joining of new friends in social networks, the publication of new papers, and the construction of new airports in the aviation network.

Priority connection mode: New nodes tend to be connected to existing nodes with more connections when they join. For example, new web pages will generally have connections to well-known web sites, new members of the community will want to get to know well-known people in the community, new papers will tend to quote well-known literature that has been widely cited, and new airports will give priority to establishing routes with large airports.

Based on these assumptions, the specific structure of the BA network model is as follows:

Growth: Starting from a smaller network $G_0$, with $n_0$ nodes and $E_0$ edges, add one new node at each step and connect to $n$ existing nodes, $n \le n_0$,

Priority connection: The connection mode is to give priority to nodes which have a high degree value. The probability of establishing a connection between the new node and the existing node $i$ is $\prod_i = k_i \Big/ \sum_j k_j$ .

The basic metrics of these three network models are shown in Table 2.1. Among them, in the small-world network, $\kappa$ is the number of neighbors of each node in the initial regular network. When $\mu \ll 1$, $f(\mu)$ is a constant, and when $\mu \gg 1$, $f(\mu) = \ln(\mu)/\mu$.

Table 2.1 Basic measures of three network models. (from Ref. [2])

| Measurement | Random network | Small-world network | Scale-free network |
|---|---|---|---|
| Degree distribution | $p(k)=\frac{e^{-\langle k\rangle}\langle k\rangle^{k}}{k!}$ | $p(k)=\sum_{i=1}^{\min(k-\kappa,\kappa)}\binom{\kappa}{i}(1-p)^{i}p^{k-i}\frac{(p\kappa)^{k-\kappa-i}}{(k-\kappa-i)!}e^{-p\kappa}$ | $p(k)\sim k^{-3}$ |
| Average degree | $\langle k\rangle=p(N-1)$ | $\langle k\rangle=2\kappa$ | $\langle k\rangle=2m$ |
| Clustering coefficient | $C=p$ | $C(p)\sim\frac{3(\kappa-1)}{2(2\kappa-1)}(1-p)^{3}$ | $C\sim N^{-0.75}$ |
| Average path length | $l\sim\ln N/\ln\langle k\rangle$ | $l(N,p)\sim p^{T}f(Np^{T})$ | $l\sim\frac{\log N}{\log(\log N)}$ |

**(4) Weighted scale-free networks**

The above network model can reflect the attributes of nodes in a network and the simple connection between those nodes, but not the diversity and difference of the interactions between the actual network nodes, for example, the strength, the tightness and the distance between the nodes. Therefore, the weighted network model was introduced to describe the actual network. Barrat, Barthélemy and Vespignani proposed the famous weighted scale-free network model BBV model [95]-[98], where the results show that the node-weight, node-degree and edge-weight satisfy the power-law distribution. The algorithm for generating the BBV model is as follows:

- The initial network is fully connected with $N_0$ seed nodes, and the weight of each edge is set at $w_0$,
- One node is added in each step, and the nodes added in each step send out $m$ edges, which are connected to the existing node $i$ according to the priority principle of node strength. The preference probability of the new node being connected to an existing node $i$ is $\prod_{new\to i}=s_i/\sum_j s_j$,
- The weight of each new edge is $w_0$, and the new edge will cause the change of the network's weight: when the new node is connected to node $i$, the node strength of node $i$ will increase by a constant parameter $\delta$, and the node strength of node $i$ will change to $s_i=s_i+w_0+\delta$, as shown in Fig 2.2,
- The increase $\delta$ is allocated to other sides of node $i$ according to the following rules: $w_{ij}\to w_{ij}+\Delta w_{ij}$, where $\Delta w_{ij}=\delta\frac{w_{ij}}{s_i}$,
- Repeat the above steps until the total number of network nodes reaches $N$.

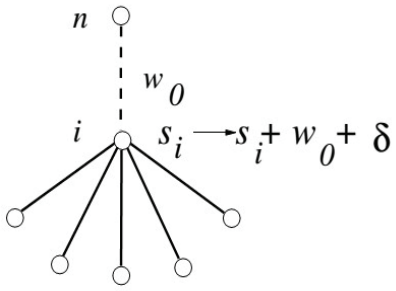



Fig 2.2 Generation diagram of BBV network. (from Ref. [95])

The measurement of complex networks has become an important research direction of complex network dynamics. By revealing some simple dynamic rules or network characteristics, such as the priority connection characteristics in scale-free networks and the high clustering characteristics in small-world networks, complex network topology can be generated. These network metrics are not only important tools for generating network models, but also important factors to shape real-world networks such as the internet and social networks.

## 2.2 Propagation Dynamic Processes

Besides the network measurements, another focus of complex network research is the dynamic behaviors on a network. Research shows that complex network measurements have an important impact on the dynamic changes of nodes. For example, the epidemic threshold of virus spread in a small-world network is larger than in a scale-free network, but the infection scale at the steady state in a scale-free network is smaller than in a small-world network. This section briefly introduces two kinds of propagation processes and some conclusions which relate to this thesis.

### 2.2.1 Epidemic Spread

For more than 200 years in the history of epidemiological research, the mathematical model of epidemic spreading [99]-[102] has been developing into a research field spanning the field of mathematical biology and other disciplines. The epidemic model describes the dynamic evolution process of health, infection and other states in populations. In epidemic models, the population can be divided into different categories according to the stage of the disease, for example, Susceptible refers to the healthy population that can be infected. Infected refers to the population that has been infected and is infectious. Recovered refers to the population that has recovered from the disease. There are other classifications that can be used to represent other possible states of an individual about a disease, such as an Immune individual. This framework can be extended to consider diseases that are transmitted through contact with external vectors, such as malaria viruses which are carried by mosquitoes. In order to understand the dynamic evolution of the number of infected individuals over time, it is necessary to define a basic individual-level-based process to control individuals from one state to another.

The simplest definition of epidemic dynamics assumes that the total population in the system is fixed, composed of $N$ individuals, and ignores any other demographic processes (migration, birth, death,

etc.). One of the simplest two-state partition models is the SIS model, which contains two state transition processes: 1) $S \to I$, when a healthy individual interacts with an infected individual, there is a certain probability they will become infected . 2) $I \to S$, there is a certain probability the infected individuals will recover to full health. In the process of simplified modeling, the probability of state transformation is assumed to be a constant. Epidemics can be described as a random reaction diffusion process. Individuals of different states correspond to individuals of different populations. These individuals evolve according to a set of interaction rules and represent different transformation probabilities between different states. In this way, the SIS model can be expressed as

$$\begin{aligned} S(i) + I(j) &\xrightarrow{\beta} I(i) + I(j) \\ I(i) &\xrightarrow{\mu} S(i) \end{aligned}, \tag{2.14}$$

where $\beta$ and $\mu$ are the probability of infection and recovery respectively. When $\mu/\beta$ is small enough, the infection process in the model will continue.

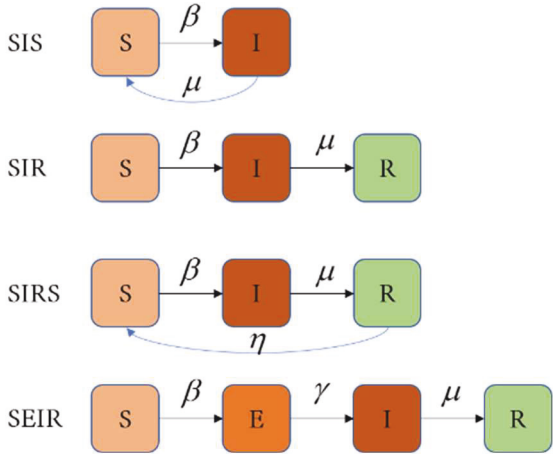



Fig 2.3 Typical epidemic spread models.

Similar to the SIS model, there are many other epidemic models that can be defined in this way. Fig 2.3 shows different epidemic spread models from the perspective of the reaction diffusion process. Squares represent different states, while arrows represent transitions between the states, which occur according to their respective probabilities. In the SIR model, an S-state node is infected by an I-state neighbor with probability $\beta$, while an I-state node will restore to the R-state with probability $\mu$ and this state will not be changed again. Different to the SIR model, the R-state node in the SIRS model will change to an S-state with probability $\eta$ and continue with the whole state evolution process. With consideration of the incubation period of an epidemic, the SEIR model was proposed by researchers. In the model, an S-state node will change to an E (exposed) state with probability $\beta$, and an E-state node will be infected by an I-state neighbor with probability $\gamma$, while an I-state node will return to the R-state with probability $\mu$, after which it will not change state any more. On the one hand, these models can be used to describe different types of infectious diseases. For example, the SI model can simulate the spread of HIV, the SIS model can represent the spread of most influenza viruses, the SIR model can

represent the spread of the smallpox virus, and the SEIR model can represent some virus transmission behaviors with a latent period. On the other hand, they can be used to describe and study other types of network propagation behaviors, such as information diffusion, cascading failure and so on.

In the study of complex networks, one of the simplest methods for dynamic behavior analysis is the mean-field approximation [32], [33]. The mean-field method is widely used in the analysis and prediction of epidemic transmission dynamics, and its analytical results can better reflect actual situations. In short, the mean-field method is to deal with the effect of environment on the object collectively, and replace the sum of single effects with the average effect. This section focuses on the application of the mean-field method in the analysis of full contact SIS models and gives basic conclusions by taking a small-world network and a scale-free network as examples.

The earlier epidemic model was proposed by Reed and Frost in an unpublished paper in 1920, and was also the first time differential equations were used to describe dynamics. In this paper, the differential equations of the SIS propagation model and some conclusions are introduced.

For a homogeneous network, random network or small-world network, etc., let the average degree of nodes in the network be <*k*>, the infection rate of an epidemic be $\beta$, the recovery probability of the nodes be $\mu$, and $S(t)$ and $I(t)$ represent the node density of the S-state and the I-state, respectively. Thus, the differential equation of SIS model is as follows

$$\frac{dI(t)}{dt} = \beta \langle k \rangle I(t)(1-I(t)) - \mu I(t) \cdot \tag{2.15}$$

This equation describes the changing of the density of I-state nodes at time $t$, in which the term $\beta \langle k \rangle I(t)(1-I(t))$ describes the new I-state individuals after the S-state nodes are infected, and the term $\mu I(t)$ describes the I-state individuals reduced by the restoration of I-state to S-state. The steady-state equation of (2.15) is $\beta \langle k \rangle I_\infty(\tau)(1-I_\infty(\tau)) - \mu I_\infty(\tau)=0$, so the steady-state solution satisfies

$$I_\infty(\tau)=\begin{cases} 0, \tau < \tau_c \\ 1-\dfrac{1}{\tau \langle k \rangle}, \tau \geq \tau_c \end{cases}, \tag{2.16}$$

where the epidemic threshold is $\tau_c = 1/\langle k \rangle$. When the effective infection probability is less than $\tau_c$, the epidemic in the network dies out, and when the effective propagation probability is greater than $\tau_c$, the epidemic spreads widely in the network, that is, there will always be some infected nodes in the network.

For heterogeneous networks, the assumption that the degree of nodes is the same will not be able to describe real-world situations. In order to more accurately describe the epidemic spread process on heterogeneous networks, Pastor-Satorras and Vespignani proposed the mean-field method of the heterogeneous network [103]. The mean-field equation of the SIS virus model on heterogeneous networks is shown as follows

$$\frac{dI_k(t)}{dt} = -\mu I_k(t) + \beta k(1 - I_k(t)) \sum_{k'} P(k' \mid k) I_{k'}(t) \quad . \tag{2.17}$$

The first term at the right side of (2.17) describes the recovery process of the infected node, and the second term represents the infection process of the healthy node, $\sum_{k'} P(k' \mid k) I_{k'}(t)$ is the infected neighbor of the node with degree *k*. According to (2.17), the epidemic threshold of the heterogeneous network is $\tau_c = \langle k \rangle / \langle k^2 \rangle$. When the network size $N \to \infty$, there is $\langle k^2 \rangle \to \infty$, therefore $\tau_c \to 0$, the epidemic threshold is 0. Fig 2.4 compares the relationship between the infection density at the steady state $I_\infty(\tau)$ and the effective infection rate $\tau$ of the SIS model on an Erdös-Rényi network (ER network) and a Preferential attachment network (PA network) [27], [30]-[31]. It can be seen from the figure that the epidemic threshold of the heterogeneous network is obviously smaller than that of the homogeneous network, while the infection scale at the steady state of the homogeneous network is larger than that of the heterogeneous network.

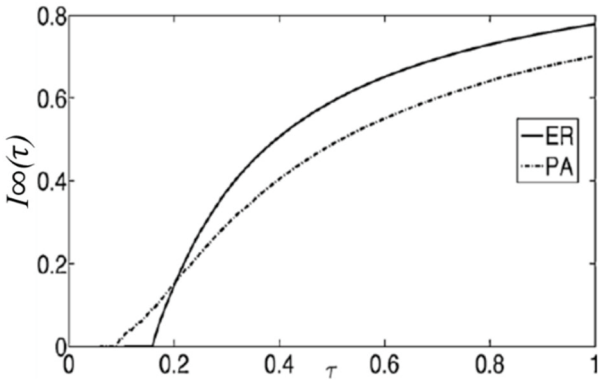



Fig 2.4 The relationship between the infection density at the steady state $I_\infty(\tau)$ and effective infect rate $\tau$ (SIS model).

### 2.2.2 Cascading Failure

Most real-world networks, because of their own network functions, will bear certain loads, such as traffic flow load in traffic networks, voltage load on power networks and information flow load on the internet. This requires that the nodes/edges in the network have a certain load capacity and a load processing capacity. When the load allocated to the nodes/edges exceeds their processing capacity, the nodes/edges will be overloaded and fail. After a failure, the node/edge redistributes its load to other nodes/edges, and this kind of cyclic behavior will continue until there are no new failed nodes/edges. This process is called a cascading failure. In short, the failure of one or a few nodes/edges in the network will lead to the failure of other nodes/edges through the coupling relationship between the nodes, and then produce a cascading effect, eventually leading to the collapse of a considerable number of nodes or even the whole network.

Based on the mechanisms of a cascading failure, many network cascading failure models are proposed, such as the Sandpile model [79], the CASCADE model [78], the OPA model [80], the Motter-

Lai (ML) model [75]-[77]. Based on the summary of these models, it can be found that there are three main factors affecting cascading failure:

**(1) Initial load**

The initial research assumed that the initial load of nodes in a network is the same, but this is not consistent with the initial load distribution in real-world networks. Later, many models used the degree or betweenness of nodes as the initial load. The betweenness describes the ability of nodes to control the network flow along the shortest path. The larger the number of shortest paths through the node, the greater the initial load of the node. While this definition is reasonable, the complexity of the calculations is relatively high because global information is needed to calculate the betweenness of a node. In order to reduce the computational complexity, many researchers have set the initial load as a function related to the node's degree, for example, considering the node's own degree and the neighbors' degree.

**(2) Load capacity**

The definition of load capacity can be roughly divided into three categories: the first is that the load capacity of each node is independent of the initial load of the node. Load capacity is defined as a statistical distribution, which increases with the increase of network size (ICA mode) or defined as a constant (ECA mode). The second definition of load capacity is proportional to the initial load of the node, which is currently the most widely used. The third is that the load capacity is non-linear to the initial load. Based on the data of real-world networks, it is found that the load capacity of nodes is not linear with the initial load, as some nodes have a very small initial load but a very large residual capacity. This discovery has triggered many new thinking and research on the relationship between the initial load and the load capacity.

**(3) Load redistribution strategy**

Load redistribution refers to the behavior of redistributing the load to other nodes in order to ensure the normal operation of network functions after a node's failure. An effective load redistribution strategy can reduce the probability of cascading failures and the scale of a network's failure. At present, load redistribution strategies include: the uniform redistribution strategy, the random redistribution strategy, the global redistribution strategy, the local optimal redistribution strategy, and the adjustable load redistribution strategy.

## 2.3 Adaptive Networks

In most real-world networks, the evolution of a network structure is always related to the states of the network nodes, and vice versa. For example, in the crowd contact network, the healthy people are "far away" from the infected individuals by wearing masks and avoiding contact. In such a process, the network structure changes due to the state of the nodes. In turn, due to the adjustments of the network structure, the original transmission paths change, and the transmission process will also be affected. For a transportation network, if a road is often blocked, the road is likely to be widened or a new road built to alleviate the congestion. That is, the flow of traffic and the network structure influence each other. In

this way, a cooperative evolution process between network topology and propagation dynamics is formed in the network. The network with the above process is called an adaptive network, combining the evolution of network structure, node state dynamics and the interaction between them together. Adaptive networks exist widely in real life, such as technology networks, transportation networks, biological networks and social networks.

Considering that individuals change their behavior adaptively in order to avoid possible infection in the outbreak of a disease, Gross et al. introduced the rewiring rate in [36], and proposed the concept of the adaptive network model for the first time. The model is described as follows: Assume that there are $N$ nodes, $L$ undirected edges in the random graph, and the node state is susceptible (S) or infected (I). Every I-state node returns to an S-state with probability $r$. For an SI-edge, an I-state node infects an S-state node with probability p, which makes the S-state node become an I-state node. At the same time, the S-state node disconnects the SI edge with probability $w$, and then randomly selects an S-node to establish a new edge. Assuming that the disease breaks out in a static homogeneous network, the average number of people an I-node can infect in the infection cycle, i.e., the basic reproduction number, is $R_0=p\langle k\rangle / r$, so the epidemic threshold can be written as $\mathrm{p}^*=r/\langle k\rangle$. When there is an adaptive rewiring process in the network, the expectation of the node's degree can be written as $k(t)=\langle k\rangle \exp(-wt)$, where $t$ refers to the duration that the node has been infected. By averaging the number of infections during the infection period $1/r$, the epidemic threshold is obtained, $\mathrm{p}^*=w/\langle k\rangle(1-\exp(-w/r))$.

Gross defined in reference [36] that the densities of the susceptible and infected nodes are $i$ and $s$ respectively, and the densities of SS-links, II-links and SI-links of each node are $l_{SS}$, $l_{II}$ and $l_{SI}$ respectively, so $s+i=1$, $l_{SS}+l_{II}+l_{SI}=\langle k\rangle/2$. $l_{abc}$ represents the density of a triple, where $a,b,c\in[S,I]$. In order to study the dynamics caused by the mechanism of adaptive rewiring, three equations of the changes of $i$, $l_{SS}$ and $l_{II}$ are applied to describe the dynamic state and topological structure of the network.

$$\frac{di}{dt}=pl_{SI}-ri, \tag{2.18a}$$

$$\frac{dl_{II}}{dt}=pl_{SI}\left(\frac{l_{ISI}}{l_{SI}}+1\right)-2rl_{II}, \tag{2.18b}$$

$$\frac{dl_{SS}}{dt}=(r+w)l_{SI}-pl_{SSI}. \tag{2.18c}$$

(2.18a) describes the change of nodes in state I, the increase of I-state nodes caused by infection is $pl_{SI}$. As the I-state nodes recover at probability $r$, the decreased number of I-state nodes caused by the recovery process is $ri$. (2.18b) describes the change of II-links. In a single infection event, the virus spreads through an SI-link, changing the SI-link into an II-link with probability $p$, so at least one II-link will be generated in a single infection event. But if there are other I-state neighbors, the infection event will generate other II-links. Therefore, the total number of II-links generated by an infection event is $\frac{l_{ISI}}{l_{SI}}+1$, where "1" represents the SI-link where the infection occurs, and the first item represents the number of ISI-triples formed by this SI-link. It can be concluded that the generation rate of an II-link is

$pl_{SI}(\frac{l_{ISI}}{l_{SI}}+1)=p(l_{ISI}+l_{SI})$. If the nodes on an II-link recover with probability *r*, then the II-link decreases. The average number of II-links formed by an I-state node is $2rl_{II}$, where "2" means that one II-link connects two I-state nodes, so the rate of decrease of II-links is $2rl_{II}$. For (2.18c), since the S-nodes in each SS-link may be connected to other I-nodes, if the S-nodes in this SS-link are infected, the SS-link will be reduced. The reduction rate of SS-link is $pl_{SSI}$, the recovery of an I-state node in an SI-link will increase the SS-link, and the corresponding rate is $rl_{SI}$, while at the same time, each process of rewiring will increase one SS-link, so the rate of SS-link increase caused by rewiring is $wl_{SI}$.

(2.18) depends on two unknown second-order moments-$l_{ISI}$ and $l_{SSI}$, and the solution of these two quantities depends on the quaternion in the network. With such a development, the equation group will be very complex and cannot be solved. Therefore, the method of moment closure approximation is used in ref. [158] to make it closed, i.e., $l_{ISI}=\frac{\langle q\rangle}{\langle k\rangle}\frac{l_{SI}l_{SI}}{s}$. When the degree distribution of the network obeys Poisson distribution, the residual average degree and average degree of the network are equal, $\frac{\langle q\rangle}{\langle k\rangle}=1$. Similarly, $l_{ISI}=2\frac{l_{SS}l_{SI}}{s}$, (2.18) can be reduced to

$$\frac{di}{dt}=pl_{SI}-ri, \tag{2.19a}$$

$$\frac{dl_{II}}{dt}=pl_{SI}(\frac{l_{SI}}{s}+1)-2rl_{II}, \tag{2.19b}$$

$$\frac{dl_{SS}}{dt}=(r+w)l_{SI}-2\frac{pl_{SS}l_{SI}}{s}. \tag{2.19c}$$

Due to the interaction between the network topology and dynamics on adaptive networks, there are many new dynamic phenomena, as shown in Fig 2.5. In the figure, the thin line represents the analytical results of (2.19), and the circle represents the simulation results. We can see from the figure that the analytical results of the differential equation are consistent with the numerical simulation results, which further proves that the model is suitable for analyzing dynamic homogeneous networks. When there is no adaptive rewiring, only one continuous phase transition occurs at the propagation threshold p*; when there is adaptive rewiring, the threshold p* increases and another lower threshold appears, that is, the persistence threshold, corresponding to saddle node bifurcation.

Although increasing the probability of adaptive rewiring can reduce the infection scale of the disease, the persistence threshold changes substantially when the probability of rewiring is large. First, the saddle node bifurcation disappears, and then a subcritical Hopf bifurcation which will lead to the instability limit cycle being replaced. When the probability of edge rewiring is higher, the subcritical Hopf bifurcation becomes the supercritical Hopf bifurcation. As the new limit cycle is stable, this supercritical Hopf bifurcation marks the third threshold, when a continuous phase transition to the oscillatory state occurs. However, such oscillations can only be observed in a relatively small range before the persistence threshold is reached, as shown in Figure 2.6. There is only a single attractor in the white and light gray

areas. The white areas are healthy and the light gray areas are morbid. In the middle gray area, the healthy state and morbid state are stable. Another small bistable region is dark gray, where a stable healthy state and a stable propagation limit cycle coexist. The transition lines between these regions correspond to a transcritical bifurcation (dotted line), saddle node bifurcation (dotted line), Hopf bifurcation (solid line) and periodic folding bifurcation (dotted line), respectively. Therefore, in general, the interaction between topology and dynamics on adaptive networks can lead to numerous dynamic phenomena.

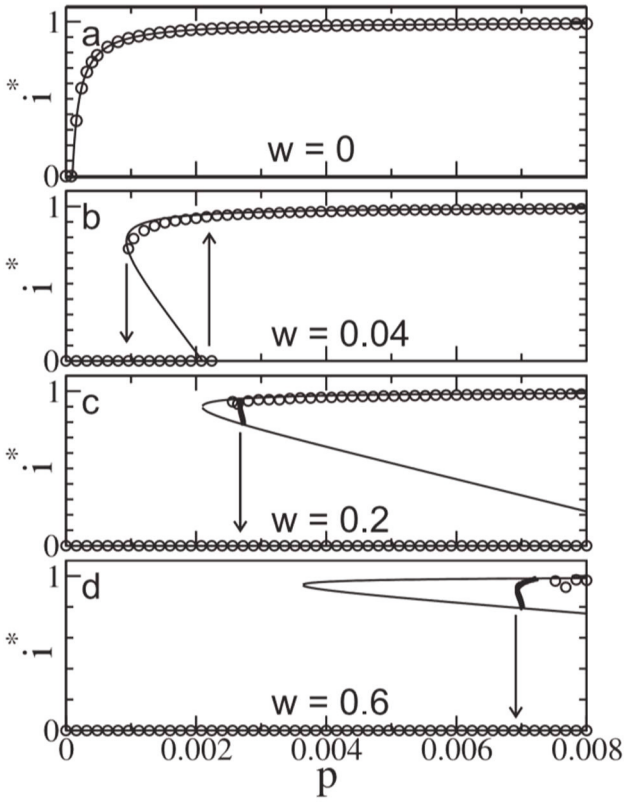



Fig 2.5 Relationship between steady-state infection density *i** and infection probability p under different adaptive rewiring rates *w.* (from Ref. [36])

Based on the adaptive network model of Gross, Shaw et al. [38] in 2008 used the adaptive mechanism to study the transmission process and characteristics of the SIRS in the adaptive network, and found the same phenomena as Gross's research, such as bistability and bifurcation. Gross et al. [104] further proposed a computational approach to the investigation of emergent properties of adaptive networks, where they have avoided the strong homogeneity assumption that is inherent in previous analytical moment closure approximations. Through their analysis, it was found that in a considerable parameter range the prevalence of the disease and the topology of the network exhibits oscillations of large amplitude.

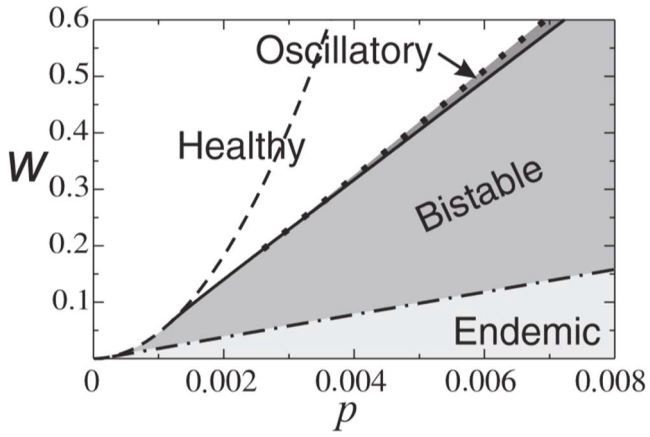



Fig 2.6 Relationship between the probability of edge break reconnection *w* and the probability of infection *p*. (from Ref. [36])

## 2.4 Summary of the Chapter

In this chapter, some basic information related to complex networks has been reviewed. Firstly, it introduced the basic concepts related to complex network, including the definition of complex networks and the definition of main network metrics. Then it introduced several network models related to the thesis - random network, small-world network and scale-free network models. Then we introduced two typical network propagation behaviors: epidemic propagation and cascading failure. Finally, the information regarding complex network structures and dynamic coevolution adaptive networks was introduced.

# Chapter 3

# Research Status

At present, the research of complex network dynamics mainly focuses on 3 aspects: the dynamics of complex networks, the dynamics on complex networks and the interactions between these two dynamics [26]. The dynamics of complex networks focuses on the complex network structure. It regards to the network itself as a complex system growing or changing according to the specific local area construction rules, and studies the evolution rules of the network and the structural characteristics of different networks [25]-[27]. The dynamics on complex networks stipulate that each node in the network has a dynamic state. The network topology remains unchanged, while the node state changes according to the local evolution rules, such as in the study of viruses, information dissemination, etc. [1]-[6], [16]. Finally, there is a very close relationship between the dynamics of complex networks and the dynamics on complex networks [36]-[39]. The network structure affects the spread dynamics in the network, for example, the influence of the selection of the initial infected seeds on the epidemics, the network degree distribution and the clustering coefficient on the propagation threshold and scale, etc. Conversely, the spreading processes also affect the network structures. For example, during an outbreak of influenza, healthy people have an awareness of the risks of becoming infected, and so their behavior of actively avoiding contact with infected individuals would cause a change of the network structure. This evasive behavior can also reduce the probability of healthy nodes becoming infected, thus inhibiting the spread of the virus. This way, a feedback loop is formed between the network topology and the dynamics on the network, which is called an adaptive network [36]-[39]. Based on the research of the above three aspects, the remainder of this chapter gives a detailed introduction to the research status of each aspect, which is more related to the research content of the thesis.

## 3.1 Measurement of Node Influence

A node is the most primary and important element of a network, and the relationship between the nodes and the positions of nodes in the networks determines the structure of the network. In complex networks, there are some special nodes or links between nodes, which, compared to other nodes in the network, can affect the structure and function of the network to a greater extent. For example, the content released by the most popular bloggers can have tens of millions of forwards in only a few minutes, or the damage of a high voltage line in a power grid will cause a large-scale blackout, while only 1% of the companies in the global economic system control 40% of the global economic lifeline. In these networks, a small percentage of the important nodes can affect most of the nodes in the network quickly. Therefore, the problem of the ranking of influence/importance of nodes and the mining problem of high influence nodes have wide application value and significance.

The ranking and mining of nodal influence has become an important area of complex network research [46]-[50]. So far, within the specific problems studied, researchers have proposed a variety of methods to rank the influence of nodes. Accordingly, these sorting methods can be roughly divided into

four types: the structural centralities [53]-[57], the iterative refinement centralities [53], [58]-[61], the node operation methods and the dynamics-sensitive methods [62], [63]. These four kinds of methods contain many more detailed sorting methods. According to their different requirements and measurement standards, these methods can effectively rank the influence of nodes in specific networks or functional environments.

Among all of these methods, the neighbor-based methods are the simplest and most intuitive methods. A straightforward and efficient algorithm is used to directly count the number of neighbors nearest to a node, resulting in the degree centrality. Chen et al. proposed the semi-local centrality [54], an improved version to the degree centrality considering the information of the multi-order neighbors (four orders). K-shell decomposition [55] can be regarded as an extension of degree centrality. It defines the influence of nodes according to their positions in the network, i.e., the node's coreness. Simply put, a node's influence is strongly correlated to its capacity to impact the behaviors of its surrounding neighbors.

The path-based sorting method can be widely used in transportation networks, communication networks and social networks, as the more influential nodes play an important role in the transmission process of information packets and traffic flows. In these networks, we need to examine the ability of nodes to control the information flow, which is often closely related to the paths available in the network. For example, the closeness of the nodes [49] measures the influence of the nodes by calculating the average distance between each node and all other nodes in the network. The smaller the average distance between one node and the other nodes in the network, the greater the closeness of the node is. Closeness centrality can also be understood as the average time for information to completely spread throughout a network, starting from a single node. According to the betweenness centrality [57], the larger the number of the shortest paths through a node, the more important the node is. The betweenness centrality describes the control ability of nodes to the network flow along the shortest paths.

The previous methods are all measures of the influence of nodes based on the number of neighbors to each node. However, the method based on an eigenvector not only considers the number of neighbors a node has, but also considers the impact of their quality on the importance of the nodes. For example, with regards to the network of the internet, the famous PageRank algorithm [58] gives the score of a node increase along the access path by simulating the process of users browsing the web, which is used to identify the importance of each web page. In HITS algorithm [59], the automatic information aggregation algorithm [60] and the SALSA algorithm [61], the dual roles of nodes, authority and hub, and their interactions are considered. These methods have received great attention in the field of physics, chemistry, biology and business, and have important reference significance.

In the node removal and contraction method, the most significant feature is that in the process of ranking the influence of nodes, the structure of the network will be in a dynamic change, and the importance of the nodes is often reflected in the destruction of the network after the node is removed. From the perspective of measuring the robustness of the network [65], [66], once some nodes fail or are removed, the network may fall into paralysis or be divided into several disconnected subnets. Many

infrastructure networks in real life, such as power transmission networks, transportation networks, water and natural gas supply networks, etc., all have the risk of "one point of failure, whole network paralysis" [45]. In order to reduce or avoid the risk, researchers have proposed many methods to study the changes of network structures and functions after node shrinkage or removal, in order to provide the basis for the design and construction of new systems.

At present, the main method to evaluate the advantages and disadvantages of the above sorting algorithms can be described as follows. Take the important nodes from the sorting algorithm as the research object, and judge whether the sorting is appropriate by examining the influence of these nodes on some structure and function of the network and on the state of the other nodes. The commonly used methods to evaluate these ranking algorithms are the network-based robustness and vulnerability methods [65], [66], and the network-based propagation dynamics model [6]-[8]. For example, if a single node is selected as the initial propagation source, the larger the propagation scale, the greater the influence the node.

### 3.2 Propagation Dynamics

Another major topic for the research of complex networks is propagation dynamics [64], in which every node in the network is regarded as a dynamic system. In this process, the topology of the network remains static while the state of the nodes is constantly changing. One important piece of progress in this field refers to the dynamic process of network propagation. Dynamic behaviors exist widely in real-world networks, and they are different in networks with different functions. This section introduces the dynamic behaviors in the networks related to this thesis, which include typical propagation models and important conclusions and results related to network structure measurements and dynamic behaviors.

**(1) Epidemic Spread Model**

As a typical dynamic process, epidemic spreading has become a research hotspot in the field of network science for nearly 20 years, and many important breakthroughs have been obtained [1], [7], [8]. In the early stage of epidemic model research, the compartmental model [81]-[83] was mostly used, including the most typical and simple SI (Susceptible-Infected) model, the SIS (Susceptible-Infected-Susceptible) model and the SIR (Susceptible-Infected-Recovered) model. These models assume that the nodes in the network are uniformly mixed, that is, each individual has the same chance to come into contact with any other individual within a unit of time. The SI model is the simplest case of transmission, which assumes that an individual is always in an infectious state after being infected, such as HIV [82]. In the real world, many infectious diseases are curable, so there are more realistic epidemic spreading models such as SIS and SIR. The compartmental model is based on the completely mixed hypothesis, but in reality, individuals can only come in contact with a limited number of people in a crowd. Therefore, further study on the virus propagation models and analyses on homogeneous networks and heterogeneous networks has been conducted, considering of the impact of the real-world network structure measurements on propagation behaviors. It has been concluded that there is an important impact of the network structure measurements on the propagation velocity/speed, infection scale and epidemic

threshold.

In the study of the epidemic spreading process, the outbreak threshold of an epidemic has a very important impact on the early warning and selection of control strategies for the outbreak and spread of the epidemic. It was found that the research method of the macroscopic emergence such as the epidemic outbreak is closely related to the non-equilibrium phase transition in statistical physics. Inspired by this, in the past decade, many analytical methods of epidemic propagation have been proposed. These range from the classical mean-field method [32], [33] to the more rigorous quantitative numerical analytical methods [7]. At present, there are many methods for threshold analysis and various conclusions about the epidemic threshold in the research of complex networks. For example, in homogeneous networks, there is a limited epidemic threshold $\tau_c = 1/\langle k \rangle$, where $\langle k \rangle$ is the average degree of the network. When the infection rate is higher than the threshold, epidemics can exist in the network for a long time, but in the contrary situation, the viruses die out rapidly [70]. In heterogenous networks, the epidemic threshold is $\tau_c = \langle k \rangle / \langle k^2 \rangle$, assuming the network size $N \to \infty$, there is $\langle k^2 \rangle \to \infty$, therefore, $\tau_c \to 0$, the virus is more likely to be prevalent in such networks and exist for a long time [69], [70]. Another important measurement of the epidemic process is the infection scale, which is generally defined as the density of infection in the network at a stable state or within the same time period. The research has shown that, although homogeneous networks have a larger epidemic threshold than in heterogeneous networks, once the epidemic spreads in the network, the virus spreads faster and easier in homogeneous networks than in heterogeneous networks.

In order to explain and predict some important problems in the spread of epidemics, computer simulations have been carried out to predict virus spread accurately as it is extremely difficult or impossible to get the general analytical solutions even for the analysis of the propagation dynamics on a static network, and further the internal laws of network dynamics evolution are found. The complexity factors, e.g., network hierarchy and aggregation, have also forced researchers to develop appropriate mathematical analytical methods and calculation models. It is worth mentioning that, although the specific mechanisms of different dynamics are different, the epidemic propagation dynamics model and its analytical methods are also used for reference in other commonly existing dynamics behavior studies, such as information, rumor diffusion [14]-[16], [84]-[86], consensus [87]-[89] and cascading failure [18]-[20].

**(2) Cascading Failure Model**

In March 2019, from the afternoon of the 7th local time, a large-scale power outage occurred throughout Venezuela, including Caracas, until the evening of the 8th, when the power supply was partially restored. Venezuelan President Maduro said at a rally that about 70% of the power supply in the areas affected by the power outage had been restored, but at noon, the power system was hit by a new round of cyber-attacks, leading to another collapse. Due to the deliberate destruction of the Guri hydropower station in Venezuela, the power of the whole country fell into a state of collapse in just one day. The blackout was initially caused by a power plant fault, however, this initial fault triggered a series of cascading failures, including circuit overloads, which spread across the whole power grid. In addition,

the blackout also indirectly led to other losses, such as communication interruption, water and gas supply paralysis, traffic congestion and so on. The losses caused by this cascading failure were huge and unpredictable. Therefore, it is very important to explore the cascading failure of this network to propose effective reliability measures.

A cascading failure is a kind of network failure mode that can cause a series of failures and, ultimately, can lead to the propagation of a whole network failure [17]-[20]. Cascading failures exist widely in many real-world networks, such as power grids, transportation networks, the internet and economic networks. Therefore, many fields invest a lot of energy to explore cascading failures, so as to propose new measures to avoid such risks and reduce the catastrophic consequences of such cascading failures. At present, the related research of cascading failures mainly focuses on the analysis of the critical condition and failure mechanism, and the empirical research conducted on real-world network failures [71].

The fault process of a network can be regarded as a typical phase transition process, and the critical point of the phase transition indicates the overall failure of the network. The criticality analysis of network faults is generally based on the percolation theory [72]-[74]. By establishing the fault interdependence among the network nodes, the tolerance ability of the whole network function is analyzed by the tolerance of each node to risk. The study of critical conditions usually assumes that the failure depends on the topology of the network, and quantifies the impact of network structure measurement factors on network resilience.

The research of fault mechanisms is mainly based on the principle of fault propagation to construct network cascading failure models, which can better describe the actual network cascading failure process. For example, an overload is a common cascading failure mechanism, which mainly occurs in power networks, transportation networks and other networks with transfer tasks. In these networks, if the traffic on one node of the network exceeds its threshold value, it will lead to the redistribution of traffic on the node, which will aggravate the load pressure of other nodes, and may lead to an overload failure of these nodes, and finally form fault propagation. In the study of complex networks, the degree/ betweenness of nodes is closely related to its load capacity and load distribution, and then the cascading failure caused by the overload is modeled. The existed typical models include the Motter Lai (ML) model [75]-[77], the Crucitti model [78] and the OPA model [79]. Among these models, the ML model is the most widely used to simulate the cascading failure processes of various networks. The OPA model is mainly aimed at simulating the cascading failure processes in power grids considering many characteristics of power grid operations.

Compared with the theoretical analysis and model construction of cascading failures, the empirical research based on real-world failure data can achieve more reliable analysis conclusions. As a result, cascading failures in real-world networks have been widely concerned. Especially in the actual fault data of power grids, it has been found that due to the dual requirements of economy and reliability, power grids are required to always operate below the critical point of its own failure to meet the reliability requirements while get the maximum economic benefits. Under such conditions, once the cascading

failure occurs in the power network, the fault scale presents a power-law distribution, which greatly increases the probability of a large-scale failure [80].

### 3.3 Cooperative Evolution of Complex Network Dynamics

The above research on complex networks divides into two directions mainly, one is to pay attention to the measurement of network structures and reveal the structural characteristics of networks, such as small-world characteristics [26] and scale-free properties [27], while the other is to study the dynamics of networks, that is, each node in the network has a dynamic state, which is autonomous. Under a specific network topology, the states of nodes change according to the local evolution rules, such as epidemic spread, information diffusion, cascading failures, etc. However, in most real-world networks, the network structure is not unchangeable, but evolves with the states of the nodes in the network. For example, in the crowd contact network, the healthy individuals have the awareness of being vigilant, and they can avoid contacting infected individuals by changing their connections, and so in turn, the individuals' evasive behaviors can reduce the probability that the healthy individuals become infected, thus inhibiting the spread of the virus. In mobile communication networks, when a node fails, in order to ensure that the communication process continues, the communicating nodes adjust the relationship with the surrounding nodes adaptively, which leads to the change of the network structure, and the communication process between the individuals also changes.

Therefore, there is a close interaction between the structure of complex networks and their propagation dynamics. The change of nodes' states in the network will affect the evolution of a network's structure, while the change of network structure also impacts the propagation behavior in the network. In this way, a feedback loop is formed between the topology and node state of the network. This feedback loop is a complex interaction between a time-varying network structure and node dynamics. This kind of network with a feedback loop is called a cooperative evolutionary network or adaptive network [36]-[38].

Gross et al. [36] first proposed an adaptive network model based on the SIS epidemic model and introduced a new important parameter, the rewiring rate $w$. Because the adaptive behavior of individuals can change their connection, which will have a very important impact on the process of disease transmission, a lot of dynamic phenomena appear in the adaptive network. Among them, the bistable and oscillatory states in the adaptive network are characteristics that do not appear in static networks. With the development of adaptive network research, many adaptive network models based on real-world network characteristics have been proposed, e.g., the adaptive network model based on the Susceptible-Infected-Recovered-Susceptible (SIRS) epidemic model [37]-[41], and the adaptive SIS model which considers community characteristics [90], [91], etc. Based on the analysis and research of these models under different network structures, the researchers proposed a variety of immunological and isolating strategies [90]-[93] for virus spread to put forward more targeted control measures, in order to better prevent the spread of disease.

Besides the processes of virus propagation in adaptive networks, network structures and many other

propagation behaviors have this kind of adaptive cooperative evolution relationship, such as adaptive information diffusion [56], adaptive election [83], etc. On the one hand, the spreading of a disease in an adaptive network is similar to other propagation mechanisms, such as the transmission being limited to the local neighborhood, the dynamic processes of the model being random, and the topology changes with the update of node state. On the other hand, there are great differences between them, for example, the state in the disease model is asymmetric, as only the infected node can spread a disease, while in the opinion formation model, different opinions are equal, so the evolution mechanism is symmetric. Therefore, considering the characteristics of dynamics, we can use the adaptive network virus propagation model to analyze the propagation process of other dynamics.

# Chapter 4

# Study on the Rapid Identification of High-influence Nodes in Complex Networks

## 4.1 Introduction

The high influential individuals in complex systems play a critical role in the dynamics on complex systems, for instance, the target population that need to be immune after an outbreak of an infectious disease, the optimal spreaders for information diffusion in online social networks [49], priority protection should be given to the important breakers and power units to prevent power outages caused by large-scale cascading failure in power grid and so on. Identifying those influential nodes and finding effective ranking methods of node influence are significant subjects in the field of complex networks.

How to design a ranking measure considering both universality and the effectiveness in complex networks has become one of the core problems affecting nodes study. In the last few years, various methods [46] have been introduced for studying the "importance" of nodes within complex network structures, mainly including the ranking methods based on nodes neighbors, such as degree centrality [53] and k-shell decomposition [54], the ranking methods based on path, such as betweenness centrality [57] and closeness centrality [50], the ranking methods based on eigenvector, such as the eigenvector centrality [53], [58], PageRank [58]and HITs algorithm [59] and the ranking methods based on nodes removal and shrink, such as the shortest distance method of node deletion [62] and node contraction method [63]. Those methods have been proved to be effective on nodes ranking through a large number of experiments. In light of these classic methods, many new improved methods were proposed, such as the semi-local centrality [47], LeaderRank [48] and some new ones based on the network topology [105]-[107]. Measures to evaluate the ranking methods are usually based on the dynamics [16] and the robustness and fragility of networks. The influence of nodes is judged by its impacts on some network structures, functions and other nodes status in the network. Anyway, designing an effective sorting method is one of the most crucial problems in complex networks research.

Currently, most of the researchers are focusing on the accuracy of the method to identify the most influential nodes, namely to find the most effective method of node ranking. In fact, now the networks we are facing demonstrate many characteristics which can be ignored in the past: 1) Extensive. With the development and progress of science and technology, we are now in a big data age, the scales of networks are growing fast. In this case, even the simplest ranking method will be time-consuming. 2) Adaptive. Many real-world networks are characterized by adaptive changes in their topology depending on the state of their nodes [36], the nodes order we ranked in the previous time may fail to be sorted in the current time. 3) Time-sensitive. Dynamic behaviors in real networks are time-sensitive, which means we need to find the influential nodes in a limited time to apply it into the dynamics. In addition, the main purpose of study on node identification is to find a fraction of top influential nodes for application, so mostly it is not necessary to rank all nodes. Based on the reasons above, instead of ranking all nodes in the network, we focus on finding a fraction of high-influence nodes in networks.

In this chapter, we proposed a method in light of node degree and network structure to identify high-influence nodes rapidly in networks. Compared with the results gained from existed ranking methods, our method is proved to be better on identifying the fraction of high-influence nodes. Moreover, we use the SIR model, in which the high-influence nodes are regarded as spreaders and protected to evaluate the performance of our method. The simulations on different networks show that our method performs well on identifying high-influence nodes.

Following parts are organized as follows. First, we introduce our new measure in Section 4.2, and in order to detect and analyze our measure, the definition of some classic centrality measures in Section 4.2 has been briefly reviewed. The data description is presented in Section 4.3, and the effectiveness of our measures is discussed and analyzed, and then we use the SIR model to evaluate the performance of our method in different networks. Conclusions are given in Section 4.4.

## 4.2 Methods

Among those different centrality measures, each one has its scope of application and their respective characteristic. Here we briefly introduce some classic and authoritative methods, which will be used in experimental analysis in Section 4.4.

A simple one is degree centrality, namely, the larger degree it is, the higher influence it will get. Compared with the centrality measure based on nodes local properties, methods considering the global information give better ranking results, such as betweenness centrality and closeness centrality as introduced in Chapter 2.

As a local property, degree considers only local information of the node itself. So the degree centrality has lower computational complexity but also is low-relevant. Comparing with degree centrality, betweenness and closeness centrality measures considering the global information can better quantify the influence of node, but they are more time-consuming. Local centrality measure is proposed as a tradeoff between low-relevant degree centrality and other time-consuming measures. It considers both the nearest and the next nearest neighbors. The local centrality $CL_i$ of node $i$ is defined as

$$Q_j = \sum_{k \in \Gamma_j} N_k \,, \tag{4.1}$$

$$CL_i = \sum_{j \in \Gamma_i} Q_j \,, \tag{4.2}$$

where $\Gamma_j$ is the set of the nearest neighbors of node $j$, $N_k$ is the number of the nearest and the next nearest neighbors of node $k$.

With the deepening of the research on node influence, a growing number of methods have been proposed. Most of these methods are based on the previous methods for improvement and get the desired effect.

Instead of ranking all nodes, we are focused on identifying a fraction of high-influence nodes. Considering the facts of uncertainty of network scale and topology, the timeliness of dynamic behaviors in real networks, we propose a method to rapidly identify high-influence nodes based on nodes local properties. Our method is introduced as follows:

1) $m$ nodes will be randomly chosen as the first source nodes, denoted as $s_i$ , $i$ = 1, 2, …, $m$,
2) find the highest-degree neighbor of each $s_i$ as the second source nodes, denoted as $s_{i\_N(1)}$ , $i$ = 1, 2, …$m$, and then find the highest-degree neighbor of each $s_{i\_N(1)}$ as the third source nodes, according to this way, we will stop until we find the $k$ + 1 source nodes, denoted as $s_{i\_N(k)}$ , $i$ = 1, 2, …, $m$, which means we will find $m \times k$ nodes, denoted by $s_{i\_N(K)}$ , $i$ = 1, 2, …, $m$, $K$ = 1, 2, …, $k$. the first source nodes $s_i$ will not be counted,
3) rank the $m \times k$ nodes we found according to degree, choose the top-$j$ nodes $T_j$ as the target nodes ( $T_j \in s_{i\_N(K)}$ ).

Considering that degree is one of the simplest but most important parameters on describing local characteristics of nodes, we can implement our method without knowing the global information of the network. In addition, the close relationship among hubs (high-degree nodes) makes our method work.

### 4.3 Experimental Analysis

#### 4.3.1 Data

In experiments, a set of artificial networks are used to test the effectiveness of our method. Furthermore, two real networks are used to evaluate the performance of our method applying in real networks [26]: (i) Facebook - the complete Facebook network data (from a single-time snapshot in September 2005) of Caltech. Only intra-college links are included. There are in total 769 individuals in this Facebook network. We here consider the largest component with 762 individuals. (ii) Email - the network of Email interchanges between members of the University Rovira i Virgili (Tarragona). The basic topological properties of the empirical networks and the artificial networks are shown in Table 4.1 and Table 4.2.

Table 4.1 The artificial networks we study and their basic properties.

| Networks | *N* | <*k*> | *k_max* | *C* | <*d*> | *r* |
|---|---|---|---|---|---|---|
| Scale-free network with clustering1 | 500 | 6 | 65 | 0.05 | 3.2166 | -0.092 |
| Scale-free network with clustering2 | 500 | 6 | 104 | 0.27 | 3.2573 | -0.080 |
| Scale-free network with clustering3 | 500 | 6 | 77 | 0.55 | 3.4794 | -0.109 |
| Scale-free network with assortativity1 | 500 | 4 | 48 | 0.0325 | 3.8629 | -0.1 |
| Scale-free network with assortativity2 | 500 | 4 | 48 | 0.0333 | 3.8893 | 0 |
| Scale-free network with assortativity3 | 500 | 4 | 48 | 0.0295 | 4.2205 | 0.2 |

Table 4.2 The empirical networks we study and their basic properties. *N* and *L* are the total numbers of nodes and links, respectively. <*k*> and *k_max* denote the average and the maximum degree. <*d*> is the average shortest distance. *C* and *r* are the clustering coefficient [26] and assortative coefficient [110], respectively.

| Networks | *N* | *L* | <*k*> | *k_max* | *C* | <*d*> | *r* |
|---|---|---|---|---|---|---|---|
| Email | 1133 | 5451 | 9.62 | 71 | 0.0096 | 3.6060 | 0.0783 |
| Facebook | 762 | 16651 | 43.70 | 248 | 0.4093 | 2.3378 | -0.0662 |

### 4.3.2 SIR Spreading Model

Spreading models measuring the dynamics and the robustness and fragility of networks are used to evaluate the performance of ranking methods. The influence of nodes is judged by its impacts on network structures, functions and other nodes status in the network. Most of the researchers are using network models to evaluate the ranking methods. For example, calculate the nodes influence by SIS/SIR model or judge the nodes importance by attacking.

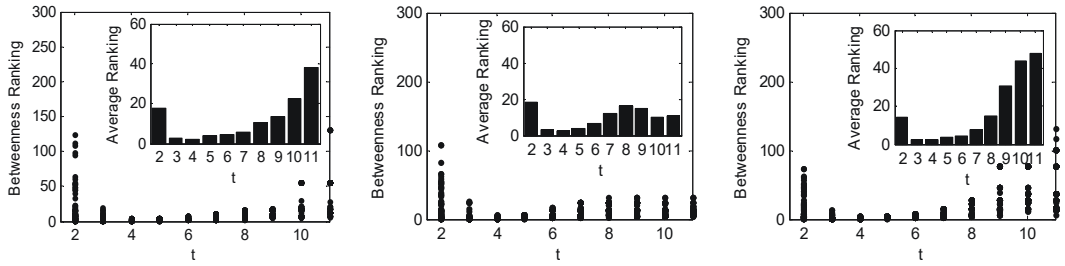
We use the SIR model to examine the performance of our method in different networks. In SIR model, there are three states [1]: (i) Susceptible(S), (ii) Infected(I), (iii) Recovered(R). S individuals are susceptible to (not yet infected) the disease, I individuals have been infected and is able to spread the disease to susceptible individuals and R individuals have been recovered and will never be infected again.

To investigate the influence of target nodes we found in the network, these nodes are first set to be infected or to be protected initially. The proportion of infected and recovered nodes at time *t*, denoted by *f(t)*, can be considered as an indicator to evaluate the influence of the initially infected nodes at time *t*. *f(t)* increases with *t*, and finally gets stable when there is no infected node in the network. Thus both of the spreading velocity and the spreading scale evaluate the influence of the initially infected nodes.

### 4.3.3 Effectiveness

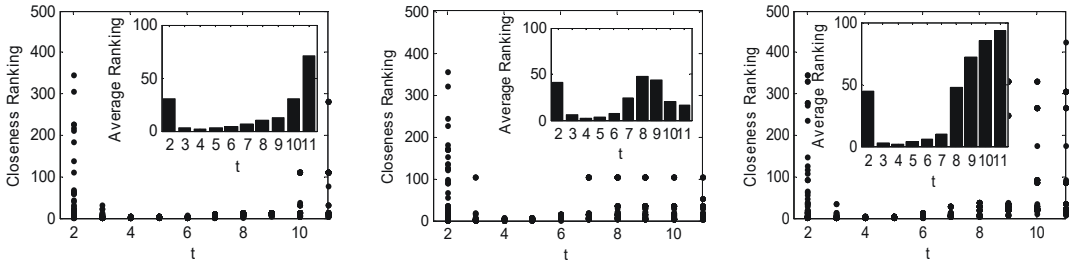
We use different ranking methods to verify the effectiveness of our method. At first, in order to show the results clearly, we simplify our method that in each implementation only one node is randomly selected as the first source node, and 10 nodes will be found according to our method, which means $m = 1$, $k = 10$. After *n* implementations (each node is randomly selected as the first source node once and only once), the results of node's ranking and the average ranking of *n* implementations on each step t by their corresponding degree centrality (DC), closeness centrality (CC), betweenness centrality (BC) and the local centrality (CL) in different artificial networks and real networks are shown in Fig 4.1-Fig 4.10, including scale-free (SF) networks with different values of the clustering coefficient *C* (Fig 4.1-Fig 4.4) and the coefficient of assortativity *r* (Fig 4.5-Fig 4.8), Email network (Fig 4.9) and Caltech Facebook network (Fig 4.10).

We can see from the Figs that the top-ranking nodes can be identified before step 8, and the nodes' average ranking in step 3-step 7 by different centralities is smaller than at other steps in different network structures, which means the nodes in step 3-step 7 are the most influential nodes we find in 10 steps.



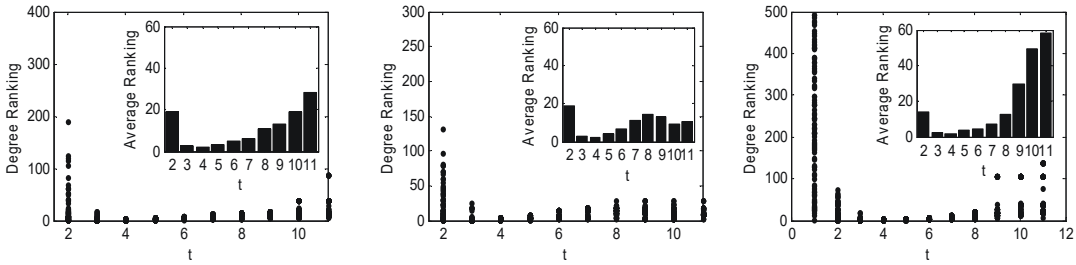

(a) (b) (c)

Fig 4.1 The nodes' ranking and the average ranking of $n$ implementations by betweenness centrality (BC) for artificial networks with exponential degree distributions, with $n$ = 100, and different values of the clustering coefficient $C$. (a) $C$ = 0.05, (b) $C$ = 0.27, (c) $C$ = 0.55, respectively.



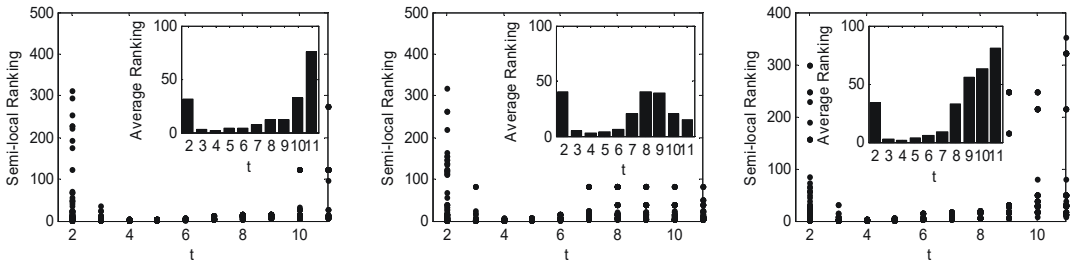

(a) (b) (c)

Fig 4.2 The node's ranking and the average ranking of $n$ implementations by closeness centrality (CC) for artificial networks with exponential degree distributions, with $n$ = 100, and different values of the clustering coefficient $C$. (a) $C$ = 0.05, (b) $C$ = 0.27, (c) $C$ = 0.55, respectively.




(a) (b) (c)

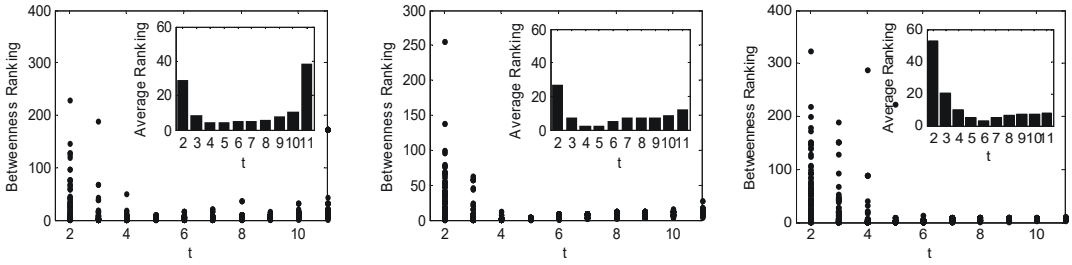
Fig 4.3 The node's ranking and the average ranking of $n$ implementations by degree centrality (DC) for artificial networks with exponential degree distributions, with $n$ = 100, and different values of the clustering coefficient $C$. (a) $C$ = 0.05, (b) $C$ = 0.27, (c) $C$ = 0.55, respectively.




(a) (b) (c)

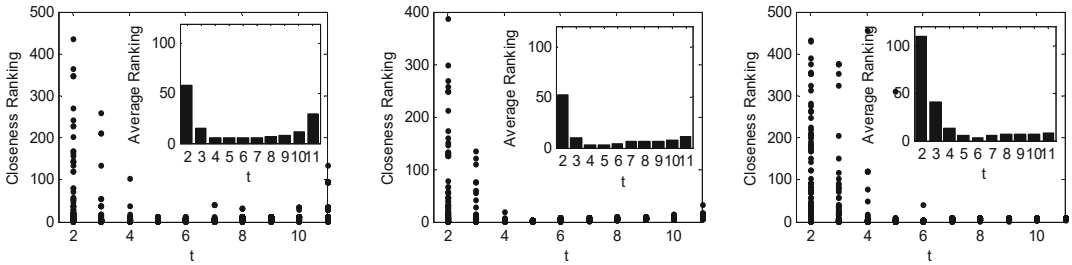
Fig 4.4 The node's ranking and the average ranking of $n$ implementations by semi-local centrality (CL) for

artificial networks with exponential degree distributions, with $n$ = 100, and different values of the clustering coefficient $C$. (a) $C$ = 0.05, (b) $C$ = 0.27, (c) $C$ = 0.55, respectively.



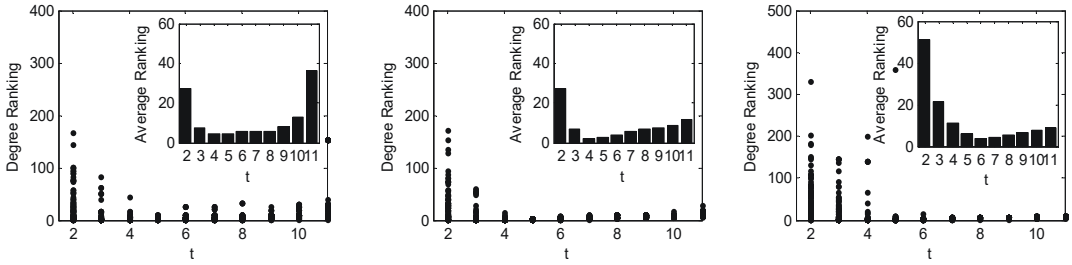

(a) (b) (c)

Fig 4.5 The node's ranking and the average ranking of $n$ implementations by betweenness centrality (BC) for artificial networks with exponential degree distributions, with $n$ = 100, and different values of the coefficient of assortativity $r$. (a) $r$ = -0.1, (b) $r$ = 0, (c) $r$ = 0.2, respectively.



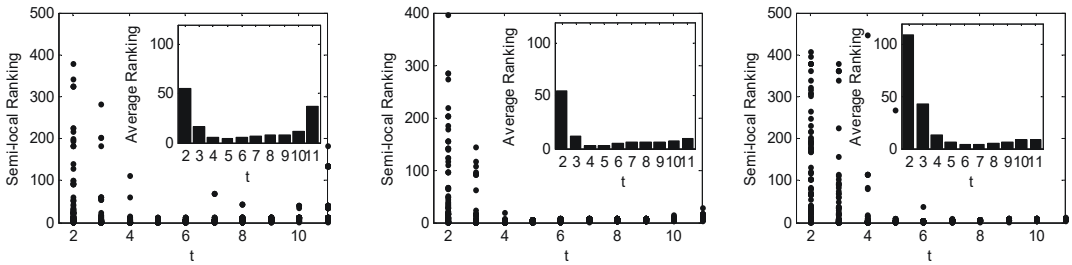

(a) (b) (c)

Fig 4.6 The node's ranking and the average ranking of $n$ implementations by closeness centrality (CC) for artificial networks with exponential degree distributions, with $n$ = 100, and different values of the coefficient of assortativity $r$. (a) $r$ = -0.1, (b) $r$ = 0, (c) $r$ = 0.2, respectively.



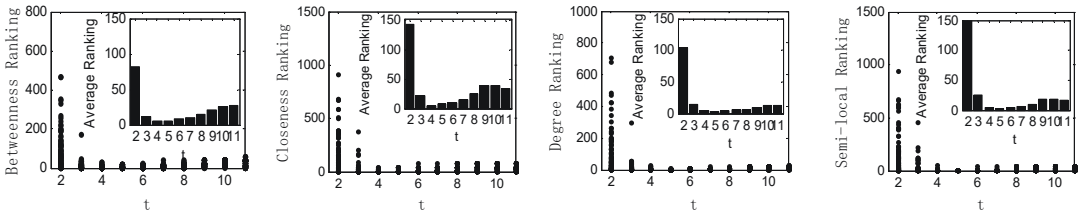

(a) (b) (c)

Fig 4.7 The node's ranking and the average ranking of $n$ implementations by degree centrality (DC) for artificial networks with exponential degree distributions, with $n$ = 100, and different values of the coefficient of assortativity $r$. (a) $r$ = -0.1, (b) $r$ = 0, (c) $r$ = 0.2, respectively.



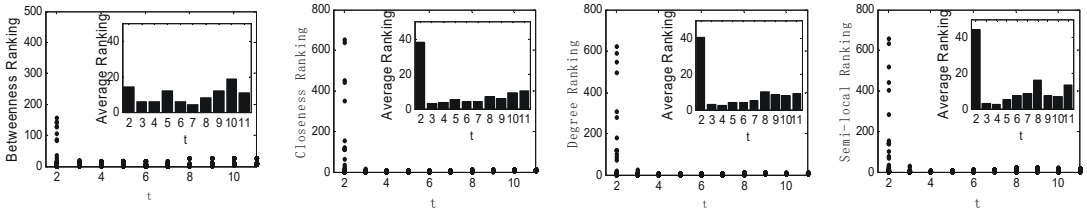
(a) (b) (c)

Fig 4.8 The node's ranking and the average ranking of $n$ implementations by semi-local centrality (CL) for artificial networks with exponential degree distributions, with $n$ = 100, and different values of the coefficient of assortativity $r$. (a) $r$ = -0.1, (b) $r$ = 0, (c) $r$ = 0.2, respectively.




(a) (b) (c) (d)

Fig 4.9 The node's ranking and the average ranking of $n$ implementations for email network by four different centralities, (a) BC, (b) CC, (c) DC, (d) CL, respectively, with $n$ = 100.




(a) (b) (c) (d)

Fig 4.10 The node's ranking and the average ranking of $n$ implementations for Caltech Facebook network by four different centralities, (a) BC, (b) CC, (c) DC, (d) CL, respectively, with $n$ = 100.

We can see from figures above that the results among different ranking measures on identifying high-influence nodes are high-related. For example, the high-influence nodes under degree centrality will be relatively high-influence under other ranking measures. The results are shown in Table 4.3-Table 4.5.

Table 4.3 The top-10 ranked nodes by degree centrality and their corresponding ranks by betweenness, closeness and semi-local centralities in SF networks with different assortativities.

| | SF network with assortativity1 | | | | SF network with assortativity2 | | | | SF network with assortativity3 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| N | DC | BC | CC | CL | DC | BC | CC | CL | DC | BC | CC | CL |
| 5 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 |
| 1 | 2 | 2 | 3 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 |
| 25 | 3 | 5 | 6 | 6 | 3 | 4 | 4 | 5 | 3 | 3 | 4 | 5 |
| 7 | 4 | 3 | 4 | 5 | 4 | 3 | 3 | 3 | 4 | 4 | 3 | 3 |
| 8 | 5 | 4 | 5 | 4 | 5 | 5 | 5 | 4 | 5 | 5 | 5 | 4 |
| 12 | 6 | 8 | 8 | 7 | 6 | 10 | 10 | 6 | 6 | 10 | 9 | 9 |
| 11 | 7 | 9 | 12 | 13 | 7 | 6 | 6 | 7 | 7 | 7 | 9 | 9 |
| 37 | 8 | 7 | 7 | 9 | 8 | 7 | 8 | 10 | 8 | 6 | 8 | 10 |
| 29 | 9 | 11 | 19 | 24 | 9 | 9 | 9 | 9 | 9 | 9 | 12 | 16 |
| 6 | 10 | 6 | 2 | 3 | 10 | 8 | 7 | 8 | 10 | 8 | 6 | 7 |

Table 4.4 The top-10 ranked nodes by degree centrality and their corresponding ranks by betweenness, closeness and semi-local centralities in SF networks with different clustering coefficients.

| SF network with clustering1 | | | | | SF network with clustering2 | | | | | SF network with clustering3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| N | DC | BC | CC | CL | N | DC | BC | CC | CL | N | DC | BC | CC | CL |
| 6 | 1 | 1 | 2 | 1 | 4 | 1 | 1 | 1 | 1 | 10 | 1 | 2 | 2 | 2 |
| 4 | 2 | 2 | 1 | 2 | 3 | 2 | 2 | 2 | 2 | 3 | 2 | 1 | 1 | 1 |
| 8 | 3 | 4 | 4 | 4 | 11 | 3 | 4 | 5 | 5 | 4 | 3 | 3 | 4 | 3 |
| 7 | 4 | 3 | 5 | 5 | 8 | 4 | 3 | 3 | 3 | 6 | 4 | 4 | 5 | 5 |
| 1 | 5 | 5 | 3 | 3 | 12 | 5 | 6 | 7 | 6 | 5 | 5 | 5 | 3 | 4 |
| 18 | 6 | 6 | 7 | 7 | 7 | 6 | 5 | 6 | 7 | 12 | 6 | 6 | 7 | 6 |
| 5 | 7 | 7 | 8 | 8 | 5 | 7 | 7 | 4 | 4 | 22 | 7 | 7 | 9 | 8 |
| 3 | 8 | 9 | 9 | 9 | 17 | 8 | 9 | 9 | 9 | 20 | 8 | 8 | 15 | 16 |
| 9 | 9 | 8 | 6 | 6 | 6 | 9 | 8 | 8 | 8 | 15 | 9 | 10 | 11 | 7 |
| 2 | 10 | 10 | 10 | 10 | 8 | 10 | 11 | 11 | 10 | 1 | 10 | 18 | 10 | 25 |

Table 4.5 The top-10 ranked nodes by degree centrality and their corresponding ranks by betweenness, closeness and semi-local centralities in real networks.

| Email network | | | | | Caltech Facebook network | | | | |
|---|---|---|---|---|---|---|---|---|---|
| N | DC | BC | CC | CL | N | DC | BC | CC | CL |
| 105 | 1 | 2 | 3 | 1 | 619 | 1 | 1 | 1 | 1 |
| 333 | 2 | 1 | 1 | 3 | 561 | 2 | 4 | 2 | 2 |
| 16 | 3 | 22 | 40 | 4 | 205 | 3 | 2 | 3 | 6 |
| 23 | 4 | 3 | 2 | 5 | 403 | 4 | 7 | 6 | 5 |
| 42 | 5 | 10 | 4 | 2 | 409 | 5 | 16 | 7 | 4 |
| 41 | 6 | 8 | 5 | 8 | 60 | 6 | 3 | 4 | 12 |
| 196 | 7 | 15 | 19 | 7 | 124 | 7 | 26 | 10 | 3 |
| 233 | 8 | 6 | 7 | 22 | 455 | 8 | 12 | 5 | 7 |
| 21 | 9 | 16 | 20 | 9 | 82 | 9 | 11 | 11 | 10 |
| 76 | 10 | 5 | 6 | 19 | 644 | 10 | 8 | 9 | 11 |

Based on the above results, we set $m = 5$, $k = 10$, in each $k + 1$ step, nodes in step 3-step 7 are chosen, then the 25 nodes we found are ranking according to degree, the top 5 nodes are chosen as the target nodes. We show the target nodes ranking rate in top 5/10/15 individually under different centralities in artificial networks and real networks in Table 4.6.

Table 4.6 The target nodes ranking rate in top 5/10/15 individually under different centralities in artificial networks and real networks. The results are the average ranking of 200 implementations.

| Networks | Rate in DC Top 5/10/15 | Rate in BC Top 5/10/15 | Rate in CC Top 5/10/15 | Rate in CL Top5/10/15 |
|---|---|---|---|---|
| SF network with clustering1 | 0.9872/1/1 | 0.9912/1/1 | 0.7955/1/1 | 0.7972/1/1 |
| SF network with clustering2 | 0.9645/0.9798/1 | 0.7617/0.9862/1 | 0.7649/0.9856/0.97 | 0.7607/0.9813/0.9836 |
| SF network with clustering3 | 1/1/1 | 1/1/1 | 1/1/1 | 1/1/1 |
| SF network with assortativity1 | 0.9852/1/1 | 0.9918/1/1 | 0.7950/1/1 | 0.8/1/1 |
| SF network with assortativity2 | 0.9632/1/1 | 0.9571/1/1 | 0.9564/1/1 | 0.9507/1/1 |

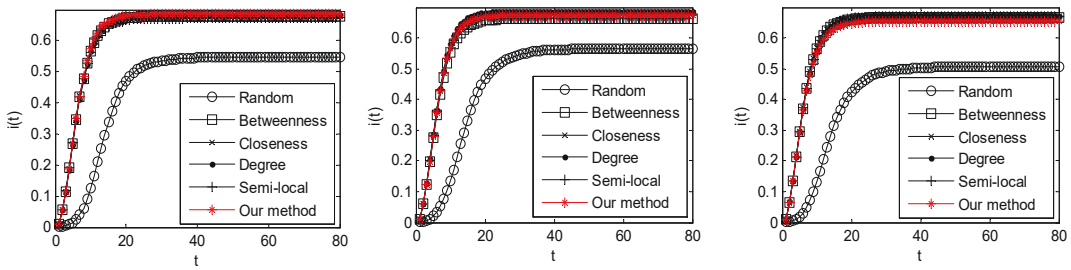
| | | | | |
|---|---|---|---|---|
| SF network with assortativity3 | 0.9951/1/1 | 1/1/1 | 1/1/1 | 1/1/1 |
| Email | 0.9956/1/1 | 0.6/0.8/0.8132 | 0.8019/0.8/0.8081 | 0.9913/1/1 |
| Facebook | 1/1/1 | 0.6/0.8/0.8 | 0/0.6/1 | 0.8/1/1 |

### 4.4 Evaluation in SIR Epidemic Model

To evaluate the performance of our method, we use the SIR model to examine the spreading influence of the target nodes. At each step, susceptible neighbor gets infected with probability $\alpha$ (here we set $\alpha=0.1$). Infected nodes recover with probability $\beta$ at each step (here we set $\beta=0.1$). To investigate the influence of target nodes we found in the network, we set these nodes to be infected or to be protected initially. The proportion of infected and recovered nodes at time $t$, denoted by $i(t)$, can be considered as an indicator to evaluate the influence of the target nodes at time $t$.

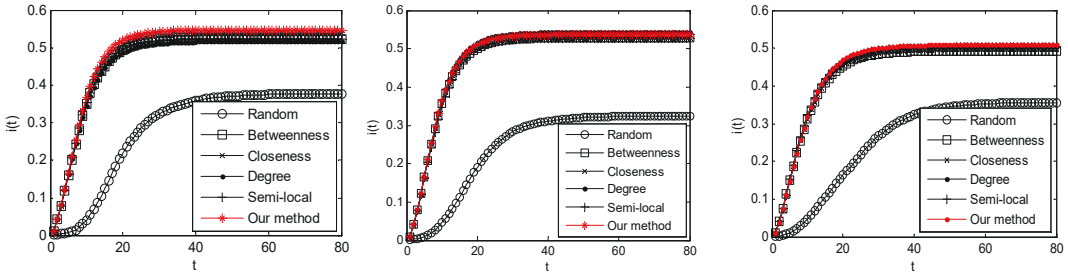
The results are shown in Fig 4.11-Fig 4.16. As the infection resources, show the proportion of infected and recovered nodes as a function of time, with the initially infected nodes we find by our method. In Fig 4.11-Fig 4.13, compared with those appear in the top-5 nodes list by different centralities in different network structures, our method can also effectively identify the nodes that lead to faster and wider spreading than the case that infection resources are randomly chosen. We can see from the results that, when multiple spreaders are considered simultaneously, the effectiveness of different ranking methods are almost the same in different network structures.

Different results are shown in the case of nodes being protected. In Fig 4.14-Fig 4.16, we can see that, compared with those appear in the top-5 nodes list by different centralities in different network structures, our method can also effectively identify the nodes that retard the spreading behavior than the case that of the protected nodes are randomly chosen. But the effectiveness differs from different ranking methods.



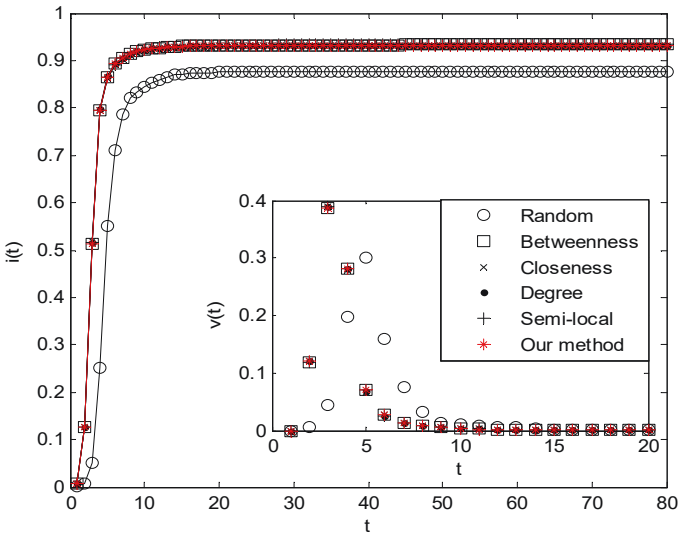

(a) (b) (c)

Fig 4.11 The spreading process as a function of time, with the initially infected nodes we find by our method, compared with those appear in the top-5 list by different centralities in artificial networks with exponential degree distributions, and different values of the clustering coefficient $C$. (a) $C = 0.05$, (b) $C = 0.27$, (c) $C = 0.55$, respectively. Results are obtained by averaging over 100 implementations.

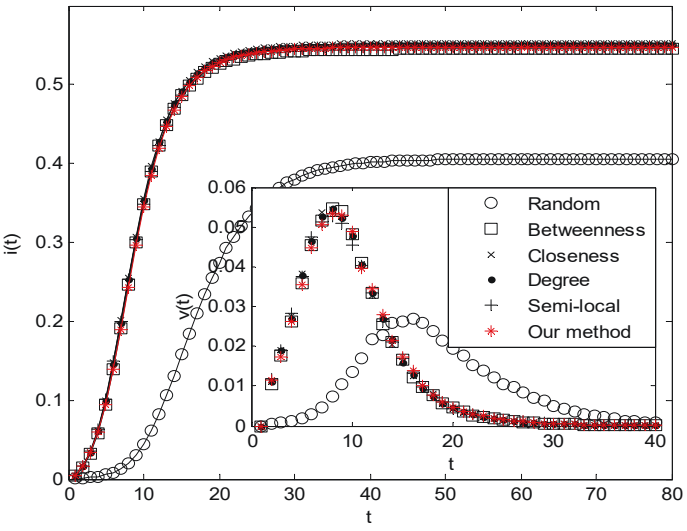



(a) (b) (c)

Fig 4.12 The spreading process as a function of time, with the initially infected nodes we find by our method, compared with those appear in the top-5 list by different centralities in artificial networks with exponential degree distributions, and different values of the coefficient of assortativity *r*. (a) *r* = -0.1, (b) *r* = 0, (c) *r* = 0.2, respectively. Results are obtained by averaging over 100 implementations.



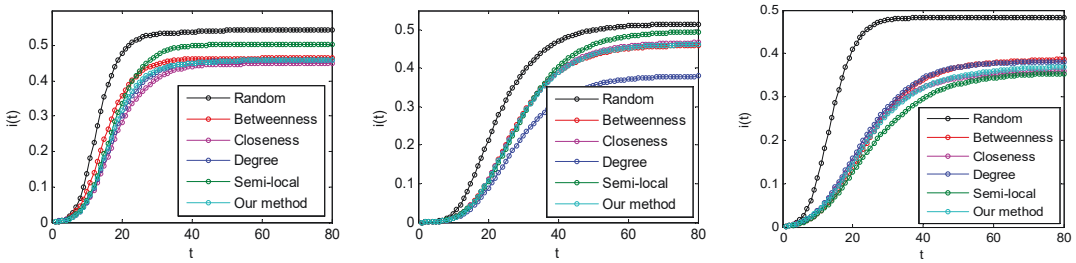

(a) Caltech Facebook network




(b) Email network

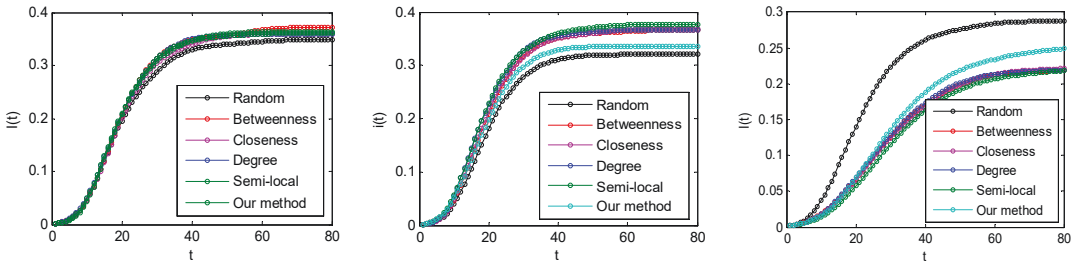
Fig 4.13 The spreading process as a function of time, with the initially infected nodes we find by our method,

compared with those appear in the top-5 list by different centralities in real networks. (a) Caltech Facebook network, (b) Email network, respectively. Results are obtained by averaging over 100 implementations.



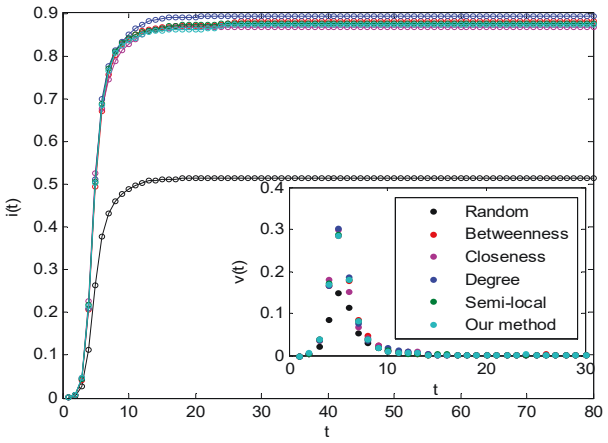

(a) (b) (c)

Fig 4.14 The spreading process as a function of time, with the protected-nodes we find by our method, compared with those appear in the top-5 list by different centralities in artificial networks with exponential degree distributions, and different values of the clustering coefficient $C$. (a) $C = 0.05$, (b) $C = 0.27$, (c) $C = 0.55$, respectively. Results are obtained by averaging over 100 implementations.

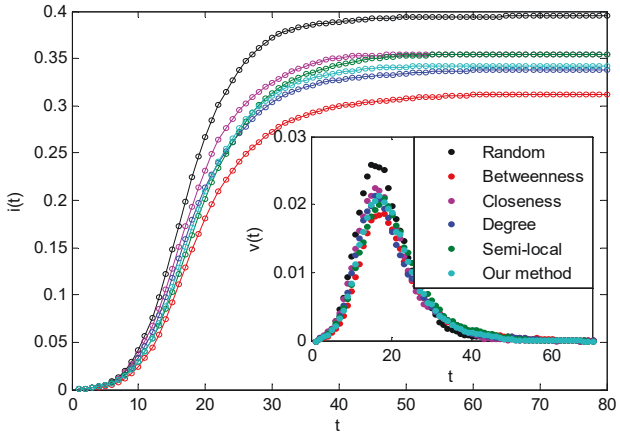



(a) (b) (c)

Fig 4.15 The spreading process as a function of time, with the protected-nodes we find by our method, compared with those appear in the top-5 list by different centralities in artificial networks with exponential degree distributions, and different values of the coefficient of assortativity $r$. (a) $r = -0.1$, (b) $r = 0$, (c) $r = 0.2$, respectively. Results are obtained by averaging over 100 implementations.




(a) Caltech Facebook network




(b) Email network

Fig 4.16 The spreading process as a function of time, with the protected-nodes we find by our method, compared with those appear in the top-5 list by different centralities in real networks. (a) Caltech Facebook network, (b) Email network, respectively. Results are obtained by averaging over 100 implementations.

## 4.5 Conclusion

This chapter considered the facts of uncertainty of network scale and topology, the timeliness of dynamic behaviors in real networks and proposed a method without ranking all nodes. To evaluate the performance, the SIR model is used to estimate the spreading influence of the target nodes we identified by our methods, comparing with the top-ranked nodes by different centrality measures. It is expected that with the influential nodes being initially infected the spreading are faster and wider than with the random nodes being initially infected, the experimental results on artificial networks (networks with different values of the clustering coefficient and the coefficient of assortativity) and real networks (the email communication network, the Caltech Facebook network) show that the proposed method can well identify high-influential nodes. The newly proposed measure performs almost as well as the well-known centrality measures, while with much lower computational complexity.

# Chapter 5

# Study on Cascading Failure of Complex Networks Considering Local Real-time Information

## 5.1 Introduction

Cascading failure [71]-[75], [78] is a very common phenomenon in real-life systems, e.g., power transmission, computer networking, finance, human body systems and transportation systems. Normally, the failure of one or several parts in the network may trigger the failure of other parts, resulting in the paralysis of large-scale network in a very short time. For example, in power grids, the failed (completely or partially) elements who are beyond their capacity become overloaded and shift their load to nearby elements in the system. The cascading process can make a large power grid collapse under certain conditions. Along with the explosive growth of real-life systems, the extensive and complex interconnection has not only brought the source sharing and optimal allocation, but also brought new challenges to the security and stability with the complexity of network dynamics. Therefore, the network robustness facing cascading failure becomes a very significant topic from the perspective of complex networks.

The real-life system can be represented by a complex network, a graph (network) with non-trivial topological features, e.g., a heavy tail in the degree distribution [27], a high clustering coefficient [26], assortativity or disassortativity among vertices [110]. Recently, based on the study of complex network, a lot of cascading failure models are proposed, which can further explore the mechanism of occurrence, prevention and control of cascading failures in depth. Motter et al. [75] studied cascading failure of scale-free networks for the first time, and proposed a capacity load cascading failure model. It was found that the robustness of scale-free networks was very weak in the face of cascading failures, that is, removing a small number of nodes with the largest load was enough to paralyze the whole network. In view of the actual cascading failure, Goh et al. [111] found that there is a long-range correlation between cascading failures, which is represented by an approximate power-law function. After that, Dobson et al. [112] derived the critical capacity value of power-law cascading failure scale. Based on the heterogeneity of the network, Peng et al. [113] studied the influence of attacking different nodes on the network robustness threshold. Liu et al. [114] considered the adjustable point capacity according to the importance of nodes, and proposed a cascading failure resistance strategy based on node capacity optimization, which effectively improved the resistance ability of scale-free networks to cascading failures.

In fact, although the existing model can well describe the cascading failure process in real-life systems, new challenges emerge due to the continuous evolution of the networks and the new features developed during the evolution. One of the most important challenges is the timeliness of network information. Network information is updated as the network evolved. Therefore, the initial information is not accurate for the current network due to the updates. The information of current network is necessary to the redistributions of excess loads. The second challenge is that along with the explosive growth of

networks, huge volumes of data have been generating [115]. In this case, it is unrealistic to redistribute the excess loads based on the global information of the current networks. Instead, it is more reasonable to redistribute the loads caused by node's failed based on local information of the current networks.

Considering the timeliness of network information and rationality of redistribution strategy, this chapter presents a new cascading failure model with the redistribution strategy based on local real-time information. The key contributions of the chapter can be summarized as follows:

1) We develop a new cascading failure model to analyze the robustness of networks under the node's failure, and redistribution strategy based on local real-time information is proposed.
2) We carry out the theoretical analysis of the robustness and the redistribution strategy, where the relationship between network robustness and redistribution strategy parameters are discussed in different cases.
3) The numerical simulations implemented in artificial networks and real-life networks confirm the validity of the robustness threshold, as well as the analysis of the redistribution strategy.

The rest of this chapter is organized as follows. In Section 5.2, the proposed cascading failure model of adaptive weighted network is presented. In Section 5.3, the theoretical analysis of the model is carries out, followed by the discussion of the network robustness in different cases. In Section 5.4, numerical results are provided, followed by conclusions in Section 5.5.

## 5.2 Cascading Failure Model

Consider a network of $N$ nodes connected by $M$ links, each node has an initial load $L(0)$ and load capability $C$, At the initial stage, $L(0) < C$, the network is in the steady state. After the node is removed (attacked or mechanical failure), the load on the failed node is redistributed to its neighbors. When the load capacity of the neighbor is not enough to handle the extra loads, the neighbor would collapse. The continuation of this process creates cascading failures in the network. Assume that the potential cascading failure is caused by the removal of one single node, then we pay attention to the following dynamic process of the network. Our model is briefly described as follows,

The initial load of node $i$ is set as

$$L_i(0) = \beta_1 k_i^{\alpha_1}, i = 1, 2, ..., N \ , \tag{5.1}$$

where $k_i$ is the degree of node $i$. $\alpha_1$ and $\beta_1$ are the adjustable parameters that govern the strength of initial loads, $\alpha_1 \geq 0$ and $\beta_1 \geq 0$. The initial load of the node is proportional to its degree, which means the nodes with larger degree can take more load. For example, the transportation hubs with more routes across can accommodate more passengers.

At the initial stage, the network is in steady state, $L_i$ is lower than its load capability $C_i$, as given by

$$C_i = (1+\beta_m)L_i(0), \tag{5.2}$$

where $\beta_m$ ($\beta_m > 0$) is the tolerance parameter and suggests the upper bound of the ability that node $i$

withstands risks, i.e., the network shows better robustness against cascading failures as $\beta_m$ grows. However, the continuous grow of load capacity is unreasonable, considering the cost of real networks. Therefore, we aim to minimum the upper bound of $\beta_m$, i.e., the critical threshold $\beta_m^*$, to reach the global robustness. If $\beta_m < \beta_m^*$, the cascading failure happens. Otherwise, the network is robust.

When the node realizes that it is malfunctioning or is being attacked, it will transfer his load to his neighbor. It is very important to develop a reasonable redistribution strategy as quickly as possible to ensure network functionality. On the one hand, it is unrealistic to get the global information of the network considering the scale and complexity of the network. Instead, we only consider local information to ensure the timeliness of the strategy. On the other hand, network information is updating as the network evolving, it is unreasonable to set the redistribution the same strategy as the initial allocation. Considering the timeliness of redistribution strategies and the updating of information, the probability that the failed node *i* transfer the load to the neighbor *j* is set as

$$P_{i\to j} = \frac{k_j^{\alpha_2}}{\sum_{m\in V_i} k_m^{\alpha_2}}, \tag{5.3}$$

where $V_i$ is the set of node *i*'s neighbors, and $j \in V_i$. The transfer proportion depends on the degree, which is a local property of the network. The larger the degree is, more risks the node will be redistributed. $\alpha_2$ ($\alpha_2 \geq 0$) is a tunable parameter and governs the power of load redistribution. $\alpha_2$ is different from $\alpha_1$ due to the fact of information updating and lack of information of initial case. The load transferred from *i* to *j* at time *t* is set as

$$\Delta L_{i\to j} = L_i(t-1)P_{i\to j}. \tag{5.4}$$

The equations of (5.1)-(5.4) above form the new cascade failure model we proposed. We can describe cascading failures as a propagation behavior: When node *i* breaks down, its load (risk) $L_i(0)$ propagate along with the collaboration relations. Its neighbors take the risks. After taking the extra risk $\Delta L_{i\to j}$ from node *i*, node *j* fails if the updated risk goes beyond the capacity ($L_j(t) > C_j$). In turn, risks of node *j* will propagate, following the same rules. The cascading failure stops until the whole network collapses or there are no more failed nodes.

### 5.3 Theoretical Analysis

In this section, we proceed to derive the critical threshold of $\beta_m$ denoted by $\beta_m*$. If $\beta_m < \beta_m*$, the cascading failure happens. Otherwise, the network is robust. The smaller $\beta_m*$ is, the more resilient the network is, e.g., against cascading failures. To derive $\beta_m*$, we analyze the critical condition of cascading failure. To guarantee the global robustness, the model ought to satisfy

$$L_j(1) = L_j(0) + \Delta L_{i\to j}(1) = L_j(0) + L_i(0)P_{i\to j} < C_j(1). \tag{5.5}$$

By substituting (5.1)-(5.3) into (5.5), (5.5) can be rewritten as

$$\beta_1 k_j^{\alpha_1} + \beta_1 k_i^{\alpha_1} \frac{k_j^{\alpha_2}}{\sum_{m \in V_i} k_m^{\alpha_2}} < (1+\beta_m)\beta_1 k_j^{\alpha_1}, \tag{5.6}$$

then we have

$$k_i^{\alpha_1} \frac{k_j^{\alpha_2}}{\sum_{m \in V_i} k_m^{\alpha_2}} < \beta_m k_j^{\alpha_1}. \tag{5.7}$$

Therefore, $\beta_m$ satisfies

$$\beta_m > k_i^{\alpha_1} \frac{k_j^{\alpha_2-\alpha_1}}{\sum_{m \in V_i} k_m^{\alpha_2}}. \tag{5.8}$$

As shown in [116], $\sum_{m \in V_i} k_m^{\alpha_2} = \sum_{k'=k_{\min}}^{k_{\max}} k_i P(k'|k_i) k'^{\alpha_2}$, (5.8) can be rewritten as

$$\beta_m > k_i^{\alpha_1} \frac{k_j^{\alpha_2-\alpha_1}}{\sum_{k'=k_{\min}}^{k_{\max}} k_i P(k'|k_i) k'^{\alpha_2}} = \frac{k_i^{\alpha_1-1} k_j^{\alpha_2-\alpha_1}}{\sum_{k'=k_{\min}}^{k_{\max}} P(k'|k_i) k'^{\alpha_2}}. \tag{5.9}$$

Assuming that the degrees are irrelevance, $P(k'|k_i) = \frac{k'P(k')}{\langle k \rangle}$, we can have

$$\beta_m > \frac{k_i^{\alpha_1-1} k_j^{\alpha_2-\alpha_1}}{\sum_{k'=k_{\min}}^{k_{\max}} \frac{k'P(k')}{\langle k \rangle} k'^{\alpha_2}} = \frac{k_i^{\alpha_1-1} k_j^{\alpha_2-\alpha_1}}{\frac{\langle k^{\alpha_2+1} \rangle}{\langle k \rangle}} = \frac{k_i^{\alpha_1-1} k_j^{\alpha_2-\alpha_1} \langle k \rangle}{\langle k^{\alpha_2+1} \rangle}. \tag{5.10}$$

From (5.10) we can obtain the critical threshold $\beta_m^*$ by adjusting $\alpha_1$ and $\alpha_2$. Here we focus on the situation where the risk occurs and discuss the impact of redistribution parameter $\alpha_2$ on the cascading failure.

When $\alpha_2 < \alpha_1$, the critical threshold satisfies

$$\beta_m^* = \begin{cases} \frac{k_{\min}^{\alpha_2-1} < k >}{< k^{\alpha_2+1} >}, \alpha_1 < 1 \\ \frac{k_{\min}^{\alpha_2-1} < k >}{< k^{\alpha_2+1} >}, \alpha_1 = 1 \\ \frac{k_{\max}^{\alpha_1-1} k_{\min}^{\alpha_2-\alpha_1} < k >}{< k^{\alpha_2+1} >}, \alpha_1 > 1 \end{cases}. \tag{5.11}$$

When $\alpha_2 > \alpha_1$, the critical threshold satisfies

$$\beta_m^* = \begin{cases} \dfrac{k_{\min}^{\alpha_1-1} k_{\max}^{\alpha_2-\alpha_1} < k >}{< k^{\alpha_2+1} >}, \alpha_1 < 1 \\ \dfrac{k_{\max}^{\alpha_2-1} < k >}{< k^{\alpha_2+1} >}, \alpha_1 = 1 \\ \dfrac{k_{\max}^{\alpha_2-1} < k >}{< k^{\alpha_2+1} >}, \alpha_1 > 1 \end{cases} . \tag{5.12}$$

Specially, when $\alpha_2 = \alpha_1$, the critical threshold satisfies

$$\beta_m^* = \begin{cases} \dfrac{k_{\min}^{\alpha_1-1} < k >}{< k^{\alpha_2+1} >}, \alpha_1 < 1 \\ \dfrac{< k >}{< k^2 >}, \alpha_1 = 1 \\ \dfrac{k_{\max}^{\alpha_1-1} < k >}{< k^{\alpha_2+1} >}, \alpha_1 > 1 \end{cases} . \tag{5.13}$$

We can find in (5.11)-(5.13) that the critical threshold can be different in different cases. We proceed to discuss the relationship between $\beta_m^*$ and the load-redistribution parameter $\alpha_2$.

(1) $\alpha_1$=1

When $\alpha_2 < \alpha_1$,

$$\beta_m^* = \frac{k_{\min}^{\alpha_2-1} < k >}{< k^{\alpha_2+1} >} = \frac{k_{\min}^{\alpha_2-1} < k >}{\frac{1}{N}\sum_{i=1}^{N} k_i^{\alpha_2+1}} = \frac{< k >}{\frac{1}{N}\sum_{i=1}^{N} k_i^2 (\frac{k_i}{k_{\min}})^{\alpha_2-1}} \geq \frac{< k >}{\frac{1}{N}\sum_{i=1}^{N} k_i^2} = \frac{< k >}{< k^2 >} . \tag{5.14}$$

In the same way, we can deduce that when $\alpha_2 > \alpha_1$,

$$\beta_m^* = \frac{k_{\min}^{\alpha_1-1} k_{\max}^{\alpha_2-\alpha_1} < k >}{< k^{\alpha_2+1} >} \geq \frac{k_{\min}^{\alpha_2-1} < k >}{< k^{\alpha_2+1} >} = \frac{k_{\min}^{\alpha_2-1} < k >}{\frac{1}{N}\sum_{i=1}^{N} k_i^{\alpha_2+1}} = \frac{< k >}{\frac{1}{N}\sum_{i=1}^{N} k_i^2 (\frac{k_i}{k_{\min}})^{\alpha_2-1}} \geq \frac{< k >}{\frac{1}{N}\sum_{i=1}^{N} k_i^2} = \frac{< k >}{< k^2 >} . \tag{5.15}$$

In this case, we find that when $\alpha_2 = \alpha_1 = 1$, we can get the minimum value of $\beta_m^*$, and $(\beta_m^*)_{\min} = \frac{< k >}{< k^2 >}$. This concludes that when the initial load-distribution and load-redistribution strategy of the network are both linearly proportional to degree, the network shows better robustness under cascading failure.

(2) $\alpha_1 < 1$

When $\alpha_2 < \alpha_1$,

$$\beta_m^* = \frac{k_{\min}^{\alpha_2-1} < k >}{< k^{\alpha_2+1} >} = \frac{k_{\min}^{\alpha_2-1} < k >}{\frac{1}{N}\sum_{i=1}^{N} k_i^{\alpha_2+1}} = \frac{N < k >}{\sum_{i=1}^{N} k_i^2 (\frac{k_i}{k_{\min}})^{\alpha_2-1}} = \frac{N < k >}{\sum_{i=1}^{N} k_i^2 f_1(\alpha_2)} , \tag{5.16}$$

where $f_1(\alpha_2)=(\frac{k_i}{k_{\min}})^{\alpha_2-1}$. $f_1(\alpha_2)'=(\frac{k_i}{k_{\min}})^{\alpha_2-1}\ln(\frac{k_i}{k_{\min}})\geq 0$, $f_1(\alpha_2)$ is monotonically increasing, and $\beta_m^*$ is monotonically decreasing since $f_1(\alpha_2)$ is on the denominator.

In the same way, when $\alpha_2>\alpha_1$,

$$\beta_m^*=\frac{k_{\min}^{\alpha_1-1}k_{\max}^{\alpha_2-\alpha_1}<k>}{<k^{\alpha_2+1}>}=\frac{N<k>}{\sum_{i=1}^{N}k_i^2(\frac{k_i}{k_{\min}})^{\alpha_1-1}(\frac{k_i}{k_{\max}})^{\alpha_2-\alpha_1}}=\frac{N<k>}{\sum_{i=1}^{N}k_i^2(\frac{k_i}{k_{\min}})^{\alpha_1-1}f_2(\alpha_2)}, \tag{5.17}$$

where $f_2(\alpha_2)=(\frac{k_i}{k_{\max}})^{\alpha_2-\alpha_1}$. $f_2(\alpha_2)'=(\frac{k_i}{k_{\max}})^{\alpha_2-\alpha_1}\ln(\frac{k_i}{k_{\max}})\leq 0$, $f_2(\alpha_2)$ is monotonically decreasing, and $\beta_m^*$ is monotonically increasing since $f_2(\alpha_2)$ is on the denominator.

We find that when $\alpha_2=\alpha_1$, we can get the minimum value of $\beta_m^*$. The same conclusion as the case $\alpha_1$=1 is when the initial load-distribution and load-redistribution strategy of the network are the same, the network shows better robustness under cascading failure.

(3) $\alpha_1>1$

When $\alpha_2<\alpha_1$,

$$\beta_m^*=\frac{k_{\max}^{\alpha_1-1}k_{\min}^{\alpha_2-\alpha_1}<k>}{<k^{\alpha_2+1}>}=\frac{N<k>}{\sum_{i=1}^{N}k_i^2(\frac{k_i}{k_{\max}})^{\alpha_1-1}(\frac{k_i}{k_{\min}})^{\alpha_2-\alpha_1}}=\frac{N<k>}{\sum_{i=1}^{N}k_i^2(\frac{k_i}{k_{\max}})^{\alpha_1-1}f_3(\alpha_2)}, \tag{5.18}$$

where $f_3(\alpha_2)=(\frac{k_i}{k_{\min}})^{\alpha_2-\alpha_1}$. $f_3(\alpha_2)'=(\frac{k_i}{k_{\min}})^{\alpha_2-\alpha_1}\ln(\frac{k_i}{k_{\min}})\geq 0$, $f_3(\alpha_2)$ is monotonically increasing, and $\beta_m^*$ is monotonically decreasing since $f_3(\alpha_2)$ is on the denominator.

when $\alpha_2>\alpha_1$,

$$\beta_m^*=\frac{k_{\max}^{\alpha_2-1}<k>}{<k^{\alpha_2+1}>}==\frac{k_{\max}^{\alpha_2-1}<k>}{\frac{1}{N}\sum_{i=1}^{N}k_i^{\alpha_2+1}}=\frac{N<k>}{\sum_{i=1}^{N}k_i^2(\frac{k_i}{k_{\max}})^{\alpha_2-1}}=\frac{N<k>}{\sum_{i=1}^{N}k_i^2f_4(\alpha_2)}, \tag{5.19}$$

where $f_4(\alpha_2)=(\frac{k_i}{k_{\max}})^{\alpha_2-1}$. $f_4(\alpha_2)'=(\frac{k_i}{k_{\max}})^{\alpha_2-1}\ln(\frac{k_i}{k_{\max}})\leq 0$, $f_4(\alpha_2)$ is monotonically decreasing, and $\beta_m^*$ is monotonically increasing since $f_4(\alpha_2)$ is on the denominator.

We find that when $\alpha_2=\alpha_1$, we can get the minimum value of $\beta_m^*$.

From our analysis we can conclude that when $\alpha_2=\alpha_1$, i.e., the initial load-distribution and load-redistribution strategy of the network are the same, we can get the minimum value of $\beta_m^*$. In the following section, we apply numerical simulations in artificial networks and real-world networks to visually observe the results of our discussion and analysis.

## 5.4 Simulation Results

In this section, numerical simulation results are provided to validate our proposed model and theoretical analysis. As discussed in Section 5.3, when $\alpha_2 = \alpha_1$, we can get the minimum value of $\beta_m^*$, i.e., the network is the most robustness. We first validate our model on a 1000-node scale-free network, the average degree is 6. Fig 5.1 plots the robustness threshold $\beta_m^*$, with the grows of $\alpha_2$. The simulation results in Fig 5.1 show intuitively that the value of $\beta_m^*$ is the smallest when $\alpha_2 = \alpha_1$ in each case.

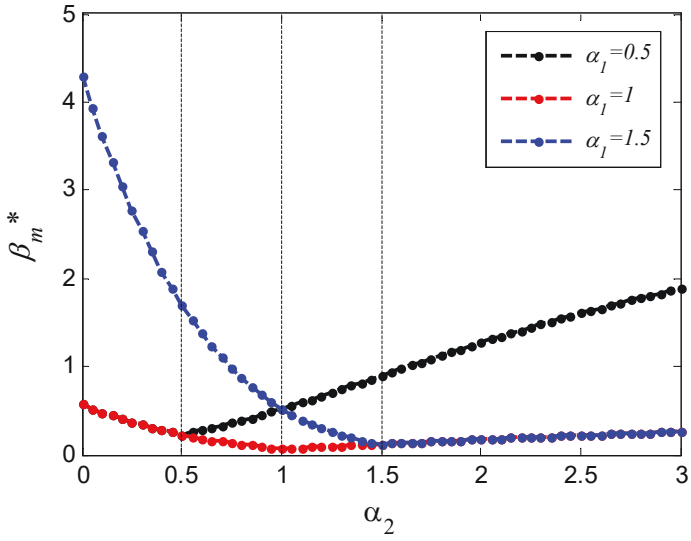



Fig 5.1. The relation between critical threshold $\beta_m^*$ of artificial network model and $\alpha_2$.

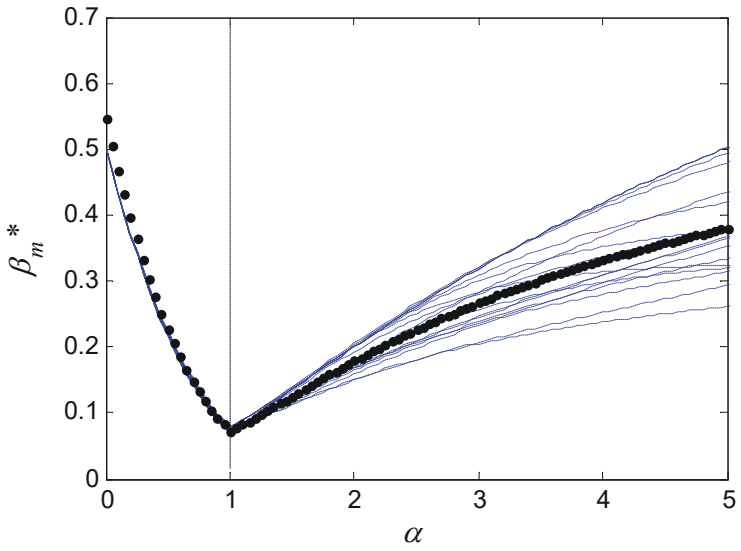
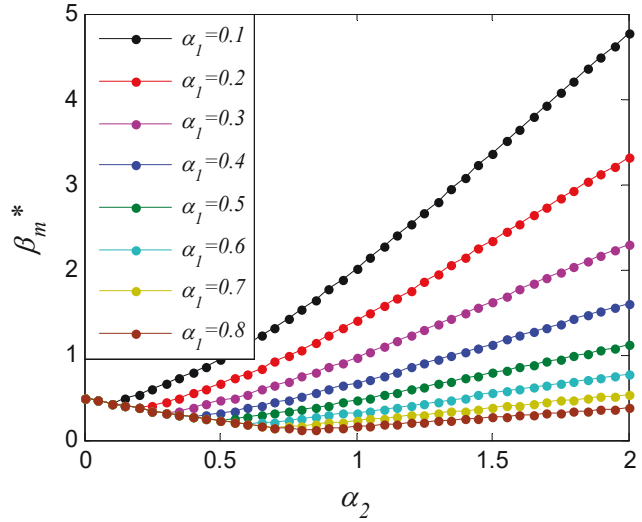



Fig 5.2. The relation between critical threshold $\beta_m^*$ and $\alpha$ under the special case: $\alpha_2 = \alpha_1 = \alpha$.

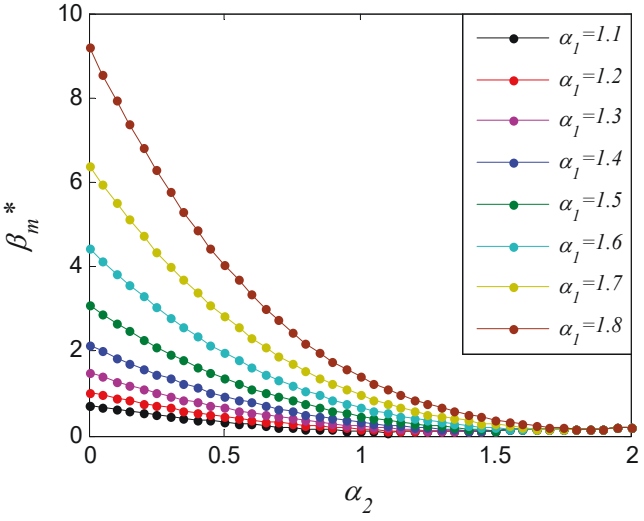
It is also shown that when $\alpha_2 = \alpha_1 = 1$, the value of $\beta_m^*$ is the smallest in all 3 cases, which is

further validated through Fig 5.2. Simulations are carried out in 200 different networks. Fig 5.2 shows the robustness threshold $\beta_m^*$, with the grows of $\alpha$ . The solid point (black) in Fig 5.2 is the average result of 200 different networks, and dashed lines (blue) represent the results in each of the 10 networks we randomly chosen.

An interesting finding is that when $\alpha_1$=1 , the value of $\beta_m^*$ is smaller than the other 2 cases. That means when the initial load distribution is linearly proportional to the degree, the network shows the greatest robustness against the node's failure. We further notice from Fig 5.3 that, when $\alpha_1 < 1$, the value of $\beta_m^*$ becomes smaller as $\alpha_2$ decreases, and when $\alpha_1 > 1$ , the value of $\beta_m^*$ becomes smaller as $\alpha_2$ grows. Therefore, without the knowledge of initial load information, it is effective to improve the network robustness with the load-redistribution linearly proportional to the degree.



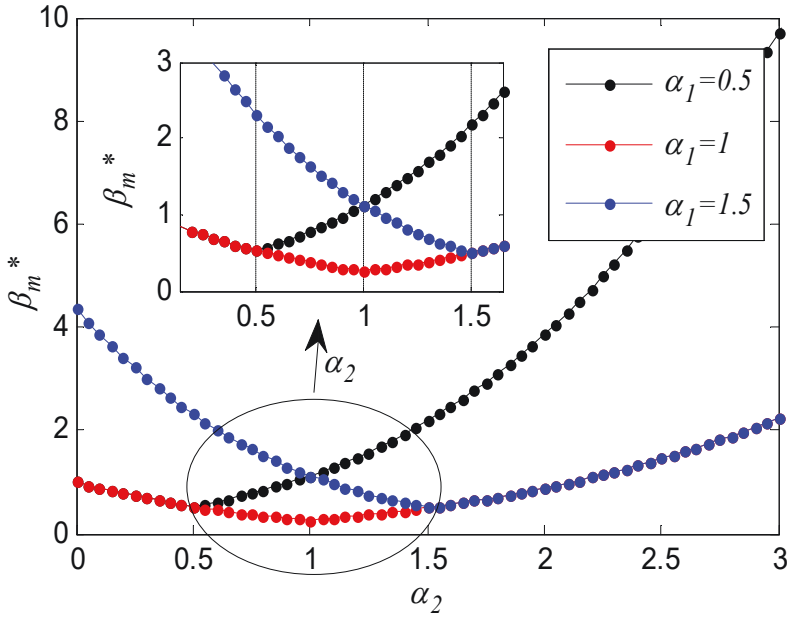

(a) $\alpha_1 < 1$



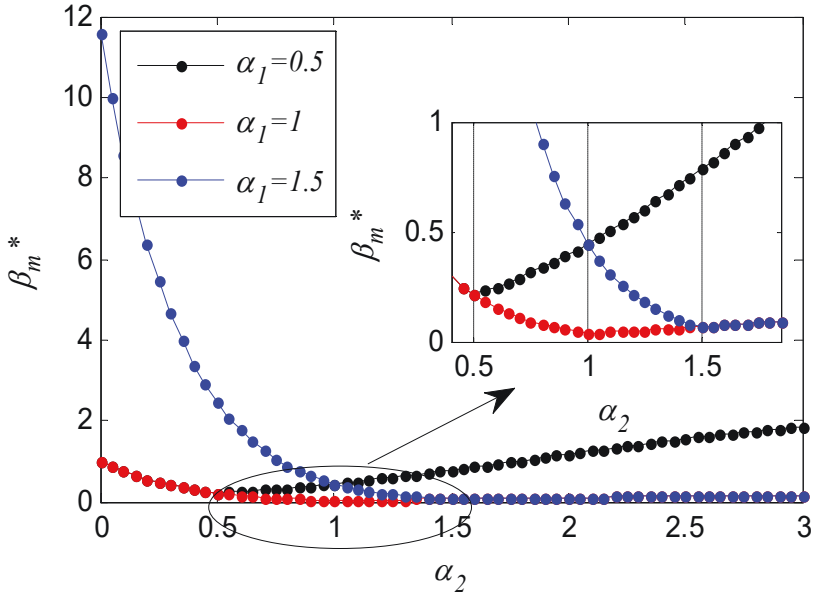

(b) $\alpha_1 > 1$

Fig 5.3. The relation between critical threshold $\beta_m^*$ of artificial network model and $\alpha_2$ under two different cases: (a) $\alpha_1 < 1$ , (b) $\alpha_1 > 1$ .

In practice, cascading failures show in many real systems as we discussed. We proceed to carry out numerical simulations on Neural network and power grid [26], as shown in Fig 5.4 and Fig 5.5. We can see that the value of critical threshold $\beta_m^*$ is the smallest when $\alpha_2 = \alpha_1$ in both networks. And when $\alpha_1$=1 , the value of $\beta_m^*$ is smaller than the other 2 cases.



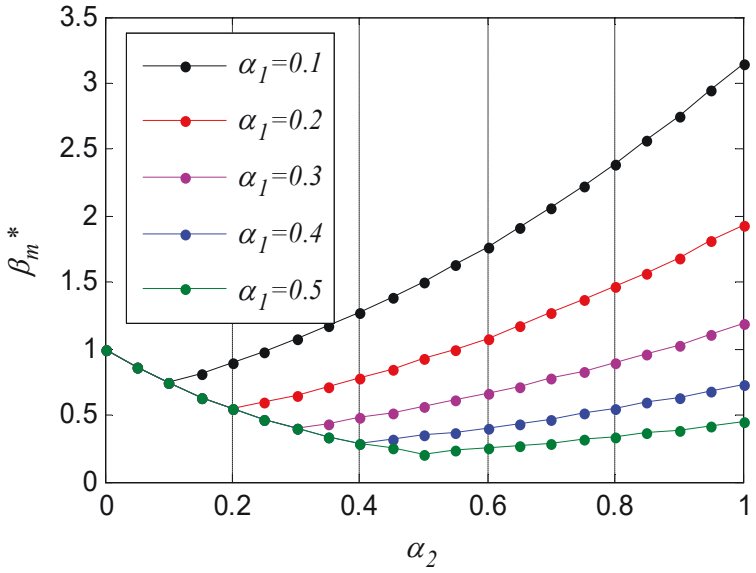

(a) Neural network



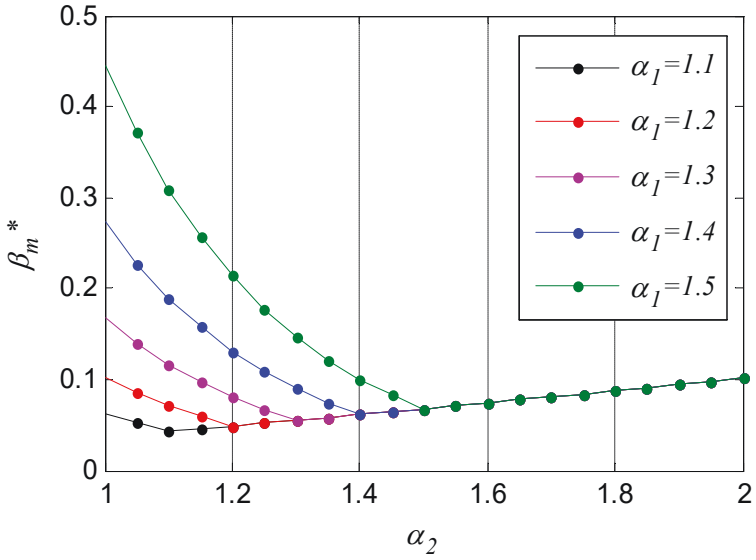

(b) Power grid

Fig 5.4. The relation between critical threshold $\beta_m^*$ of real networks and $\alpha_2$ , (a) Neural network, (b) Power grid.

In general, our simulation results show that bond exists between the initial load and load-redistribution. The network shows better robustness against cascading failure when the initial load-distribution and load-redistribution strategy of the network are the same. Specially, when both of the initial load-distribution and load-redistribution are linearly proportional to the node's degree, the network shows the greatest robustness against the cascading failure. In addition, regardless of the initial load, the network also shows good robustness against cascading failure when load-redistribution is linearly proportional to the node's degree. Our simulation results show a very good guiding role to study the network robustness optimization strategy in the future.




(a) $\alpha_1 < 1$




(b) $\alpha_1 > 1$

Fig 5.5. The relation between critical threshold $\beta_m^*$ of real networks and $\alpha_2$ under two different cases: (a) $\alpha_1 < 1$, (b) $\alpha_1 > 1$.

## 5.5 Conclusion

In this chapter, a new cascading failure model to analyze the robustness of networks under the node's failure is developed, and redistribution strategy based on local real-time information is analyzed. Theoretical analysis of the robustness and the redistribution strategy are carried out, and the relationship between network robustness and redistribution strategy parameters are discussed in different cases. Our analysis shows that when the initial load-distribution and load-redistribution strategy of the network are the same and linearly proportional to degrees, the network shows better robustness under cascading failure. Furthermore, without the knowledge of initial load information, it is effective to improve the network robustness with the load-redistribution linearly proportional to the degree. The numerical simulations implemented in artificial networks and real-life networks confirm the validity of the robustness threshold, as well as the analysis of the redistribution strategy.

# Chapter 6

# Study on the Quantification of Social Network Robustness under the Virus Attacks

## 6.1 Introduction

As one of the typical scenarios of attacks, epidemic spreading in social network has exterted an influence on our daily activities. At present, the COVID-19 has spread all over the world within six months [117]-[120]. Based on the Situation Report-133 of World Health Organization (WHO) [121], the data as received by WHO from national authorities by 10:00 CEST, 01 June 2020 has shown that there are 6057853 cases confirmed as infected globally and 371166 people dead. The robustness of complex networks against virus attacks has become one of the most concerned topics in complex network study [122]-[127].

The study on epidemics has a very long history and classic epidemic models are built to describe the virus spreading, such as the SI model, SIS model, and the SIR model [124]. Besides the epidemics, plenty of other cyberattacks can also be found in almost any kind of networks, for example, the DDoS attacks on the Internet [128], [129], the cascading failures in the power grids [130] and the epidemic spread in social networks [124]. As a malicious and deliberate attempt by an individual (organization) to breach the information system of another individual (organization), the attacks spread from one node to another in the network. Once one or some components are attacked, it may cause incalculable losses to the entire network and other related networks. Recently, the epidemic models have been adapted to study the spread of the above cyberattacks.

Robustness of the network refers to the ability of the network to maintain a certain degree of structural integrity and function after being subjected to a fault or attack [131]. Epidemic spreading models have been used to study the network robustness with respect to the virus attacks. Epidemic threshold is the most commonly used measure for the network robustness in regard to virus attacks, i.e., the larger the epidemic threshold, the more robust a network is against the spread [127], [132]. Recently, Mina Youssef [125]-[127] put forward a new measure to make an evaluation on the robustness of complex networks with respect to the spread of SIS epidemics. The results demonstrated that the proposed measure of network robustness with respect to the virus attacks is of great effectiveness for the epidemics with different final infection scales.

Since the network robustness is measured in the epidemic threshold and the final infection scale, the spread velocity [133]-[135], as another important indicator describing the epidemics, should also be taken into account to measure network robustness. On the one hand, many epidemics will eventually achieve network-wide infection or immunity, such as the epidemic process described by SI or SIR model. In this case, the final infection scale of the epidemics remains the same in different networks. On the other hand, network structures have a great effect on the spread velocity, i.e., the virus spread velocity differs in different networks. In addition, differences can be observed in the trends of spread velocity, epidemic threshold and the final infection scale in the network. Therefore, the spread velocity is one of

the key factors that cannot be ignored to measure the network robustness with respect to the epidemics.

In this chapter, a novel metric, combined with the spread velocity, the infection scale at the steady state and the epidemic threshold, is proposed to measure the robustness in regard to virus attacks in social networks. First, we show some examples of networks where the epidemic threshold and/or infection scale at the steady state fail to assess their robustness. Then the new robustness metric is introduced, based on which the network robustness concerning virus spreading are analyzed. In addition, it is demonstrated by the simulation results that as the average degree grows in both homogeneous and heterogeneous networks, the network becomes more vulnerable to the attacks. Moreover, in homogeneous networks, the network robustness improves due to the increasing numbers of random-connected links.

The rest of this chapter is organized as follows. In Section 5.2, we analyze the necessity of putting forward the new metric, and then introduce the novel metric to quantify the network robustness based on epidemic spread in Section 5.3. We present the simulation results in Section 5.4, and the main conclusions and future work are summarized in Section 5.5.

## 6.2 The Network Robustness with respect to Epidemic Spread

Epidemic threshold is the most commonly used measure for the network robustness with respect to the spread of epidemic, i.e., the larger the epidemic threshold, the more robust a network is against the epidemics. The existed literatures shown that large BA networks [27] consequently are more vulnerable to epidemic spreading than WS networks [26] based on the epidemic threshold. Then the researchers found that the epidemic threshold may fails to assess the network robustness, a new metric to quantify the network robustness considering both epidemic threshold and fraction of infection at steady state was proposed in SIS epidemic model [126].

In fact, it is not comprehensive to use epidemic threshold and/or infection scale to measure network robustness with respect to the virus attacks. For example, Fig 6.1 shows the SIS epidemic spreading process in 3 different networks. Supposing that in the SIS epidemic model, the rate of a susceptible node being infected by a single infected neighbor is $\beta$ , and the infected node recovered with the rate $\delta$. We observe that the final density of the infection nodes in BA network is smaller than the final density of the infection nodes in WS network and Regular network, i.e., $I_{Regular} > I_{WS} > I_{BA}$ . From the perspective of the infection scale of the steady state, BA network is more robustness than WS network and Regular network. However, comparison of the spread velocities shows that the spread velocity in BA network is the fastest, and that of regular network is much slower than in other two networks, i.e., $V(t)_{BA} > V(t)_{WS} > V(t)_{Regular}$ . That is, one single indicator, for example, the fraction of infection at steady state or the spread velocity, cannot accurately measure the network robustness with respect to virus attacks.

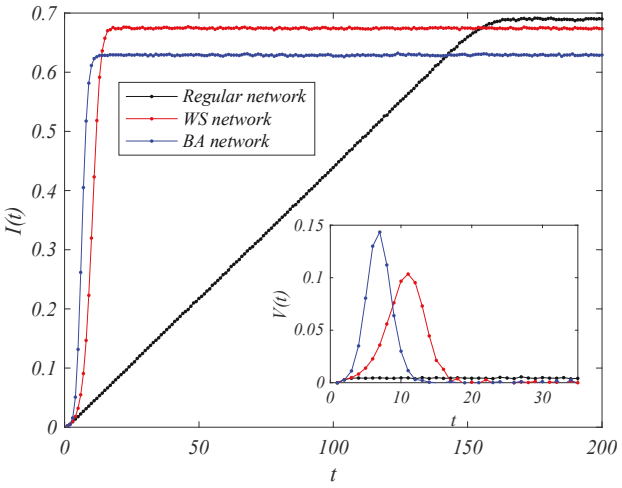



Fig 6.1. The SIS epidemic spreading process ( $\beta=\gamma=0.3$ ).

And when the final infection densities at the steady state are the same, such as in the SI model shown in Fig 6.2, we can see that the same results of the spread velocity of different networks in SIS epidemic model, i.e., $V(t)_{BA} > V(t)_{WS} > V(t)_{Regular}$ . As we all know, the epidemic threshold in the homogeneous network is larger than in the heterogeneous network. Therefore, even in the models with the same infection scale at the steady state, it is still inappropriate to apply one single indicator to measure the network robustness.

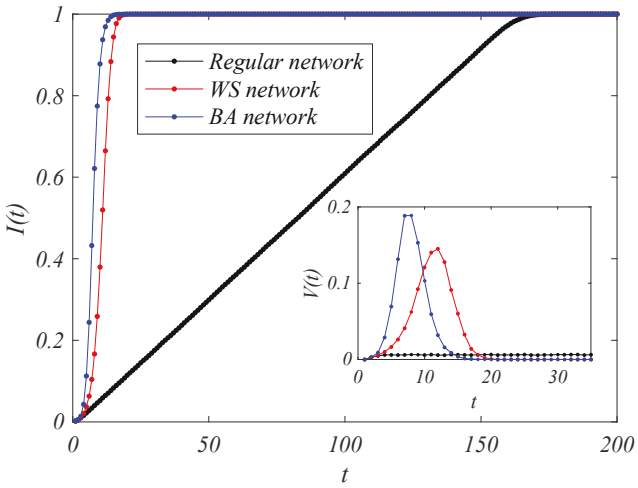



Fig 6.2. The SI epidemic spreading process ( $\beta=0.3$ ).

Moreover, due to the difference of the spread velocity, the time to reach a stable state of epidemic spread is different. We did simulations on WS network and BA network. Fig 6.3 counts the time points when the epidemics reach steady state ($T(i\_max)$) under different infect rates $\beta$ in BA and WS networks. We can see more intuitively from Fig 5.3 that the time points of reaching the steady states are different in the 2 networks, and as the infection rate decreases, this gap becomes more and more obvious. Therefore, under the condition of low infection probability, the difference of spread velocity in the networks is very large.

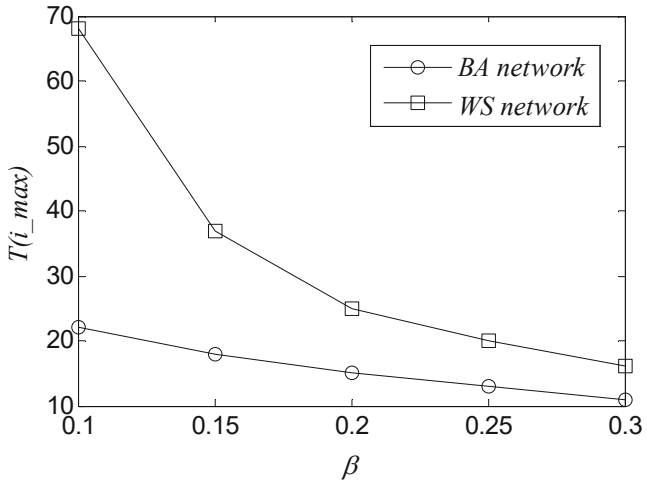



Fig 6.3. The time of reaching the steady states of different networks.

It can be seen from the above analysis that epidemic threshold, fraction of infection at steady state and spreading velocity are all important factors to measure the network robustness under virus attacks. Comparing the propagation processes in the three networks, we can conclude that the spread velocity in the BA network is the fastest, but the final infection scale is the smallest. In the WS network and the regular network, although the spread velocity is slower, the final infection scale is larger. Therefore, it is necessary to propose a network robustness measurement considering of multiple indicators with respect to virus attacks.

### 6.3 The Novel Metric to Quantify the Network Robustness under the Virus Attacks

Combining the epidemic threshold, propagation scale and spreading velocity together, we propose a multiple-indicator-based measurement to quantify the network robustness with respect to virus attacks. We first set the parameters in SI and SIS epidemic models. For the sake of simplicity, we set $\delta = 1$ in the rest of this chapter, and then the effective infection rate can be defined as $\tau = \beta / \delta = \beta$ in SI and SIS model. The density of infected nodes at time $t$ is described as $i(t)$, the steady state of the infection under the effective infection rate $\tau$ can be written as

$$i_\infty(\tau) = \lim_{t\to\infty} i_t(\tau).$$

We first define the cumulative infection $C_{i_t}(\tau)$ as the sum of infection density at each time slot under the effective infection rate $\tau$,

$$C_{i_t}(\tau) = \sum_{t'=0}^{t} i_{t'}(\tau).$$

Taking into account of all values of $\tau$, supposing that $t_s$ is the first time the network reaches the steady state, the new robustness measure with respect to virus attack, $R_{VA}$, can be written as

$$R_{VA}=\frac{1}{t}\int_{\tau_s}^{\infty}C_{i_t}(\tau)d\tau=\frac{1}{t}\int_{\tau_s}^{\infty}\sum_{t'=0}^{t}i_{t'}(\tau)d\tau,\quad t\le t_s \tag{6.1}$$

where $\tau_s$ is the epidemic threshold.

Introducing the effective cure rate $s=1/\tau$ in [126], (6.1) can be rewritten as

$$R_{VA}=\frac{1}{t}\int_{0}^{\rho}C_{i_t}(s)ds=\frac{1}{t}\int_{0}^{\rho}\sum_{t'=0}^{t}i_{t'}(s)ds,\quad t\le t_s. \tag{6.2}$$

When $t = t_s$, the robustness of network $G$ can be written as

$$R_{VA}{}^{G}=\frac{1}{t_s}\int_{0}^{\rho}C_{i_{t_s}}(s)ds=\frac{1}{t_s}\int_{0}^{\rho}\sum_{t'=0}^{t_s}i_{t'}(s)ds. \tag{6.3}$$

In (6.3), the length of $t_s$ represents the spreading velocity, and $i_t(s)$ represents the density of infected nodes at time $t$. Then the proposed $R_{VA}{}^{G}$ considering of epidemic threshold, the infection scale and the spread velocity. Since $t_s$ changes under different effective cure rate, it is hard to make statistics of $t_s$. In order to avoid the impact of the density of stable infection within the time period $[t_s, t_\infty]$ on the network robustness, we use the $i_{max}(s)-i_t(s)$ instead of $i_t(s)$ in (6.3). Therefore, the network robustness at $t$-time can be written as

$$R_{VA}(t)=\frac{1}{t}\int_{0}^{\rho}\sum_{t'=0}^{t}(i_{max}(s)-i_{t'}(s))ds, \tag{6.4}$$

the robustness of network $G$ can be written as

$$R_{VA}{}^{G}=\frac{1}{t_s}\int_{0}^{\rho}\sum_{t'=0}^{t_s}(i_{max}(s)-i_{t'}(s))ds. \tag{6.5}$$

We can see from (6.4) and (6.5) that, when $t<t_s$, the network robustness at $t$-time, $R_{VA}(t)$, depends on the epidemic threshold and the infection scale at time $t'$, $t'\in[0,t]$. While the robustness of network $G$ depends on the length of $t_s$, besides the epidemic threshold and the infection scale at time $t'$, $t'\in[0,t_s]$. Therefore, the shorter time it takes to reach steady state and/or the larger the infection scale at time $t'$, the larger the $R_{VA}{}^{G}$ is, and the more vulnerable of the network under the virus attack, accordingly.

We choose two epidemic models to analyze the network robustness, one is the epidemic process that the whole network is infected finally, i.e., the SI epidemic model, the other one is that the infection density is stable at a non-1 value, i.e., the SIS epidemic model.

**Case 1. The robustness of homogeneous network with respect of SI epidemic spreading**

The state of each node in the SI model is infected or healthy, and the change of infected individuals over time can be described as

$$\frac{di}{dt} = \beta \langle k \rangle i(1-i) . \tag{6.6}$$

By separating variables, (6.6) can be written as

$$\frac{di}{i(1-i)} = \beta \langle k \rangle dt , \tag{6.7}$$

integrating both sides of (6.7), we can obtain

$$\ln \frac{1-i(t)}{i(t)} = -\beta \langle k \rangle t + c .$$

The density of the infected nodes at time $t$ can be written as

$$i(t) = \frac{1}{1+(1/i_0 - 1)e^{-\beta \langle k \rangle t}} . \tag{6.8}$$

The final density of the infection of SI model equals to 1, i.e., $i_\infty = 1$. Based on (6.8), the robustness of homogeneous network $G$ with respect of SI epidemic spreading can be written as

$$\begin{aligned} R_{VA}^{SI} &= \frac{1}{t_s}\int_{\tau_s}^{\infty}\sum_{t'=0}^{t_s}(i_\infty(\tau) - i_t(\tau))d\tau \\ &= \frac{1}{t_s}\int_{\tau_s}^{\infty}\sum_{t'=0}^{t_s}(1 - \frac{1}{1+(1/i_0 - 1)e^{-\beta \langle k \rangle t'}})d\tau \end{aligned} . \tag{6.9}$$

**Case 2. The robustness of homogeneous network with respect of SIS epidemic spreading**

Ignoring the degree correlations of nodes in homogeneous network, the density of infected nodes at time $t$, i.e., $i(t)$, satisfies (6.10)

$$\frac{di}{dt} = -i + \beta \langle k \rangle i(1-i) . \tag{6.10}$$

Integrating both sides of (6.10),

$$\int_0^t dt = \int_{i_0}^{i(t)} \frac{1}{-i + \beta \langle k \rangle i(1-i)} di , \tag{6.11}$$

then (6.11) can be rewritten as

$$t = \frac{1}{\beta \langle k \rangle - 1}\int_{i_0}^{i(t)} \frac{1}{i} di + \frac{\beta \langle k \rangle}{\beta \langle k \rangle - 1}\int_{i_0}^{i(t)} \frac{1}{\beta \langle k \rangle - \beta \langle k \rangle i - 1} di , \tag{6.12}$$

we can obtain that

$$e^{(\beta \langle k \rangle - 1)t} = \frac{i(t)}{\beta \langle k \rangle - \beta \langle k \rangle i(t) - 1} \Big/ \frac{i_0}{\beta \langle k \rangle - \beta \langle k \rangle i_0 - 1} ,$$

$$\frac{i(t)}{\beta\langle k\rangle-\beta\langle k\rangle i(t)-1}=\frac{i_0 e^{(\beta\langle k\rangle-1)t}}{\beta\langle k\rangle-\beta\langle k\rangle i_0-1},$$

$$i(t)(\beta\langle k\rangle-\beta\langle k\rangle i_0-1)=i_0 e^{(\beta\langle k\rangle-1)t}(\beta\langle k\rangle-\beta\langle k\rangle i(t)-1).$$

The density of the infected nodes at time $t$ can be written as

$$i(t)=\frac{(\beta\langle k\rangle-1)i_0 e^{(\beta\langle k\rangle-1)t}}{\beta\langle k\rangle-\beta\langle k\rangle i_0-1+i_0\beta\langle k\rangle e^{(\beta\langle k\rangle-1)t}}. \tag{6.13}$$

Let (6.10) equals to 0, we can get

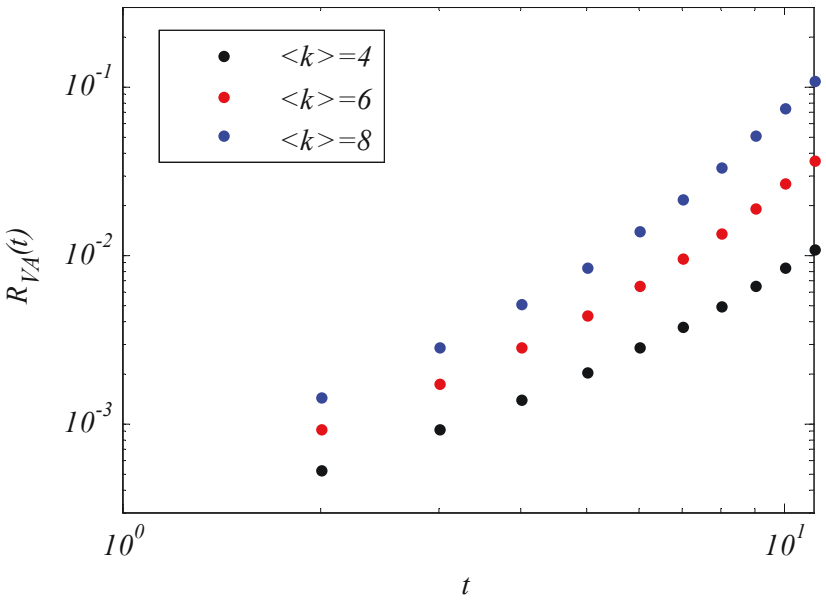
$$\frac{di}{dt}=-i+\beta\langle k\rangle i(1-i)=0,$$

when $\tau=\frac{\beta}{\delta}=\beta>\tau_c$, the infection density of the final stable state is

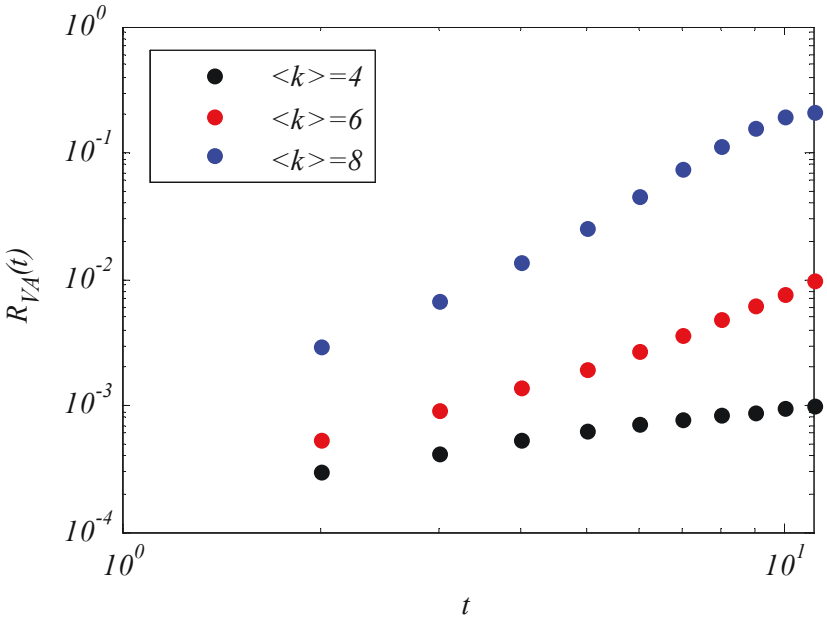
$$i_\infty=1-\frac{1}{\beta\langle k\rangle}. \tag{6.14}$$

Based on (6.13) and (6.14), the robustness of homogeneous network $G$ with respect of SIS epidemic spreading can be written as

$$\begin{aligned}R_{VA}^{SIS}&=\frac{1}{t_s}\int_{\tau_s}^{\infty}\sum_{t'=0}^{t_s}(i_\infty(\tau)-i_t(\tau))d\tau\\&=\frac{1}{t_s}\int_{\tau_s}^{\infty}\sum_{t'=0}^{t_s}(1-\frac{1}{\beta\langle k\rangle}-\frac{(\tau\langle k\rangle-1)i_0 e^{(\tau\langle k\rangle-1)t'}}{\tau\langle k\rangle-\tau\langle k\rangle i_0-1+i_0\tau\langle k\rangle e^{(\tau\langle k\rangle-1)t'}})d\tau\end{aligned}. \tag{6.15}$$




(a) SI model




(b) SIS model

Fig 6.4. The robustness of homogeneous networks at $t$-time with respect to SI epidemic spreading (Fig 6.4(a)) and SIS epidemic spreading (Fig 6.4(b)) ( $\beta = 0.2,\ \delta=1$ ).

We present the numerinal solutions of (6.9) and (6.13) in Fig 6.4 and Table 6.1. We can see that as the average degree of the network grows, the network becomes more vulnerable to the virus attack. Fig 6.4 shows the robustness at $t$-time ($t < t_s$) in homogeneous network with respect to SI and SIS epidemic spreading.

Table 6.1. The robustness of homogeneous network $G$ with respect to SI and SIS epidemic spreading.

| *Network* / *Robustness* | $<k>=4$ | $<k>=6$ | $<k>=8$ |
|---|---|---|---|
| $R_{VA}^{SI}$ | 0.1454 | 0.2798 | 0.3502 |
| $R_{VA}^{SIS}$ | 0.1261 | 0.2106 | 0.2777 |

## 6.4 Simulations

In this section, Monte-Carlo simulations are used to further explore the robustness of different networks with respect to the virus attacks. Simulations are carried out in different networks with $N$=500 nodes. All the simulations are averaged over 500 runs.

First, the simulations are carried out in WS small world networks with the rewiring rate $p$. Based on the construction algorithm of the WS model, at the beginning, the network is a regular graph, and then randomly reconnects each edge in the network with probability $p$, that is, one endpoint of the edge remains unchanged, and the other endpoint is taken as the network. In the above model, $p$=0 corresponds to a completely regular network, $p$=1 corresponds to a completely random network, and the transition from a completely regular network to a completely random network can be controlled by adjusting the $p$

value. We capture a set of networks where the rewiring rate $p$ changes from 0 to 1 and analyze the robustness of these networks. The examples of networks are of the same average degree, i.e., $<k> = 6$, and almost have the same epidemic thresholds, in which the epidemic threshold hardly works on measuring the network robustness.

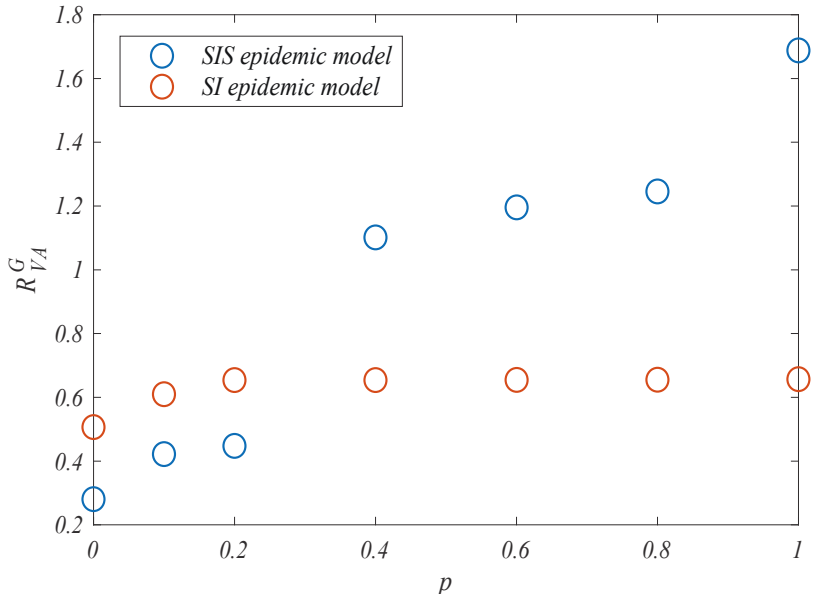
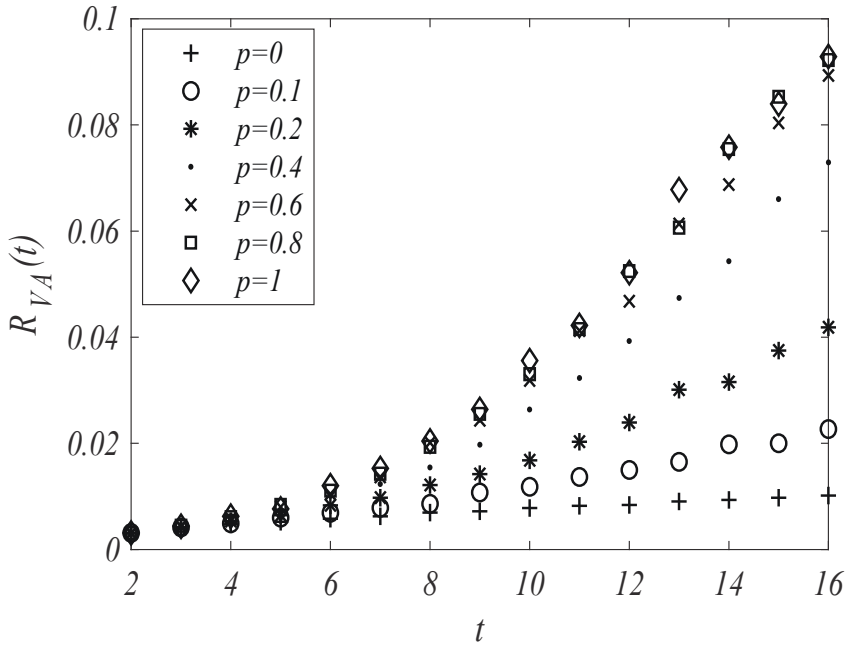


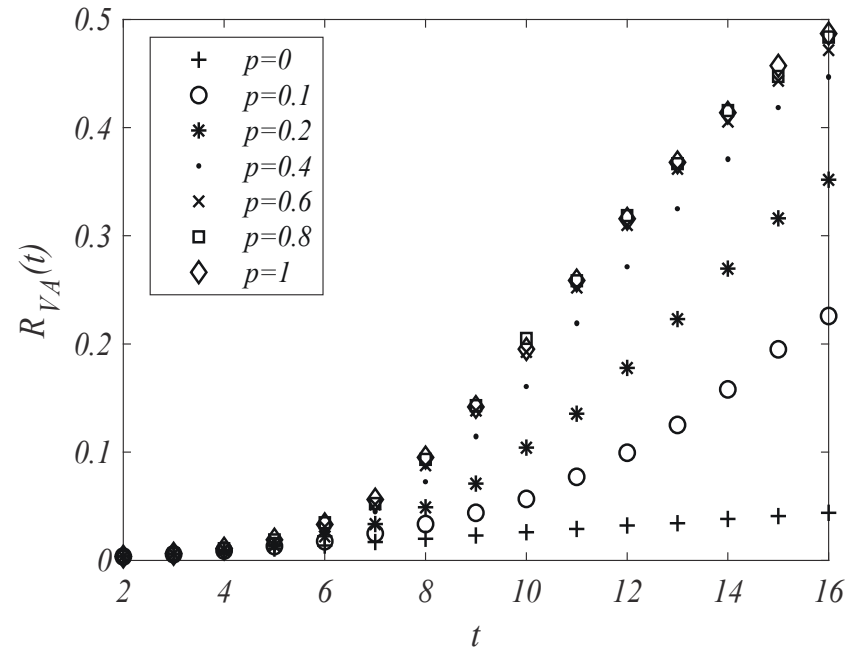

Fig 6.5. The robustness of WS networks with respect of SIS/SI epidemic model.

Fig 6.5 shows the network robustness $R_{VA}{}^{G}$ of different networks. In SIS epidemic model (red circle), we can see that as $p$ grows, $R_{VA}$ becomes larger, that is, the network becomes more vulnerable. When $p = 0$, the network has a small value of $R_{VA}$, that is because in regular network ($p = 0$), the virus spreads very slowly, as shown in Fig 6.2, that is to say, the spread velocity plays a greater role on measuring the network robustness than the steady infection. In SI model (blue circle), both the epidemic threshold and the steady infection are the same, then our robustness measure of network is only related to the spread velocity. We can see that from Fig 6.5 that the robustness of regular network is smaller than other networks due to the slow spread velocity. However, the WS networks and random network are almost have the same robustness as the difference of spread velocity is small in these networks.

We further count the network robustness at time $t$ ($t < t_s$), the result in Fig 6.6 shows a better robustness at time $t$ in regular network. As $p$ increases, the robustness of network at time $t$ increases, i.e., the network becomes more fragile.




(a) SIS epidemic model



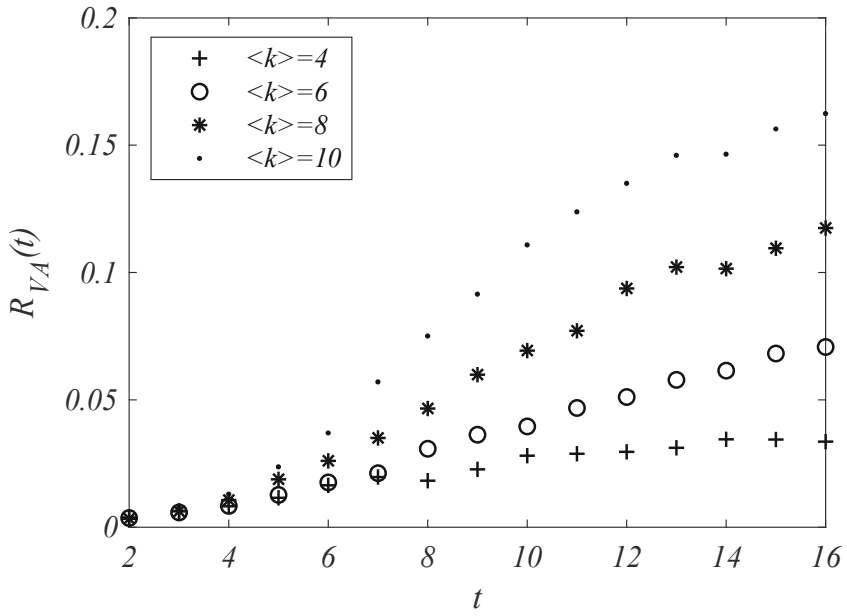

(b) SI epidemic model

Fig 6.6. The robustness of WS networks at *t*-time with respect to SI epidemic spreading (Fig 6.6(a)) and SIS epidemic spreading (Fig 6.6(b)) $\beta = 0.25$.

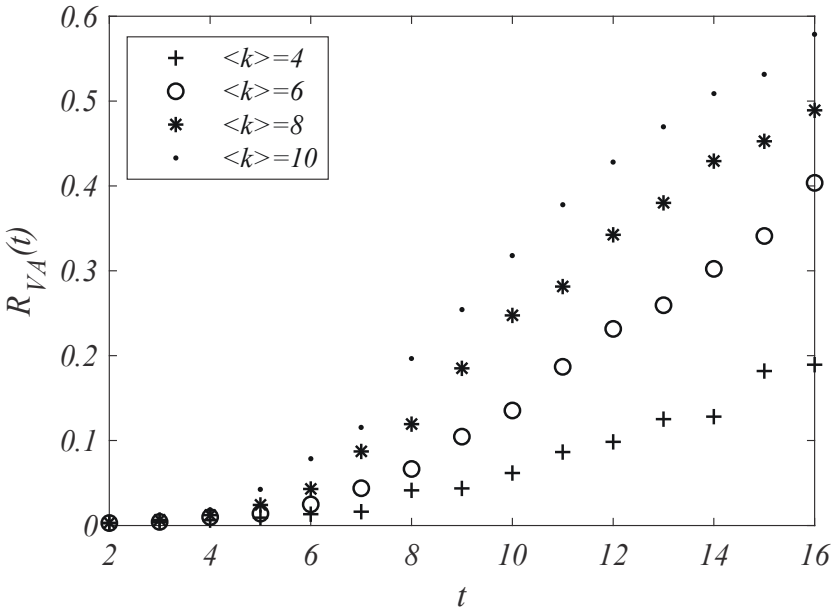
The simulations are also carried out in BA networks with different average degree <*k*>, we can see from Table 6.2 that BA network becomes more vulnerable to the virus attacks as the average degree of the network grows. Fig 6.7 shows the robustness at *t*-time ($t < t_s$) in BA networks with respect to SI and SIS epidemic spreading.

Table 6.2. The robustness of BA network with respect to SI and SIS epidemic spreading.

| Network / Robustness | <k>=4 | <k>=6 | <k>=8 | <k>=10 |
|---|---|---|---|---|
| $R_{VA}^{SI}$ | 0.5589 | 0.5594 | 0.6035 | 0.7053 |
| $R_{VA}^{SIS}$ | 0.0101 | 0.0115 | 0.1223 | 0.1643 |




(a) SIS epidemic model




(b) SI epidemic model

Fig 6.7. The robustness of BA networks at *t*-time with respect to SI epidemic spreading (Fig 6.7(a)) and SIS epidemic spreading (Fig 6.7(b)) $\beta = 0.15$.

Our simulation results show that in both homogeneous and heterogeneous networks, the network becomes more vulnerable as the average degree grows. In homogeneous networks, the robustness of

networks with respect to virus spreading decreases as $p$ increases, i.e., the regualr network shows better robustness than other homogeneous networks.

## 6.5 Conclusion

Considering the spread velocity, the epidemic threshold and steady infection, a new robustness measure with respect to virus attacks in social networks is proposed. The robustness of homogeneous network with respect of SI and SIS epidemic spreading are analyzed. Simulation results show that the network becomes more vulnerable to the virus attacks as the average degree of the network grows. In homogeneous networks, the network becomes more robustness improves due to the increasing numbers of random-connected links.

The research in the above three chapters is carried out in static network. As we discussed in Chapter 1-Chapter 3, in most of the real-world network, the structure is changing all the time, especially when the propagation process occurs in the network. In the following chapters, the research is carried out in dynamical networks. The relationship between the dynamic state of nodes and the network structure is explored.

# Chapter 7

# Study on the Reliability of Large-Scale Adaptive Weighted Network

## 7.1 Introduction

Allowing nodes to adaptively connect to reliable neighbors and disconnect those unreliable, a self-healing adaptive network is able to operate based on the credibility and reliability of individual nodes, and inhibit virus spread and cascading failures. Adaptive (weighted) networks have become increasingly important, as a result of the proliferation of the cloud computing [136]-[138], vehicular ad-hoc networks (VANETs) [139], and social networks [141]. Adaptive (weighted) networks are of particular interest in practice, where attacks are often strategic and responsive to defenders' actions. The networks can combat strategic attacks [142], by rewiring to bypass attacked nodes [36], [141]. As a result, the topology of the network keeps changing in response to the attacks, confusing the attackers, counteracting strategic attacks (e.g., to strategically critical nodes with high degrees in static networks), and transferring the strategic attacks to exhibit stationarity.

An example of adaptive weighted networks is network function virtualization (NFV) on cloud computing platforms, where a large number of virtual machines (VMs) are installed, running virtual network functions (VNFs) [143]-[145]. The VMs are connected through virtual links. Network services need to be processed at different VMs running different VNFs in correct orders. The VMs and virtual links can be configured in response to requests of network services, and the weight of a virtual link can indicate the workload of services that a VM partially completes and forwards to another VM for further processing. In the case where some VMs are congested due to distributed denial-of-service (DDoS) [148]-[150] attacks or infected due to computer viruses, new virtual links can be established to bypass these VMs. The weights (or in other words, the workloads) of the disconnected virtual links can be transferred to the new links. The VMs that are neither attacked nor infected can check their routing tables, decide to rewire their virtual links. The number of the new connections can be set to be equal to the number of links disconnected, so as to maintain the consistency of workload execution and the controllability of NFV.

Studies have been carried out to design rewiring protocols and analyze rewiring effects, typically in adaptive unweighted networks, where rewiring is random and independent of the logical or geographical closeness between a specific pair of nodes. In practice, there are great potentials for a healthy node to disconnect suspicious neighbors based on the frequency of communication occurrences. A healthy node may preferentially disconnect a frequently communicated, suspicious neighbor, so as to prevent cascading failures, such as DDoS attacks and virus infection, in NFV. Alternatively, a healthy node may choose to disconnect infrequently communicated, suspicious neighbors, so as to maintain the functionality of the network for intensive urgent tasks at the cost of network failures in the long term. The conditions inhibiting and facilitating virus spread or cascading failures are important to the analysis of network reliability. Extensive studies have been carried out on the conditions in conventional networks without rewiring or weighting of network links, by using the susceptible-infected-susceptible (SIS)

models [151], [152]. To the best of our knowledge, however, there has been no rigorous analytical study on the emerging adaptive weighted networks. A key challenge is that not only can the nodes change states (as modeled in typical SIS models [7]), but the links connecting the nodes can also rewire and change over time (as opposed to the typical SIS models). Another critical challenge is that the links can be differently weighted. These challenges cannot be straightforwardly addressed by existing SIS models. Nontrivial extensions of the models are required.

This chapter presents a new mean-field model to analyze the resistance of adaptive weighted networks against cascading failures, such as DDoS attack and computer virus. As a consequence of the new challenges, new derivations are necessary to extend the SIS model and evaluate the impact of rewiring and of weighted network links on the reliability of the adaptive weighted networks:

1) A new set of differential equations are formulated to model the continuous-time Markov chain process of the rewiring of weighted links in adaptive weighted networks. The differential equations are linearized. The largest eigenvalue of the Jacobian matrix of the linearization is the key to the study of the network reliability, but is not readily achievable.

2) We judiciously decompose the Jacobian matrix, evaluate the eigenvalues of the different parts by using determinant transformations and spectral analysis, and finally unveil the range of the largest eigenvalue of the Jacobian matrix. The upper and lower bounds of the range provide the sufficient conditions for the inhibition and proliferation of virus or cascading failures in adaptive weighted networks.

3) Two case studies verify the conditions, with exponentially and log-normally distributed link weights. By exploiting Order Statistics and Taylor expansion, we reveal that the condition of proliferation of virus or cascading failures is inversely proportional to both the network degree and average link weight. Extensive simulations confirm the validity of the identified conditions, as well as the effectiveness of adaptive weighted networks in terms of suppressing cascading failures. An important finding is that the distributions of the link weights can have a strong impact on network reliability against virus spread or cascading failures in adaptive weighted networks.

This is distinctively different from the existing conclusions on current static weighted networks [153]. We also find that the higher upper bound the rewiring rates of the weighted network links have, the more robust the adaptive weighted networks are against outbreaks of virus or failures. In the case of non-uniform rewiring rates, the distributions of the rewiring rates can also have a marked impact on the network reliability.

The rest of this chapter is organized as follows. In Section 7.2, the related works are reviewed. In Section 7.3, the structure of adaptive weighted network is described. The proposed mean-field model of adaptive weighted network is presented in Section 7.4, followed by the stability analysis of the adaptive weighted networks. Two rewiring strategies are discussed and evaluated in Section 7.5. In Section 7.6, numerical and simulation results are provided, followed by conclusions in Section 7.7.

## 7.2 Adaptive Weighted Network Structure

Consider a generic network of $N$ nodes connected by $L$ weighted bidirectional links. The weights of the links are collected in $W = \{w_1, w_2, \cdots, w_M\}$, $w_i > 0$, $i = 1, 2, \cdots, M$, where $M$ is the number of different weights, measuring the closeness between two connected nodes (e.g., in terms of distance or communication frequency). The higher a weight is, the closer two connected nodes are. The probability distribution of $w_i$ is denoted by $g(w_i)$. With reference to the SIS epidemic models, each node of the network can be in either a healthy/susceptible (S) or unhealthy/infected (I) state, indicating the node is reliable or not, respectively. We assume stationary random infections or failures which are reasonable in adaptive weighted networks, as discussed in Section 7.1. Moreover, the assumption of stationary random infections has been extensively assumed in the existing SIS models, even in the case where the networks are static and could be vulnerable to strategic attacks. At any instant, for a $w$-weighted link connecting an unreliable node and a reliable node (SI/IS link), the reliable node can become unreliable with the rate $\beta_w = \tau w$. $\tau$ is a coefficient known in prior. The unreliable node can recover with rate $\gamma$ as the result of patching or antivirus software updating.

We consider that the reliable nodes can protect themselves by disconnecting from unreliable neighbors and reconnecting to other reliable nodes, thereby preserving network reliability. With probability $r_w$ for an SI link weighted $w$, the reliable node breaks the link to the unreliable one and forms a new link to another randomly selected reliable node. The rewiring rate $r_w$ is a random variable in general cases. The weights of the disconnected links can be transferred to the new links, while the weights of other links remain unchanged. $r_w$ can depend on w, e.g., the closeness of the nodes. In the case of NFV, the weight of a virtual link can indicate the workload from one VM to another, as described in Section 7.1. The virtual links to the congested/failed VMs can be rewired to other VMs, and the weights (or workloads) of the links can be transferred to the new links. We assume that the number of links is fixed, as predominantly assumed in the literature on adaptive (un)weighted networks, e.g., [36], [142]. In many cases, the assumption is reasonable and practical to maintain the connectivity and controllability of the networks. In a special case where the entire network becomes alert to threats (e.g., known virus or failures), the rewiring rate can be independent of the link weight, i.e., $r_w = r$ , $\forall w$.

Fig 7.1 presents the operations of a node in an adaptive weighted network, where DDoS attacks or computer viruses can propagate to explore vulnerabilities in the network. The weight of the link between a pair of nodes can account for the frequency the nodes interact; or in other words, the workload the nodes send to each other. A susceptible (or healthy) node is more likely to be infected by an infected neighbor it interacts frequently, than by one it interacts infrequently. Once one of its neighbors is infected or fails, the node can observe the misbehaviors of the neighbor and rewire its link to bypass the infected neighbor, thereby preventing propagation of the attacks or failures. As a result, the topology of the network keeps changing in response to attacks or failures, quarantining infected individuals and counteracting the vulnerability explorations.

Other notations are defined as follows: $[A](A \in \{S, I\})$ denotes the number of nodes in state $A$, and $[S] + [I] = N$. $[AB]_w$ denotes the number of edges weighted $w$, connecting two nodes in states A and B,

and $2[SI]_w+[SS]_w+[II]_w=kNg(w)$. $k$ is the average degree of the network, and $2[SI]_w=[SI]_w+[IS]_w$. $k[A]=[AA]+[AB]$, $[AA]=\sum_{i=1}^{M}[AA]_{w_i}$, $[AB]=\sum_{i=1}^{M}[AB]_{w_i}$. $[ABC]_{ww'}$ denotes the number of triplets *A-B-C*, with edge *AB* weighted *w* and edge *BC* weighted *w'*.

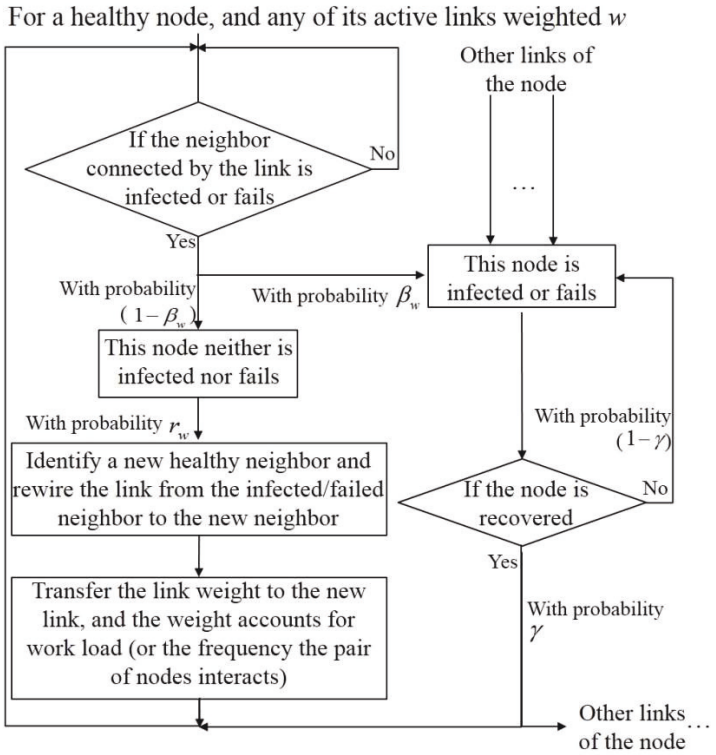



Fig 7.1. The flowchart of a node in regards of a *w*-weighted link. The model is continuous-time and therefore the flowchart runs continuously.

### 7.3 Proposed Mean-field Model of Adaptive Weighted Network

The research approach we take is to first model a new continuous-time Markov chain process to capture the rewiring and weighting of network links in adaptive weighted networks, and then analyze the conditions of the equilibriums of the model. The equilibriums considered in this chapter are: (1) a disease-free equilibrium in which virus infections or cascading failures are completely eliminated; and (2) an outbreak equilibrium in which the infections or failures are insuppressible. In other words, the whole adaptive weighted network stabilize, either free of infections/failures, or with insuppressible infections/failures. By applying the Hartman-Grobman theorem [154], the conditions of the equilibriums are analyzed by linearizing the model and evaluating the largest eigenvalue of the Jacobian matrix of the linearization. With mathematical manipulation, we derive the upper and lower bounds of the largest eigenvalue, which provide the sufficient conditions respectively for the proliferation and inhibition of virus or cascading failures in adaptive weighted networks.

Mean-field approximations are taken to improve the tractability of the continuous-time Markov Chain process. Mean-field theory studies the behavior of large and complex stochastic models where a large number of small individual components can interact with each other [103]. The mean-field approximations use a single average effect to approximate the effect of all the other individuals on any

given individual. As a result, the interactions between individuals can be decoupled for analytical tractability, and the populations of individuals with different characteristics can be studied. The mean-field approximation is suitable for large-scale networks [155].

The time-varying populations of the nodes and the links are captured by a set of differential equations, as given by (7.1).

$$\frac{d[S]}{dt} = \gamma \sum_i [I_i] - \sum_i \sum_w \beta_w [S_i I]_w \,, \tag{7.1a}$$

$$\frac{d[I]}{dt} = -\gamma \sum_i [I_i] + \sum_i \sum_w \beta_w [S_i I]_w \,, \tag{7.1b}$$

$$\begin{aligned} \frac{d[SS]_w}{dt} = {} & \gamma \sum_i \sum_j ([S_i I_j]_w + [I_i S_j]_w) + r_w \sum_i \sum_j ([S_i I_j]_w + [I_i S_j]_w) \\ & - \sum_{w'} \beta_{w'} \sum_i \sum_j ([S_i S_j I]_{ww'} + [IS_i S_j]_{w'w}) \end{aligned}, \tag{7.1c}$$

$$\begin{aligned} \frac{d[II]_w}{dt} = {} & 2\gamma \sum_i \sum_j [I_i I_j]_w + \beta_w \sum_i \sum_j ([S_i I_j]_w + [I_i S_j]_w) \\ & + \sum_{w'} \beta_{w'} \sum_i \sum_j ([I_i S_j I]_{ww'} + [IS_i I_j]_{w'w}) \end{aligned}, \tag{7.1d}$$

$$\begin{aligned} \frac{d[SI]_w}{dt} = {} & -\gamma \sum_i \sum_j ([S_i I_j]_w + [I_i I_j]_w) - \beta_w \sum_i \sum_j [S_i I_j]_w \\ & + \sum_{w'} \beta_{w'} \sum_i \sum_j ([S_i S_j I]_{ww'} - [IS_i I_j]_{w'w}) - r_w \sum_i \sum_j [S_i I_j]_w \end{aligned}, \tag{7.1e}$$

$$\begin{aligned} \frac{d[IS]_w}{dt} = {} & -\gamma \sum_i \sum_j ([I_i S_j]_w + [I_i I_j]_w) - \beta_w \sum_i \sum_j [I_i S_j]_w \\ & + \sum_{w'} \beta_{w'} \sum_i \sum_j ([IS_i S_j]_{w'w} - [I_i S_j I]_{ww'}) - r_w \sum_i \sum_j [I_i S_j]_w \end{aligned}, \tag{7.1f}$$

where $[A_i]$ represents the probability that the node $i$'s state is $A$, $[A_i B_j]$ represents the probability that a link connecting a pair of nodes in states $A$ and $B$, $\forall A, B \in \{S, I\}$. Here, (7.1a) captures the time-changing population of nodes in the healthy (or susceptible) state. The first term on the right-hand side (RHS) of (7.1a) corresponds to the part of the population recovering from the infected state with the probability of $\gamma$. The second term corresponds to the part of the population infected by their infected neighbors with the probability $\beta_w$. Likewise, (7.1b) captures the time-changing population of nodes in the infected state.

Eqs. (7.1c) and (7.1d) characterize the time-varying numbers of links weighted by different weights and connecting nodes in different states. For instance, (7.1c) captures the changing number of the $w$-weighted links connecting two healthy nodes. The first term on the RHS of (7.1c) is the increased part of the link number, resulting from the recovery of the infected ends of the links with the probability of $\gamma$. The second term is another increased part of the link number, resulting from rewiring to bypass an infected node with the probability of $r_w$. The third term is the number of the previous $w$-weighted SS links which become SI links due to the infection at one end of the links through a $w'$-weighted link with the probability of $\beta_{w'}$. Likewise, (7.1d) captures the time-changing number of $w$-weighted II links connecting a pair of infected nodes, (7.1e) and (7.1f) capture the time-changing number of $w$-weighted SI and IS links, respectively.

By taking the mean-field approximation, the expectation of infected nodes in the network can be

written as the sum of the probability that each node in the network is infected, i.e., $[A]=\sum_i [A_i] = N[A_i]$. Similarly, the expectation of $w$-weighted SI links can be written as $[AB]_w=\sum_i\sum_j [A_iB_j]_w$, by assuming all the $w$-weighted edges exhibit the same state. The temporal changes in the population of healthy nodes and infected nodes can be written as

$$\frac{d[S]}{dt} = \gamma[I] - \sum_w \beta_w [SI]_w , \tag{7.2a}$$

$$\frac{d[I]}{dt} = -\gamma[I] + \sum_w \beta_w [SI]_w , \tag{7.2b}$$

where (7.2a) captures the time-changing population of nodes in the healthy (or susceptible) state. The first term on the RHS of (7.2a) corresponds to the part of the population recovering from the infected state with the probability of $\gamma$. The second term corresponds to the part of the population infected by their infected neighbors with the probability of $\beta_w$. Likewise, (7.2b) captures the time-changing population of nodes in the infected state.

The temporal change of a link depends on its weight and the states of the nodes at both ends of the link. By taking the mean-field approximation, the expectation of the $w$-weighted $AB$ links in the network can be written as the sum of the probability that each link connecting a pair of $A$ node and $B$ node, i.e., $[AB]_w=\sum_i\sum_j [A_iB_j]_w$, $\forall A, B \in \{\mathrm{S}, \mathrm{I}\}$. The temporal changes in the numbers of links of different types can be written as

$$\frac{d[SS]_w}{dt} = \gamma([SI]_w + [IS]_w) + r_w([SI]_w + [IS]_w) - \sum_{w'} \beta_{w'}([SSI]_{ww'} + [ISS]_{w'w}) , \tag{7.3a}$$

$$\frac{d[II]_w}{dt} = -2\gamma[II]_w + \beta_w([SI]_w + [IS]_w) + \sum_{w'} \beta_{w'}([ISI]_{ww'} + [ISI]_{w'w}) , \tag{7.3b}$$

$$\frac{d[SI]_w}{dt} = -\gamma[SI]_w + \gamma[II]_w - \beta_w[SI]_w + \sum_{w'} \beta_{w'}([SSI]_{ww'} - [ISI]_{w'w}) - r_w[SI]_w , \tag{7.3c}$$

$$\frac{d[SI]_w}{dt} = -\gamma[IS]_w + \gamma[II]_w - \beta_w[IS]_w + \sum_{w'} \beta_{w'}([ISS]_{w'w} - [ISI]_{ww'}) - r_w[IS]_w . \tag{7.3d}$$

Eqs. (7.3a) and (7.3b) characterize the time-varying numbers of links weighted by different weights and connecting nodes in different states. For instance, (7.3a) captures the time-changing number of the $w$-weighted links connecting two healthy nodes. The first term on the RHS of (7.3a) results from the recovery of the infected ends of the links with the probability of $\gamma$. The second term on the RHS of (7.3a) results from rewiring to bypass an infected node with the probability of $r_w$. The third term is the number of previous $w$-weighted SS links which become SI links due to the infection at one end of the links through a $w'$-weighted link with the probability of $\beta_{w'}$. Likewise, (7.3b) captures the time-changing number of $w$-weighted II links connecting a pair of infected nodes; and (7.3c) and (7.3d) capture the time-changing number of $w$-weighted SI and IS links, respectively.

We assume that the weight of a link is symmetry, i.e., $w(i,j) = w(j,i), \forall i \neq j$. Then

$[SI]_w = [IS]_w$. (7.2) and (7.3) can be rewritten as

$$\frac{d[S]}{dt} = \gamma[I] - \sum_w \beta_w [SI]_w , \tag{7.4a}$$

$$\frac{d[I]}{dt} = -\gamma[I] + \sum_w \beta_w [SI]_w , \tag{7.4b}$$

$$\frac{d[SS]_w}{dt} = 2\gamma[SI]_w + 2r_w[SI]_w - 2\sum_{w'} \beta_{w'}[SSI]_{ww'} , \tag{7.4c}$$

$$\frac{d[II]_w}{dt} = -2\gamma[II]_w + 2\beta_w[SI]_w + 2\sum_{w'} \beta_{w'}[ISI]_{ww'} , \tag{7.4d}$$

$$\frac{d[SI]_w}{dt} = -\gamma[SI]_w + \gamma[II]_w - \beta_w[SI]_w - r_w[SI]_w - \sum_{w'} \beta_{w'}[ISI]_{w'w} + \sum_{w'} \beta_{w'}[SSI]_{ww'} , \tag{7.4e}$$

which is the mean-field approximation of the continuous-time Markov chain model for the adaptive weighted networks. The time-varying states of both the nodes and links are captured.

We note that the linear expressions in (7.4) are due to the fact that the system of interest is continuous-time. At every time instant $\Delta t \to 0$, the probability of a healthy node being infected by more than one infected neighbor approaches zero. We can apply the moment closure approximation to evaluate $[A]$ and $[AB]_w$, and the number of triplets $[ABC]_{ww'}$ can be written as [153]:

$$[ABC]_{ww'} = \xi \frac{[AB]_w [BC]_{w'}}{[B]} , \tag{7.5}$$

where $\xi = \frac{k-1}{k}$. Based on (7.5), we can have

$$[SSI]_{ww'} = \xi \frac{[SS]_w [SI]_{w'}}{[S]} , \tag{7.6}$$

$$[ISI]_{ww'} = \xi \frac{[SI]_w [SI]_{w'}}{[S]} . \tag{7.7}$$

By substituting (7.6) and (7.7) into (7.4), we can rewrite (7.4) as

$$\frac{d[S]}{dt} = \gamma[I] - \sum_w \beta_w [SI]_w , \tag{7.8a}$$

$$\frac{d[I]}{dt} = -\gamma[I] + \sum_w \beta_w [SI]_w , \tag{7.8b}$$

$$\frac{d[SS]_w}{dt} = 2(\gamma + r_w)[SI]_w - 2\xi \frac{[SI]_w}{[S]} \sum_{w'} \beta_{w'}[SI]_{w'} , \tag{7.8c}$$

$$\frac{d[II]_w}{dt} = -2(\gamma[II]_w - \beta_w[SI]_w - \xi \frac{[SI]_w}{[S]} \sum_{w'} \beta_{w'}[SI]_{w'}) , \tag{7.8d}$$

$$\frac{d[SI]_w}{dt} = -(\gamma + \beta_w + r_w)[SI]_w + \gamma[II]_w + \xi \frac{[SS]_w - [SI]_w}{[S]} \sum_{w'} \beta_{w'}[SI]_{w'} . \tag{7.8e}$$

For the purpose of cross-validation of the proposed model, we consider a special case where the rewiring rate $r_w = r$ is a constant. When $r_w = 0$, i.e., the network is static, and (7.8) can be rewritten in the

exactly same way as [156], eq.6] describing static weighted networks. In other words, (7.8) is cross-validated by the special case.

### 7.4 Stability Analysis of Adaptive Weighted Network

We proceed to derive the reliability threshold of $\tau$, denoted by $\tau^*$ based on (7.8). $\tau^*$ is the evaluation of the reliability of the adaptive weighted networks against cascading failures. If $\tau < \tau^*$, the adaptive weighted network can eventually become reliable, i.e., all nodes eventually become reliable. Otherwise, the network is unreliable, i.e., the unreliability proliferates. The larger $\tau^*$ is, the more resilient the network is, e.g., against virus spread and cascading failures. To derive $\tau^*$, we analyze the stability of the equilibrium of the adaptive weighted networks.

As stated in the Lyapunov's first method [157], the behavior of a dynamical system in a domain near an equilibrium point is qualitatively the same as the behavior of its linearization near this equilibrium point. If and only if the Jacobian matrix of the linearization has all negative eigenvalues, a nonlinear dynamical system is stable at the equilibrium. The equilibrium point of interest, also known as the disease-free equilibrium point, is $(\frac{d[II]_w}{dt}, \frac{d[SI]_w}{dt}) = (0,0)$, at which all nodes are reliable [158]. By exploiting the Lyapunov's first method, we linearize (7.8) in the vicinity of the equilibrium, evaluate the eigenvalues of the linearization at the equilibrium, and study the condition under which the Jacobian matrix of the linear system has all negative eigenvalues [159]. As a result, we are able to establish the thresholds to preserve stability or undergo instability at the equilibrium.

Based on the aforementioned condition $[S]+[I]=N$, $2[SI]_w+[SS]_w+[II]_w = kNg(w)$ and $2[SI]_w=[SI]_w+[IS]_w$, (7.8d) and (7.8e) can be respectively rewritten as

$$\frac{d[II]_w}{dt} = -2\gamma[II]_w + 2\beta_w[SI]_w + \frac{2(k-1)[SI]_w}{kN-[SI]-[II]}\sum_{w'}\beta_{w'}[SI]_{w'}, \tag{7.9a}$$

$$\frac{d[SI]_w}{dt} = (k-1)\frac{kNg(w)-3[SI]_w-[II]_w}{kN-[SI]-[II]}\sum_{w'}\beta_{w'}[SI]_{w'} - (\gamma+\beta_w+r_w)[SI]_w + \gamma[II]_w. \tag{7.9b}$$

By suppressing all higher order terms of $[II]_w$ and $[SI]_w$, (7.9) can be linearized, as given by

$$\frac{d[II]_w}{dt} \approx -2\gamma[II]_w + 2\beta_w[SI]_w, \tag{7.10a}$$

$$\frac{d[SI]_w}{dt} \approx (k-1)g(w)\sum_{w'}\beta_{w'}[SI]_{w'} - (\gamma+\beta_w+r_w)[SI]_w + \gamma[II]_w. \tag{7.10b}$$

The linear stability analysis of (7.10) is carried out in the vicinity of the equilibrium. By the Hartman-Grobman theorem [160], the behavior of the system around an equilibrium point can be evaluated through the eigenvalues of the Jacobian matrix of (7.10). Let $\mathbf{J} = [\mathbf{J}_{ij}]$ denote the Jacobian matrix of (7.10) at the equilibrium, as given by [156]

$$\mathbf{J}=\begin{pmatrix}\mathbf{J}^{11} & \mathbf{J}^{12}\\ \mathbf{J}^{21} & \mathbf{J}^{22}\end{pmatrix} \tag{7.11}$$

where

$\mathbf{J}^{11}$=diag[$-2\gamma$,$-2\gamma$,...,$-2\gamma$],

$\mathbf{J}^{12}$=diag[$2\beta_{w_1}$,$2\beta_{w_2}$,...,$2\beta_{w_M}$],

$\mathbf{J}^{21}$=diag[$\gamma$,$\gamma$,...,$\gamma$],

$$\mathbf{J}^{22}=\begin{cases}(k-1)g(w_i)\beta_{w_j}, & if\ i\neq j;\\ -(\gamma+\beta_{w_i}+r_{w_i})+(k-1)g(w_i)\beta_{w_i}, & if\ i=j.\end{cases}$$

Note that the matrix block $\mathbf{J}^{11}$ is an *M*-by-*M* diagonal matrix. By block matrix multiplication, we can get

$$\begin{pmatrix}\mathbf{J}^{11} & \mathbf{J}^{12}\\ \mathbf{J}^{21} & \mathbf{J}^{22}\end{pmatrix}\begin{pmatrix}\mathbf{I}_M & -\mathbf{J}^{11^{-1}}\mathbf{J}^{12}\\ 0 & \mathbf{I}_M\end{pmatrix}=\begin{pmatrix}\mathbf{J}^{11} & 0\\ \mathbf{J}^{21} & \mathbf{H}\end{pmatrix}, \tag{7.12}$$

where $\mathbf{I}_M$ is the identity matrix, and $\mathbf{H}=\mathbf{J}^{22}-\mathbf{J}^{21}\mathbf{J}^{11^{-1}}\mathbf{J}^{12}$.

Let row vector $\mu$ be the spectrum of the square matrix **J**, i.e., collects all eigenvalues of **J**, and $\mu_i\in\mu$ be the *i*-th (largest) eigenvalue of **J**, $i$ = 1, 2, …, 2*M*. The characteristic polynomial of **J** can be written as

$$\det(\mathbf{J}-\mu_i\mathbf{I}_{2M})=\det\left[\begin{pmatrix}\mathbf{J}^{11} & 0\\ \mathbf{J}^{21} & \mathbf{H}\end{pmatrix}-\mu_i\mathbf{I}_{2M}\right]=\det\begin{bmatrix}\mathbf{J}^{11}-\lambda_i\mathbf{I}_M & 0\\ \mathbf{J}^{21} & \mathbf{H}-\eta_i\mathbf{I}_M\end{bmatrix}, \tag{7.13}$$

where $\lambda_i$ and $\eta_i$ are the *i*-th eigenvalues of $\lambda$ and $\eta$, i.e., the spectra of the square matrices $\mathbf{J}^{11}$ and **H**, respectively. Then we have $\det(\mathbf{J}-\mu_i\mathbf{I}_{2M})=0$, so that

$$\det(\mathbf{J}^{11}-\lambda_i\mathbf{I}_M)=0,\ \ \det(\mathbf{H}-\eta_i\mathbf{I}_M)=0,$$

in other words, [$\lambda$, $\eta$] is also the spectrum of **J** from (7.13).

In [159], the linear state model (7.10) is stable at the equilibrium, if and only if the real parts of all the eigenvalues of **J** are negative. Since $\mathbf{J}^{11}$ is an *M*-by-*M* diagonal matrix with all the main diagonal entries equal to $-2\gamma$, all of the M eigenvalues of $\mathbf{J}^{11}$ are $-2\gamma<0$. To this end, we can have $\max\{\mu\}<0$, if and only if the maximum eigenvalue of **H**, denoted by $\eta_{\max}(\mathbf{H})<0$. Let $\mathbf{H}=[H_{ij}]$, we have

$$H_{ij}=\begin{cases}(k-1)g(w_i)\beta_{w_j}, & if\ i\neq j\\ -(\gamma+r_{w_i})+(k-1)g(w_i)\beta_{w_i}, & if\ i=j\end{cases}.$$

To evaluate $\eta_{\max}(\mathbf{H})$, we decouple **H** as $\mathbf{H}=\mathbf{H}^{(1)}+\mathbf{H}^{(2)}$, and rewrite $\mathbf{H}^{(1)}=[H_{ij}^{(1)}]$ and $\mathbf{H}^{(2)}=[H_{ij}^{(2)}]$. Then,

$$H_{ij} = H_{ij}^{(1)} + H_{ij}^{(2)}, \tag{7.14}$$

where

$$H_{ij}^{(1)} = \begin{cases} (k-1)g(w_i)\beta_{w_j} & if\ i \neq j \\ -\gamma + (k-1)g(w_i)\beta_{w_i} & if\ i = j \end{cases}, \tag{7.15}$$

$$H_{ij}^{(2)} = \begin{cases} 0 & if\ i \neq j \\ -r_{w_i} & if\ i = j \end{cases}. \tag{7.16}$$

By substituting (7.15), we can write $\det[\mathbf{H}^{(1)}]$, as given by

$$\det[\mathbf{H}^{(1)}] = (k-1)^M \prod_{i=1}^{M} g(w_i) \prod_{i=1}^{M} \beta_{w_i} \begin{vmatrix} 1+x_1 & \cdots & 1 \\ 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1+x_M \end{vmatrix}, \tag{7.17}$$

where for notational simplicity, we define

$$x_i = -\frac{\gamma}{(k-1)g(w_i)\beta_{w_i}}, i = 1, 2, ..., M \cdot$$

According to basic determinant transformations, we can have

$$\begin{vmatrix} 1+x_1 & 1 & \cdots & 1 \\ 1 & 1+x_2 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1+x_M \end{vmatrix} \overset{(a)}{=} \begin{vmatrix} 1+x_1 & 1 & \cdots & 1 \\ -x_1 & x_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -x_1 & 0 & \cdots & x_M \end{vmatrix} = \prod_{i=1}^{M} x_i \begin{vmatrix} 1+\frac{1}{x_1} & \frac{1}{x_2} & \cdots & \frac{1}{x_M} \\ -1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & \cdots & 1 \end{vmatrix},$$
$$\overset{(b)}{=} \prod_{i=1}^{M} x_i \begin{vmatrix} 1+\sum_{i=1}^{M} \frac{1}{x_i} & \frac{1}{x_2} & \cdots & \frac{1}{x_M} \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix} \overset{(c)}{=} (\prod_{i=1}^{M} x_i)(1 + \sum_{i=1}^{M} \frac{1}{x_i}) \tag{7.18}$$

where (*a*) is achieved by subtracting the first row from all the rest of rows; (*b*) is achieved by adding all the other columns to the first column; (*c*) is obtained by multiplying the elements of the first column to their respective minors.

As a result, (7.17) can be rewritten as

$$\begin{aligned} \det[\mathbf{H}^{(1)}] &= (k-1)^M (\prod_{i=1}^{M} g(w_i))(\prod_{i=1}^{M} \beta_{w_i})(\prod_{i=1}^{M} x_i)(1 + \sum_{i=1}^{M} \frac{1}{x_i}), \\ &= (-\gamma)^{M-1}(-\gamma + (k-1)\langle \beta_w \rangle) \end{aligned} \tag{7.19}$$

where $\langle \beta_w \rangle = \sum_{n=1}^{M} g(w_n)\beta_{w_n}$, takes the expectation of $\beta_w$.

The spectrum of $\mathbf{H}^{(1)}$, denoted accordingly by $\eta^{(1)}$, satisfies $\det[\mathbf{H}^{(1)} - \eta_i^{(1)}\mathbf{I}] = 0$. According to (7.19), we have

$$(-\gamma - \eta_i^{(1)})^{M-1}(-\gamma + (k-1)\langle\beta_w\rangle - \eta_{\max}{}^{(1)}) = 0. \tag{7.20}$$

By solving (7.20), one can obtain

$$\eta_i^{(1)} = -\gamma, i = 2,...,M,$$
$$\eta_1^{(1)} = \eta_{\max}{}^{(1)} = (k-1)\langle\beta_w\rangle - \gamma.$$

where $\eta_{\max}{}^{(1)}$ is the largest eigenvalue of $\mathbf{H}^{(1)}$ and $\eta_i{}^{(1)}$ is any other eigenvalue.

Let $\zeta = \gamma + \max_i\{r_{w_i}\} + 1$, then $\mathbf{H} + \zeta\mathbf{I} > 0$, where all the entries are positive. Since $\mathbf{H} = \mathbf{H}^{(1)} + \mathbf{H}^{(2)}$ and $\mathbf{H}^{(2)}$ is a diagonal matrix, we can get

$$0 < \mathbf{H}^{(1)} + \eta_m{}^{(2)}\mathbf{I} + \zeta\mathbf{I} \le \mathbf{H} + \zeta\mathbf{I} \le \mathbf{H}^{(1)} + \eta_1{}^{(2)}\mathbf{I} + \zeta\mathbf{I},$$

where $\eta_m{}^{(2)} = \min_i\{-r_{w_i}\}$, $\eta_1{}^{(2)} = \max_i\{-r_{w_i}\}$. Matrices $\mathbf{P} \le \mathbf{Q}$ stands for that entry of $\mathbf{P}$ is no less than the corresponding entry of $\mathbf{Q}$.

By Perron-Frobenius theorem [161], for any matrices $\mathbf{A}$ and $\mathbf{B}$ with $0 \le \mathbf{A} \le \mathbf{B}$, the spectral radii of $\mathbf{A}$ and $\mathbf{B}$ satisfy $\rho(\mathbf{A}) \le \rho(\mathbf{B})$. For $\mathbf{A} = [a_{ij}]$ with $a_{ij} > 0$, $\forall i, j$, the spectral radius $\rho(\mathbf{A})$ is equal to the largest eigenvalue. Therefore, $\eta_{\max}(\mathbf{H})$ satisfies

$$\eta_1^{(1)} + \eta_m{}^{(2)} + \zeta \le \eta_{\max}(\mathbf{H}) + \zeta \le \eta_1^{(1)} + \eta_1^{(2)} + \zeta,$$

i.e.,

$$\eta_1^{(1)} + \eta_m{}^{(2)} \le \eta_{\max}(\mathbf{H}) \le \eta_1^{(1)} + \eta_1^{(2)}. \tag{7.21}$$

From [159], the equilibrium is stable, if $\eta_{\max}(\mathbf{H}) \le \eta_1^{(1)} + \eta_1^{(2)} < 0$; and the equilibrium can be unstable, if $\eta_{\max}(\mathbf{H}) \ge \eta_1^{(1)} + \eta_m{}^{(2)} > 0$. Since $\beta w = \tau w$, $\langle\beta w\rangle = \tau\langle w\rangle$, we have

$$\eta_1^{(1)} + \eta_1^{(2)} = (k-1)\tau\langle w\rangle - \gamma + \eta_1{}^{(2)},$$
$$\eta_1^{(1)} + \eta_m{}^{(2)} = (k-1)\tau\langle w\rangle - \gamma + \eta_m{}^{(2)}.$$

The network is reliable if

$$\tau < \tau_l^* = \frac{\gamma + \min_i\{r_{w_i}\}}{(k-1)\langle w\rangle}, \tag{7.22}$$

where $\tau_l^*$ gives the lower bound of $\tau$ in reliable states.

The network is unreliable if

$$\tau > \tau_u^* = \frac{\gamma + \max_i \{r_{w_i}\}}{(k-1)\langle w \rangle}, \tag{7.23}$$

where $\tau_u^*$ gives the upper bound of $\tau$ in unreliable states.

In the special case where $r_{w_i} = r$, we have $\min_i \{r_{w_i}\} = \max_i \{r_{w_i}\} = r$, and $\tau^*$ can be written explicitly in a closed-form, as given by

$$\tau^* = \frac{\gamma + r}{(k-1)\langle w \rangle}. \tag{7.24}$$

We can see in (7.24) that $\tau^*$ increases with the growth of the rewiring rate *r*, and decreases with the growth of the average link weights $\langle w \rangle$ in the case of uniform rewiring rates. It is also shown that the distribution of weight $w_i$ has little impact on $\tau^*$ when $r_{w_i} = r$. Moreover, we can see in (7.23) that, with non-uniform rewiring rates, the bounds of $\tau^*$ depend on $r_{w_i}$ which, in turn, depends on $w_i$. As the network becomes unreliable when $\tau > \tau^*$, we conclude that the higher the upper bound of $r_{w_i}$ is, the more resistant the network is against the outbreak. To this end, in the case of non-uniform rewiring rates $r_{w_i}$, the distribution of $w_i$ can have a strong impact on the upper bound of $r_{w_i}$ and thus the resistance of the network.

We note that our analysis is distinctively different from the existing studies. As discussed in Section II, the existing study of adaptive weighted networks, i.e., [151], [152], is based on numerical evaluations and provides no analysis of the reliability of the networks. In different yet relevant contexts of (static) weighted networks and adaptive unweighted networks, analyses do avail, and may also evaluate the reliability by assessing the eigenvalues of the Jacobian. However, the analysis of weighted networks does not capture the rewiring rate $r_w$ which is key to the reliability threshold of adaptive weighted networks $\tau^*$; see (7.22) and (7.23). The analysis of adaptive unweighted networks cannot account for the non-uniform weights and the subsequent infection rates of adaptive weighted networks, which require new mathematic manipulations and lead to the new bounds of $\tau^*$. In contrast, we consider a new adaptive weighted network, where links can be rewired on-the-fly and the weights of disconnected links can be transferred to the new links. We develop a new continuous-time Markov model to characterize the changing states of the links, capturing the real-time rewiring process of the networks. With the non-trivial analysis of the Jacobian of the linearization of the model, the largest eigenvalue of the Jacobian is analyzed to specify the respective thresholds under which cascading failures can be inhibited or proliferate.

We also note that the link weights may not be symmetric in the presence of DDoS attacks. Our model can be readily applied to asymmetric link weights. As a matter of fact, (7.3) divides all the *w*-weighted links into four different types: $SS_w$, $II_w$, $SI_w$ and $IS_w$, and provides the temporal changes in the numbers of links of the different types. $[IS]_w$ does not have to be equal to $[SI]_w$, or in other words, the

link weights can be asymmetric. The above analysis, involving the linearization of differential equations, the derivation of the Jacobian of the linearized, and the evaluation of the eigenvalues of the Jacobian, can be readily based on (7.3).

### 7.5 Rewiring Strategies and Network Stability

As discussed in Section 7.4, the rewiring rate $r_{w_i}$ can be designed in different ways which can have a strong impact on the bounds of $\tau^*$. This section studies the impact by taking two different but simple linear designs of $r_{w_i}$ under two classical distributions of $w_i$ for example. Let $w_{(i)}$ denote the $i$-th smallest of $W \in R^{M\times 1}$.

The first design (Design 1) specifies the positive correlation between the rewiring rate and $w_i$, i.e., $r_{w_i} = \alpha_1 w_i, \alpha_1 \geq 0$. This is the case where a reliable node preferentially breaks its heavily loaded links with frequently interacted neighbors, especially in the case of cascading failures. The second design (Design 2) specifies the negative correlation between the rewiring rate and $w_i$, i.e., $r_{w_i} = \alpha_2(1 - \frac{w_i}{\max\{w_i\}}), \alpha_2 \geq 0$. This is the case where a reliable node preferentially breaks its infrequently used (or lightly loaded) links, especially for the purpose of alleviating interruptions to ongoing network operations.

The two example distributions of W are: (a) exponential distribution (ED)[162], [163], and (b) log-normal distribution (LD)[164]-[166]. Both distributions are non-negative and suitable to describe the nonnegative link weights of adaptive weighted networks. For the purpose of fair comparisons between the strategies, the mean of the exponential distribution is set to be equal to that of the log-normal distribution. Given the same mean, denoted by $\langle w \rangle$, the two distributions have different dispersions, and so are the expectations of $w_{(1)}$ and $w_{(M)}$ under different distributions. Order statistics[167]-[169] are exploited to evaluate $w_{(1)}$ and $w_{(M)}$, and in turn the lower bounds of $\tau^*$ in (7.23). This helps provide insight on the importance of dispersion on the reliability of the adaptive networks.

#### 7.5.1 Exponential Distribution

The probability density distribution (PDF) and cumulative distribution function (CDF) of $w_i$ $\forall i$ are $f(w_i) = \lambda e^{-\lambda w_i}$ and $F(w_i) = 1 - e^{-\lambda w_i}$, respectively. By exploring order statistics, the PDFs of $w_{(1)}$ and $w_{(M)}$ can be written as

$$f_1(w_{(1)}) = \lambda M e^{-\lambda M w_{(1)}},$$

$$f_M(w_{(M)}) = \lambda M [1 - e^{-\lambda w_{(M)}}]^{M-1} e^{-\lambda w_{(M)}}.$$

We can find that $w_{(1)}$ has an exponential distribution with parameter $\lambda M$. As a result,

$$\mathrm{E}[w_{(1)}] = \frac{1}{\lambda M} \tag{7.25}$$

The PDF of $w_{(M)}$ can be rewritten as

$$\begin{aligned} f_M(w_{(M)}) &= \lambda M[1-e^{-\lambda w_{(M)}}]^{M-1} e^{-\lambda w_{(M)}} \\ &= \lambda M e^{-\lambda w_{(M)}} \sum_{i=0}^{M-1} \binom{M-1}{i} (-e^{-\lambda w_{(M)}})^i \\ &= \sum_{i=0}^{M-1} (-1)^i M \binom{M-1}{i} \lambda e^{-(i+1)\lambda w_{(M)}} \end{aligned} . \tag{7.26}$$

By exploiting order statistics, the expectation of $w_{(M)}$ is given by

$$\mathrm{E}[w_{(M)}] = \frac{1}{\lambda} \sum_{i=1}^{M} \frac{1}{i} . \tag{7.27}$$

For Design 1 where $r_{w_i} \propto w_i$, $\mathrm{E}[\max_i \{r_{w_i}\}] \propto \mathrm{E}[\max_i \{w_i\}] = \frac{1}{\lambda} \sum_{i=1}^{M} \frac{1}{i}$, the network is unreliable if $\tau > \tau^* \propto \frac{\gamma + \frac{1}{\lambda} \sum_{i=1}^{M} \frac{1}{i}}{(k-1)\langle w \rangle}$.

For Design 2 where $r_{w_i} \propto \theta\text{-}w_i$, $\mathrm{E}[\max_i \{r_{w_i}\}] \propto \mathrm{E}[\theta - \max_i \{w_i\}] = \theta - \frac{1}{\lambda} \sum_{i=1}^{M} \frac{1}{i}$, the network is unreliable if $\tau > \tau^* \propto \frac{\gamma + \theta - \frac{1}{\lambda} \sum_{i=1}^{M} \frac{1}{i}}{(k-1)\langle w \rangle}$.

### 7.5.2 Log-normal Distribution

Let $f(w_i)$ and $F(w_i)$ be the PDF and CDF of $w_i$. Then we have

$$f(w_i) = \frac{1}{w_i \sigma \sqrt{2\pi}} e^{-\frac{(\ln w_i - \mu)^2}{2\sigma^2}} = \frac{\phi(\log w_i)}{w_i},$$

$$F(w_i) = \Phi(\log w_i),$$

where $\mu$ and $\sigma$ are the mean and the standard deviation, respectively, and $\phi(\cdot)$ and $\Phi(\cdot)$ denote the PDF and CDF of the normal distribution respectively. The mean $m$ and the variance $v$ are functions of $\mu$ and $\sigma$, as given by

$$m = e^{\mu + \frac{\sigma^2}{2}}, \quad v = e^{(2\mu + \sigma^2)(e^{\sigma^2} - 1)} . \tag{7.28}$$

By exploring order statistics, the PDF of $w_{(M)}$ can be written as

$$f_M(w_{(M)}) = M[F(w_{(M)})]^{M-1} f(w_{(M)}) = M[\Phi(\log w_{(M)})]^{M-1} \frac{\phi(\log w_{(M)})}{w_{(M)}} . \tag{7.29}$$

The expectation of $w_{(k)}$ is

$$\mathrm{E}[w_{(k)}] = \int_0^{+\infty} w_{(k)} f_k(w_{(k)}) dw_{(k)}, k = 1, 2, ..., M \text{ , } k = 1, 2, \ldots, M. \tag{7.30}$$

By submitting (7.29) to (7.30), we have

$$\mathrm{E}[w_{(M)}] = M \int_0^{+\infty} e^{y} [\Phi(y)]^{M-1} \phi(y) dy, y = \log(w_{(M)}) . \tag{7.31}$$

By exploiting order statistics, the expectation of $w_{(M)}$ is given by

$$\mathrm{E}[w_{(M)}] = \frac{M}{\sqrt{\pi}} (\frac{1}{2})^{M-1} \sum_{i=0}^{M-1} \binom{M-1}{i} I(z \mid \mu, \sigma) , \tag{7.32}$$

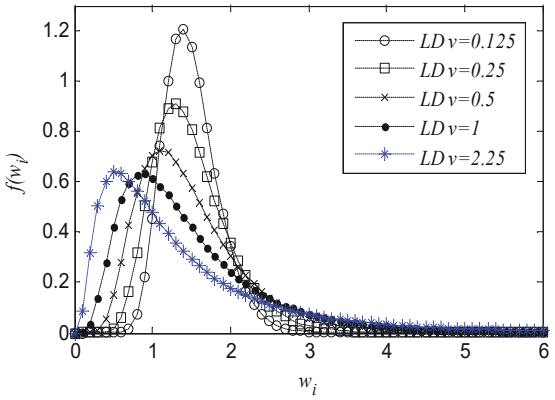
where $z = \frac{y-\mu}{\sqrt{2\sigma}}$, $I(z \mid \mu, \sigma) = (\frac{2}{\sqrt{\pi}})^{i} ( \sum_{M_1, \ldots, M_i = 0}^{+\infty} \frac{(-1)^{\sum_{l=1}^{i} M_l}}{\prod_{l=1}^{i} (2M_l + 1) \prod_{l=1}^{i} M_l !} )^{i} \times \int_{-\infty}^{+\infty} e^{z^2 + \sqrt{2}\sigma z + \mu} z^{2\sum_{l=1}^{i} M_l + i} dz$ .

For Design 1 where $r_{w_i} \propto w_i$, $\mathrm{E}[\max_i \{r_{w_i}\}] \propto \frac{M}{\sqrt{\pi}} (\frac{1}{2})^{M-1} \sum_{i=0}^{M-1} \binom{M-1}{i} I(z \mid \mu, \sigma)$, the network is unreliable if

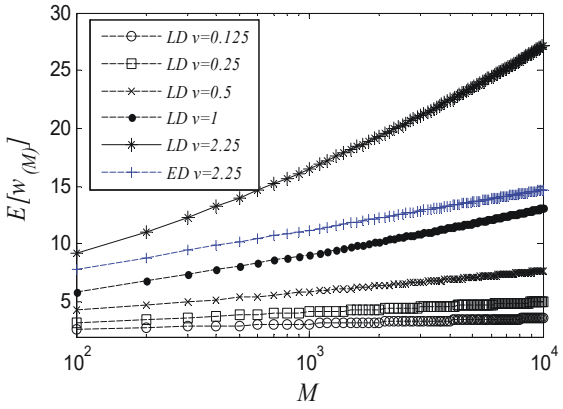
$$\tau > \tau^* \propto \frac{\gamma + \frac{M}{\sqrt{\pi}} (\frac{1}{2})^{M-1} \sum_{i=0}^{M-1} \binom{M-1}{i} I(z \mid \mu, \sigma)}{(k-1) \langle w \rangle} .$$

For Design 2 where $r_{w_i} \propto \theta - w_i$, $\mathrm{E}(\max_i \{r_{w_i}\}) \propto \mathrm{E}(\theta - \max_i \{w_i\}) = \theta - \frac{M}{\sqrt{\pi}} (\frac{1}{2})^{M-1} \sum_{i=0}^{M-1} \binom{M-1}{i} I(z \mid \mu, \sigma)$, the network is unreliable if

$$\tau > \tau^* \propto \frac{\gamma + \theta - \frac{M}{\sqrt{\pi}} (\frac{1}{2})^{M-1} \sum_{i=0}^{M-1} \binom{M-1}{i} I(z \mid \mu, \sigma)}{(k-1) \langle w \rangle} .$$




(a) The PDF of the log-normal distribution




(b) $E[w_{(M)}]$ of the log-normal distributions

Fig 7.2. The PDF and E[$w_{(M)}$] of the log-normal distribution, where the mean of the distribution is $m$ = 1.5, and $v$ is the variance of the distribution. We plot $v$ = 0.125, 0.25, 0.5, 1 and 2.25 for the log-normal distribution to show the impact of the variance on the E[$w_{(M)}$].

We note that E[$w_{(M)}$] varies with different weight distributions. Fig 7.2(a) plots the PDFs of the log-normal distributions under different variances $v$. Given the same mean, $m$, we can see that the log-normal distribution becomes increasingly dispersive, as $v$ increases. According to (7.27) and (7.30), Fig 7.2(b) plots E[$w_{(M)}$] for the log-normal distribution with the growth of $M$. We see that, as the dispersion of the distribution increases, E[$w_{(M)}$] increases accordingly. Considering Design 1 and 2, we can conclude that: (1) for Design 1, as E[$w_{(M)}$] gets larger, $\max_i\{r_{w_i}\}$ becomes larger and the threshold of $\tau$ that inhibits outbreaks becomes larger. Therefore, the more dispersive the distribution is, the more resistent the network is against outbreaks in Design 1; and (2) for Design 2, as E[$w_{(M)}$] gets larger, $\max_i\{r_{w_i}\}$ becomes smaller and the threshold of $\tau$ that inhibits outbreaks occur decreases. Therefore, the more dispersive the distribution is, the less resistent the network is against outbreaks in Design 2.

In addition to the reliability threshold $\tau^*$, another evaluation of the reliability of adaptive weighted networks against cascading failures is the steady-state density of unreliable nodes and the spreading speed of the infection/failures at an outbreak equilibrium of the adaptive weighted network. At a disease-free equilibrium point, $(\frac{d[II]_w}{dt}, \frac{d[SI]_w}{dt}) = (0,0)$, and $[I] = 0$, and the cascading failure is inhibited. The population of infected nodes becomes zero. The entire population of nodes is healthy.

At an outbreak equilibrium, the average populations of susceptible (or healthy) and infected nodes stop changing over time, i.e., $\frac{d[I]}{dt}=0$, $\frac{d[S]}{dt}=0$, and $[I] > 0$. As a result, the average number of links connecting different types of nodes, i.e., infected and susceptible nodes, stabilizes. $\frac{d[II]}{dt}=0$, $\frac{d[SI]}{dt}=0$, and $\frac{d[SS]}{dt}=0$. The populations of infected and healthy nodes are non-zero. By substituting these steady-state conditions into (7.8), we have

$$\gamma[I]-\sum_{w}\beta_{w}[SI]_{w}=0, \tag{7.33a}$$

$$-\gamma[I]+\sum_{w}\beta_{w}[SI]_{w}=0, \tag{7.33b}$$

$$(\gamma+r_{w})[SI]_{w}-\xi\frac{[SI]_{w}}{[S]}\sum_{w'}\beta_{w'}[SI]_{w'}=0, \tag{7.33c}$$

$$\gamma[II]_{w}-\beta_{w}[SI]_{w}-\xi\frac{[SI]_{w}}{[S]}\sum_{w'}\beta_{w'}[SI]_{w'}=0, \tag{7.33d}$$

$$-(\gamma+\beta_{w}+r_{w})[SI]_{w}+\gamma[II]_{w}+\xi\frac{[SS]_{w}\text{-}[SI]_{w}}{[S]}\sum_{w'}\beta_{w'}[SI]_{w'}=0. \tag{7.33e}$$

By rewriting (7.33a) as $\gamma[I]=\sum_{w}\beta_{w}[SI]_{w}$, and then substituting into (7.33c) and (7.33d), we obtain

$$(\gamma+r_{w})[SI]_{w}-\xi\frac{[SS]_{w}}{[S]}\gamma[I]=0, \tag{7.34a}$$

$$\gamma[II]_{w}-\beta_{w}[SI]_{w}-\xi\frac{[SI]_{w}}{[S]}\gamma[I]=0. \tag{7.34b}$$

Since $[AB]=\sum_{w}[AB]_{w}$, $\forall A,B\in\{S,I\}$, we can rewrite (7.34) as

$$\gamma[SI]+\sum_{w}r_{w}[SI]_{w}-\xi\frac{[SS]}{[S]}\gamma[I]=0, \tag{7.35a}$$

$$\gamma[II]-\sum_{w}\beta_{w}[SI]_{w}-\xi\frac{[SI]}{[S]}\gamma[I]=0. \tag{7.35b}$$

By substituting $\beta_{w}=\tau w$, $r_{w_i}=\alpha_2(1-\frac{w_i}{\max\{w_i\}})$ and $\sum_{w}\beta_{w}[SI]_{w}=\gamma[I]$ into (7.35), we can obtain

$$\gamma[SI]+\alpha_2[SI]-\frac{\gamma\alpha_2}{\max\{w\}\tau}[I]-\xi\frac{[SS]}{[S]}\gamma[I]=0, \tag{7.36a}$$

$$\gamma[II]-\gamma[I]-\xi\frac{[SI]}{[S]}\gamma[I]=0, \tag{7.36b}$$

which can be rearranged to provide the steady-state degrees of infected and healthy nodes, as given by

$$[SS]=\frac{(\alpha_2+\gamma)[SI]-\frac{\gamma\alpha_2}{\max\{w\}\tau}[I]}{\gamma([II]-[I])}[SI], \tag{7.37a}$$

$$[II]=[I]+\xi\frac{[SI]}{[S]}[I]. \tag{7.37b}$$

By substituting (7.37) and $[S]+[I]=N$ into $2[SI]+[II]+[SS]=kN$, we can obtain

$$2[SI]+\frac{(\alpha_2+\gamma)[SI]-\frac{\gamma\alpha_2}{\max\{w\}\tau}[I]}{\gamma([II]-[I])}[SI]+[I]+\xi\frac{[SI]}{[S]}[I]=kN. \tag{7.38}$$

By substituting (7.37b), then (7.38) can be rewritten as

$$2[SI]+\frac{(\alpha_2+\gamma)[SI]-\frac{\gamma\alpha_2}{\max\{w\}\tau}[I]}{\gamma\xi[I]}(N-[I])+[I]+\xi\frac{[SI]}{N-[I]}[I]=kN. \tag{7.39}$$

Together with constraints $[I] < N$ and $[SI]+[II] \le kN$, (7.39) provides the sufficient condition of the mass of infections/failures at an outbreak equilibrium of the networks. $\frac{[SI]+[II]}{[I]}$ can be accordingly evaluated from (7.39) to show the degree of infected/failed nodes at the equilibrium, indicating changes in the topology of adaptive weighted networks in response to virus spread and cascading failures, as well as the effect of rewiring of weighted links.

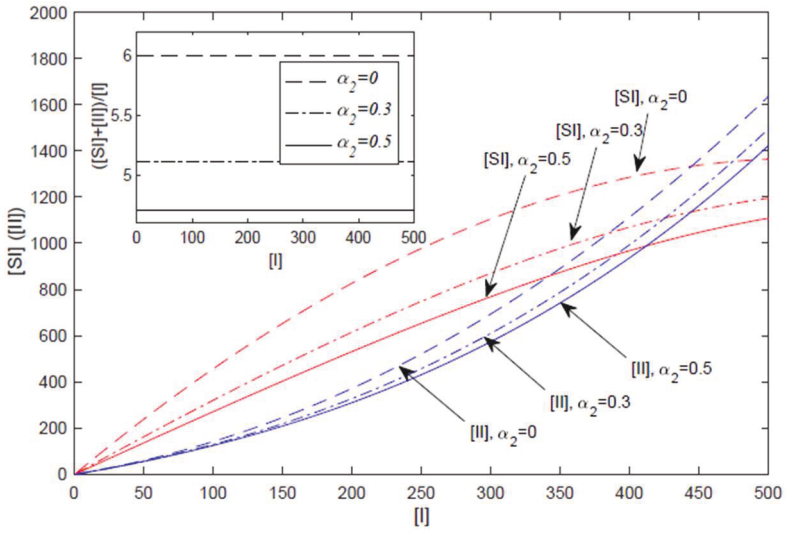



Fig 7.3. The relations between [I] and [SI] (and [II]). Plotted are the numbers of [SI] (red) and [II] (blue) with respect to [I], under different values of $\alpha_2$. $N = 1000$, $k = 6$, $\tau = 0.1$, $\gamma = 0.5$, and $\max\{w\} = 10$.

Fig 7.3 plots $[SI]$ and $[II]$ with the growth of $[I]$, where different values are tested for $\alpha_2$. We can see that [SI] exhibits concavity with regards to $[I]$, while $[II]$ exhibits convexity. In other words, the adaptive rewiring can increasingly isolate infected/failed nodes by breaking the links which can potentially infect healthy nodes. As a result, infections/failures become increasingly concentrated within the small set of infected/failed nodes. We also see that the average degree of infected/failed nodes remains consistent, as the growth of $[I]$ in an outbreak equilibrium, but the average degree does decrease with the growth of $\alpha_2$. Moreover, the average degree of infected/failed nodes is lower than the average degree of all nodes, indicating the infected/failed nodes are less connected and are prone to be separated from other nodes.

## 7.6 Numerical and Simulation Results

In this section, numerical and simulation results are provided to validate our proposed model and stability analysis. Figures are plotted based on discrete-time Monte-Carlo simulations of 100 iterations. Therefore, each data point in the figures is the average result of 100 independent runs. For each of the runs, a single infected node is randomly chosen at $t = 0$, as the initial point of infection.

As discussed in Section VI, the rewiring process is intimately associated with the closeness between nodes. Here we analyze two different linear designs of $r_{w_i}$ : namely Design 1 with $r_{w_i} = \alpha_1 w_i, \alpha_1 \ge 0$;

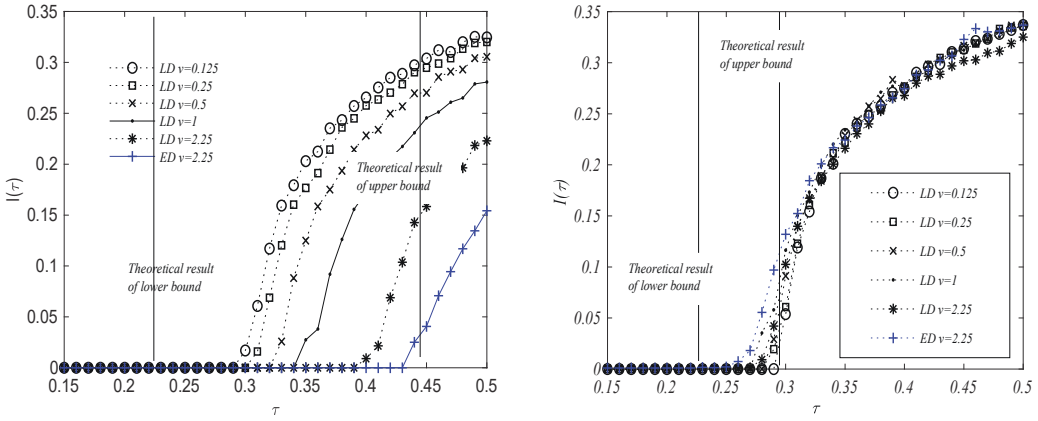
and Design 2 with $r_{w_i}=\alpha_2(1-\frac{w_i}{\max\{w_i\}}),\alpha_2\geq 0$. In the simulations, only one of the designs is taken across the network. Two distributions of the link weights $w_i$ are compared: the exponential distribution and the log-normal distribution. For comparison fairness, the means of the exponential and the log-normal distributions are both set to be $1/\lambda=m=0.5$ (so that the average value of $r_{w_i}$ is identical in both designs), and their variances are both set to be 2.25 (by configuring $m=e^{\mu+\frac{\sigma^2}{2}}$ and $v=e^{(2\mu+\sigma^2)(e^{\sigma^2}-1)}$ for the log-normal distribution; see (7.28) in Section 7.5). $\alpha_1$ and $\alpha_2$ are preconfigurable coefficients. We set $\alpha_1$=0.2 and $\alpha_2$=0.3326 to ensure the average value of $r_{w_i}$ is identical in both designs. In addition, we also plot the curves where the variance of the log-normal distribution is $v$=0.125, 0.25, 0.5 and 1 to show the impact of the variance on the propagation of cascading failure or virus spread. The simulations are carried out in an ER random network [25] of 1000 nodes connected by randomly generated 1997 links, where the weights of the links follow the exponential or log-normal distributions, ED and LD, as discussed in Section 7.5, respectively. For fair comparison, the distributions of **W** have identical mean $\langle w\rangle$. Other properties of the random network are summarized in Table 7.1.

Table 7.1 basic properties of the random network with two exemplary distributions of link weights.

| *Distribution* | $v$ | $k$ | $\langle w\rangle$ | $w_{\min}$ | $w_{\max}$ |
|---|---|---|---|---|---|
| *Log-normal(LD)* | 0.125 | 3.994 | 1.5 | 0.6757 | 3.0985 |
| | 0.25 | 3.994 | 1.5 | 0.4082 | 4.4738 |
| | 0.5 | 3.994 | 1.5 | 0.2035 | 7.6161 |
| | 1 | 3.994 | 1.5 | 0.1579 | 10.2048 |
| | 2.25 | 3.994 | 1.5 | 0.0626 | 15.0701 |
| *Exponential(ED)* | 2.25 | 3.994 | 1.5 | 0.00085 | 15.296 |

Fig 7.4 plots the percentile of unreliable nodes $I$ with the growth of $\tau$ in the steady-state network, where both the two rewiring strategies are presented. We can see that $\tau^*$ increases in Design 1 as the dispersion of the link weights increases; see Fig 7.4(a), and in Design 2, $\tau^*$ increases as the dispersion of the weights decreases; see Fig 7.4(b). Moreover, the simulation results are consistent with the analysis in Section VI. Our analysis is validated with accuracy. As the weight distribution becomes more dispersive, the maximum value of the link weights in the network becomes increasingly larger, and the minimum value of the weights becomes smaller. To this end, the preferential disconnections of the links with frequently communicated neighbors are likely to take place on the SI links with large weights, with the growth of the diversity of the weights. This can defer the outbreak of cascading failures or virus spread as the diversity of the weights grows in Design 1. On the contrary, in Design 2, the preferential disconnections of links with infrequently communicated neighbors are likely to take place on links with small weights, with the growth of the diversity of the weights. This can defer the outbreak of cascading failures or virus spread as the diversity of the weights decreases in Design 2. Furthermore, based on (7.22), the network can eventually become reliable if $\tau$ < 0.223 in Designs 1 and

2, consistent with the analytical results of the lower bound as shown in Fig 6.4. Based on (7.23), the outbreaks occur if $\tau > 0.445$ in Design 1 and $\tau > 0.297$ in Design 2, consistent with our analytical results of the upper bound as shown in Fig 7.4. In both designs, the network is reliable if τ is smaller than the analytical lower bound, and the network is unreliable if $\tau$ is larger than the analytical upper bound.




(a) $r_{w_i} = \alpha_1 w_i$ (b) $r_{w_i} = \alpha_2 (1 - \frac{w_i}{\max\{w_i\}})$

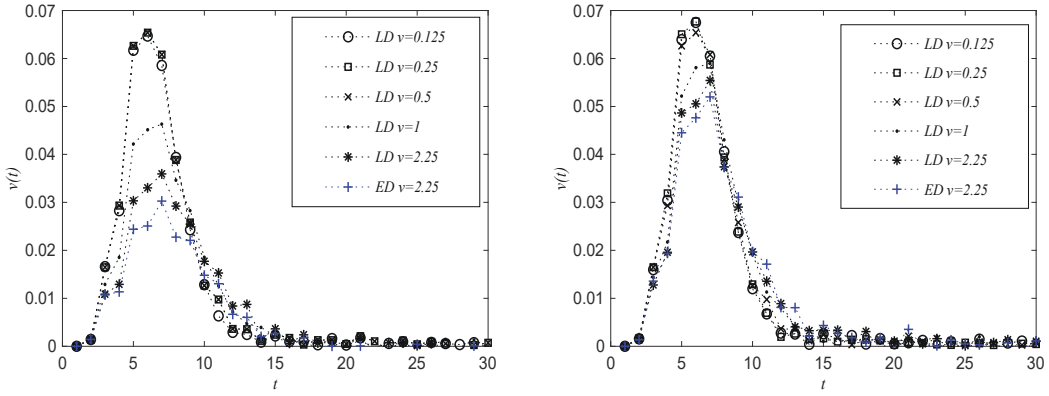
Fig 7.4 The steady-state density of unreliable nodes *I* as a function of $\tau$ under non-uniform rewiring rate, where

(a) $r_{w_i} = \alpha_1 w_i$ , (b) $r_{w_i} = \alpha_2 (1 - \frac{w_i}{\max\{w_i\}})$ with $\alpha_1 = 0.2$ and $\alpha_2 = 0.3326$.

We note that our reliability analysis provides an upper bound for $\tau$ in reliable states (denoted by $\tau_u^*$) and a lower bound in unreliable states (denoted by $\tau_l^*$). Under the condition of $\tau < \tau_l^*$, the cascading failures can be eventually inhibited and the network is reliable; and under the condition of $\tau > \tau_u^*$, the cascading failures would proliferate and the network is deemed to be unreliable. These conditions are the sufficient conditions of the network reliability and unreliability, and may not be the necessary conditions. Confirmed by extensive simulations, we demonstrate that these sufficient conditions are effective, even though they can be loose in some circumstances, as shown for Design 1. In other circumstances, the sufficient conditions can be very tight, as shown for Design 2.

An interesting finding is that our designs can have a strong impact on the steady-state density of unreliable nodes in the network. In Fig 7.4(a), we see that the steady-state density of unreliable nodes decreases, as the dispersion of the link weights increases. In the case that the SI links with large weights are disconnected preferentially, only the SI links with small weights are left intact, leading to the reduction of the average transmission rate of the network. As a result, the steady-state density of unreliable nodes declines in Design 1. In contrast, in the case that the SI links with small weights are disconnected preferentially, the SI links with large weights are left intact and this can increase the rate of turning reliable nodes to be unreliable. In other words, the steady-state density of unreliable nodes increases, as the dispersion of the weights grows. Finally, by assessing Fig 7.4, we can notice that, given the same mean and variance, an exponential distribution of the links weights is preferred over the log-normal distribution in regards of network reliability under Design 1; and the other way around under

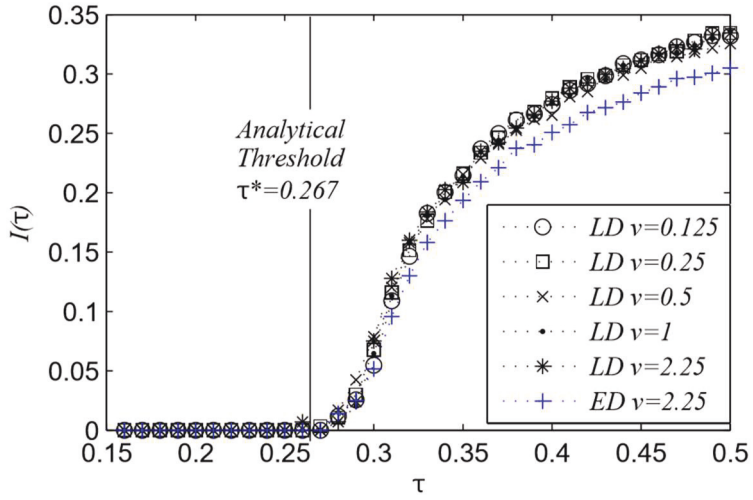
Design 2.



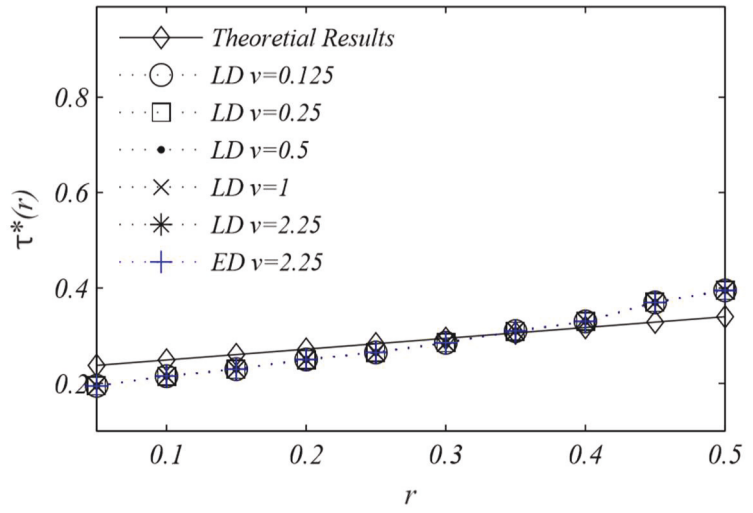

(a) $r_{w_i}=\alpha_1 w_i$ (b) $r_{w_i}=\alpha_2(1-\frac{w_i}{\max\{w_i\}})$

Fig 7.5. The spreading velocity of infection $v(t)$ at each time slot t under the two rewiring designs, where (a) Design 1: $r_{w_i}=\alpha_1 w_i$, (b) Design 2: $r_{w_i}=\alpha_2(1-\frac{w_i}{\max\{w_i\}})$, with $\alpha_1 = 0.2$, $\alpha_2 = 0.3326$ and $\tau = 0.5$.

The spreading velocity of the virus or failure is an important measure of the designs. We define the spreading velocity as the difference of infection density between consecutive time slots, denoted by $v(t)=i(t)-i(t-1)$. Fig 7.5 plots the spreading velocity under two designs, as the time elapses. We can see that in both designs, the more dispersive the distribution of the link weights is, the lower the velocity peak is. This indicates larger dispersion of the link weights can result in slower spreading. In Design 1, the disconnections of SI links with large weights remove the fast propagation paths of virus or failures, hence slowing down the propagation of virus or failures. And in Design 2, although the SI links with small weights are disconnected preferentially, the density of those links decreases as the dispersion of the link weights grows. As a result, large dispersion of the weights can also reduce the spreading velocity of virus or failures in Design 2. In the special case where $r_{w_i}=r$, we can calculate the accurate value of $\tau$ by using (7.24). Fig 7.6 shows the impact of the rewiring process and the weight distribution on the reliability threshold $\tau^*$ in the special case. The figure confirms the validity of the analytic results of $\tau^*$ from (7.24) by comparing with Monte-Carlo simulations. With the identical value of $\langle w\rangle$, we can see that the distribution of the weight $w_i$ has little impact on $\tau^*$, and hence validates our analysis. We also see that $\tau^*$ increases with the growth of the rewiring rate $r$. That is because, as the rewiring rate $r$ grows, the SI links can be increasingly likely to be disconnected. This leads to the reduction of the transmission paths. Therefore, the interruption of infection by the rewiring process can make the transmission increasingly difficult, or in other words, inhibits the transmission.




(a) The steady-state density of infected nodes $I$ as a function of $\tau$ in random networks, where $r = 0.2$, $\gamma = 1$.




(b) The reliability threshold $\tau^*(r)$ as a function of rewiring rate r in random networks.

Fig 7.6. The special case of uniform rewiring rate, where the theoretical results of reliability threshold $\tau^*$ are given by (6.24).

In practice, networks can display a small-world effect [26] and a scale-free property [27]. These networks are particularly relevant to NFV. As a matter of fact, these networks have been widely used to portray actual virtual network characteristics. It has also been proposed to construct virtual networks to comply with scale-free or small-world models, in attempts to reduce network average path length and to simplify NFV [170]-[172]. In this sense, our simulation settings align with the virtual network characteristics. We proceed to carry out Monte-Carlo simulations on weighted networks with small-world effect and scale-free property, respectively. The properties of the two types of networks with 500 nodes connected by 1500 links are summarized in Table 7.2. Fig 7.7 shows the density of infected nodes $i(t)$ under the types of two sets of networks. It is clear that, in both WS small-world and BA scale-free networks [26], [27], increasing the dispersion of the link weights can lead to a decline of the infected

population in the steady-state network in Design 1, while decreasing the dispersion can do so in Design 2.

Table 7.2 Basic properties of WS network and BA network.

| Network | Distribution | $v$ | $k$ | $\langle w \rangle$ | $w_{\min}$ | $w_{\max}$ |
|---|---|---|---|---|---|---|
| WS Network | Log-normal | 0.125 | 6 | 1.5 | 0.6365 | 3.067 |
| | | 0.25 | 6 | 1.5 | 0.4512 | 4.4048 |
| | | 0.5 | 6 | 1.5 | 0.2993 | 6.5457 |
| | | 1 | 6 | 1.5 | 0.1915 | 8.4185 |
| | | 2.25 | 6 | 1.5 | 0.0863 | 17.1523 |
| | Exponential | 2.25 | 6 | 1.5 | 0.00081 | 10.5253 |
| BA Network | Log-normal | 0.125 | 6 | 1.5 | 0.7015 | 3.1527 |
| | | 0.25 | 6 | 1.5 | 0.5291 | 4.4506 |
| | | 0.5 | 6 | 1.5 | 0.2548 | 6.4403 |
| | | 1 | 6 | 1.5 | 0.1284 | 8.4293 |
| | | 2.25 | 6 | 1.5 | 0.0694 | 15.2284 |
| | Exponential | 2.25 | 6 | 1.5 | 0.0005 | 11.2972 |

In general, our simulation results show that the reliability threshold $\tau^*$ depends on the distribution of the link weights and the specific rewiring strategy in the adaptive weighted networks. Preferentially disconnecting links to unreliable neighbors can effectively inhibit the spread of virus or failures, e.g., by increasing the reliability threshold, and reducing the steady-state population of unreliable nodes, and the spreading velocity of instability. The conclusion drawn is that the larger the dispersion of the link weights is, the more effectively the instability can be prevented from proliferation. On the other hand, preferential disconnections of the links with small weights can inhibit the spread as the dispersion of the weights decreases, e.g., increasing the reliability threshold and reducing the steady-state population of unreliable nodes. Unexpectedly, the dispersion of the link weights slows down the spread velocity as the links with small weights are preferentially disconnected.

## 7.7 Conclusion

In this chapter, a mean-field approximated dynamic system is proposed to model the time-varying populations of failed nodes and risky links in adaptive weighted networks. A linear stability analysis was conducted upon the dynamic system, and the threshold was identified for the network to inhibit failures and remain reliable in the steady state. Validated by simulations, our analysis revealed that the threshold depends on both the distribution of the link weights and the adopted rewiring strategy. It is also shown that preferentially disconnecting frequently communicated, suspicious peers can effectively inhibit failures and virus spread. As cascading failures, DDoS, computer virus and malware, can be potentially analyzed by using our analysis which is generic with an emphasis on theoretical insights and understanding. The presented analysis is not closely coupled with real behaviors of specific vulnerability exploration of particular attacks and viruses though. In the future, we will take the anatomy of different attacks into account and evaluate network reliability under specific types of attacks.

# Chapter 8

# Study on the Dynamical Rewiring in Adaptive Weighted Heterogeneous Networks

## 8.1 Introduction

Virus spreading on networks is one of the main issues in complex network research. Both of the biological virus and computer virus spreading have attracted many studies through using the classic infectious disease models [6], [30], [31]. These studies have made it clear that certain topological properties have a strong impact on virus spreading. In fact, the dynamic process in turn will affect the network structure. For example, humans tend to respond to the emergence of an epidemic by avoiding contacts with infected individuals in a social network. In this way, a feedback loop between the state and topology of the network begins to take shape and networks that exhibit such an interplay between the dynamics and topology are called adaptive networks.

Based on the SIS epidemic model, the adaptive network was first proposed by Gross et al [36]. In this model, the network structure changed all the time according to the states of neighboring nodes and this in turn affected epidemic dynamics. An extensive research on the coupling between epidemic processes and the underlying network topology has been carried out after the seminal-work of Gross et al. Shaw et al. studied epidemic dynamics on an adaptive network based on the Susceptible–Infected–Recovered–Susceptible (SIRS) model [37]. Vincent et al. studied the interplay and outcomes of disease and topology on adaptive networks with various initial configurations by introducing an improved compartmental formalism [141]. Ilker et al. studied the spread of an epidemic on an adaptive network with community structure [90]. Based on the dynamic interaction, many effective epidemic spreading control strategies were proposed in adaptive networks. Shaw et al. studied vaccine control for disease spread on an adaptive network and shown that vaccine control is much more effective in adaptive networks than in static networks due to feedback interaction between the adaptive network rewiring and the vaccine application [93]. Yang et al. paid attention to the emergence of community structure in the transient process and the effects of community-based control strategies on epidemic spreading [91]. Song et al. introduced a new preferentially reconnecting edge strategy of adaptive networks depending on spatial distance [92].

For simplicity, most of the existing studies on adaptive networks ignore the link weights between nodes and assume that the weights are uniform. However, the relationship between individuals in real networks, e.g., the distance between the two cities in traffic networks, the intimacy between individuals in social networks and the communication radius in communication networks, are significantly different. Therefore, the various relationships between individuals are closely related to the virus spreading behaviors in many real networks. Considering the fact that the contact strengths among individuals are diverse, novel weighted adaptive network models were developed recently [151], [152]. Many interesting results were found in the study of virus spreading on weighted adaptive networks. Zhou et al. found that this weight adaption process could significantly aggravate the prevalence of an epidemic and examined

the effectiveness of the link-removal strategy with their model, and the results shown that the weight adaption process may weaken the efficiency of the strategy [151]. Chao et al. demonstrate that the rewiring strategy has a close relationship with the epidemic spreading, and this strategy cannot always suppress the disease, which is different from some previous studies [152]. Yun et al. considered epidemic spreading on a weighted adaptive network in which the network topology varies according to the global and local infective information of individuals [173]. It was found that greater interacting strength could effectively inhibit virus spreading.

As those weighted adaptive network models have greatly enriched the existing real network models, many challenges arise and need to be solved. Existing studies have shown that a lot of real networks are of heterogeneous weight distribution, such as friendship networks, scientists collaboration networks, technical networks, which can strongly affect the epidemic spreading in static networks [153]-[156]. Therefore, the first challenge is the influence of weight distribution on epidemic spreading in a weighted adaptive network. The second challenge is the rewiring process. Due to the individual difference, e.g., relationships, individuals' knowledge level on the disease, the adaptability of individuals dynamically in response to the epidemics should be different from each other. Correspondingly, in our model, the individual behaviors are heterogeneous according to the link weights. In addition, the network structures are always dynamically evolving, how to design more feasible and reasonable rewiring strategies to inhibit epidemics is also a very meaningful challenge.

Considering the diversity of interpersonal relationships and the ability to adapt the network topology dynamically in response to the dynamic state of nodes, a novel SIS model on a weighted adaptive heterogeneous network is proposed in this chapter. In our model, the epidemic dynamics on weighted adaptive heterogeneous networks is studied and the influence of network structure on epidemics is discussed. Furthermore, we design effective rewiring strategies based on the individual behaviors for the inhibition of epidemics.

The main contributions of this study include three aspects. First, we propose a novel SIS model on a weighted adaptive heterogeneous network, where network structure, weights and virus spreading behavior are dynamically interacted. Second, individuals' behaviors and interpersonal relation have significant influence on epidemic spreading. Our analysis shows that larger dispersion of weight of initial networks leads to slower spreading in a weighted adaptive network while the adaptive rewiring process during the dynamics can also effectively inhibit epidemic spreading. Third, based on individuals' behaviors, the rewiring strategies dynamically related to the real-time link weights are proposed for the inhibition of epidemics and the results show the effectiveness of the rewiring strategies with our models.

The rest of this chapter is organized as follows. In Section 8.2, we introduce our model in detail, including nodes states dynamics, network topology dynamics and present the mathematical description. Simulation results are given in Section 8.3. In Section 8.4, the influence of individual spontaneous behaviors on epidemics are analyzed, rewiring strategies based on spontaneous behaviors are presented. Finally, we conclude this chapter in Section 8.5.

## 8.2 Adaptive Weighted Heterogeneous Network Model

Consider a generic network of $N$ nodes connected by $L$ weighted bidirectional links. We can separate the whole dynamic process into two sub-processes: 1) the dynamics on networks and 2) the dynamics of networks. The first process specifies epidemic spreading. The nodes are either susceptible (S) or infected (I). In every time step and for every link connecting an I node with a S one (SI-link), the susceptible becomes infected with the probability $\beta$ . The infected recovers from the virus with probability $\mu$ , becoming susceptible again. $\tau = \beta / \mu$ is defined as the effective infection rate. In weighted networks, we assume that the probability of a susceptible node getting infected through an SI-link with weight $w$ is $\beta_w \sim w\beta$. When there are more than one infected neighbor, the probability for node $i$ being infected is $\beta_i(t) = 1 - \prod_{j \in Nb(i)} (1 - \beta_j(t))$, $Nb(i)$ respects the number of infected neighbors of node $i$.

The second dynamics specifies rewiring process. When epidemics outbreaks, awareness of the presence of the virus will prompt individuals to change their behavior to avoid being infected. Correspondingly, in our model, we allow susceptible individuals to protect themselves by rewiring their links. With probability $r_w$, one susceptible node breaks the link to infected and forms a new link randomly select susceptible. Double connections and self-connections are not allowed to form in this way. After the rewiring, the weight of the broken link is transferred to the new-formed link while the weights of the other links remain constant. During the rewiring process, the weights are transferring locally due to the disconnection and reconnection of the links.

Many real networks are not homogeneous networks, instead, their connectivity is heterogeneous. Both of their node-degree distributions and the link-weight distributions have a power-law form, and are independent of the connectivity scale. An example of adaptive weighted heterogeneous networks is NFV on cloud computing platforms. The VMs are connected through virtual links, the link numbers of each VM and the weights of links are different due to the different positions and workloads of services. When some VMs are congested due to DDoS attacks or infected due to computer viruses, new virtual links can be established to bypass these VMs. The weights (or the workloads) of the old virtual links to the bypassed VMs can be transferred to the new links. To achieve this, the VMs that are neither attacked or infected can check their routing tables, and spontaneously decide and activate the new virtual links. Meanwhile, the establishment of a new link (or the vanish of an existed link) brings a total increase (or decrease) of traffic and the weight redistribution.

We apply the pair-based mean-field approach to describe the changes of the number of nodes and edges in different states in the adaptive weighted heterogeneous network.

### 8.2.1 The Dynamics on Network

$$\frac{d[S_k]}{dt} = \mu[I_k] - \sum_w \beta_w [S_k I]_w \quad ; \tag{8.1a}$$

$$\frac{d[I_k]}{dt} = -\mu[I_k] + \sum_w \beta_w [S_k I]_w ; \tag{8.1b}$$

$$\frac{d[S_k S_{k'}]_w}{dt} = \mu([S_k I_{k'}]_w + [I_k S_{k'}]_w) - \sum_{w'} \beta_{w'} ([S_k S_{k'} I]_{ww'} + [I S_k S_{k'}]_{w'w}) ; \tag{8.1c}$$

$$\frac{d[I_k I_{k'}]_w}{dt} = \beta_w ([S_k I_{k'}]_w + [I_k S_{k'}]_w) - 2\mu[I_k I_{k'}]_w + \sum_{w'} \beta_{w'} ([I_k S_{k'} I]_{ww'} + [I S_k I_{k'}]_{w'w}) ; \tag{8.1d}$$

$$\frac{d[S_k I_{k'}]_w}{dt} = -\beta_w [S_k I_{k'}]_w - \mu[S_k I_{k'}]_w + \mu[I_k I_{k'}]_w + \sum_{w'} \beta_{w'} ([S_k S_{k'} I]_{ww'} - [I S_k S_{k'}]_{w'w}) ; \tag{8.1e}$$

$$\frac{d[I_k S_{k'}]_w}{dt} = -\beta_w [I_k S_{k'}]_w - \mu[I_k S_{k'}]_w + \mu[I_k I_{k'}]_w + \sum_{w'} \beta_{w'} ([I S_k S_{k'} I]_{w'w} - [I_k S_{k'} S]_{w'w}) . \tag{8.1f}$$

where $[A_k B]_w = \sum_{k'} [A_k B_{k'}]_w$, $[A_k B_{k'} C]_{ww'} = \sum_{k''} [A_k B_{k'} C_{k''}]_{ww'}$. $[A_k]$ ( $A \in \{S, I\}$ ) denotes the number of nodes with degree $k$ in state $A$, and $[S] = \sum_k [S_k]$, $[I] = \sum_k [I_k]$ and $[S]+[I] = N$. $[A_k B_{k'}]_w$ denotes the number of edges weighted $w$, connecting two nodes in states $A_k$ and $B_{k'}$. $[A_k B_{k'} C_{k''}]_{ww'}$ denotes the number of triplets $A_k$-$B_{k'}$-$C_{k''}$, with edge $A_k B_{k'}$ weighted $w$ and $B_{k'} C_{k''}$ weighted $w'$.

Eqs (8.1*a*)-(8.1*f*) capture the time-changing population of nodes in the healthy and infected state and links weighted by different weights and connecting nodes in different states due to the infection and recovery. In (8.1*a*) and (8.1*b*), the first term at the right-hand side (RHS) corresponds to the part of the population which recover from the infected state with the probability of $\mu$. The second term corresponds to the part of the population which become infected by their infected neighbors with the probability $\beta_w$. (8.1*c*) captures the changing population of the $w$-weighted links connecting two healthy nodes $S_k$ and $S_{k'}$. The first term at the RHS of (8.1*c*) is the increased part of the population, resulting from the recovery of the infected ends of the links with probability of $\mu$. The second term is the number of previous $w$-weighted $S_k S_{k'}$ links which become $S_k I_{k'}$ ( $I_k S_{k'}$ ) links due to the infection at one end of the links through a $w'$-weighted link with the probability of $\beta_{w'}$. Likewise, (8.1*d*) captures the changing population of $w$-weighted $I_k I_{k'}$ links connecting a pair of infected nodes, and (8.1*e*)-(8.1*f*) captures the changing population of $w$-weighted $S_k I_{k'}$ and $I_k S_{k'}$ links.

### 8.2.2 The Dynamics of Network

For the dynamics of networks, the evolution of weights is accompanied by the rewiring process. On the one hand, the weight of the former SI-link is transferred to the new SS-link, which can be called as weight replacement process. On the other hand, the weights of links connected to the nodes whose degrees changed are redistributed, which can be called weight redistribution process. The two processes can be described as follows:

$$\frac{d[S_k S_{k'}]_w}{dt} = r_w ([S_k I_{k'}]_w + [I_k S_{k'}]_w) + \Delta[S_k S_{k'}]_w ; \tag{8.2a}$$

$$\frac{d[S_k I_{k'}]_w}{dt} = -r_w[S_k I_{k'}]_w + \Delta[S_k I_{k'}]_w ; \tag{8.2b}$$

$$\frac{d[I_k S_{k'}]_w}{dt} = -r_w[I_k S_{k'}]_w + \Delta[I_k S_{k'}]_w . \tag{8.2c}$$

In Eqs 8.2, the first term of each equation at the RHS corresponds to the link rewiring process accompanying weight transferring, i.e., the exchange between $[S_k S_{k'}]_w$ and $[S_k I_{k'}]_w$ ($[I_k S_{k'}]_w$). The second term $\Delta[A_k B_{k'}]_w$ represents the exchanges of $w$-weighted $A_k B_{k'}$ links and $w'$-weighted $A_k B_{k'}$ links due to the weight redistribution, $A, B \in \{S, I\}$ and $w' \neq w$, for example, the exchange between $[S_k S_{k'}]_w$ and $[S_k S_{k'}]_{w'}$.

The rules of weight redistribution are as follows: When an edge linked to an existing vertex $i$ disappears, the local rearrangement of weights between $i$ and its neighbor $j$ according to the simple rule

$$w_{ij} \to w_{ij} - \Delta w_{ij}, \tag{8.3}$$

meanwhile, when a new edge links to vertex $i$, the local rearrangement of weights between $i$ and its neighbor $j$ according to the rule

$$w_{ij} \to w_{ij} + \Delta w_{ij}, \tag{8.4}$$

where

$$\Delta w_{ij} = \delta \frac{w_{ij}}{s_i}. \tag{8.5}$$

$s_i$ represents the strength of node $i$, expressed by $s_i = \sum_j w_{ij}$. The rules consider that the establishment of a new edge (or the vanish of an existed edge) of weight $w$ with the vertex $i$ induces a total increase (or decrease) of traffic $\delta$ that is proportionally distributed among the edges departing from the vertex according to their weights.

The rewiring process depends on the dimensionless parameter $\delta$, that is the fraction of weight which is caused by the vanish of existed links and creation of new links. The value of $\delta$ represents the intensity of the node's reaction to its edge increasing or decreasing. In the case of $\delta \geq 1$, we mimic situations in which an appreciable fraction of traffic generated by the connection newly added or disappeared will be dispatched in the already existing connections. For example, in the airport networks where the transit traffic is rather relevant in hubs. When $\delta < 1$, we face situations such as the scientific collaboration network where it is reasonable to consider that the birth of a new collaboration (co-authorship) or the disappear of an existed collaboration is not triggering a more intense activity on previous collaborations. After one rewiring process, the average weight of the network remains unchanged.

A simple network with 6 nodes and 4 edges marked $w_1$-$w_6$ as links' weights at time $t$ is shown in Fig 8.1. From time $t$ to $t$+1, some dynamics appear in the network: Node $F$ is infected by node $B$. To keep

healthy, node *A* escapes from node *B* with rate $r_w$, and then reconnects to node *E*. During the process, the weight (the workloads) of the link between *A* and *B* is transferred to the new link between *A* and *E*. Meanwhile, as the numbers of neighbors of both nodes *B* and *E* change, the traffic burden increases on *E* while decreases on node *B*. The weight redistribution occurs on the links of the nodes *B* and *E*.

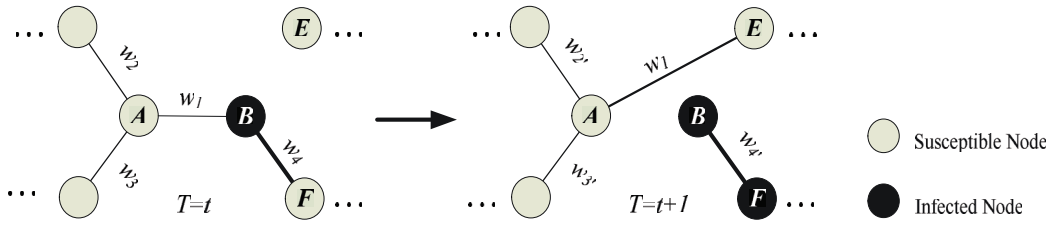



Fig 8.1. A simple example of dynamic processes in a weighted adaptive heterogeneous network.

## 8.3 Dynamical Rewiring Strategies

The rewiring process is composed of two parts: link-disconnection and link-reconnection. During the process, the weights of the links change. The adaptabilities of individuals dynamically in response to the epidemics are different from each other due to individual differences in social networks. For example, if one of your families gets flu, you will not disconnect the connection since you two contact each other very closely. By contraries, it is unlikely that you will keep the contact with the unfamiliar infected neighbor. Based on these dynamic changes of links and weights, we discuss the real situations and make some reasonable strategies to inhibit the virus spreading.

### 8.3.1 Link-disconnecting Strategies

The first design specifies the positive correlation between the rewiring rate and *w*, i.e., $r_w = a_1 w$, $a_1 > 0$. This is the case the adaptabilities of individuals dynamically in response to the epidemics are different from each other due to individual differences in social networks. For example, if one of your families gets flu, you will not disconnect the connection since you two contact each other very closely. By contraries, it is unlikely that you will keep the contact with the unfamiliar infected neighbor. This behavior of healthy individuals is defined as a ***Spontaneous Defense Behavior*** (SDB) here. Meanwhile, infected individuals try to avoid contacting with healthy ones, especially with the familiar individuals. This behavior leads to another kind of link-break rule: SI-link will be cut off easier when the individuals in both ends are more familiar, which is called ***Spontaneous Quarantine Behavior*** (SQB). Correspondingly, the rewiring rates are heterogeneous in our models. We assume that $r_{w_l} \sim \frac{1}{w_l}(w_l)$ as the probability of disconnecting the link with weight $w_l$, representing SDB (SQB).

### 8.3.2 Link-reconnecting Strategies

After the disconnecting, the susceptible node forms a new link to another susceptible one, which is generally chosen randomly in homogeneous networks. Since the nodes are divided into different categories based on their degree in heterogeneous networks, we propose a preferential link-reconnecting

strategy based on the degrees of nodes. After link-disconnecting process, the healthy node $i^p$ first choose a susceptible neighbor $i_N^p$ randomly, then preferentially choose $i_N^p$'s neighbor with largest degree to reconnect. Make sure it is not connected with $i^p$ in the current time. During the whole process, all we need to know is the local information of the neighbors of nodes $i^p$ and $i_N^p$, which reduce the computational complexity to the utmost and ensure the feasibility of our link-reconnecting strategy in large-scale networks.

### 8.4 Simulations

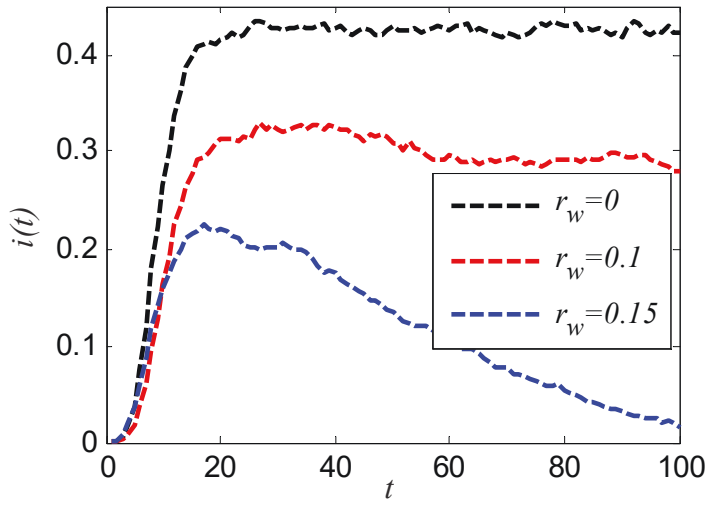
Based on the model we proposed and the rewiring strategies we design, we apply the Monte-Carlo simulation to explore the impact of dynamical changes of network structure, including links and weights, on epidemic spreading in this section. We investigate the impact of heterogeneity of weight distribution and the inhibitory effects of different rewiring strategies on epidemic spreading. Without specific statement, we use the density of infected nodes and the epidemic threshold to denote the epidemic spreading. $i(t)$ and $I$ are the density of the infection at time $t$ and in the steady state, respectively. The larger the $I/i(t)$ is, the severer the disease is. The epidemic threshold of $\tau$ is denoted by $\tau^*$. If $\tau < \tau^*$, the adaptive weighted network can eventually become healthy, i.e., all nodes eventually become healthy. Otherwise, the network is not healthy, i.e., the epidemic breaks out. Each data is obtained by averaging over 100 independent runs.

The weighted SF network model used in our simulations is one of the most well-known models introduced by Barrat, Barthelemy, and Vespignani (BBV networks) [95], whose degree, strength and weight distributions are power-law distributions with heavy tails. The basic properties of the artificial network models in our simulations are presented in Table 8.1. As $\delta$ grows, the dispersion of weights becomes larger.

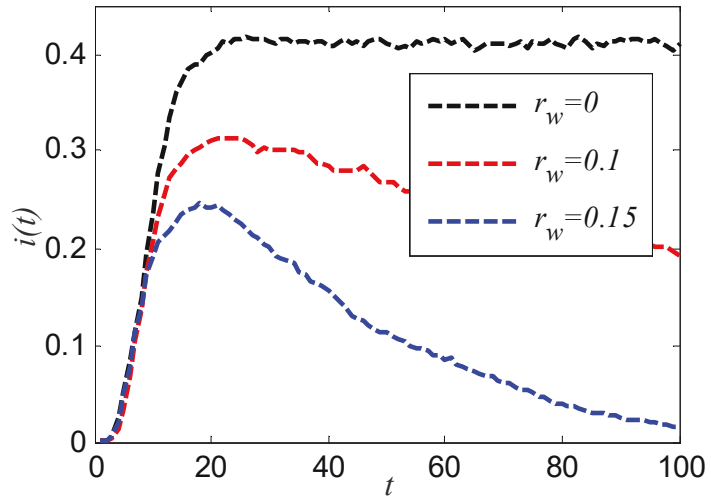
Table 8.1. Basic properties of BBV network models.

| Network | $\delta$ | $N$ | $M$ | $k$ | $\langle w \rangle$ | Max_$w$ |
|---|---|---|---|---|---|---|
| BBV Network1 | 0.1 | 1000 | 1997 | 3.994 | 1.4764 | 16.687 |
| BBV Network2 | 0.5 | 1000 | 1997 | 3.994 | 1.4764 | 22.567 |
| BBV Network3 | 2 | 1000 | 1997 | 3.994 | 1.4764 | 77.431 |

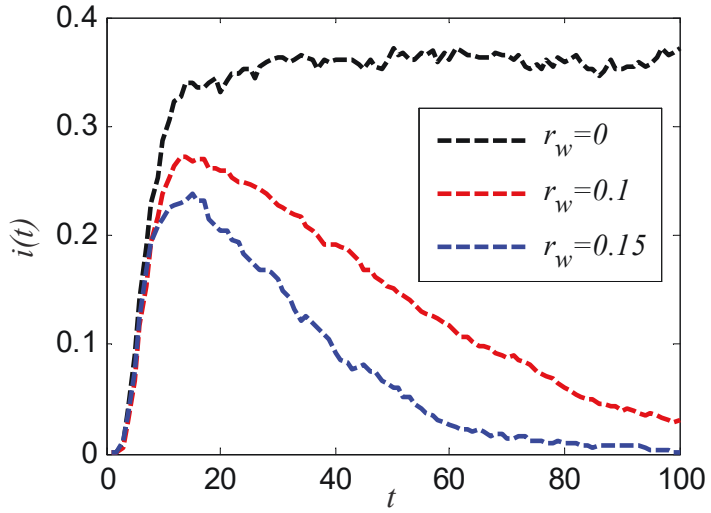
First, we study the impact of rewiring rate on virus spreading in BBV networks. We can see from Fig 8.2 that, as the rewiring rate increases, the epidemic size at time $t$ has been decreasing, which means the rewiring process can significantly inhibit the epidemic spreading. The rewiring process promotes the isolation of infected individuals, which can significantly inhibit the infection.




(a) BBV network1




(b) BBV network2

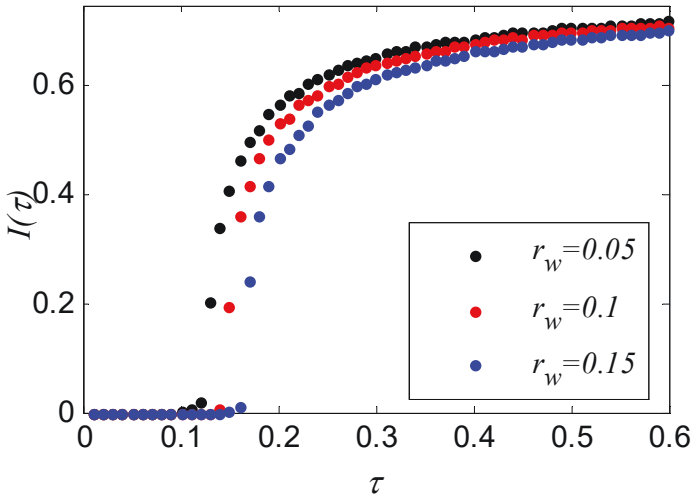



(c) BBV network3

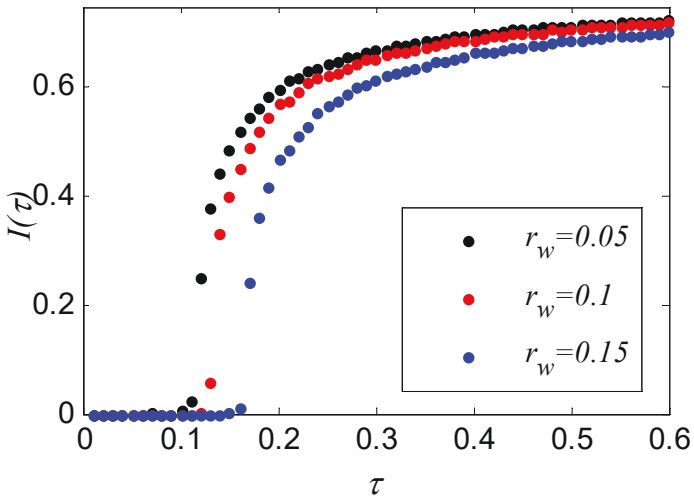
Fig 8.2. The evolution of the fraction of infected nodes $i(t)$ under different cases of $r_w$ in weighted adaptive networks.

As the probability of rewiring rate increases, the probability of isolation of the infected individual

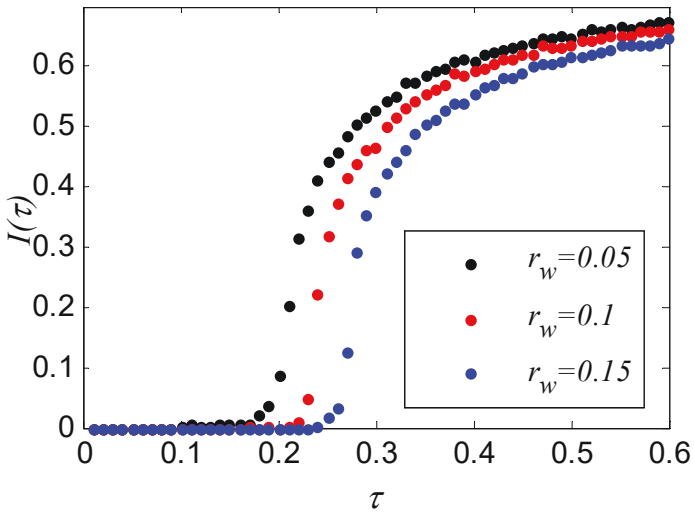
is also increasing. At the beginning of the infection, a small fraction of infected nodes is scattered in the network, who can be isolated more easily with a larger rewiring rate. Therefore, the outbreak of infection becomes harder as $r_w$ grows. Moreover, our simulation results are consistent with the analysis. The infection density at steady state under different cases of $r_w$ are shown in Fig 8.3. We can find that, as $r_w$ increases, $\tau^*$ becomes larger, which makes the infection less likely to erupt. Besides, in each network model, the final fraction of infection $I(\tau)$ becomes smaller. As a result, the results in Fig 8.2 verify our conclusions in Fig 8.3 and further prove that the increase of rewiring rate can significantly increase the epidemic threshold and reduce the infection density, i.e., inhibit the epidemic spreading.




(a) BBV network1




(b) BBV network2




(c) BBV network3

Fig 8.3. The infection density at steady state under different cases of $r_w$ in weighted adaptive heterogeneous networks.

It is shown in Table 8.1 that larger value of $\delta$ induces larger dispersion of weight of networks. Then an important question is that how the value of $\delta$ impacts epidemic spreading. We notice that by comparing the results from the three networks in Fig 8.2, the final size under the same rewiring rate decreases as the dispersion of weights becomes larger. To further confirm this conclusion, we calculate the infection scale in the BBV networks under different rewiring rates. Fig 8.4 plots the percentile of infected nodes $I$ with the growth of $r_w$ in BBV network models. We can see that the final infection decreases as the dispersion of weights grows.

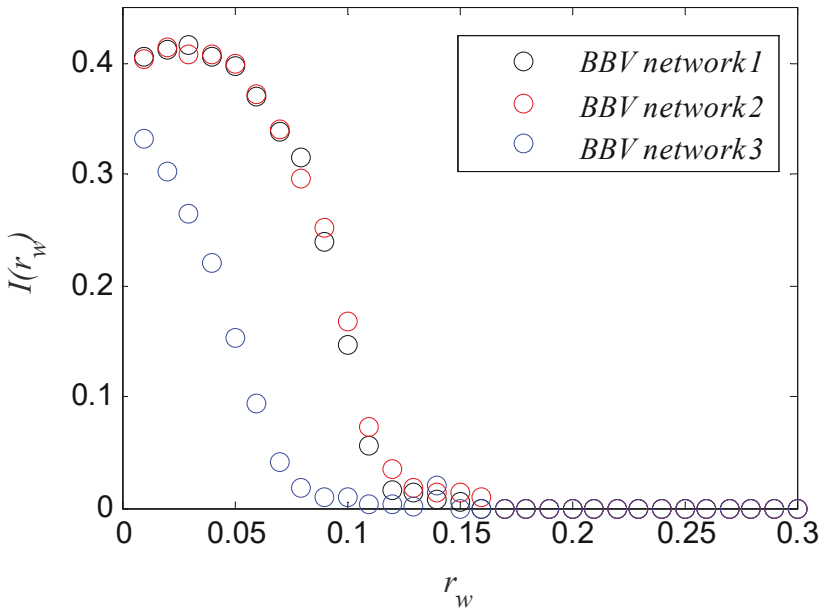



Fig 8.4. The final fraction of infected nodes $I$ as a function of rewiring rate $r_w$ in BBV network models.

The simulation results show that the weight distribution and rewiring process have a very important impact on virus spreading. In fact, the rewiring process is closely related to the weights as we discussed

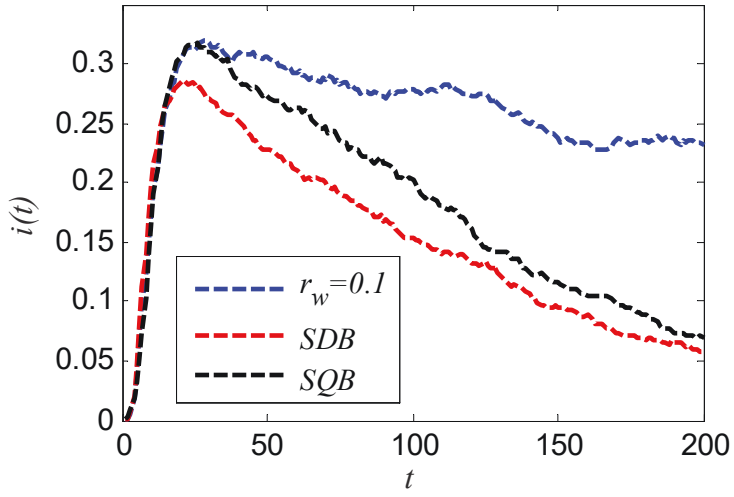
above. Based on the link-disconnecting and link-reconnecting process, we propose strategies of adjusting the individual behaviors to dynamically respond to the epidemics. In the following part, we study the impact of the rewiring strategies on epidemic spreading by taking different but simple designs of rewiring process.

- **Link-disconnecting Strategy**

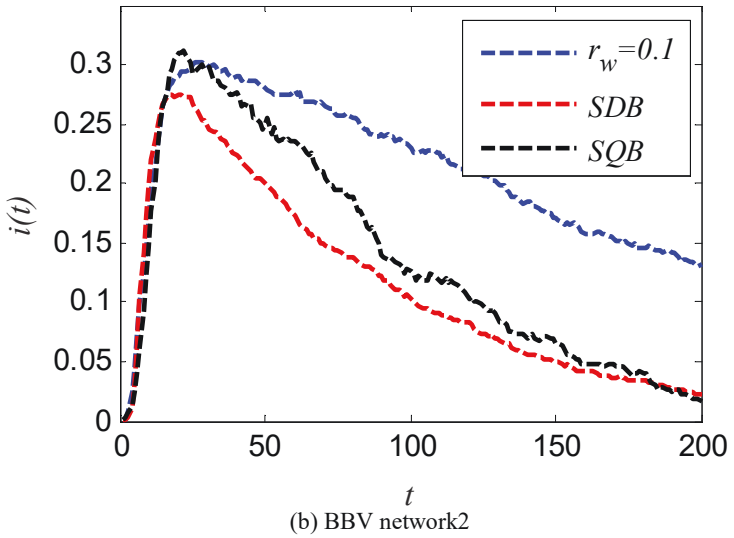
Since the weights between individuals are different, the adaptability of individuals should be different from each other. SDB and SQB in Section 8.3 are defined to describe the disconnecting behaviors of healthy and infected individuals respectively. For healthy individuals, it is easier to cut off the links with less familiar infected neighbors, while the infected individuals try to avoid contacting with closer healthy individuals. Therefore, we design two disconnecting strategies based on the weights to represent SDB (SQB).

The first design (SDB) is specifies the negative correlation between the rewiring rate and $w$, i.e., $r_w = \frac{a_1}{w}$, $a_1 > 0$. This is the case where a healthy node preferentially breaks its unfamiliar links to protect itself. The second design (SQB) specifies the positive correlation between the rewiring rate and $w$, i.e., $r_w = a_2 \frac{w}{w_m}$, $a_2 > 0$, $w_m$ is the maximum weight value of the networks. This is the case where an infected node preferentially breaks its high-weight links with frequently interacted neighbors. Without loss of generality, $a_1$ and $a_2$ are preconfigurable coefficients. In our simulation, the average value of rewiring rate $\langle r_w \rangle$ is identical in both designs, $\langle r_w \rangle = 0.1$, $a_1 = 0.14764$ and $a_2 = 5.2446$.

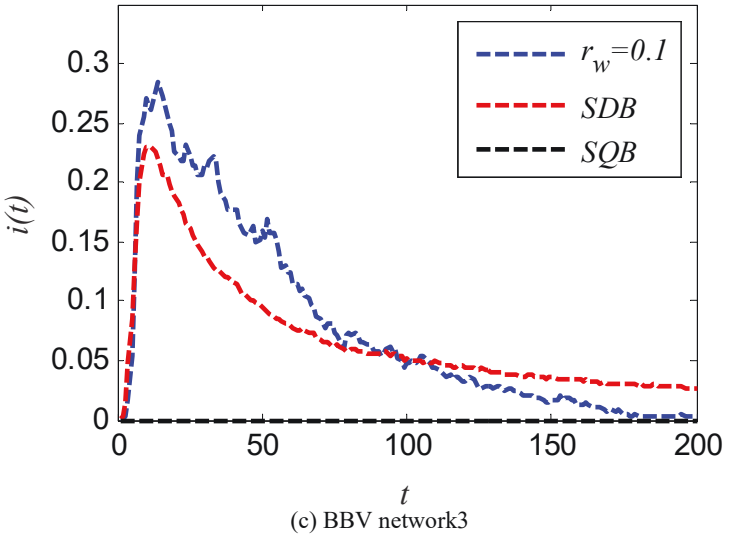
By using Monte-Carlo simulation, we analyze the dynamics in weighted adaptive network models with heterogeneous rewiring rates. Fig 8.5 shows the density of infected nodes $i(t)$ in the BBV networks, where both of the two disconnecting strategies are presented. It is clear that both of our disconnection strategies can effectively inhibit the epidemic spreading, the heterogeneous disconnection process leads to a significant decline of the infected population in the steady-state networks by comparing with the homogeneous rewiring rate, $r_w = 0.1$. An interesting finding is that Fig 8.5 shows that the second design (SQB) can better inhibit the virus spreading than the first one (SDB) as the dispersion of weights becomes larger. That is because, as the dispersion of weights grows, links with extremely large value of weight appear in networks, who would be disconnected easier under the second design. The virus spreading behavior is well inhibited since there are only links with lower value of weight left in the networks.




(a) BBV network1




(b) BBV network2

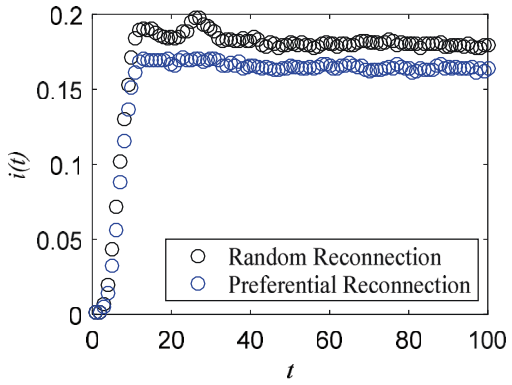



(c) BBV network3

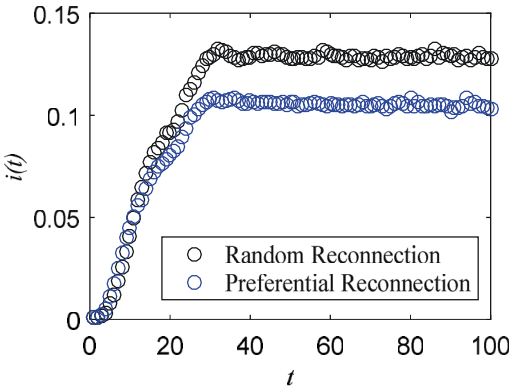
Fig 8.5. Epidemic spreading in weighted adaptive heterogeneous networks under heterogeneous link-disconnecting strategies.

- **Link-reconnecting Strategy**

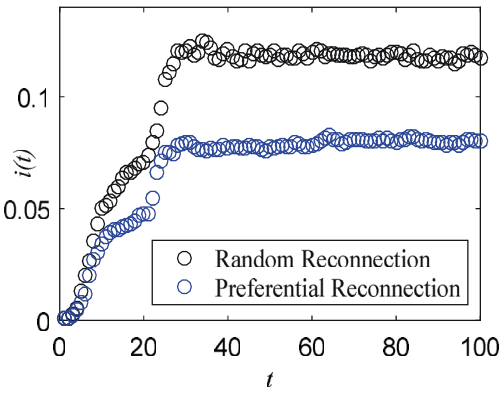
Following the rules made for our link-reconnecting strategy in Section 8.3, simulations are carried out in the artificial network models. Fig 8.6 shows the constraint of link-reconnecting strategies on epidemic propagation. We can find that our strategy is more effective on inhibiting the epidemic spreading, especially in the network with larger dispersion of weights. Our priority link-reconnecting strategy is based on the degree of each node. In BBV networks, the more dispersed the weight distribution, the more dispersed the degree of the network. Since the degree distribution and the strength distribution are directly proportional, the inhibition of virus spreading is more effective in networks with larger dispersion of weight.




(a) BBV network1




(b) BBV network2




(c) BBV network3

Fig 8.6. Epidemic spreading in weighted heterogeneous adaptive networks under link-reconnecting strategies.

In general, our simulation results show that the epidemic spreading depends on the distribution of the link weights and the specific rewiring strategy in the adaptive weighted networks. The weight

distribution and rewiring process have a very important impact on virus spreading. The conclusion drawn is that the larger the dispersion of the link weights or/and the rewiring rate is, the more effectively the infection can be prevented from proliferation. Our rewiring strategies can effectively inhibit the spread of virus, i.e., by reducing the steady-state population of infected nodes.

## 8.5 Conclusion

When infectious virus outbreak in the social networks, humans tend to respond to the emergence of an epidemic by avoiding contacts with infected individuals. These behavioral responses to epidemic situation have significant impacts on virus spreading. Considering the diversity of interpersonal relationships, the interaction between network topology and epidemic spreading, we propose a novel SIS model on a weighted adaptive heterogeneous network, where we describe the interplay between dynamic behaviors and epidemic spreading. Based on the dynamical model, effective rewiring strategies are proposed for inhibition of epidemics. Our results show that network topology and individuals' behaviors have significant influence on virus spreading in a weighted adaptive heterogeneous network. The adaptive rewiring process during the dynamics can effectively inhibit virus spreading. Large dispersion of weights can significantly decrease the final epidemic size. Moreover, our proposed rewiring strategies are effective for the inhibition of epidemics.

# Chapter 9

# Contributions and Future Work

## 9.1 Work Summary

The thesis focuses on the application of complex network structure metrics in propagation dynamics and their interaction. As important tools study on complex network theory, Markov chain method, mean-field method and Monte-Carlo simulation are used to study the influence of network structure measurement on propagation dynamics. The influence of node degree on these applications are studied from two applications, node influence identification and cascading failure, and a novel metric based on the network dynamical behavior measurements is proposed to quantify the network robustness in regard to virus attacks. In addition, the thesis also studies the co-evolution process between network structure and propagation dynamics, and proposes dynamic adaptive strategies to suppress virus and fault propagation in the network.

The main research contents and achievements are summarized as follows:

**(1) Study on the rapid identification of high-influence nodes in complex networks**

Considering the uncertainty of network scale, topology structure and the timeliness of dynamic behavior in the actual network, a method based on node degree value to quickly find high impact nodes is proposed. By this method, a small number of high impact nodes in the network can be quickly found to further control the network. The results show that the proposed method can find high impact nodes in the network in a quick and effective way. In addition, the propagation process in different network structures is also studied. The results show that the proposed method has a very positive impact on the propagation process in the network.

**(2) Study on cascading failure of complex networks considering local real-time information**

Considering the real-time performance of network information and the rationality of load redistribution strategy, a new cascading failure model is proposed to analyze the robustness of network under node failure, and a load redistribution strategy based on local real-time information is proposed. The influence of load redistribution strategy on network robustness is analyzed theoretically, and the relationship between network robustness and parameters of redistribution strategy under different conditions is discussed. The analysis shows that when the initial load allocation and load redistribution strategies remain the same and have a linear relationship with the degree value of nodes, the network shows better robustness under cascading failures. In addition, the initial load information remains unknown. It is an effective load redistribution strategy to distribute load proportion according to node degree value to improve network robustness.

**(3) Study on the quantification of social network robustness under the virus attacks**

A new robustness measure with respect to virus attacks in social networks is proposed in view of the spread velocity, the epidemic threshold and the infection scale under the steady state. Simulation

results show that the network becomes more vulnerable to the virus attacks as the average degree of the network grows in both homogeneous networks and heterogeneous networks. Our new measure confirms that the irregularity in node degrees decreases the robustness of the homogeneous networks.

**(4) Study on the reliability of large-scale adaptive weighted network**

Considering the real-time characteristics of network information and the rationality of load redistribution strategy, a mean-field approximate dynamic system is proposed to simulate the time-varying population of failure nodes and risk links in adaptive weighted networks. An analysis on the linear stability of the dynamic system is carried out, and the threshold value of the network is determined to suppress the fault and maintain the reliability in the steady state. The simulation results show that the network reliability threshold depends on the distribution of edge connection weight and the strategy of edge broken reconnection. The results also show that the priority to disconnect suspicious individuals with frequent communication can inhibit the transmission of faults and viruses in an effective way.

**(5) Study on the dynamical rewiring in adaptive weighted heterogeneous networks**

Based on SIS propagation model, a new adaptive weighted non-uniform network is proposed, where the network structure, weight and virus propagation behavior are dynamically interactive. Personal behavior and interpersonal relationship in the process of virus transmission exerted profound effects on the spread of the epidemic. The analysis shows that the greater the degree of weight dispersion of the initial network, the slower the propagation in the adaptive weighted network, and the adaptive broken edge reconnection process can also suppress the epidemic transmission in an effective way. Finally, in accordance with the individual behavior, the edge breaking and reconnecting strategy produced, which is dynamically related to the weight of real-time edge connection, so as to suppress the epidemic. The results show the effectiveness of the strategy and model.

## 9.2 Research Prospects

Although the work of the thesis has achieved some phased results, in the process of research, we also found more work to be explored. According to the research content of the thesis, the following further key research contents are put forward:

**(1) Research on node influence prediction method based on big data**

With the increasing complexity of the real network, the network presents the characteristics of big data [174]-[177]. First of all, the number of nodes and connected edges in the network increases dramatically, and the nodes and connected edges are becoming increasingly diversified. Secondly, the network functionality is becoming increasingly powerful, and there are more and more complex dynamic behaviors running in the network. In addition, the close coupling among networks makes the network with higher complexity. Therefore, the prediction or ranking of node influence based on big data network becomes a very difficult problem to solve. New challenges are nowhere from the data collection of node information to the selection and extraction of node features, and then to the evaluation of the effectiveness of node importance methods. In particular, in recent years, the rise of machine learning [178]-[180] and deep learning [181], [182] provides solutions for these problems and challenges. Therefore, the research

of node influence based on network big data is a very noteworthy issue.

**(2) Research on cascading failures in multilayer networks**

As mentioned in the previous chapter, when a cascading failure occurs in a network, it will not only spread to the whole network, but also to other related networks. For example, when the power network is damaged and cascaded failure happens, all the power related networks will be damaged in varying degrees, such as the Internet, traffic network, etc. This cascading failure across the network will bring catastrophic losses. It is of great practical significance to study cascading failures in a variety of coupled networks, analyze the propagation mechanism and propagation rules, and propose effective strategies to resist cascading failures.

**(3) Modeling and analysis of communication dynamics**

Nowadays, the real-world networks show amazing changes, including network structure and dynamic behaviors on the networks, especially the emergence of online networks, showing many new features that the traditional networks don’t have. Therefore, the traditional modeling methods based on the former are not enough to describe the new characteristics of network structure and dynamic behaviors. Based on machine learning and other methods, the new evolution rules of network structure are worth exploring. In light of these more realistic network models and real network data, a more realistic communication dynamics model can be established and the propagation dynamics processes in the networks can be explored.

# References


[1] Wang X F, Li X, Chen G R. Introduction to network science[M]. Beijing: Higher Education Press, 2012. (In Chinese)

[2] Costa L D F, Rodrigues F A, Travieso G, et al. Characterization of complex networks: A survey of measurements[J]. Advances in Physics, 2007, 56(1): 167-242.

[3] Chen G R, Wang X F, Li X. Introduction to complex networks: model, structure and dynamics[M]. Beijing: Higher Education Press, 2014. (In Chinese)

[4] Strogatz S H. Exploring complex networks[J]. Nature, 2001, 410(6825): 268.

[5] Van Der Hofstad R. Random graphs and complex networks[M]. Cambridge University Press, 2016.

[6] Zhou J, Liu Z H. Epidemic spreading in complex networks[J]. Frontiers of Physics in China, 2008, 3(3): 331-348.

[7] Pastor-Satorras R, Castellano C, Mieghem P V, et al. Epidemic processes in complex networks[J]. Review of Modern Physics, 2014, 87(3): 120-131.

[8] Zhou J, Liu Z H. Epidemic spreading in complex networks[M]. Statistical Mechanics of Complex Networks. Springer Berlin Heidelberg, 2003: 127-147.

[9] Boguñá M, Pastor-Satorras R. Epidemic spreading in correlated complex networks[J]. Physical Review E, 2002, 66(4 Pt 2): 047104.

[10] Valler N C, Prakash B A, Tong H, et al. Epidemic spread in mobile Ad-Hoc networks: Determining the tipping point[C]. International Conference on Research in Networking. Springer Berlin Heidelberg, 2011: 266-280.

[11] Ren H, Wang Y Q, Song Y R. Epidemic spreading on complex networks considering communication flow[J]. Journal of Nanjing University of Posts & Telecommunications, 2011, 31(6): 85-89.

[12] Venkatraman A, Mukhija D, Kumar N, et al. Zika virus misinformation on the internet[J]. Travel Medicine and Infectious Disease, 2016, 14(4): 421-422.

[13] Yang L X, Yang X. The impact of nonlinear infection rate on the spread of computer virus[J]. Nonlinear Dynamics, 2015, 82(1-2): 85-95.

[14] Xia L L, Jiang G P, Song B, Song Y R, Rumor spreading model considering hesitating mechanism in complex social networks[J]. Physica A: Statistical Mechanics and its Applications, 2015, 437: 295-303.

[15] Xia L L, Song Y R, Li C C, Jiang G P, Improved targeted immunization strategies based on two rounds of selection[J]. Physica A: Statistical Mechanics and its Applications, 2018, 496: 540-547.

[16] Dey P, Pyne S, Roy S. Information spreading in online social networks: A case study on Twitter network[C]. ACM International Symposium on Mobile Ad Hoc Networking and Computing. ACM, 2017: 34.

[17] Rahnamay-Naeini M, Wang Z, Ghani N, et al. Stochastic analysis of cascading-failure dynamics in power grids[J]. IEEE Transactions on Power Systems, 2014, 29(4): 1767-1779.

[18] Zhu Y, Yan J, Sun Y, et al. Revealing cascading failure vulnerability in power grids using risk-Graph[J]. IEEE Transactions on Parallel & Distributed Systems, 2014, 25(12): 3274-3284.

[19] Pahwa S, Scoglio C, Scala A. Abruptness of cascade failures in power grids[J]. Scientific Report, 2014, 4(1): 3694.

[20] Huang X. Cascading Failures in interdependent systems and financial networks[J]. International Journal of Robust & Nonlinear Control, 2015, 22(16): 1837-1852.

[21] Zhang D. Big data security and privacy protection[C]. 8th International Conference on Management and Computer Science (ICMCS 2018). Atlantis Press, 2018.

[22] Zyskind G, Nathan O. Decentralizing privacy: Using blockchain to protect personal data[C]. 2015 IEEE Security and Privacy Workshops. IEEE, 2015: 180-184.

[23] Jungnickel D. Basic graph theory[J]. Algorithms & Computation in Mathematics, 2013: 1-31.

[24] Adams C. Leonhard Euler and the seven bridges of Königsberg[J]. Mathematical Intelligencer, 2011, 33(4): 18-20.

[25] Erdös P, Rényi A, On random graphs I[J], Publicationes Mathematicae Debrecen, 6 (1959). 290–297.

[26] Watts D J, Strogatz S H. Collective dynamics of “small-world” networks[J]. Nature, 1998, 393: 440-442.

[27] Barabási A L, Albert R. Emergence of scaling in random Networks[J]. Science, 1999, 286(5439): 509-512.

[28] National Research Council. Network science committee on network science for future army applications[J]. 2005.

[29] Brandes U, Robins G, McCranie A, et al. What is network science?[J]. Network Science, 2013, 1(1): 1-15.

[30] Barabási A L. Network science[M]. Cambridge University Press, 2016.

[31] Lewis T G. Network science: Theory and applications[M]. John Wiley & Sons, 2011.

[32] Barabási A L, Albert R, Jeong H. Mean-field theory for scale-free random networks[J]. Physica A: Statistical Mechanics and its Applications, 1999, 272(1–2): 173-187.

[33] Flyvbjerg H, Sneppen K, Bak P. Mean-field theory for a simple model of evolution[J]. Physical Review Letters, 1993, 71(24): 4087-4090.

[34] Dai R W, Li Y D. Researches on hall for workshop of meta synthetic engineering and system complexity[J]. Complex Systems and Complex Science, 2004, 1(4): 1-10. (In Chinese)

[35] Meyer C. Matrix analysis and applied linear algebra[M]. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2000.

[36] Gross T, D'Lima C J D, Blasius B. Epidemic dynamics on an adaptive network[J]. Physical Review Letters, 2006, 96(20): 208701.

[37] Schwartz I B, Shaw L B. Rewiring for adaptation[J]. Physics, 2010, 3(17).

[38] Shaw L B, Schwartz I B. Fluctuating epidemics on adaptive networks[J]. Physical Review E, 2008, 77(6): 066101.

[39] Zhu G, Chen G, Xu X J, et al. Epidemic spreading on contact networks with adaptive weights[J]. Journal of Theoretical Biology, 2013, 317: 133-139.

[40] Liu Z, Hu B. Epidemic spreading in community networks[J]. Europhysics Letters, 2005, 72(2): 315.

[41] Wu X, Liu Z. How community structure influences epidemic spread in social networks[J]. Physica A: Statistical Mechanics and its Applications, 2008, 387(2-3): 623-630.

[42] Huang W, Li C. Epidemic spreading in scale-free networks with community structure[J]. Journal of Statistical Mechanics: Theory and Experiment, 2007(01): P01014.

[43] Ren G, Wang X. Epidemic spreading in time-varying community networks[J]. Chaos: An Interdisciplinary Journal of Nonlinear Science, 2014, 24(2): 023116.

[44] Boccaletti S, Bianconi G, Criado R, et al. The structure and dynamics of multilayer networks[J]. Physics Reports, 2014, 544(1): 1-122.

[45] Kivelä M, Arenas A, Barthelemy M, et al. Multilayer networks[J]. Ssrn Electronic Journal, 2013, 2(3): 261- 268.

[46] Ren X L, Lü L Y. Review of ranking nodes in complex networks[J]. Chinese Science Bulletin, 2014, (13). (In Chinese)

[47] Chen D B, Lü L Y, Shang M S, et al. Identifying influential nodes in complex networks[J]. Physica A: Statistical Mechanics and its Applications, 2012, 391: 1777-1787.

[48] Lü L, Zhang Y C, Yeung C H, et al. Leaders in social networks, the delicious case[J]. PloS One, 2011, 6(6): e21202.

[49] Wu X D, Li Y, Li L. Influence analysis of online social networks[J]. Chinese journal of computers, 2014, (4):735-752. (In Chinese)

[50] Li M, Zhang Q, Deng Y. Evidential identification of influential nodes in network of networks[J]. Chaos, Solitons & Fractals, 2018, 117: 283-296.

[51] Zareie A, Sheikhahmadi A. EHC: Extended H-index centrality measure for identification of users' spreading influence in complex networks[J]. Physica A: Statistical Mechanics and its Applications, 2019, 514: 141-155.

[52] Aral S, Dhillon P S. Social influence maximization under empirical influence models[J]. Nature Human Behaviour, 2018, 2(6): 375.

[53] Bonacich P. Factoring and weighting approaches to status scores and clique identification[J]. Journal of Mathematical Sociology, 1972, 2: 113–120.

[54] Chen D B, Lü L, Shang M S, et al. Identifying influential nodes in complex networks[J]. Physica A: Statistical Mechanics and its Applications, 2012, 391: 1777–1787.

[55] Kitsak M, Gallos L K, Havlin S, et al. Identification of influential spreaders in complex networks[J]. Nature Physics, 2010, 6: 888–893.

[56] Cheng X Q, Ren F X, Shen H W, et al. Bridgeness: A local index on edge significance in maintaining global connectivity[J]. Journal of Statistical Mechanics-Theory and Experiment, 2010, 2010: P10011.

[57] Freeman L C. A set of measures of centrality based on betweenness[J]. Sociometry, 1977, 40: 35–41.

[58] Brin S, Page L. The anatomy of a large-scale hypertextual web search engine[J]. Computer Networks and ISDN Systems, 1998, 30: 107–117.

[59] Kleinberg J M. Authoritative sources in a hyperlinked environment[J]. Journal of Applied and Computational Mechanics, 1999, 46: 604–632.

[60] Chakrabarti S, Dom B, Raghavan P, et al. Automatic resource compilation by analyzing hyperlink structure and associated text[J]. Computer Networks and ISDN Systems, 1998, 30: 65–74.

[61] Lempel R, Moran S. The stochastic approach for link-structure analysis (SALSA) and the TKC effect[J]. Computer Networks, 2000, 33: 387–401.

[62] Chen Y, Hu A Q, Hu X. Evaluation method for node importance in communication networks [J]. Journal of China Institute of Communications, 2004, 25: 129–134. (In Chinese)

[63] Tan Y J, Wu J, Deng H Z. Evaluation method for node importance based on node contraction in complex networks [J]. Systems Engineering-Theory & Practice, 2006, 26: 79–83. (In Chinese)

[64] Weisbuch G. Complex systems dynamics[M]. CRC Press, 2018.

[65] Dereich S, Mörters P. Random networks with sublinear preferential attachment: The giant component[J]. Annals of Probability, 2013, 41: 329–384.

[66] Schneider C M, Moreira A A, Andrade J S, et al. Mitigation of malicious attacks on networks[J]. Proceedings of the National Academy of Sciences, 2011, 108: 3838–3841.

[67] Bailey, Norman T J. The mathematical theory of infectious diseases and its applications[M]. Charles Griffin & Company Ltd 5a Crendon Street, High Wycombe, Bucks HP13 6LE., 1975.

[68] Anderson R M, Robert M M. Infectious diseases of humans: dynamics and control[M]. Oxford University Press, 1992.

[69] Pastor-Satorras, R, Alessandro V. Handbook of graphs and networks: from the genome to the Internet[M]. John Wiley & Sons, 2006: 113-132.

[70] Boguá M, Pastor-Satorras R, Alessandro V. Epidemic spreading in complex networks with degree correlations[J]. Statistical mechanics of complex networks. Springer, Berlin, Heidelberg, 2003, 127-147.

[71] Xu X J, Li D Q, Jiang Y N, Kang R. Review of recent progress on propagation of cascading failures in complex networks [J]. Electronic Science & Technology, 2015(6): 697-701. (In Chinese)

[72] Bunde A, Havlin S. Fractals and disordered systems[M]. Fractals and Disordered Systems. 1991.

[73] Parshani R, Buldyrev S V, and Havlin S, Critical effect of dependency groups on the function of networks[J]. Proceedings of the National Academy of Sciences, 2011, 108(3): 1007-1010.

[74] Bashan A, Parshani R, Havlin S, Percolation in networks composed of connectivity and dependency links[J]. Physical Review E, 2011, 83(5): 051127.

[75] Motter A E, Lai Y C, Cascade-based attacks on complex networks[J]. Physical Review E, 2002. 66(6): 065102.

[76] Witthaut D, Timme M. Nonlocal effects and countermeasures in cascading failures[J]. Physical Review E, 2015, 92(3): 032809.

[77] Zhang L, Du W, Ying W, et al. Optimal resource allocation in interdependent networks[J]. Physica A: Statistical Mechanics and its Applications, 2018, 508: 104-110.

[78] Crucitti P, Latora V, Marchiori M, Model for cascading failures in complex networks[J]. Physical Review E, 2004. 69(4): 045104.

[79] Dobson I B A, Carreras, Lynch V E, et al. An initial model fo complex dynamics in electric power system blackouts[C]. Hawaii International Conference on System Sciences. IEEE Computer Society, 2001.

[80] Carreras, B.A, Newman D E, Dobson I, et al. Evidence for self-organized criticality in a time series of electric power system blackouts[J]. IEEE Transactions on Circuits and Systems I: Regular Papers, 2004, 51(9): 1733-1740.

[81] Trajanovski S, Guo D, Van M P. From epidemics to information propagation: Striking differences in structurally similar adaptive network models[J]. Phys Rev E, 2015, 92(3-1): 030801.

[82] Ghosh I, Tiwari P K, Samanta S, et al. A simple SI-type model for HIV/AIDS with media and self-imposed psychological fear[J]. Mathematical Biosciences, 2018, 306: 160-169.

[83] Wiedermann M, Donges J F, Heitzig J, et al. Macroscopic description of complex adaptive networks coevolving with dynamic node states[J]. Physical Review E, 2015, 91(5).

[84] Lazer D M J, Baum M A, Benkler Y, et al. The science of fake news[J]. Science, 2018, 359(6380): 1094-1096.

[85] McNair B. Fake news: Falsehood, fabrication and fantasy in journalism[M]. Routledge, 2017.

[86] Kucharski A. Post-truth: Study epidemiology of fake news[J]. Nature, 2016, 540(7634): 525.

[87] Parsegov S E, Proskurnikov A V, Tempo R, et al. Novel multidimensional models of opinion dynamics in social networks[J]. IEEE Transactions on Automatic Control, 2017, 62(5): 2270-2285.

[88] Proskurnikov A V, Matveev A S, Cao M. Opinion dynamics in social networks with hostile camps: Consensus vs. polarization[J]. IEEE Transactions on Automatic Control, 2016, 61(6): 1524-1536.

[89] Jia P, MirTabatabaei A, Friedkin N E, et al. Opinion dynamics and the evolution of social power in influence networks[J]. SIAM review, 2015, 57(3): 367-397.

[90] Tunc I, Shaw L B. Effects of community structure on epidemic spread in an adaptive network[J]. Physical Review E, 2014, 90(2): 022801.

[91] Yang H, Tang M, Zhang H F. Efficient community-based control strategies in adaptive networks[J]. New Journal of Physics, 2012, 14(12): 123017.

[92] Song Y R, Jiang G P, Gong Y W. Epidemic propagation on adaptive coevolutionary networks with preferential local-world reconnecting strategy[J]. Chinese Physics B, 2013, 22(4): 040205.

[93] Shaw L B, Schwartz I B. Enhanced vaccine control of epidemics in adaptive networks[J]. Physical Review E, 2010, 81(4): 046120.

[94] Erdős P, Rényi A. Asymmetric graphs[J]. Acta Mathematica Academiae Scientiarum Hungarica, 1963, 14(3-4): 295-315.

[95] Barrat A, Barthelemy M, Vespignani A. Modeling the evolution of weighted networks[J]. Physical Review E Statistical Nonlinear & Soft Matter Physics, 2004, 70(2): 066149.

[96] Gang Y, Tao Z, Jie W, et al. Epidemic spread in weighted scale-free networks[J]. Chinese Physics Letters, 2005, 22(2): 510.

[97] Duan W, Song Z, Qiu X. Heterogeneous edge weights promote epidemic diffusion in weighted evolving networks[J]. Modern Physics Letters B, 2016, 30(21): 1650300.

[98] Tang C, Li A, Li X. Asymmetric game: A silver bullet to weighted vertex cover of networks[J]. IEEE Transactions on Cybernetics, 2017 (99): 1-12.

[99] Pourbohloul B. Modeling infectious diseases in humans and animals[M]. Modeling Infectious Diseases in Humans and Animals. 2008: 864-865.

[100] Andersson H, Britton T. Stochastic epidemic models and their statistical analysis[M]. Springer US, New York, 2000.

[101] Brauer F. Mathematical models in population biology and epidemiology[J]. American Mathematical Monthly, 2003, 40(3): 267-291.

[102] Diekmann O, Heesterbeek H, Britton T. Mathematical tools for understanding infectious disease dynamics[M], Princeton University Press, Princeton, USA, 2012.

[103] Pastor-Satorras R, Vespignani A. Epidemic dynamics and endemic states in complex networks [J]. Physical Review E, 2001, 63(6 Pt 2): 066117.

[104] Gross T, Kevrekidis I G. Robust oscillations in SIS epidemics on adaptive networks: Coarse-graining by automated moment closure[J]. Europhysics Letters, 2007, 82(3): 38004.

[105] Su X P, Song Y R. Leveraging neighborhood “structural holes” to identifying key spreaders in social networks [J]. Acta Physica Sinica, 2015, 64(2). (In Chinese)

[106] Zhao Z Y, Yu H, Zhu Z L. Identifying influential spreaders based on network community structure[J]. Chinese Journal of Computers, 2014, (4): 753-766. (In Chinese)

[107] Hu Q C, Yin Y S, Ma P F, et al. A new approach to identify influential spreaders in complex networks[J]. Acta Physica Sinica, 2013, 62(14). (In Chinese)

[108] Zachary W W. An information flow model for conflict and fission in small groups[J]. Journal of Anthropological Research, 1977, 33(4): 452-473.

[109] Han Z, Tan X, Chen Y, et al. NCSS: an effective and efficient complex network community detection algorithm[J]. Scientia Sinica Informationis, 2016, 46(4): 431-444.

[110] Newman M E J. Assortative mixing in networks[J]. Physical Review Letters, 2002, 89(20): 208701.

[111] Goh K I, Lee D S, Kahng B, et al. Cascading toppling dynamics on scale-free networks[J]. Physica A: Statistical Mechanics and its Applications, 2005, 346(1/2): 93-103.

[112] Dobson L, Carreras B A, Newman D E. A probabilistic loading-dependent model of cascading failure and possible implications for blackouts[J]. System Sciences, 2003, 65(1): 6-9.

[113] Peng X Z, Yao H, Du J, et al. Invulnerability of scale-free network against critical node failures based on a renewed cascading failure model[J]. Physica A: Statistical Mechanics and its Applications, 2015, 421(3): 69-77.

[114] Liu Y N, Li X, Chen S Z, et al. Model for cascading network failures based on the nodes with different tolerance parameter[J]. Journal of China Universities, 2011, 18(5): 95-101.

[115] Dehmer M, Holzinger A, Pickl S. Big data of complex networks[C]. Chapman & Hall/CRC, 2016.

[116] Duan D. Cascading failure of complex networks based on load local preferential redistribution rule[J]. Complex Systems and Complexity Science, 2015, 12(1): 33-39.

[117] Lancet T. The gendered dimensions of COVID-19[J]. The Lancet, 2020, 395(10231): 1168.

[118] Dave M, Seoudi N, Coulthard P. Urgent dental care for patients during the COVID-19 pandemic[J]. The Lancet, 2020, 395(10232): 1257.

[119] Cameron H, Brian O, Nicholas S, et al. Will COVID-19 fiscal recovery packages accelerate or retard progress on climate change?[J]. Oxford Review of Economic Policy, 2020.

[120] Odeh N D, Babkair H, Abu-Hammad S, et al. COVID-19: Present and future challenges for dental practice[J]. International Journal of Environmental Research and Public Health, 2020, 19(9): 3151.

[121] https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200601-covid-19-sitrep-133.pdf?sfvrsn=9a56f2ac_4

[122] Wesley C, Mata Angélica S, Ferreira S C. Robustness and fragility of the susceptible-infected-susceptible epidemic models on complex networks[J]. Physical Review E, 2018, 98(1):012310.

[123] Jiang Y, Hu A, Huang J. Importance-based entropy measures of complex networks' robustness to attacks[J]. Cluster Computing, 2018, 22: 3981-3988.

[124] Pastor-Satorras R, Castellano C, Mieghem P V, et al. Epidemic processes in complex networks[J]. Review of Modern Physics, 2014, 87(3): 120-131.

[125] Mieghem P V. The viral conductance of a network[J]. Computer communications, 2012, 35(12): 1494-1506.

[126] Mina Youssef, Robert Kooij, Caterina Scoglio. Viral conductance: Quantifying the robustness of networks with respect to spread of epidemics[J]. Journal of Computational Science, 2 (2011): 286-298.

[127] Socievole A, De Rango F, Scoglio C, et al. Assessing network robustness under SIS epidemics: The relationship between epidemic threshold and viral conductance[J]. Computer Networks, 2016, 103(jul.5): 196-206.

[128] Zargar S T, Joshi J, Tipper D. A survey of defense mechanisms against distributed denial of service (DDoS) flooding attacks[J]. IEEE Communications Surveys & Tutorials, 2013, 15(4): 2046-2069.

[129] Song B, Wang X, Ni W, et al. Reliability analysis of large-scale adaptive weighted networks[J]. IEEE Transactions on Information Forensics and Security, 2019, PP(99): 1-1.

[130] Song B, Zhang Z H, Song Y R, et al. Preferential redistribution in cascading failure by considering local real-time information[J]. Physica A: Statal Mechanics and its Applications, 2019, 532: 121729.

[131] Guillaume D, Diana A, Alexander D, et al. Molecular mechanisms of paralogous compensation and the robustness of cellular networks[J]. Journal of Experimental Zoology Part B Molecular & Developmental Evolution, 2014, 322(7): 488-499.

[132] Banerjee S, Chatterjee A, Shakkottai S. Epidemic thresholds with external agents[J]. Proceedings IEEE Infocom, 2013.

[133] Zhu G H, Chen G R, Zhang H F, et al. Propagation dynamics of an epidemic model with infective media connecting two separated networks of populations[J]. Communications in Nonlinear Science and Numerical Simulation, 2015, 20(1): 240-249.

[134] Perasso A. Global stability and uniform persistence for an infection load-structured SI model with exponential growth velocity[J]. Communications on Pure & Applied Analysis, 2019, 18(1): 15-32.

[135] Li C C, Jiang G P, Song Y R, et al. Modeling and analysis of epidemic spreading on community networks with heterogeneity[J]. Journal of Parallel & Distributed Computing, 2018, 119(SEP.): 136-145.

[136] Abdulrahman A, Yong D, Guiyi W, et al. Collaborative security in vehicular cloud computing: A game theoretic view[J]. IEEE Network, 2018, 32(3): 72-77.

[137] Hayes B. Cloud computing[J]. Communications of the Acm, 2008, 51(7): 9-11.

[138] Mijumbi R, Serrat J, Gorricho J L, et al. Network function virtualization: State-of-the-art and research challenges[J]. IEEE Communications Surveys & Tutorials, 2015, 18(1): 236-262.

[139] Cha X, Ni W, Zheng K, et al. Collaborative authentication in decentralized dense mobile networks with key predistribution[J]. IEEE Transactions on Information Forensics and Security, 2017: 1-1.

[140] Gross T, Blasius B. Adaptive coevolutionary networks: A review[J]. Journal of the Royal Society Interface, 2007, 5(20): 259.

[141] Marceau V, Pierre-Andr´e N, H´ebert-Dufresne L, Antoine A, Louis J D. Adaptive networks: coevolution of disease and topology[J]. Physical Review E, 2010, 82(2): 036116.

[142] Zargar S T, Joshi J, Tipper D. A survey of defense mechanisms against distributed denial of service (DDoS) flooding attacks[J]. IEEE Communications Surveys & Tutorials, 2013, 15(4):2046-2069.

[143] Smith J E, Nair R. The architecture of virtual machines[J]. Computer, 2005, 38(5): 32-38.

[144] Tao F, Li C, Liao T W, et al. BGM-BLA: a new algorithm for dynamic migration of virtual machines in cloud computing[J]. IEEE Transactions on Services Computing, 2016, 9(6): 910-925.

[145] Xu M, Tian W, Buyya R. A survey on load balancing algorithms for virtual machines placement in cloud computing[J]. Concurrency and Computation: Practice and Experience, 2017, 29(12): e4123.

[146] ETSI N F V. Network Functions Virtualisation (NFV)[J]. Management and Orchestration, 2014, 1: V1.

[147] Herrera J G, Botero J F. Resource allocation in NFV: A comprehensive survey[J]. IEEE Transactions on Network and Service Management, 2016, 13(3): 518-532.

[148] Yan Q, Yu F R, Gong Q, et al. Software-defined networking (SDN) and distributed denial of service (DDoS) attacks in cloud computing environments: A survey, some research issues, and challenges[J]. IEEE Communications Surveys & Tutorials, 2016, 18(1): 602-622.

[149] Somani G, Gaur M S, Sanghi D, et al. DDoS attacks in cloud computing: Issues, taxonomy, and future directions[J]. Computer Communications, 2017, 107: 30-48.

[150] Hoque N, Bhattacharyya D K, Kalita J K. Botnet in DDoS attacks: trends and challenges[J]. IEEE Communications Surveys & Tutorials, 2015, 17(4): 2242-2270.

[151] Zhou Y, Xia Y. Epidemic spreading on weighted adaptive networks[J]. Physica A: Statistical Mechanics & Its Applications, 2014, 399(4): 16-23.

[152] Dong C, Yin Q, Liu W, et al. Can rewiring strategy control the epidemic spreading[J]. Physica A: Statistical Mechanics and its Applications, 2015, 438: 169-177.

[153] Rattana P, Blyuss K B, Eames K T D, et al. A class of pairwise models for epidemic dynamics on weighted networks[J]. Bulletin of Mathematical Biology, 2013, 75(3).

[154] Arrowsmith D K, Place C M. Dynamical systems: differential equations, maps and chaotic behaviour[M]. Chapman & Hall, 1992.

[155] Cohn I, Elhay T, Friedman N, et al. Mean-field variational approximation for continuous-time bayesian networks[J]. Journal of Machine Learning Research, 2010, 11(5): 2745-2783.

[156] Wu Q, Zhang F. Threshold conditions for SIS epidemic models on edge-weighted networks[J]. Physica A: Statistical Mechanics and its Applications, 2016, 453(1): 77-83.

[157] Lyapunov A M. The general problem of the stability of motion[J]. International Journal of Control, 1992, 55(3): 531-534.

[158] Preciado V M, Zargham M, Enyioha C, et al. Optimal resource allocation for network protection against spreading processes[J]. IEEE Transactions on Control of Network Systems, 2014, 1(1): 99-108.

[159] Dreessche P, Watmough J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission[J]. Mathematical Biosciences, 2002, 180(1-2): 29-48.

[160] Agarwal R, Banerjee A, Gauthier V, Becker M, Yeo C, Lee B S. Achieving small-world properties using bio-inspired techniques in wireless networks[J]. Computer Journal, 2012, 55(8): 909–931.

[161] Meyer C D. Matrix analysis and applied linear algebra[M]. Matrix Analysis and Applied Linear Algebra Book and Solutions Manual. 2000.

[162] Singh A K. The exponential distribution-theory, methods and applications[J]. Technometrics, 1997, 39(3):341-341.

[163] Marshall A W, Olkin I. A multivariate exponential distribution[J]. Journal of the American Statistical Association, 1967, 62(317): 30-44.

[164] Mitchell R L. Permanence of the log-normal distribution[J]. Journal of the Optical Society of America, 1968, 58(9): 1267-1272.

[165] Limpert E, Stahel W A. The log-normal distribution[J]. Significance, 2017, 14(1): 8-9.

[166] Monteiro M J. Fitting molecular weight distributions using a log-normal distribution model[J]. European Polymer Journal, 2015, 65: 197-201.

[167] David H A, Nagaraja H N. Order statistics[J]. Encyclopedia of Statistical Sciences, 2004, 9.

[168] Han L, Gao F, Li Z, et al. Low complexity automatic modulation classification based on order-statistics[J]. IEEE Transactions on Wireless Communications, 2017, 16(1): 400-411.

[169] Murray M K. Differential geometry and statistics[M]. Madras Chapman and Hall, 1993.

[170] Esch M, Tobias E. Decentralized scale-free network construction and load balancing in Massive Multiuser Virtual Environments[C]. 6th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom 2010). IEEE, 2011.

[171] Jia W, Hu Y, Shou G, et al. Constructing limited scale-free topologies for virtual networks[C]. IEEE International Conference on Computer Communication & the Internet. IEEE, 2016: 270-274.

[172] Lin R, Wu B, Zhao Y, et al. Critical nodes detecting in virtual networking environment[C]. Services. IEEE, 2014.

[173] Feng Y, Ding L, Huang Y H, et al. Epidemic spreading on weighted networks with adaptive topology based on infective information [J]. Physica A: Statistical Mechanics & Its Applications, 2016, 463: 493-502.

[174] Gandomi A, Haider M. Beyond the hype: Big data concepts, methods, and analytics[J]. International Journal of Information Management, 2015, 35(2): 137-144.

[175] Tankard C. Big data security[J]. Network Ssecurity, 2012, 2012(7): 5-8.

[176] Sagiroglu S, Sinanc D. Big data: A review[C]. 2013 International Conference on Collaboration Technologies and Systems (CTS). IEEE, 2013: 42-47.

[177] Cárdenas A A, Manadhata P K, Rajan S P. Big data analytics for security[J]. IEEE Security & Privacy, 2013, 11(6): 74-76.

[178] Witten I H, Frank E, Hall M A, et al. Data mining: Practical machine learning tools and techniques[M]. Morgan Kaufmann, 2016.

[179] Alpaydin E. Introduction to machine learning[M]. MIT press, 2009.

[180] Jordan M I, Mitchell T M. Machine learning: Trends, perspectives, and prospects[J]. Science, 2015, 349(6245): 255-260.

[181] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553): 436.

[182] Schmidhuber J. Deep learning in neural networks: An overview[J]. Neural Networks, 2015, 61: 85-117.