2015

Robustness and structure of complex networks

BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

ROBUSTNESS AND STRUCTURE OF COMPLEX NETWORKS

by

SHUAI SHAO

B.Sc., University of Science and Technology of China, 2010
M.A., Boston University, 2012

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
2015

Approved by

First Reader

H. Eugene Stanley, Ph.D.
University Professor and Professor of Physics
Boston University

Second Reader

Shlomo Havlin, Ph.D.
Professor of Physics
Bar-Ilan University

To my mom and dad,

the best parents in the world.

Acknowledgments

First of all, I would like to thank my advisor Professor H. Eugene Stanley, for his support and inspiration during my years with him. Gene is an amazingly energetic physicist, who offered me his perpetual encouragement, continuous trust, and constant support throughout all my research development. Moreover, he is a kind and warm friend, who brought me to this wonderful academic family and provided me the unparalleled opportunity to work with so many wonderful scholars.

I would like to thank my collaborators, Professor Shlomo Havlin and Dr. Xuqing Huang. Shlomo offered me the best insights and guidance as my co-advisor, without whom I would never have gotten to where I am today. Also, his passion for pursuing the real beauty of science and his kindness for always being willing to help really influenced me. Xuqing helped me get started in the group and passed me his experience and persistence in scientific research. It is my privilege to work with you both.

I would like to thank all the faculties and colleagues at Boston University, for being a unique family that has supported me during my doctoral studies. Jiang Li, Han Han, Kun Geng, Lingyue Cao, Na Xu, Xuqing Huang, Qian Li, Songbo Jin, Wei Li, Zhiqiang Su, Kang Liu, Duan Wang, Di Zhou, Xin Yuan, thank you for sharing my joy and pain, and every cherishable moment together. Mirtha Cabello, Bob Tomposki, and Guoan Hu, thank you for making my life a whole lot easier.

I would like to thank all my special ones here in Boston. Mengran Yang, Ren Wang, Aibo Zhao, Yuanrui Li, and all other dear friends. Because of you, this city becomes so special to me.

At last, I would like to thank the best parents in the world, my mother, Mrs. Shifeng Wang, and my father, Mr. Xuejin Shao. Thank you for bringing me to this world and supporting me in all my endeavors; thank you for standing by me every step of the way and loving me unconditionally, despite my success or failure.

ROBUSTNESS AND STRUCTURE OF COMPLEX NETWORKS

(Order No.)

SHUAI SHAO

Boston University, Graduate School of Arts and Sciences, 2015

Major Professor: H. Eugene Stanley, Professor of Physics

ABSTRACT

This dissertation covers the two major parts of my PhD research on statistical physics and complex networks: i) modeling a new type of attack – localized attack, and investigating robustness of complex networks under this type of attack; ii) discovering the clustering structure in complex networks and its influence on the robustness of coupled networks.

Complex networks appear in every aspect of our daily life and are widely studied in Physics, Mathematics, Biology, and Computer Science. One important property of complex networks is their robustness under attacks, which depends crucially on the nature of attacks and the structure of the networks themselves. Previous studies have focused on two types of attack: random attack and targeted attack, which, however, are insufficient to describe many real-world damages. Here we propose a new type of attack – localized attack, and study the robustness of complex networks under this type of attack, both analytically and via simulation. On the other hand, we also study the clustering structure in the network, and its influence on the robustness of a complex network system.

In the first part, we propose a theoretical framework to study the robustness of complex networks under localized attack based on percolation theory and generating function method. We investigate the percolation properties, including the critical threshold of the phase transition p_c and the size of the giant component P_∞. We compare localized attack with random attack and find that while random regular (RR) networks are more robust against localized attack, Erdős-Rényi (ER) networks are equally robust under both types of

attacks. As for scale-free (SF) networks, their robustness depends crucially on the degree exponent λ. The simulation results show perfect agreement with theoretical predictions. We also test our model on two real-world networks: a peer-to-peer computer network and an airline network, and find that the real-world networks are much more vulnerable to localized attack compared with random attack.

In the second part, we extend the tree-like generating function method to incorporating clustering structure in complex networks. We study the robustness of a complex network system, especially a network of networks (NON) with clustering structure in each network. We find that the system becomes less robust as we increase the clustering coefficient of each network. For a partially dependent network system, we also find that the influence of the clustering coefficient on network robustness decreases as we decrease the coupling strength, and the critical coupling strength q_c, at which the first-order phase transition changes to second-order, increases as we increase the clustering coefficient.

Contents

List of Figures

List of Abbreviations

DPD	Double–poisson Distribution
ER	Erdös – Rényi
FDD	Fixed–degree Distribution
LA	Localized Attack
NOI	Number of Interactions
NON	Network of Networks
RA	Random Attack
RR	Random Regular
SF	Scale–free
WWW	World–Wide–Web

Chapter 1

Introduction

1.1 Complex Networks and Graph Theory

Networks are present in almost every aspect of our life [1–21]. A network is a set of nodes connected by a set of links. Nodes represent the fundamental units of the system in question, and links establish which of the nodes are connected to others. Systems taking the form of networks include the Internet, social networks, financial networks, biological networks, infrastructure networks, and many others. For examples, airports and flights form an airline network, in which two nodes representing two airports are connected when an airline exists between these two airports. Similarly, buses and stops form a ground transportation network. The network of friendship between individuals, working relations, or network of business relations between firms are examples of social and economic networks.

Graph theory is used for describing mathematical concepts in networks [21]. The very first famous example of graph theory and network topology is the problem of bridges of Königsberg, where people had been wondering for years whether all seven bridges connecting the different parts of the town could be traveled, without passing any of them twice. Euler realized the only important factor of the problem is the topology of the network itself, and therefore solved the problem by concluding that to fulfill the requirement every node in the graph should be connected by an even number of bridges. In the 1960s, two mathematicians, Paul Erdös and Alfred Rényi, introduced the first probability model of networks– random

graph model. In this model, every pair of nodes in the network are randomly connected by a link with the same probability. The study of random graphs has led to ideas very similar to those of statistical physics. Concepts like percolation, scaling, order parameters, renormalization, self-similarity, phase transitions, and critical exponents from statistical physics are all present in the field of random graphs, and are used in the study of complex networks.

The random graph model, or Erdös – Rényi network (ER), has been widely studied for decades. However, with the development of science and technology and the availability of more large-scale data, it is revealed that for many real-world systems, ER model fails to describe their properties. At the end of the twentieth century, the work of Barabási and Albert on the World-Wide-Web (WWW) network made clear that the link connection of these and many other networks is not completely random, and it cannot be described by ER model. The study of these new types of networks leads to novel physical laws, which arise owing to the new topology. By using graph theory, we can describe different types of networks based on different topologies.

1.1.1 Erdös – Rényi (ER) networks

In an ER network, each pair of nodes in the network are randomly connected with the same probability p. The number of nodes in the network is usually expressed with N. Degree of a node, which is expressed in k, means number of links of a certain node. Average degree $\langle k \rangle$ represents the average number of links connected to nodes in the network. The degree distribution is the probability distribution of degrees for each node in the whole network, which is expressed with $P(k)$.

For an ER network, if each pair of nodes are connected with the same probability p, it turns out that the degree distribution follows a Poisson distribution ($P(k) = e^{-\langle k \rangle} \langle k \rangle^k / k!$). The characteristic of an ER network is that most of the nodes have about the same number of degrees around average degree $\langle k \rangle$.

1.1.2 Scale-free (SF) networks

As mentioned above, it is found that the ER network model fails to describe many real-world systems. For examples, ER model cannot help to explain why computer viruses are able to survive in the Internet for a very long time. In 1999, Barabási investigated several real-world networks and found that their degree distributions follow power-law distribution, and he proposed the scale-free (SF) network model to explain networks with power-law distribution and thus do not have a typical scale of degree.

For a SF network, the degree distribution follows a power law distribution ($P(k) \sim k^{-\lambda}$). λ, which is called the power law exponent, is a constant for a network. The property of a SF network is that most of the nodes have very low degrees, while a few nodes have very high degrees. Those nodes with high degrees are called hubs. It was found that for most real-world networks, like the Internet, friendship social networks, world-wide-web, scientific citation networks, gene-regulation network, airline networks, protein-protein interaction networks, and so on, the power law exponent λ lies between 2 and 3, which suggests some universal mechanisms behind the formation of these networks.

1.2 Generating Function and Percolation Model

1.2.1 Generating function method

A component in the network is a group of nodes connected internally, but disconnected from other components. For many degree distributions the network is composed of many separate components. In each such component there exists a path between any two nodes in the component, but no path exists between nodes in different components. The size of a component is the number of nodes in the component. When there is a component with size proportional to the size of the entire network, it is called the giant component (sometimes also called percolating cluster or giant cluster).

The generating function method is a general and useful method for determining the existence and the size of the giant component in the network. This powerful approach was

first developed by Newman, Watts, and Strogatz to study the structure and properties of complex networks [3]. They used the generating function method to study the size of the giant component as well as the component size distribution.

For a complex network with degree distribution $P(k)$, the generating function of the degree distribution is

$$G_0(x) = \sum_{k=0}^{\infty} P(k)x^k. \tag{1.1}$$

The probability of reaching a node with degree k by following a specific link is $kP(k)/\langle k\rangle$, and the corresponding generating function of those probabilities is :

$$G_1(x) = \frac{\sum kP(k)x^{k-1}}{\sum kP(k)} = \frac{d}{dx}G_0(x)/\langle k\rangle. \tag{1.2}$$

$H_1(x)$, the generating function for the probability of reaching a branch of a given size by following a link, satisfies a self-consistent equation:

$$H_1(x) = xG_1(H_1(x)). \tag{1.3}$$

Since $G_0(x)$ is the generating function for the degree of a node, the generating function for the probability that a node belongs to an n-node component is:

$$H_0(x) = xG_0(H_1(x)). \tag{1.4}$$

The size of the giant component is $P_\infty = 1 - H_0(1)$.

1.2.2 Percolation model

It is well known that in grids and other organized lattices, in any dimension larger than one, a percolation transition occurs. The percolation model assumes that sites(nodes) or bonds(links) in the lattice are occupied with some probability p. The system is considered percolating if a path exists from one side of the lattice to the other. The percolation phase transition occurs at some critical density p_c that depends on the type and dimensionality of the lattice.

For complex networks, the ideas of percolation theory can still be applied to obtain useful results. The only difference is that instead of using a spanning cluster that spans

over the whole lattice, we now use a giant component, whose size is proportional to the size of the entire network, to characterize the condition of percolation. The condition of the existence of a giant component above the percolation threshold and its absence below the threshold also applies to lattices, and therefore can be considered as more general than the spanning property.

Percolation model can be used to study the robustness of the network. We assume that only nodes belong to the giant component are functional in the network. When some nodes in the network are attacked and removed for some reason, other nodes that are connected to the network through those nodes will also be disconnected. If the attack is severe enough, many nodes will be disconnected and the entire network might breakdown and stop functioning. As one keeping attacking more and more nodes, phase transition will occur at some critical threshold below which a giant component exists in the network (thus the network is functional), and above which the network will collapse. The robustness of the network under attacks depends crucially on the nature of the attack and the structure of the network.

For example, the attack could be random, which means each node in the network is attacked with the same probability $1 - p$ (thus the probability for each node to survive the attack is p). If we begin with a distribution of degrees $P_0(k_0)$, the new distribution of degrees in the network after the attack will be:

$$P(k) = \sum_{k_0=k}^{\infty} P_0(k_0) C_{k_0}^k p^k (1-p)^{k_0-k}. \tag{1.5}$$

Calculating the first two moments for this distribution, given $\langle k_0 \rangle$ and $\langle k_0^2 \rangle$ for the original distribution before attack, we have:

$$\langle k \rangle = \sum_{k=0}^{\infty} P(k) k = p \langle k_0 \rangle. \tag{1.6}$$

Similarly, we have:

$$\langle k^2 \rangle = \sum_{k=0}^{\infty} P(k) k^2 = p^2 \langle k_0^2 \rangle + p(1-p) \langle k_0 \rangle. \tag{1.7}$$

And the criterion for criticality is

$$\frac{\langle k^2 \rangle}{\langle k \rangle} = \frac{p^2 \langle k_0^2 \rangle + p(1-p)\langle k_0 \rangle}{p\langle k_0 \rangle} = 2. \tag{1.8}$$

Reorganizing the above equation gives us the critical threshold for percolation:

$$p_c = \frac{1}{\langle k_0^2 \rangle / \langle k_0 \rangle - 1}. \tag{1.9}$$

1.3 A Network of Networks

Previous work in network research has focused primarily on analyzing single networks that do not interact with other networks, despite the fact that this is not the case for many real-world scenarios. In 2010, an analytical framework for studying the percolation properties of interacting networks has been introduced [10]. The percolation properties of a network of networks differ greatly from those of single isolated networks. In particular, although networks with broad degree distributions, e.g., scale-free networks, are robust when analyzed as single networks, they become vulnerable in a network of networks (NON) [14, 15]. Moreover, because the constituent networks of an NON are connected by node dependencies, an NON is subject to cascading failure. When there is strong interdependent coupling between networks, the percolation transition is discontinuous (is a first-order transition), unlike the well-known continuous second-order transition in single isolated networks.

1.3.1 The model

Because previous models deal almost exclusively with individual networks treated as isolated systems, many challenges remain. In many real-world systems an individual network is one component within a much larger complex multi-level network (is part of a network of networks). Node failure in one network will cause the failure of dependent nodes in other networks, and vice-versa. This recursive process can lead to a cascade of failures throughout the network of the networks system. The study of individual particles has enables physicists to understand the properties of a gas, but in order to describe a liquid or a solid the interactions between the particles also need to be understood. This is similar in network

theory. To adequately model most real-world systems, understanding the interdependence of networks and its effect on the structural and functional behaviors of the system is significant.

In order to model interdependent networks, we consider two networks, A and B [15] in which the functionality of a node in network A is dependent upon the functionality of one or more nodes in network B (see Figure 1.1), and vice-versa. The direction of a dependency link specifies the dependency of the nodes it connects. For example, link $A_i \rightarrow B_j$ provides a critical resource from node A_i to node B_j. If node A_i stops functioning due to attack or failure, node B_j will stop functioning as well but not vice-versa. Similarly, link $B_i \rightarrow A_j$ provides a critical resource from node B_i to node A_j.

We begin our attacking process by removing a fraction $1 - p$ of network A nodes and all A-edges connected to these nodes. All the nodes in network B that are connected to the removed A-nodes by $A \rightarrow B$ links are also removed since they depend on the removed nodes in network A. As mentioned above, when those nodes in network A are removed and stop functioning, those nodes in network B that depend on them will also stop functioning. The B-edges of these nodes in network B are also further removed, which will cause the removal of additional nodes in network A that are connected to the removed B-nodes by $B \rightarrow A$ links. As a result, a cascade of failures occurs in the network of networks system. As nodes and edges are removed, each network breaks down into connected components. Our assumption based on percolation theory is that only nodes in the giant component will keep functioning, and nodes belonging to small component will become non-functional. Thus in interdependent networks only the giant mutually-connected component is of interest.

1.3.2 Theoretical framework

Without losing generality, we now consider two partially-interdependent networks. This framework consists of two networks A and B with the number of nodes N_A and N_B, respectively. The nodes are randomly connected with degree distribution $P_A(k)$ in network A and $P_B(k)$ in network B. In addition, a fraction q_A of network A nodes depend on nodes in network B and a fraction q_B of network B nodes depend on nodes in network A. Here, we

assume that a node from one network depends on no more than one node from the other network. Also, we assume that if A_i depends on B_j, and B_j depends on A_k, then $k = i$ (the "no-feedback" condition).

Initially, we attack the network A by removing a fraction $1 - p$ of nodes (see Figure 1.2 [15]). Thus the remaining fraction of network A nodes after the initial attack is $\psi_1^{'} \equiv p$. Once a fraction $1 - p$ of nodes is randomly attacked and removed from a network, the generating function of the network remains the same, but must be computed from a new argument $z \equiv px + 1 - p$, where x is the argument for the original network. Thus the fraction of nodes that belongs to the giant component in network i is given by

$$P_{\infty,i} = pg_i(p), \tag{1.10}$$

where

$$g_i(p) = 1 - G_i[pf_i(p) + 1 - p], \tag{1.11}$$

and $f_i(p)$ satisfies

$$f_i(p) = H_i[pf_i(p) + 1 - p]. \tag{1.12}$$

In the case of two networks, the remaining functional part of network A contains a fraction $\psi_1 = \psi_1^{'} g_A(\psi_1^{'})$ of the network nodes, where $g_A(\psi_1^{'})$ is defined by Eqs.(1.11) and (1.12). Since a fraction q_B of nodes in network B depends on nodes in network A, the number of nodes in network B that become nonfunctional is $(1 - \psi_1)q_B = q_B[1 - \psi_1^{'} g_A(\psi_1^{'})]$. So the remaining fraction of network B nodes is $\phi_1^{'} = 1 - q_B[1 - \psi_1^{'} g_A(\psi_1^{'})]$, and the fraction of nodes in the giant component of network B is $\phi_1 = \phi_1^{'} g_B(\phi_1^{'})$.

Following this process, we can obtain $\psi_t^{'}$ and $\phi_t^{'}$, the remaining fraction of nods at each stage of the cascade of failures:

$$\begin{aligned}
&\psi_1^{'} \equiv p, \\
&\phi_1^{'} = 1 - q_B[1 - pg_A(\psi_1^{'})], \\
&\psi_t^{'} = p[1 - q_A(1 - g_B(\phi_{t-1}^{'}))], \\
&\phi_t^{'} = 1 - q_B[1 - pg_A(\psi_{t-1}^{'})].
\end{aligned} \tag{1.13}$$

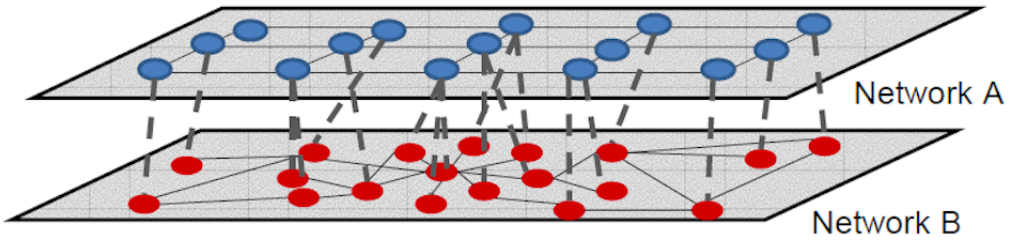


Figure 1.1: Example of two interdependent networks. Nodes in network B (computer network) are dependent on nodes in network A (power grid) for power; on the other hand, nodes in network A are also dependent on network B for control information.

To determine the state of the system at the end of the cascade process, we set $x = \psi_t^{'} = \psi'_{t+1}$ and $y = \phi_t^{'} = \phi_{t+1}^{'}$. This is because when system arrives at the stationary state, eventually the clusters stop fragmenting and the fractions of randomly removed nodes at step t and $t+1$ are equal. So when we arrive at the stationary state, we have two equations with two unknowns:

$$\begin{aligned} x &= p\{1 - q_A[1 - g_B(y)]\}, \\ y &= 1 - q_B[1 - g_A(x)p]. \end{aligned} \tag{1.14}$$

The giant component of networks A and B at the end of the cascade of failures are, respectively, $P_{\infty,A} = \psi_\infty = xg_A(x)$ and $P_{\infty,B} = \phi_\infty = yg_B(y)$. The equations can be solved numerically or graphically and the size of the giant component for each network can be obtained for each p.

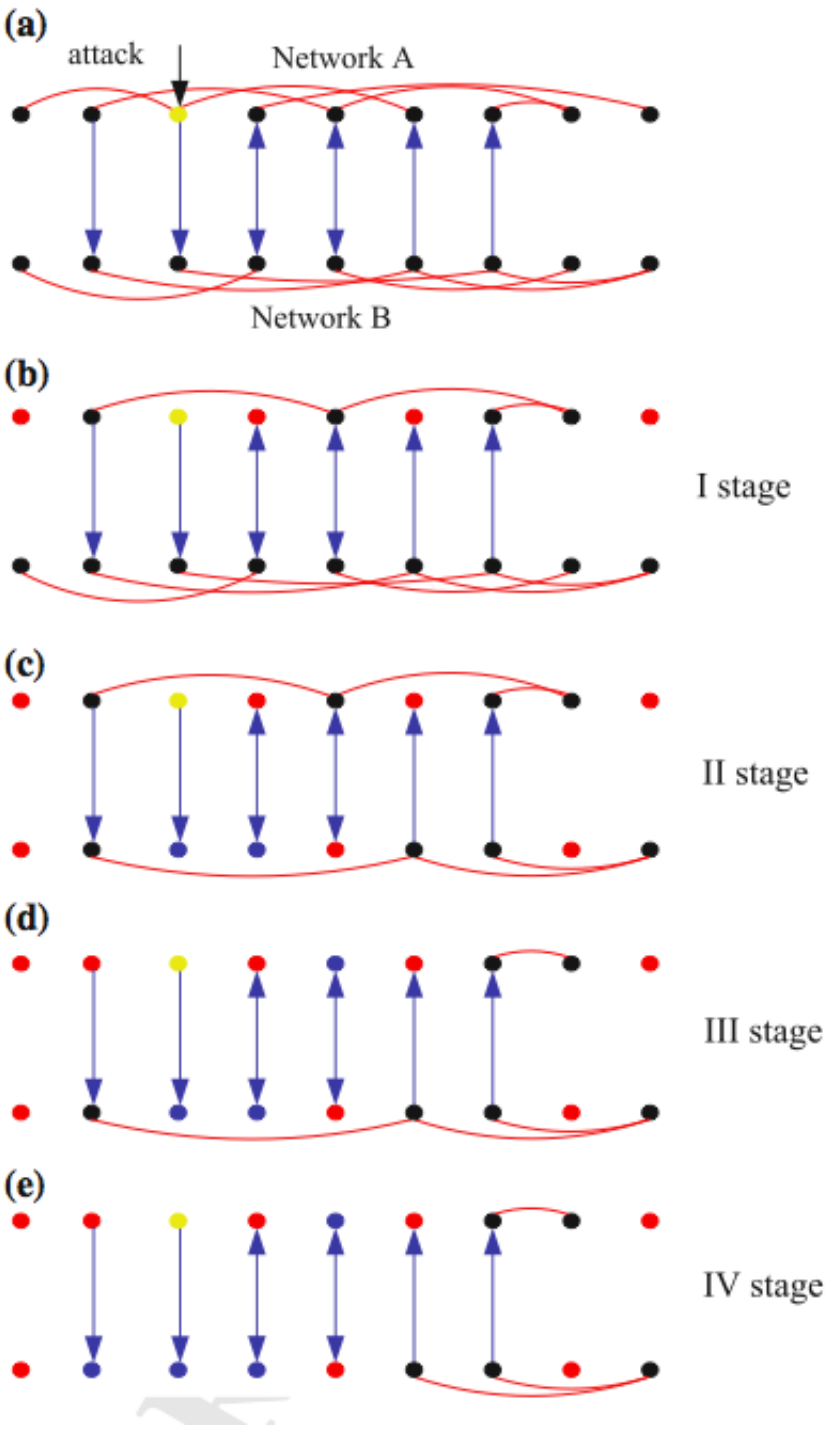


Figure 1.2: Description of the dynamic process of cascading failures on two partially interdependent networks. In the figure, the black nodes are the nodes that are alive, the yellow nodes are the initially attacked nodes, the red nodes are the nodes removed because they do not belong to the giant component, and the blue nodes are the nodes removed because the nodes they depend upon in the other network are removed.

Chapter 2

Robustness of Complex Networks under Localized Attacks

2.1 Introduction

The functioning of complex networks such as the Internet, airline routes, and social networks is crucially dependent upon the interconnections between network nodes. These interconnections are such that when some nodes in the network fail, others connected through them to the network will also be disabled and the entire network may collapse. In order to understand network robustness and design resilient complex systems, one needs to know whether a complex network can continue to function after a fraction of its nodes have been removed either through node failure or malicious attack [1–21]. This question is dealt with in percolation theory [21–24] in which the percolation phase transition occurs at some critical occupation probability p_c. Above p_c, a giant component, defined as a cluster whose size is proportional to that of the entire network, exists; below p_c the giant component is absent and the entire network collapses. Only nodes in the giant component continue to function after the node-removal process.

The robustness of complex networks under attack is dependent upon the structure of the underlying network and the nature of the attack. Previous research has focused on two types of initial attack: random attack and hub-targeted attack. In a random attack each node in the network is attacked with the same probability [1–3, 8, 10, 21]. In a hub-targeted attack

the probability that high-degree nodes will be attacked is higher than that for low-degree nodes [1, 3, 4, 7, 12]. An important feature of the network structure is its degree distribution, $P(k)$, which describes the probability that each node has a specific degree k. Networks with different degree distributions behave very differently under different types of attack. For instance, the Internet, which shows a power law degree distribution, is extremely robust against random attack but vulnerable to hub-targeted attack [1, 2].

However these two types of attack—random attack and hub-targeted attack—do not adequately describe many real-world scenarios in which complex networks suffer from damage that is localized, i.e., a node is affected, then its neighbors, and then their neighbors, and so on (see Fig. 2.1). Examples include the effects of earthquakes, floods, or military attacks on infrastructure networks and the effects of a computer virus or malware on computer networks. Recent occurrences of the latter include attacks carried out by cybercriminals who create a "botnet", a cluster of neighboring "zombie computers" in a computer network and, by using them, are able to damage the entire network. An understanding of the effect of this kind of attack on the functioning of a network is still lacking.

Here we will analyze the robustness of complex networks sustaining this kind of localized attack in order to determine how much damage a network can sustain before it collapses, i.e., to find the percolation threshold p_c. We also want to predict the fraction of nodes that keep functioning after an initial attack of a fraction of $1 - p$ nodes, i.e., the relative size of the giant component (the order parameter), P_∞. Note that localized attack has been studied only on specific network structures [25] or on interdependent spatially embedded networks [26], but a general theoretical formalism for studying localized attacks on complex networks is currently missing.

Here we develop a mathematical framework for studying localized attacks on complex networks with arbitrary degree distribution and we find exact solutions for percolation properties such as the critical threshold p_c and the relative size of the giant component P_∞. In particular, we apply our framework to study and compare the robustness of three types of random networks, (i) Erdős-Rényi (ER) networks with a Poissonian degree distribution

($P(k) = e^{-\langle k \rangle}\langle k \rangle^k / k!$) [27], (ii) random-regular (RR) networks with a Kronecker delta degree distribution ($P(k) = \delta_{k,k_0}$), and (iii) scale-free (SF) networks with a power law degree distribution ($P(k) \sim k^{-\lambda}$) [5]. We find that the effect of a localized attack on an ER network is identical to that of a random attack. For an RR network, we find that the p_c of a localized attack is always smaller (i.e., more robust) than that of a random attack. However, the robustness of a SF network against localized attack is found to be critically dependent upon the power law exponent λ. Surprisingly, a critical exponent λ_c exists such that when $\lambda < \lambda_c$, for localized attack the network is significantly more vulnerable compared to random attack, with p_c being larger. While for $\lambda > \lambda_c$, the opposite is true.

2.2 The Model

Consider a random network with a degree distribution $P(k)$, which indicates the probability that a node in the network has k neighbors. The generating function of the degree distribution is defined as $G_0(x) = \sum_{k=0}^{\infty} P(k)x^k$ [28, 29]. We start from a randomly chosen "root" node. All nodes in the random network are listed in ascending order of their distances from this root node (see Fig. 2.1(a)). The shell l is defined as the set of nodes that are at distance l from the root node [30, 31]. Within the same shell, all nodes are at the same distance from the root node and are positioned randomly.

We initiate the localized attack process by removing the root node, then the nodes in the first shell, and so on. We remove nodes in the ascending order of their distances from the root node. Within the same shell we remove nodes randomly and, after nodes in shell l are fully removed, we begin removing nodes in shell $l + 1$. We continue the localized attack process until a fraction $1 - p$ of nodes in the entire network are removed. Thus a "hole" of attacked nodes forms around the root node. The remaining p fraction of nodes in the network are those at greater distances from the root node (see Fig. 2.1(b)). After the initial removal of $1 - p$ fraction of the network nodes and all links connected to them, the remaining network fragments into connected clusters. As in percolation theory [22, 23], only nodes in the giant component (the largest cluster) are still functional. Nodes belonging to

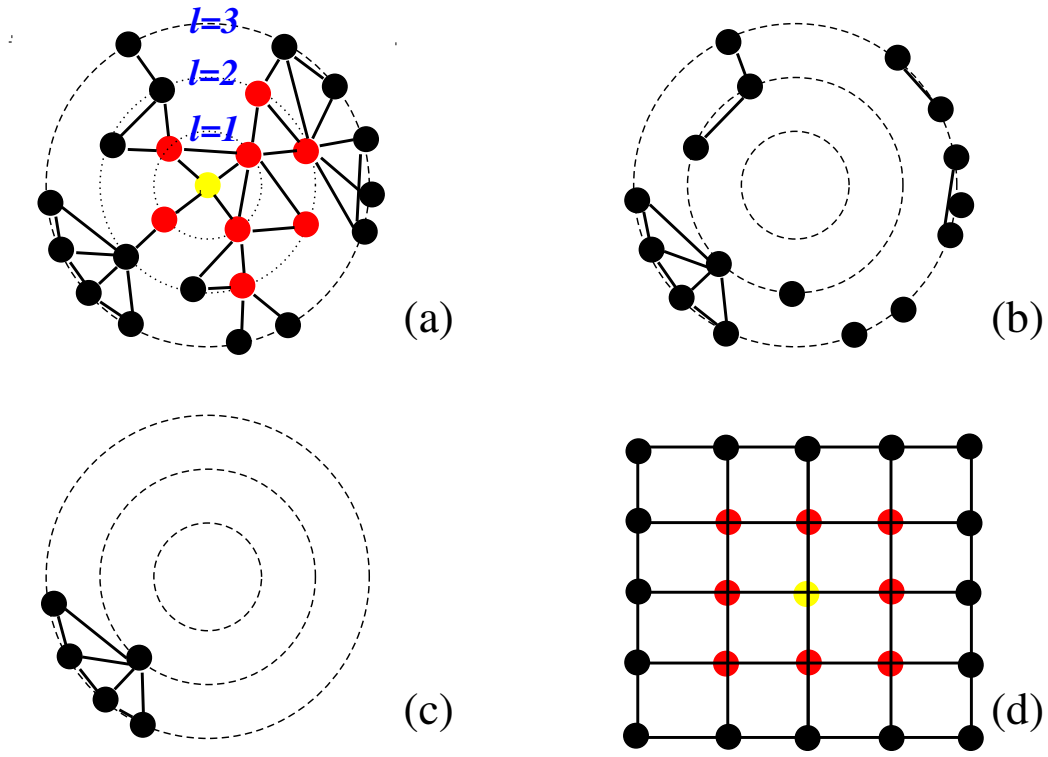


Figure 2.1: Schematic illustration of the localized attack process. (a) A fraction $1-p$ of the nodes are chosen to be removed, starting from the root node, its nearest neighbors, next nearest neighbors, and so on (yellow represents the root node, red the other nodes to be removed). (b) Remove the chosen nodes and the links. An attacked "hole" centered around the root node is formed. (c) Only nodes in the giant component (largest cluster) keep functioning and are left in the network. (d) Localized attack on regular lattice (here, square lattice). For a regular lattice with $N \to \infty$, one needs to attack all nodes in order to collapse the network, i.e., $p_c \to 0$.

other small clusters are considered non-functional and are also removed (see Fig. 2.1(c)). Note that for localized attack on a regular lattice, as the number of network nodes $N \to \infty$, $p_c \to 0$, i.e., one has to attack all nodes in the regular lattice in order to collapse the lattice (see Fig. 2.1(d)).

2.3 Theoretical Framework

Consider a random network with arbitrary degree distribution $P(k)$, which represents the probability of a node in the network to have k links. The corresponding generating function is defined as

$$G_0(x) = \sum_{k=0}^{\infty} P(k)x^k. \tag{2.1}$$

We separate the process of a localized attack into two stages: (i) at the first stage, we remove all the nodes belonging to the attacked area but keep the links connecting the removed nodes to the remaining nodes; (ii) at the second stage, we remove those links. Now consider the degree distribution $P_p(k)$ of the remaining nodes after the first stage. Following Ref. [31] and letting $A_p(k)$ be the number of nodes with degree k in the remaining network, we have

$$P_p(k) = \frac{A_p(k)}{pN}. \tag{2.2}$$

With one more node being removed, $A_p(k)$ changes as

$$A_{(p-1/N)}(k) = A_p(k) - \frac{P_p(k)k}{\langle k(p)\rangle}, \tag{2.3}$$

where $\langle k(p)\rangle \equiv \sum P_p(k)k$. In the limit $N \to \infty$, Eq. (2.3) can be presented in terms of a derivative of $A_p(k)$ with respect to p,

$$\frac{dA_p(k)}{dp} = N\frac{P_p(k)k}{\langle k(p)\rangle}. \tag{2.4}$$

By differentiating Eq. (2.2) with respect to p and plugging it into Eq. (2.4), we have

$$p\frac{dP_p(k)}{dp} + P_p(k) - \frac{P_p(k)k}{\langle k(p)\rangle} = 0. \tag{2.5}$$

The solution of Eq. (2.5) can be expressed as

$$P_p(k) = P(k)\frac{f^k}{G_0(f)}, \tag{2.6}$$

and the average degree of the remaining network is

$$\langle k(f)\rangle = \frac{fG_0^{'}(f)}{G_0(f)}, \tag{2.7}$$

where $f \equiv G_0^{-1}(p)$. Thus the generating function of $P_p(k)$ is

$$G_a(x) \equiv \sum_k P_p(k)x^k = \frac{G_0(fx)}{G_0(f)}. \tag{2.8}$$

Now consider the second stage of removing the links of the remaining nodes which lead to the removed nodes. The number of links belonging to the nodes on the outer shell of the attacked hole, $L(f)$, can be expressed as [31],

$$L(f) = N(G_0^{'}(1)f^2 - G_0^{'}(f)f). \tag{2.9}$$

Since loops are allowed in the random network model, those links can be connected either to the remaining nodes or to other nodes on the same outer shell of the attacked hole. The number of links of the remaining nodes which lead to the removed nodes is

$$\tilde{L}(f) = L(f)\frac{Np\langle k(f)\rangle}{Np\langle k(f)\rangle + L(f)} = N[fG_0^{'}(f) - \frac{G_0^{'}(f)^2}{G_0^{'}(1)}]. \tag{2.10}$$

The probability that a link in the remaining network will end at an unremoved node is equal to

$$\tilde{p} = 1 - \frac{\tilde{L}(f)}{pN\langle k(f)\rangle} = \frac{G_0^{'}(f)}{G_0^{'}(1)f}. \tag{2.11}$$

Because the network is randomly connected, removing the links that end at the removed nodes is equivalent to randomly removing a $1-\tilde{p}$ fraction of links of the remaining network. The generating function of the remaining network after the random removal of a $1-\tilde{p}$ fraction of links is equal to [32]

$$G_0^p(x) \equiv G_a(1-\tilde{p}+\tilde{p}x) = \frac{1}{G_0(f)}G_0[f + \frac{G_0^{'}(f)}{G_0^{'}(1)}(x-1)], \tag{2.12}$$

where $f \equiv G_0^{-1}(p)$. Note that Eq. (2.13) is the generating function of the degree distribution of the remaining network after a localized attack.

2.4 Comparison with random attack

We find that the generating function of the degree distribution of the remaining network after the localized attack is (see supplementary information)

$$G_0^p(x) = \frac{1}{G_0(f)} G_0[f + \frac{G_0'(f)}{G_0'(1)}(x-1)], \tag{2.13}$$

where p is the fraction of unremoved nodes and $f \equiv G_0^{-1}(p)$. The critical probability p_c where the network collapses and the size of the giant component $P_\infty(p)$ for $p > p_c$ can be derived analytically from Eq. (2.13). The generating function of the cluster sizes in the remaining network is $H_0^p(x) = xG_0^p(H_1^p(x))$, where $H_1^p(x)$ satisfies the transcendental equation $H_1^p(x) = xG_1^p(H_1^p(x))$ and $G_1^p(x) = G_0^{'p}(x)/G_0^{'p}(1)$ [28]. By combining Eq. (2.13) and the criterion for the network to collapse [2, 3], $G_1^{'p}(1) = 1$, we find that p_c satisfies

$$G_0''(G_0^{-1}(p_c)) = G_0'(1). \tag{2.14}$$

The size of the giant component $S(p)$ as a fraction of the remaining network satisfies

$$S(p) = 1 - G_0^p(H_1^p(1)), \tag{2.15}$$

where $H_1^p(1)$ satisfies $H_1^p(1) = G_1^p(H_1^p(1))$. The relative size of the giant component as a fraction of the original network is $P_\infty(p) = pS(p)$.

We apply the above mathematical framework to three types of complex networks: Erdős-Rényi (ER) networks, random-regular (RR) networks, and scale-free (SF) networks, and compare the results of a localized attack with those of a random attack.

For an ER network with an average degree $\langle k \rangle$, the degree distribution follows a Poissonian distribution $P(k) = e^{-\langle k \rangle}\langle k \rangle^k / k!$ and the corresponding generating function of degree distribution is $G_0(x) = e^{\langle k \rangle (x-1)}$. From Eq. (2.13) we have $G_0^p(x) = e^{p\langle k \rangle (x-1)}$, which is the same as the generating function of the degree distribution for the remaining network after a random attack. Thus the effect of a localized attack is exactly the same as that of a random attack on an ER network (see Fig. 2.2(a)), and the critical threshold is $p_c = 1/\langle k \rangle$. The size of the giant component $P_\infty(p)$ satisfies $P_\infty(p) = p(1 - e^{-\langle k \rangle P_\infty(p)})$. In an RR

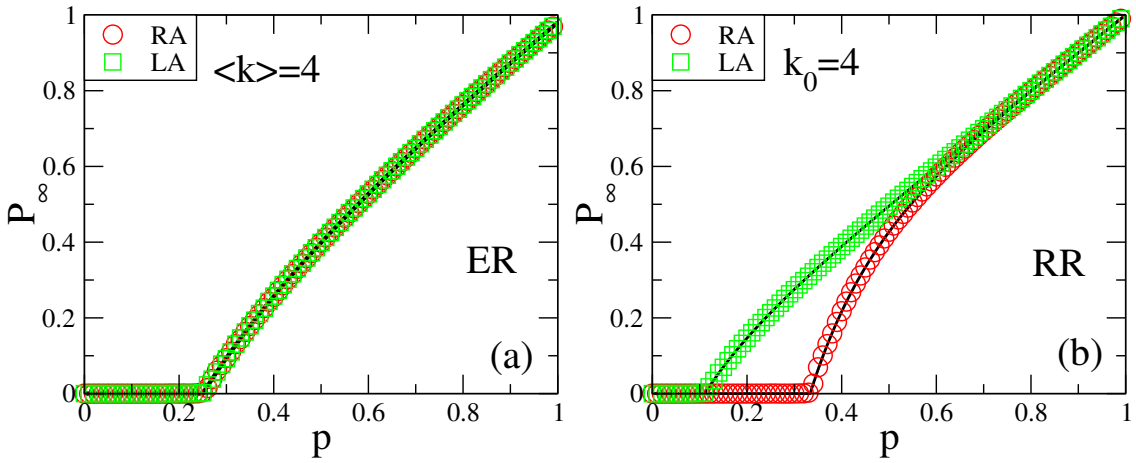


Figure 2.2: Percolation transitions for (a) an ER network and (b) an RR network under localized attack (LA) and random attack (RA), with network size $N = 10^6$, average degree $\langle k \rangle = 4$ in ER network, and $k_0 = 4$ in RR network. Theoretical results (solid lines) and simulations (symbols) agree well with each other. Note that the effect of localized attack and random attack on an ER network (see (a)) are identical (here, $p_c = 1/\langle k \rangle = 0.25$), while an RR network (see (b)) is more robust against localized attack compared to random attack.

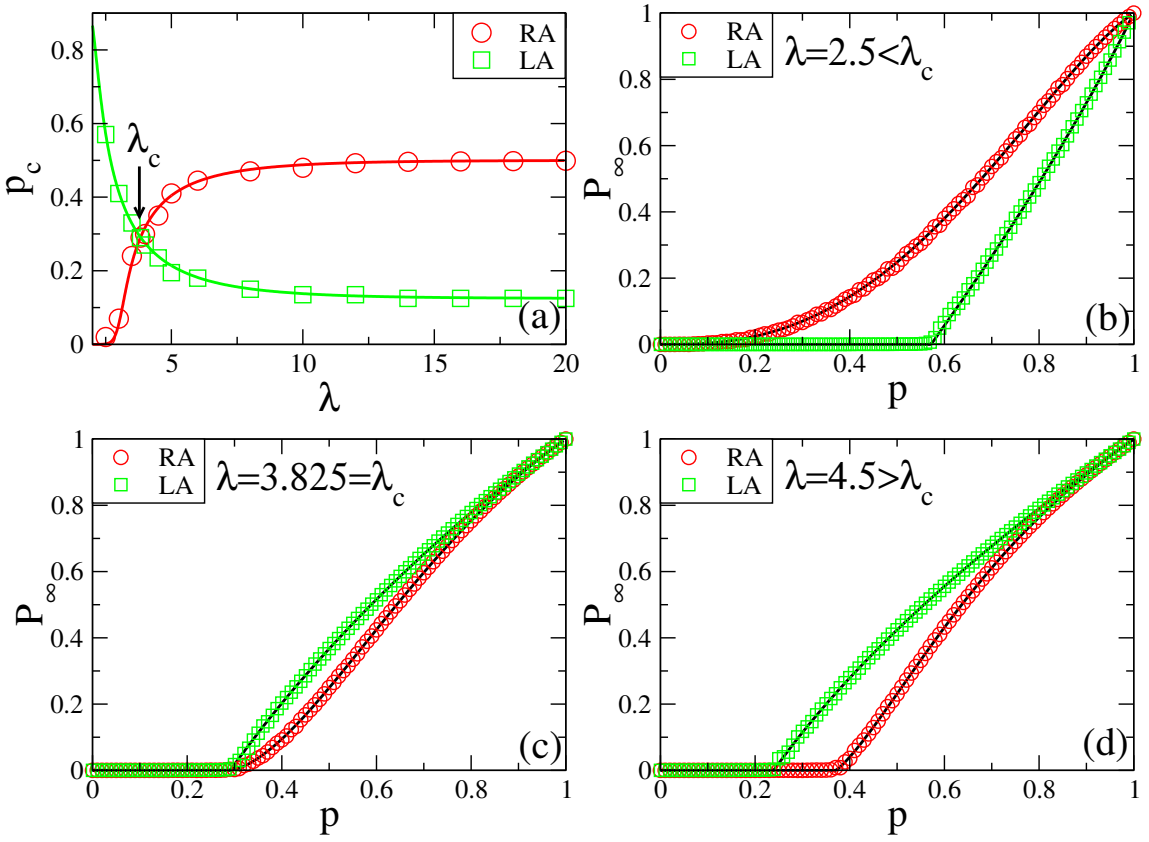


Figure 2.3: Percolation properties for a SF network under localized attack (LA) and random attack (RA). Solid lines are from theory (Eq. (2.13)) and symbols represent simulation results with $N = 10^6$, $m = 2$, and $\langle k \rangle = 3$. (a) Critical threshold p_c as a function of degree exponent λ. When $\lambda \to \infty$, the SF network converges to an RR network with $k_0 = \langle k \rangle = 3$, so $p_c(RA) \to 1/(k_0 - 1) = 0.5$ and $p_c(LA) \to (k_0 - 1)^{-\frac{k_0}{k_0 - 2}} = 0.125$, as confirmed in simulations. Note that for $2 < \lambda \leq 3$, $p_c \to 0$ in the thermodynamic limit ($N \to \infty$) for random attack [2]. (b) When $\lambda < \lambda_c$, the SF network is more vulnerable to localized attack compared to random attack. (c) When $\lambda = \lambda_c$, p_c for localized attack and for random attack are equal. (d) When $\lambda > \lambda_c$, the SF network is more robust against localized attack compared to random attack.

network each node is connected to k_0 other nodes randomly and the generating function of the degree distribution is $G_0(x) = x^{k_0}$. Using Eq. (2.14) we find that the critical threshold for a localized attack on an RR network is

$$p_c = (k_0 - 1)^{-\frac{k_0}{k_0-2}}. \tag{2.16}$$

Note that for an RR network under random attack the critical threshold is $p_c = (k_0 - 1)^{-1}$. Thus, for $k_0 > 2$, p_c under localized attack is always smaller than p_c under random attack (see Fig. 2.2(b)). This means that an RR network is more resilient against localized attack than against random attack. When $k_0 \gg 1$, random and localized attacks have the same critical threshold ($p_c = 1/(k_0 - 1)$), since in this limit every node is a neighbor of the root node and there is no difference between random and localized attacks. Since $\lim_{k_0 \to 2} p_c = e^{-2} \approx 0.135$ and $\lim_{k_0 \to \infty} p_c = 0$, one can see that p_c for a localized attack on an RR network is always within the range $(0, e^{-2})$ for all $k_0 > 2$. For $p > p_c$, from Eq. (2.15), the relative size of the giant component $P_\infty(p)$ satisfies

$$(p - P_\infty(p))^{\frac{1}{k_0}} - p^{\frac{1}{k_0}} = (p - P_\infty(p))^{\frac{k_0-1}{k_0}} - p^{\frac{k_0-1}{k_0}}. \tag{2.17}$$

For a SF network the degree distribution is $P(k) \sim k^{-\lambda}$ ($m \leq k \leq M$), where m and M are the lower and upper bound of the degree, respectively, and λ is the power exponent. The critical threshold p_c and the size of the giant component $P_\infty(p)$ are solved numerically by using the theoretical framework developed in Eq. (2.13) (see Fig. 2.3). We find that the degree heterogeneity plays an important role in the robustness of SF networks against localized attack. The critical threshold p_c and the size of the giant component $P_\infty(p)$ for the percolation transition of the SF network under localized attack depends on λ. We find that in a SF network there is a critical value λ_c below which a localized attack is significantly more severe than a random attack, but when $\lambda > \lambda_c$ a random attack is more severe. Indeed, as seen in Fig. 2.3(a), for $\lambda < \lambda_c$, p_c for a localized attack is significantly higher than for a random attack. As λ increases and the network becomes less heterogeneous, p_c decreases and the network becomes more robust against localized attacks. The specific value of λ_c

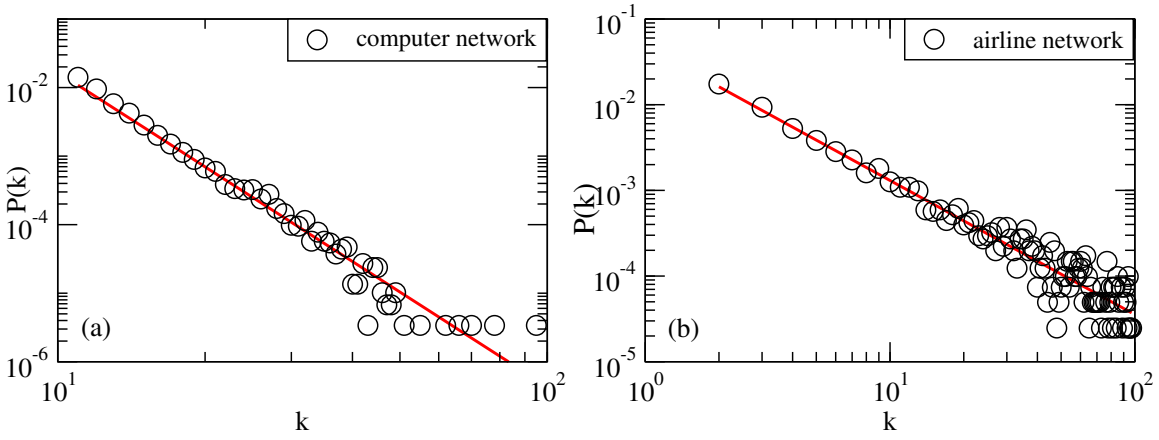


Figure 2.4: Degree distribution of (a) the peer-to-peer computer network with N=62586, $\langle k \rangle = 4.73$ and $\lambda = 4.59$, and (b) the global airline route network with N=3308, $\langle k \rangle = 12.2$ and $\lambda = 1.57$. The degree distribution of both network approximately follow power law distribution.

depends on other parameters, such as m, M, and $\langle k \rangle$. In Fig. 2.3(b)−(d), we plot the size of the giant component $P_\infty(p)$ as a function of p and compare the results of a localized attack with those of a random attack. One intuitive explanation for the dependence of network robustness on λ is that, on the one hand, there is a higher probability that higher degree nodes will be within the attacked hole, which accelerates the fragmentation of the SF network; on the other, only nodes on the surface of the attacked hole are connected to the remaining network and contribute to its breakdown, which mitigates the fragmentation process. The total impact of the localized attack is the result of the competition between these two effects. As λ increases and the SF network becomes less heterogeneous, the first effect becomes less dominant and the network becomes more robust. Our analytical analysis shows that for an ER network these two effects always compensate each other and yield equal effects from both localized attack and random attack. For an RR network, on the other hand, the degrees are all the same and therefore only the second effect exists, and the underlying network becomes more robust against localized attack than against random attack.

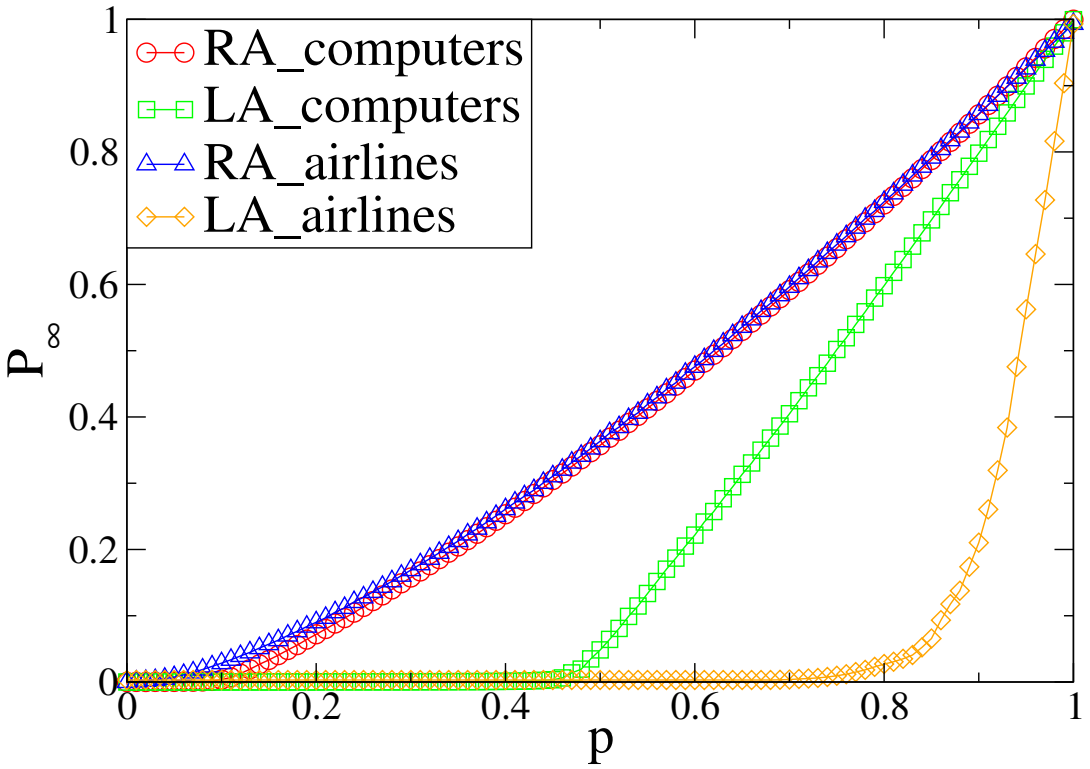


Figure 2.5: Robustness of real-world networks against localized attack (LA) and random attack (RA). A comparison of localized attack and random attack on a peer-to-peer computer network and a global airline route network [33, 34]. The size of the giant component $P_\infty(p)$ after locally attacking $1-p$ fraction of the whole network, versus p. The circles (red) and squares (green) represent simulation results of the peer-to-peer computer network ($N = 62586$, $\langle k\rangle = 4.73$ and $\lambda = 4.59$) under random attack and localized attack respectively. The triangles (blue) and the diamonds (orange) represent simulation results of the global airline route network ($N = 3308$, $\langle k\rangle = 12.2$ and $\lambda = 1.57$) under random attack and localized attack respectively. The simulation results are the average over 100 and 1000 realizations for the computer network and the airline network respectively.

2.5 Localized Attack on Real-world Networks

We test and compare the robustness of real-world networks against localized attack and random attack using a peer-to-peer computer network [33] and a global airline route network [34]. The degree distributions of both networks approximately follow power law (see Figure 2.4). Figure 2.5 shows that, for both real-world networks, localized attack can collapse the network much more easily: a node failure of 30% in the global airline route network and 55% in the peer-to-peer computer network can disable the total network. When the attack is random, however, a node failure of 98% in the global airline route network and 90% in the peer-to-peer computer network must occur before the network collapses. This shows that a localized attack is significantly more harmful to real-world SF networks than a random attack, supporting our theoretical results for SF networks with $\lambda < \lambda_c$.

2.6 Localized Attack on Interdependent Networks

Here, without losing generality, we consider two interdependent networks A and B with the same number of nodes N [10]. Within each network, the nodes are randomly connected with degree distribution $P_A(k)$ and $P_B(k)$, respectively. Each node in network A depends on a random corresponding node in network B, and vice versa. This means if the node in network B upon which the node in network A depends stop functioning, the corresponding node in network A will also stop functioning. Besides, we assume here that if a node i in network A depends on a node j in network B and j depends on a node l in network A, then $l = i$ (no-feedback condition [14–16]). We start our localized attack process by initially removing a fraction $1 - p$ of nodes in network A shell by shell, and remove all the links that connect to those removed nodes. Network A starts to fragment into connected components as nodes and links are removed and nodes that are not connected to the giant component are considered inactive and are also removed. Owing to the dependency, all the nodes in network B that depend on the removed nodes in network A are also removed. Network B also starts to fragment and only nodes in the giant component are kept. Then network B

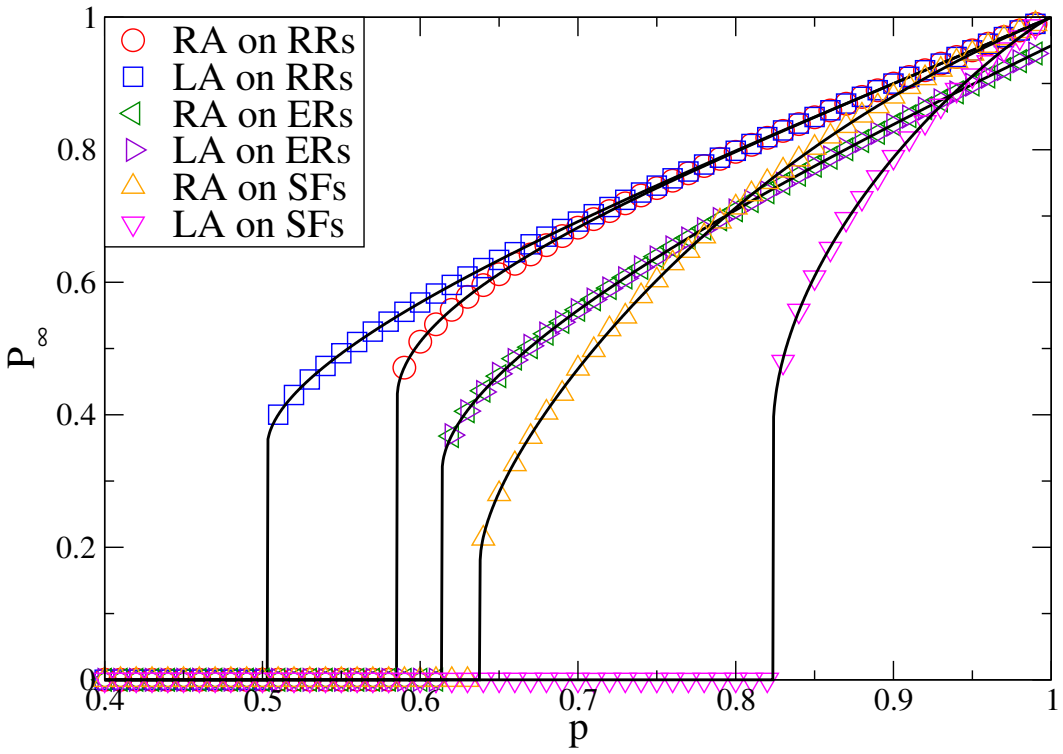


Figure 2.6: Comparison of percolation transitions for a pair of interdependent networks under localized attack and random attack. For RR networks, each node within the networks is randomly connected to $k_0 = 4$ other nodes. For ER networks, the average degree of the Poisson distribution is $\langle k \rangle = 4$. For SF networks, the lower and upper bound of the degree distribution are $m = 2$ and $M = 1000$, respectively. The power exponent of the degree distribution is $\lambda = 2.5$. Solid lines are from theoretical predictions and symbols represent simulations with network size $N = 10^6$. Note that the simulations results are in good agreement with the theory.

spreads damage back to network A, and back and forth, until the two networks completely fragment or form a mutually connected giant component [10].

The only difference between the cascading process under a localized attack and the case under a random attack is the type of initial attack on network A. After that, all the processes are similar. If we find a network $\tilde{A}$ with generating function $\tilde{G}_0(x)$, such that after a random attack with removing $1-p$ fraction of nodes, the generating function of the remaining network is the same as $G_0^p(x)$, then the localized attack problem on network A and B can be mapped to a random attack problem on network $\tilde{A}$ and B. By using $\tilde{G}_0(1-p+px) = G_0^p(x)$ and from Eq. (2.13), we have

$$\tilde{G}_0(x) = \frac{1}{G_0(f)} G_0[f + \frac{G_0'(f)}{G_0'(1)G_0(f)}(x-1)], \tag{2.18}$$

where $f \equiv G_0^{-1}(p)$.

Next by using the framework developed in Ref. [10], we introduce a function for network A

$$g_A(p) = 1 - \tilde{G}_0[1 - p(1 - f_A(p))], \tag{2.19}$$

where $f_A(p)$ satisfies a transcendental equation

$$f_A(p) = \tilde{G}_1[1 - p(1 - f_A(p))], \tag{2.20}$$

and analogous equations hold for network B. After the system of the interdependent networks reaches stationarity, the fraction of nodes in the mutually giant component is P_∞, which satisfies

$$P_\infty = x g_B(x) = y g_A(y), \tag{2.21}$$

where x and y satisfy

$$x = p g_A(y), \quad y = p g_B(x). \tag{2.22}$$

By eliminating y from the equations, we obtain

$$x = p g_A[p g_B(x)]. \tag{2.23}$$

The critical case ($x = x_c, p = p_c$) emerges when the derivatives of both sides in Eq. (2.23) with respect to x equal each other,

$$1 = p^2 \frac{dg_A}{dx}[pg_B(x)]\frac{dg_B}{dx}(x) \mid_{x=x_c,p=p_c}, \tag{2.24}$$

which, together with Eq. (2.23), yields the solution for p_c and the critical size of the giant mutually connected component, $P_\infty(p_c) = x_c g_B(x_c)$.

We can solve the above equations numerically and compare with simulation results. As shown in Fig. 2.6, the size of the mutual giant component $P_\infty(p)$ are plotted for a pair of interdependent networks under a localized attack and a random attack. Note that the behavior of the phase transition is first order (abrupt) in contrast of being second order (continuous) for a single network. As expected, similar conclusion can be drawn from the comparison of the robustness of interdependent networks system under localized attacks and random attacks. While a pair of interdependent RR networks are more robust against localized attack, a pair of interdependent ER networks show the same robustness under two types of attacks. The robustness of a pair of interdependent SF networks is dependent on the heterogeneity of the degree distribution, i.e., λ. Most real-world coupled networks ($2 < \lambda \leq 3$) are easier to collapse under a localized attack than under a random attack.

2.7 Summary

To conclude, we have developed a mathematical framework for studying the percolation of localized attacks on complex networks with an arbitrary degree distribution. Using generating function methods, we have solved exactly for the percolation properties of random networks under localized node removal. Our results show that the effects of localized attack and random attack on an Erdős-Rényi network are identical. While a random-regular network is more robust against localized attack than against random attack, the robustness of a scale-free network depends on the heterogeneity of the degree distribution. When $\lambda < \lambda_c$, the SF network is found to be significantly more vulnerable with respect to localized attack compared to random attack. When $\lambda > \lambda_c$, the opposite is true. Our results can provide

insight into understanding the robustness of complex systems and facilitate the design of resilient infrastructures.

Chapter 3

Robustness of Fully Interdependent Networks with Clustering

3.1 Introduction

In a system of interdependent networks, the functioning of nodes in one network is dependent upon the functioning of nodes in other networks of the system. The failure of nodes in one network can cause nodes in other networks to fail, which in turn can cause further damage to the first network, leading to cascading failures and catastrophic consequences. Power blackouts across entire countries have been caused by cascading failures between the interdependent communication and power grid systems [35, 36]. Because infrastructures in our modern society are becoming increasingly interdependent, understanding how systemic robustness is affected by these interdependencies is essential if we are to design infrastructures that are resilient [37–40]. In addition to research carried out on specific systems [13, 41–46], a mathematical framework [10] and its generalizations [11, 12, 47] have been developed recently. These studies use a percolation approach to analyze a system of two or more interdependent networks subject to cascading failure [48, 49]. It was found that interdependent networks are significantly more vulnerable than their stand-alone counterparts. The dynamics of cascading failure are strongly affected by the structure patterns of network components and by the interaction between networks. This research has focused almost exclusively on random interdependent networks in which clustering within compo-

nent networks is small or approaches zero. Clustering quantifies the propensity for two neighbors of the same vertex to also be neighbors of each other, forming triangle-shaped configurations in the network [50–52]. Unlike random networks in which there is very little or no clustering, real-world networks exhibit significant clustering. Recent studies have shown that, for single networks, both bond percolation and site percolation in clustered networks have higher epidemic thresholds compared to the unclustered networks [53–58].

Here we present a mathematical framework for understanding how the robustness of interdependent networks is affected by clustering within the network components. We extend the percolation method developed by Newman [53] for single clustered networks to coupled clustered networks. We find that interdependent networks that exhibit significant clustering are more vulnerable to random node failure than networks without significant clustering. We are able to simplify our interdependent networks model—without losing its general applicability—by reducing its size to two networks, A and B, each having the same number of nodes N. The N nodes in A and B have bidirectional dependency links to each other, establishing a one-to-one correspondence. Thus the functioning of a node in network A depends on the functioning of the corresponding node in network B and vice versa. Each network is defined by a joint distribution P_{st} (generating function $G_0(x,y) = \sum_{s,t=0}^{\infty} P_{st} x^s y^t$) that specifies the fraction of nodes connected to s single edges and t triangles [53]. The conventional degree of each node is thus $k = s + 2t$. The clustering coefficient c is

$$\begin{aligned} c &= \frac{3 \times (\text{number of triangles in network})}{\text{number of connected triples}} \\ &= \frac{N \sum_{st} t P_{st}}{N \sum_k \binom{k}{2} P_k}. \end{aligned} \tag{3.1}$$

3.2 Site Percolation of Single Clustered Networks

We begin by studying the generating function of remaining nodes after a fraction of $(1-p)$ nodes is randomly removed from one clustered network. After the nodes are removed, we define t'_i to be the number of triangles of which node i is a part, d'_i to be the number

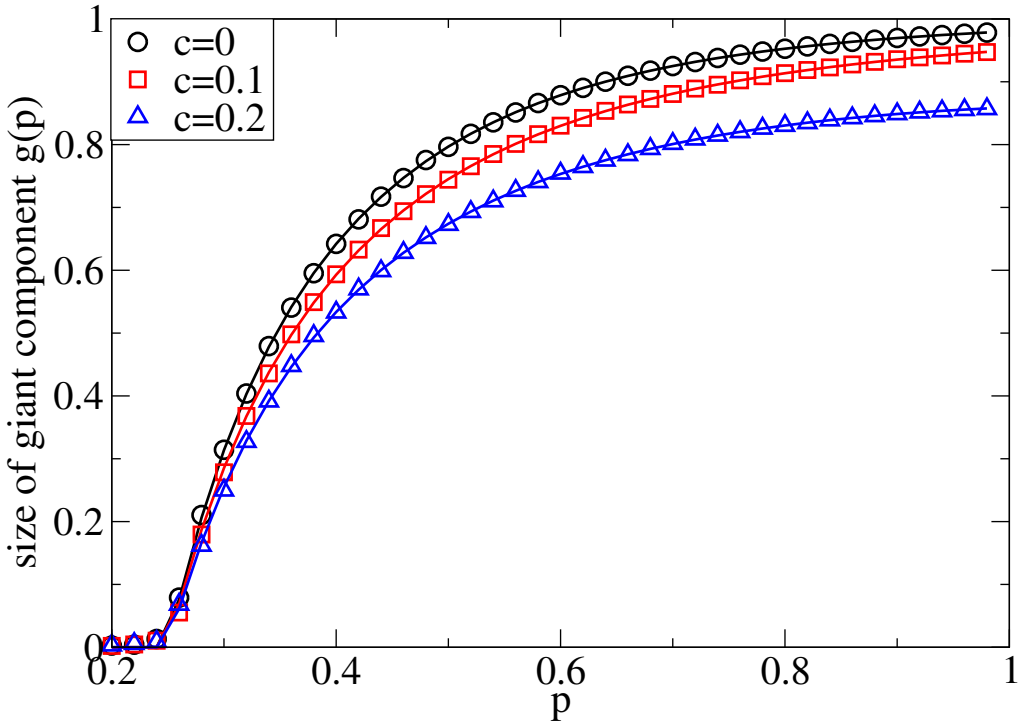


Figure 3.1: Size of giant component $g(p)$ in single networks with degree distribution Eq. (3.7) and average degree $\langle k \rangle = 4$, as a function of p, the fraction of remaining nodes after random removal of nodes. Curves are from theory Eq. 3.8, symbols are from simulation.

of single edges that form triangles prior to attack, and n_i' to be the number of stand-alone single edges prior to attack. This network is thus defined by the joint distribution $P_{n',t',d'}$. The probability that a node has n' single edges from single edges is the sum of all the probabilities that nodes with more than n' single edges will have exactly n' edges remaining, which is $Q_1(n') \equiv \sum_{s=n'}^{\infty} \left(\begin{array}{c} s \\ n' \end{array} \right) p^{n'}(1-p)^{s-n'}$. Similarly, the probability that a node has t' triangles is the sum of all the probabilities that nodes with more than t' triangles will have exactly t' triangles remaining. Since the probability that a triangle will survive is p^2, the sum is $Q_2(t') \equiv \sum_{t=t'}^{\infty} \left(\begin{array}{c} t \\ t' \end{array} \right) p^{2t'}(1-p^2)^{t-t'}$. The probability that a triangle corner will have one edge broken is $\frac{2p(1-p)}{1-p^2}$ and the probability that it will have both edges broken is $\frac{(1-p)^2}{1-p^2}$. Thus the probability that a node had d' single edges forming triangles prior to

their destruction is $Q_3(d') \equiv \begin{pmatrix} t-t' \\ d' \end{pmatrix} [\frac{2p(1-p)}{1-p^2}]^{d'} [\frac{(1-p)^2}{1-p^2}]^{t-t'-d'}$. Combining these three, we have the corresponding generating function

$$\begin{aligned} G(x,y,z,p) &= \sum_{n',t',d'} P_{n',t',d'} x^{n'} y^{t'} z^{d'} \\ &= \sum_{n'=0}^{\infty} x^{n'} Q_1(n') \sum_{t'=0}^{\infty} y^{t'} Q_2(t') \sum_{d'=0}^{t-t'} z^{d'} Q_3(d') P_{s,t} \\ &= G_0(xp+1-p, yp^2 + 2zp(1-p) + (1-p)^2). \end{aligned} \tag{3.2}$$

We define $s' = n' + d'$ to be the total number of single links of a node after attack. The joint degree distribution after attack is $P'_{s',t'}$ which satisfies $P'_{s',t'} = \sum_{n'=0}^{s'} P_{n',t',d'}$, with $d' = s' - n'$. The generating function of $P'_{s',t'}$ is

$$\begin{aligned} G_0(x,y,p) &= \sum_{s',t'} P'_{s',t'} x^{s'} y^{t'} \\ &= \sum_{s'=0}^{\infty} \sum_{n'=0}^{s'} \sum_{t'} P_{n',t',d'} x^{s'} y^{t'} \\ &= \sum_{n',d',t'} P_{n',t',d'} x^{n'} y^{t'} x^{d'} \\ &= G(x,y,x,p). \end{aligned} \tag{3.3}$$

Therefore, the generating function of the remaining network after attack is

$$G_0(x,y,p) = G_0(xp+1-p, yp^2 + 2xp(1-p) + (1-p)^2). \tag{3.4}$$

The size of the giant component $g(p)$ of the remaining network according to Ref. [53] is

$$g(p) = 1 - G_0(u, v^2, p), \tag{3.5}$$

where

$$\begin{aligned} u &= G_q(u, v^2, p), \\ v &= G_r(u, v^2, p), \end{aligned} \tag{3.6}$$

and $G_q(x,y,p) = \frac{1}{\mu} \frac{\partial G_0(x,y,p)}{\partial x}$, $G_r(x,y,p) = \frac{1}{\nu} \frac{\partial G_0(x,y,p)}{\partial y}$ where μ and ν are the average number of single links and triangles per node, respectively.

As an example, consider the case when $(1-p)$ fraction of nodes are removed randomly from a network with doubly Poisson degree distribution

$$P_{st} = e^{-\mu}\frac{\mu^s}{s!}e^{-\nu}\frac{\nu^t}{t!}, \tag{3.7}$$

where the parameters μ and ν are the average numbers of single edges and triangles per vertex, respectively. According to Eq. (3.1), the clustering coefficient is $c = \frac{2\nu}{2\nu+(\mu+2\nu)^2}$. Then, $G_0(x,y) = e^{\mu(x-1)}e^{\nu(y-1)}$ and $G_0(x,y,p) = G_q(x,y,p) = G_r(x,y,p) = e^{[\mu p+2p(1-p)\nu](x-1)}e^{\nu p^2(y-1)}$, and $u = v = 1-g(p)$, leading to

$$g(p) = 1 - e^{[\mu p+2p(1-p)\nu]g(p)}e^{\nu p^2(g(p)^2-2g(p))}. \tag{3.8}$$

This equation is a closed-form solution for the giant component $g(p)$ and can be solved numerically. The critical case appears when the derivatives of the both sides of Eq. (3.8) are equal. That leads to the critical condition $\langle k\rangle p_c = 1$, which is independent of clustering. However the degree distribution of the doubly Poisson model changes as we keep the average degree and change the clustering coefficient. When the degree distribution is fixed, the critical threshold actually increases as clustering increases [56, 57]. Furthermore, Fig. 3.1 shows the resulting giant component as a function of p. Note that single networks with higher clustering have smaller giant components.

3.3 Degree-Degree Correlation

When constructing clustering in a network, it is usually impossible to avoid generating degree-degree correlations. To better understand the effect of clustering on degree-degree correlations, we present an analytical expression of degree correlation as a function of the clustering coefficient for a doubly Poisson-clustered network—see Eq. (3.7).

The degree-degree correlation [59] can be expressed as

$$\rho_D = \frac{N_1N_3 - N_2^2}{N_1\sum_{i=1}^{N} d_i^3 N_2^2} \tag{3.9}$$

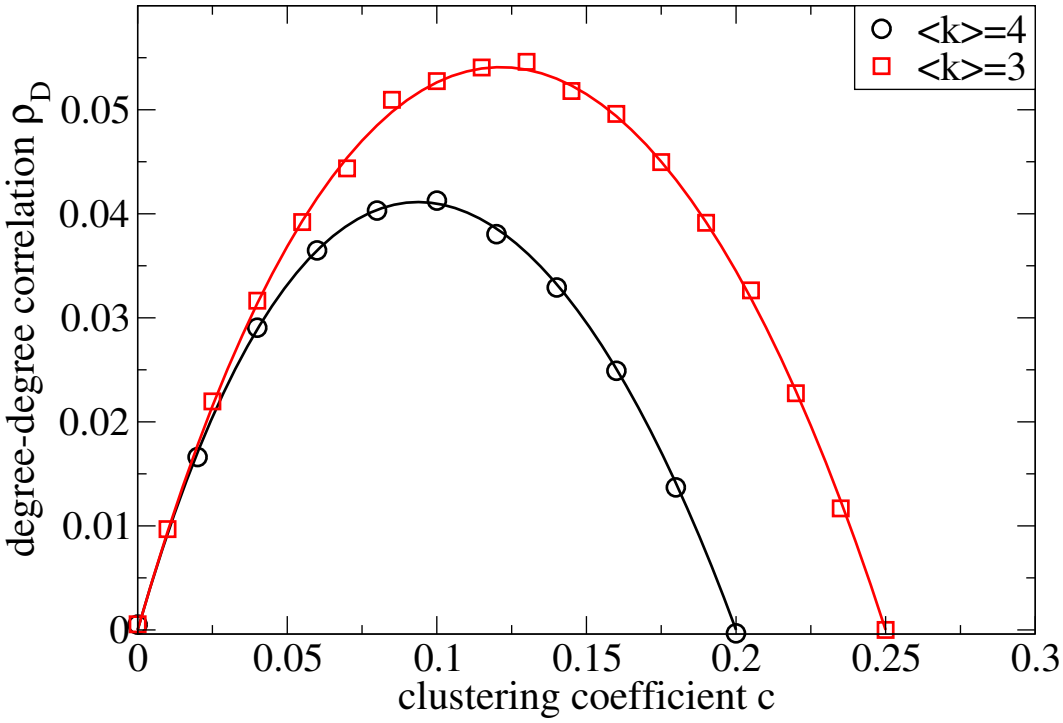


Figure 3.2: Degree-degrees correlation as a function of the clustering coefficient for Poisson network (Eq. (3.7)) with average degree $\langle k\rangle = 3$ and 4. Curves are from theory (Eq. 3.10) and symbols from simulations.

where N_m is the total number of m hop walks between all possible node pairs (i, j) including cases $i = j$.

The generating function of the degree of a node in the network is $\sum_{s,t=0}^{\infty} P_{st} z^{s+2t} = G_0(z, z^2)$. Let q_{st} be the fraction of nodes with s single edges and t triangles that are reached by traversing a random single link, where s includes the traversed link and r_{st} is the fraction of nodes with s single edges and t triangles reached by traversing a link of a triangle, $q_{st} = \frac{sP_{s,t}}{\langle s\rangle}, r_{st} = \frac{tP_{s,t}}{\langle t\rangle}$. Their corresponding generating functions are $G_q(x, y) = \frac{1}{\langle s\rangle}\frac{\partial G_0(x,y)}{\partial x}x$ and $G_r(x, y) = \frac{1}{\langle t\rangle}\frac{\partial G_0(x,y)}{\partial y}y$. Moreover, $N_3 = \sum_i \sum_j a_{ij} N_2(j)$, where $N_2(j)$ is the total number of two-hop walks starting from node j. The number of three-hop walks from a node i is equal to the total number of two-hop walks starting from all of its neighbors. Thus, $N_3 = \sum_j k_j N_2(j)$, where the number of two-hop walks starting from a node j with degree k_j

will be counted k_j times in N_3. Equivalently, $N_3 = N\sum_{st}(s+2t)P_{s,t}N_2(s,t)$, where $N_2(s,t)$ is the number of two hop walks from a node with s single edges and t triangles. The generating function of the number of single edges and of triangles reached in two hops from a random node is $G_2(x,y) = \sum_{st} P_{s,t} \cdot G_q^s(x,y) \cdot G_r^{2t}(x,y)$. The generating function of the total number of links and of triangles reached within three hops starting from all nodes is $G_3(x,y) = N\sum_{st} P_{s,t} \cdot (G_q(x,y))^{s(s+2t)} \cdot (G_r(x,y))^{2t(s+2t)}$. The number N_k of k-hop walks can be approximated by its mean in a large network

$$\begin{aligned} N_1 &= N\langle k\rangle, \\ N_2 &= N\frac{\partial G_2}{\partial x}|x=1,y=1 + 2N\frac{\partial G_2}{\partial y}|x=1,y=1 \\ N_3 &= \frac{\partial G_3}{\partial x}|x=1,y=1 + 2\frac{\partial G_3}{\partial y}|x=1,y=1 \end{aligned}$$

When both s and t follow a Poisson distribution,

$$\begin{aligned} G_0(x,y) &= e^{\mu(x-1)}e^{\nu(y-1)} \\ G_q(x,y) &= G_0(x,y)x \\ G_r(x,y) &= G_0(x,y)y. \end{aligned}$$

In this case,

$$\begin{aligned} N_1 &= N\langle k\rangle \\ N_2 &= N\langle k\rangle\left(\frac{\langle k\rangle}{1-c}+1\right) \\ N_3 &= \left(\langle k\rangle^3 + 2\langle k\rangle^2 + 4\nu\langle k\rangle + \langle k\rangle + 6\nu\right)N \\ \sum_{i=1}^{N} d_i^3 &= \left(\langle k\rangle^3 + 3\langle k\rangle^2 + (6\nu+1)\langle k\rangle + 6\nu\right)N, \end{aligned}$$

which together with Eq. (3.9) leads to

$$\rho_D = \frac{c - c^2 - \langle k\rangle c^2}{1 - c + \langle k\rangle c - 2\langle k\rangle c^2}, \tag{3.10}$$

where c is the clustering coefficient, Eq. (3.1).

Figure 3.2 shows the relation between the degree correlation and the clustering coefficient c for a Poissonian network [see Eq. (3.7)], for two given average degrees ($\langle k \rangle = 3$ and 4). The figure shows a positive degree-degree correlation across the entire range, which means the model is assortative [56]. The degree-degree correlation increases until c achieves half of its maximum and then decreases to zero when c reaches its maximum. When c is 0 or the maximum, the nodes connect to either all single links or all triangles, respectively.

3.4 Percolation on Interdependent Clustered Networks

To study how clustering within interdependent networks affects a system's robustness, we apply the interdependent networks framework [10]. In interdependent networks A and B, a fraction $(1 - p)$ of nodes is first removed from network A. Then the size of the giant components of networks A and B in each cascading failure step is defined to be p_1, p_2, ..., p_n, which are calculated iteratively

$$\begin{aligned} p_n &= \mu_{n-1} g_A(\mu_{n-1}), \text{n is odd}, \\ p_n &= \mu_n g_B(\mu_n), \text{n is even}, \end{aligned} \tag{3.11}$$

where $\mu_0 = p$ and μ_n are intermediate variables that satisfy

$$\begin{aligned} \mu_n &= p g_A(\mu_{n-1}), \text{n is odd}, \\ \mu_n &= p g_B(\mu_{n-1}), \text{n is even}. \end{aligned} \tag{3.12}$$

As interdependent networks A and B form a stable mutually-connected giant component, $n \to \infty$ and $\mu_n = \mu_{n-2}$, the fraction of nodes left in the giant component is p_∞. This system satisfies

$$\begin{aligned} x &= p g_A(y), \\ y &= p g_B(x), \end{aligned} \tag{3.13}$$

where the two unknown variables x and y can be used to calculate $p_\infty = x g_B(x) = y g_A(y)$. Eliminating y from these equations, we obtain a single equation

$$x = p g_A[p g_B(x)]. \tag{3.14}$$

The critical case ($p = p_c$) emerges when both sides of this equation have equal derivatives,

$$1 = p^2 \frac{dg_A}{dx}[p g_B(x)] \frac{dg_B}{dx}(x)|_{x=x_c, p=p_c}, \tag{3.15}$$

which, together with Eq. (3.14), yields the solution for p_c and the critical size of the giant mutually-connected component, $p_\infty(p_c) = x_c g_B(x_c)$.

Consider for example the case in which each network has doubly-Poisson degree distributions as in Eq. (3.7). From Eq. (3.13), we have $x = p(1 - u_A)$, $y = p(1 - u_B)$, where

$$u_A = v_A = e^{[\mu_A y + 2y(1-y)\mu_A](u_A - 1) + \nu_A p^2 (v_A^2 - 1)},$$
$$u_B = v_B = e^{[\mu_B x + 2x(1-x)\mu_B](u_B - 1) + \nu_B p^2 (v_B^2 - 1)}.$$

If the two networks have the same clustering, $\mu \equiv \mu_A = \mu_B$ and $\nu \equiv \nu_A = \nu_B$, p_∞ is then

$$p_\infty = p(1 - e^{\nu p_\infty^2 - (\mu + 2\nu) p_\infty})^2. \tag{3.16}$$

The giant component, p_∞, for interdependent clustered networks can thus be obtained by solving Eq. (3.16). Note that when $\nu = 0$ we obtain from Eq. (3.16) the result obtained in Ref. [10] for random interdependent ER networks. Figure 3.3a, using numerical simulation, compares the size of the giant component after n stages of cascading failure with the theoretical prediction of Eq. (3.11). When $p = 0.7$ and $p = 0.64$, which are not near the critical threshold ($p_c = 0.6609$), the agreement with simulation is perfect. Below and near the critical threshold, the simulation initially agrees with the theoretical prediction but then deviates for large n due to the random fluctuations of structure in different realizations [10]. By solving Eq. (3.16), we have p_∞ as a function of p in Fig. 3.3b for a given average degree and several values of clustering coefficients and in Fig. 3.4a for a given clustering and for different average degree values. As the figure shows, when higher clustering within a network is introduced, the percolation transition yields a higher value of p_c (see inset of Fig. 3.3b).

When clustering changes in this doubly Poisson distribution model, degree distribution and degree-degree correlation also change. First, to address the influence of the degree distribution, we study the critical thresholds of shuffled clustered networks. Shuffled clustered networks have neither clustering nor degree-degree distribution but keep the same degree distribution as the original clustered networks. The brown dashed curve in Fig. 3.3b

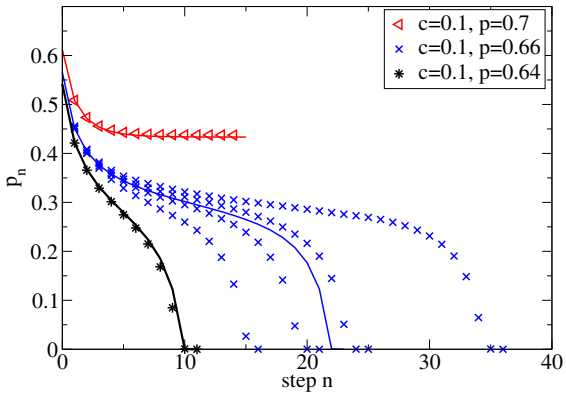
represents the giant component of interdependent shuffled clustered networks with original clustering $c = 0.2$. The figure shows that the difference in p_c between the $c = 0$ network and the shuffled $c = 0.2$ network is only 0.01, while the difference between the $c = 0$ and the $c = 0.2$ networks is 0.12. In addtion, $c = 0.2$ clustered networks has no degree-degree correlation (Fig. 3.2), which means the 0.12 shift of p_c is due to clustering and not to a change in degree distribution. We also show the critical thresholds of interdependent shuffled clustered networks as the red dashed line in the inset of Fig. 3.3b. Note that the change of degree distribution barely shifts the critical threshold. We next discuss the effect of the degree-degree correlation on the change of critical threshold. From Ref. [61], the degree assortativity alone monotonously increases the percolation critical threshold of interdependent networks. Because in our case degree-degree correlation first increases and then decreases (see Fig. 3.2), while critical the threshold of interdependent networks increases monotonously as clustering increases, we conclude that clustering alone increases the value of p_c. Thus clustering within networks reduces the robustness of interdependent networks. This probably occurs because clustered networks contain some links in triangles that do not contribute to the giant component, and in each stage of cascading failure the giant component will be smaller than in the unclustered case.

We also study the effect of the mean degree $\langle k \rangle$ on the percolation critical point. Figures 3.4a and 3.4b both show that, when clustering is fixed, the percolation critical point of interdependent networks decreases as the average degree $\langle k \rangle$ of network increases, making the system more robust. Figure 3.4b also shows that a larger minimum average degree is needed to maintain the network against collapse without any node removal as clustering increases.

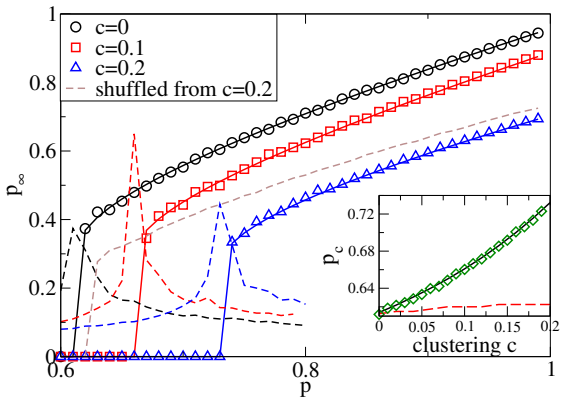
3.5 Summary

To conclude, based on Newman's single network clustering model, we present a generating-function formalism solution for site percolation on both single and interdependent clustered networks. We also derive an analytical expression, Eq. (3.10), for degree-degree correlation

as a function of the clustering coefficient for a doubly-Poisson network. Our results help us better understand the effect of clustering on the percolation of interdependent networks. We discuss the influence of a change of degree distribution and the degree-degree correlation associated with clustering in the model on the critical threshold of interdependent networks and conclude that p_c for interdependent networks increases when networks are more highly clustered.

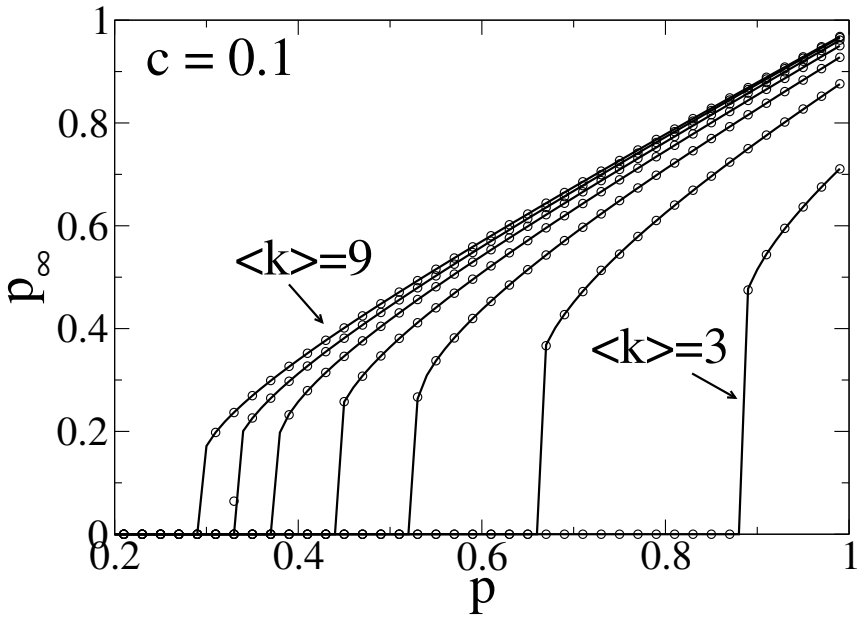


(a)

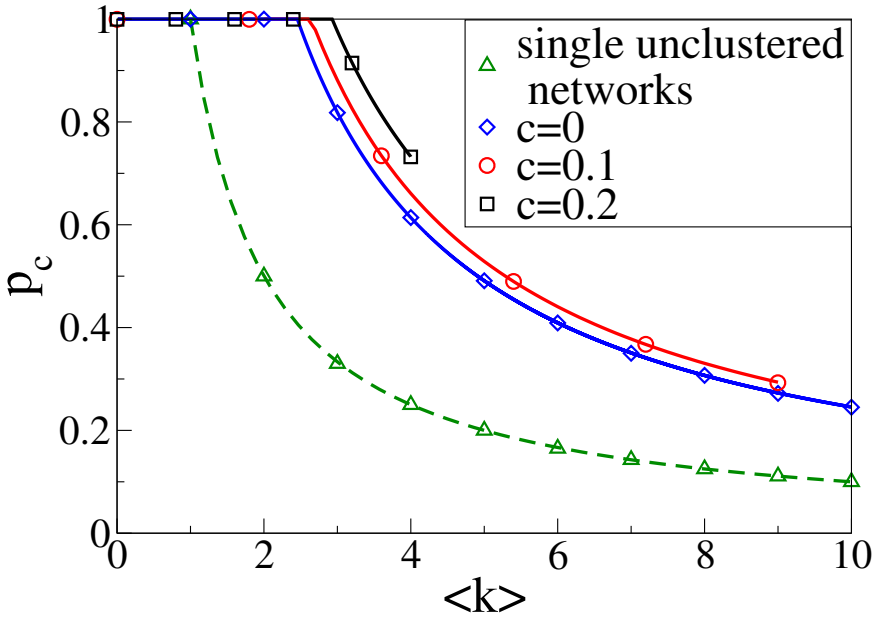


(b)

Figure 3.3: (a) Size of mutually connected giant component as a function of cascading failure steps n. Results are for $c = 1$, $p = 0.64$ (below p_c), $p = 0.66$ (at p_c) and $p = 0.7$ (above p_c). Lines represent theory (Eqs. (3.11) and (3.12)) and dots are from simulations. Note that at p_c there are large fluctuations. (b) Size of giant component, p_∞, in interdependent networks with both networks having clustering via degree distribution Eq. (3.7) and average degree $\langle k \rangle = 4$, as a function of p. Dashed lines are number of interactions (NOI) before cascading failure stops obtained by simulation [60]. The star curve is for shuffled $c = 0.2$ network, which keeps the same degree distribution but without clustering and without degree-degree correlation. Inset: Green squares and solid line represents critical thresholds, p_c, of interdependent networks as a function of clustering coefficient c. Red dashed line represents critical threshold of shuffled interdependent networks which originally has clustering coefficient c. The shuffled networks have zero clustering and degree-degree correlation, but has the same degree distribution as the original clustered networks. In all figures, symbols and dashed lines represent simulation, solid curves represent theoretical results.



(a)



(b)

Figure 3.4: (a) Size of giant component as a function of p for fixed clustering coefficient $c = 0.1$ and different average degrees. From right to left $\langle k \rangle = 3, 4, 5, ..., 9$. (b) Critical threshold p_c as a function of average degree for different clustering coefficients. The solid curves are for interdependent networks and the dashed curve is for single networks. Symbols and curves represent simulation and theoretical predictions respectively.

Chapter 4

Robustness of a Partially Interdependent Network of Clustered Networks

4.1 Introduction

Clustering, the propensity of two neighbors of the same node to be also neighbors of each other, has been observed in many real-world networks [50, 51, 62, 63]. For example, in a social network, if B and C are friends of A, they have a high probability of also being each other's friends. The average of this probability over the whole network is called the clustering coefficient. Empirical studies show that in many real-world networks, e.g., the Internet, scientific collaboration networks, metabolic and protein networks, and movie actor networks, the measured clustering coefficient is of the order of 10%, significantly higher than that of random networks [52].

Many computational models have been proposed to generate the clustering coefficient in networks, but all have been limited to numerical analysis [64–68]. Newman recently developed an analytical approach that incorporates clustering into random graphs by extending the generating function method, a widely used analytical tool in network research [53]. He considered two properties for each node—single links and triangles—and constructed a joint distribution for both. The clustering coefficient can be tuned by changing the ratio between the average number of single links and triangles. This approach enables us to evaluate analytically many properties of the resulting networks, such as component size, emergence

and size of a giant component, and other percolation properties.

Previous studies of clustering have focused on single network analysis, but real-world networks interact with and depend on other networks. In 2010, Buldyrev et al. [69] developed a theoretical framework for studying percolation in two fully interdependent networks and observed an unusual first-order (abrupt) percolation transition that differed from the known second-order (continous) phase transition in a single network. Parshani et al. [11] generalized this framework to partially-interdependent networks and found a change from a first-order to a second-order phase transition when the coupling strength was reduced below a critical value. Since 2010, there have been many studies of interdependent networks, sometimes called "networks of networks" [49, 70–86]. With respect to percolation properties, when interdependent nodes in the network of networks are treated as identical, the special cases is the multiplex network (from dynamical point of view, on the other hand, these two could be very different) [87–89]. Recently, Huang et al. [90] developed an approach to site percolation on clustered networks and studied the robustness of a pair of *fully* interdependent networks with clustering within each network.

Here we generalize the framework of Huang et al. [90] and extend it (i) to the study of percolation in two *partially* interdependent networks with clustering within each network and (ii) to the study of a network of clustered networks (NON), i.e., a network consisting of more than two interdependent clustered networks. We study how clustering within the networks influences such percolation properties as the critical shreshold p_c at which the giant component collapses, the sizes of the giant components ψ_∞ and ϕ_∞ in the two networks, the critical coupling q_c at which the first-order phase transition changes to a second-order phase transition, and the dynamics of cascading failure between two clustered networks. Simulation results agree well with theoretical results in all cases.

In Sec. V we also examine two joint distribution models for incorporating clustering into random graphs, i.e., (i) the model proposed by Newman [53], in which a double-Poisson distribution (see Sec. III) is assumed for the joint degree distribution, and the average degree is kept constant while the clustering is changed, and (ii) the clustering

model developed by Hackett et al. [91], in which a different joint distribution keeps both the average degree and the degree distribution constant while the clustering is changed. We discuss the similarities and differences in the percolation properties of the networks generated by these two distribution models. The model presented by Newman is studied both analytically and via simulations (Secs. III and IV), and the model presented by Hackett et al. is studied using only simulations (Sec. V).

4.2 The Model

In our model we consider two networks A and B that have the same number of nodes N. Within each network the nodes are connected with joint degree distribution $P_A(s,t)$ and $P_B(s,t)$, which specifies the fraction of nodes connected to s single links and t triangles in networks A and B, respectively [53]. The generating functions [28, 32] of the joint degree distributions are

$$\begin{aligned} G_{A0}(x,y) &= \sum_{s,t=0}^{\infty} P_A(s,t) x^s y^t, \\ G_{B0}(x,y) &= \sum_{s,t=0}^{\infty} P_B(s,t) x^s y^t. \end{aligned} \tag{4.1}$$

The conventional degree of a node is $k = s + 2t$ and the conventional degree distributions of the networks are

$$\begin{aligned} P_A(k) &= \sum_{s,t=0}^{\infty} P_A(s,t) \delta_{k,s+2t}, \\ P_B(k) &= \sum_{s,t=0}^{\infty} P_B(s,t) \delta_{k,s+2t}. \end{aligned} \tag{4.2}$$

The clustering coefficient is defined in [28] as

$$c \equiv \frac{3 \times (\text{number of triangles in network})}{\text{number of connected triples}} = \frac{3N_\Delta}{N_3}, \tag{4.3}$$

where $3N_\Delta \equiv N \sum_{st} tP(s,t)$ and $N_3 = N \sum_k \binom{k}{2} P(k)$.

Our initial attack is the random removal of a $(1-p)$ fraction of nodes from network A.

The generating function of the resulting network is [90]

$$\begin{aligned} G'_{A0}(x,y) &\equiv G_{A0}(x,y,p) \\ &= G_{A0}(xp+1-p, p^2y + 2xp(1-p) + (1-p)^2), \end{aligned} \tag{4.4}$$

and the fraction of nodes belonging to the giant component in the remaining network is

$$g_A(p) \quad = \quad 1 - G_{A0}(u, v^2, p), \tag{4.5}$$

where u, v satisfy

$$u = G_{Aw}(u, v^2, p), \qquad v = G_{Ar}(u, v^2, p). \tag{4.6}$$

The functions $G_{Aw}(x,y,p)$ and $G_{Ar}(x,y,p)$ are defined as

$$\begin{aligned} G_{Aw}(x,y,p) &\equiv \frac{1}{\langle s' \rangle} \frac{\partial G_{A0}(x,y,p)}{\partial x}, \\ G_{Ar}(x,y,p) &\equiv \frac{1}{\langle t' \rangle} \frac{\partial G_{A0}(x,y,p)}{\partial y}, \end{aligned} \tag{4.7}$$

where $\langle s' \rangle = \frac{\partial G_{A0}(x,y,p)}{\partial x}\big|_{x=1,y=1}$ and $\langle t' \rangle = \frac{\partial G_{A0}(x,y,p)}{\partial y}\big|_{x=1,y=1}$. Similar equations hold for network B.

We next consider the interaction between clustered networks A and B [11]. Assume a q_A fraction of nodes in network A is dependent on nodes in network B and a q_B fraction of nodes in network B is dependent on nodes in network A. This means that if a node in network B upon which a node in network A depends fails, the corresponding node in network A will also fail, and vice versa. We also assume that a node from one network may be dependent on no more than one node from the other network and if a node i in network A is dependent on a node j in network B and j depends on a node l in network A, then $l = i$ (a no-feedback condition [49, 70, 71]). After n steps of cascading failures, ψ_n and ϕ_n are the fractions of nodes in the giant components of networks A and B, respectively. After the two-network system reaches staionarity, the sizes of giant components in the two networks are [11]

$$\psi_\infty = x g_A(x), \qquad \phi_\infty = y g_B(y), \tag{4.8}$$

where the two variables x and y satisfy

$$\begin{aligned} x &= p\{1 - q_A[1 - g_B(y)]\}, \\ y &= 1 - q_B[1 - pg_A(x)]. \end{aligned} \tag{4.9}$$

4.3 The Double-Poisson Distribution

As an example, consider two Erdős-Rényi (ER) networks [92–94] with clustering, in which the number of single links s and triangles t of a node obey a double-Poisson distribution $P_{st}=e^{-\langle s\rangle}\frac{\langle s\rangle^s}{s!}e^{-\langle t\rangle}\frac{\langle t\rangle^t}{t!}$ (s and t follow a Poisson distribution independently) [53]. Here $\langle s\rangle$ and $\langle t\rangle$ are the average number of single links and triangles per node, respectively. Assuming that in network A $\langle s\rangle = \langle s\rangle_A$ and $\langle t\rangle = \langle t\rangle_A$, then the generating functions in Eq. (4.4) and Eq. (4.7) become

$$\begin{aligned} &G_{A0}(x, y, p) = G_{Aw}(x, y, p) = G_{Ar}(x, y, p) \\ &= e^{[\langle s\rangle_A p+2p(1-p)\langle t\rangle_A](x-1)+\langle t\rangle_A p^2(y-1)}, \end{aligned} \tag{4.10}$$

and the same holds for network B. Denoting $f_A(x) = 1 - g_A(x)$ and $f_B(y) = 1 - g_B(y)$, we now have

$$\begin{aligned} f_A(x) &= exp\{\langle t\rangle_A x^2(1 - f_A(x))^2 - \langle k\rangle_A x(1 - f_A(x))\}, \\ f_B(y) &= exp\{\langle t\rangle_B y^2(1 - f_B(y))^2 - \langle k\rangle_B y(1 - f_B(y))\}, \end{aligned} \tag{4.11}$$

where $\langle k\rangle_A$ and $\langle k\rangle_B$ are the average degrees for networks A and B, respectively ($\langle k\rangle_A = \langle s\rangle_A + 2\langle t\rangle_A$, and $\langle k\rangle_B = \langle s\rangle_B + 2\langle t\rangle_B$). By combining Eq. (4.9) and Eq. (4.11) and eliminating x and y, we obtain two transcendental equations for f_A and f_B,

$$\begin{aligned} f_A &= e^{\langle t\rangle_A p^2(1-f_A)^2(1-q_A f_B)^2-\langle k\rangle_A p(1-f_A)(1-q_A f_B)}, \\ f_B &= e^{\langle t\rangle_B (1-f_B)^2\{1-q_B[1-p(1-f_A)]\}^2-\langle k\rangle_B (1-f_B)\{1-q_B[1-p(1-f_A)]\}}. \end{aligned} \tag{4.12}$$

By substituting the parameter vector $(\langle k\rangle_A, \langle t\rangle_A, \langle k\rangle_B, \langle t\rangle_B, q_A, q_B, p)$, we can solve for f_A and f_B, and thus find the size of the giant components in network A, ψ_∞, and network B, ϕ_∞. By substituting the double-Poisson distribution into Eq. (4.3), the clustering coeffi-

cients in the two networks become

$$c_A = \frac{2\langle t\rangle_A}{\langle k\rangle_A^2 + 2\langle t\rangle_A}, \qquad c_B = \frac{2\langle t\rangle_B}{\langle k\rangle_B^2 + 2\langle t\rangle_B}. \tag{4.13}$$

If we fix the other parameters and increase p, the fraction of nodes not removed in the initial attack, a phase transition occurs at a critical threshold p_c and a giant component appears. As we decrease the coupling strength q_A and q_B, the behavior of this phase transition will change from first-order to second-order. A first-order phase transition, denoted by I, corresponds to a scenario in which the size of one or both giant components in the two networks change discontinuously from a finite value to zero. If we plot f_A and f_B in Eqs. (4.12) on a two-dimensional graph, this corresponds to the scenario that two curves $f_A(f_B)$ and $f_B(f_A)$ are tangential with each other ($\frac{df_B(f_A)}{df_A}\frac{df_A(f_B)}{df_B} = 1$) [11]. By adding this condition into Eqs. (4.12), we can solve for $f_A = f_{A_I}$, $f_B = f_{B_I}$ and $p = p_I$. A second-order phase transition (denoted by II), corresponding to a scenario in which the size of one or both giant components decreases continuously to zero, is obtained by substituting $f_A \to 1$ or $f_B \to 1$ into Eqs. (4.12), which allows us to find $f_{A_{II}}$, $f_{B_{II}}$ and p_{II}. The critical coupling strength q_c is solved by making the conditions for both first-order and second-order phase transitions equal.

For the sake of simplicity, we now consider the symmetrical case, $\langle k\rangle = \langle k\rangle_A = \langle k\rangle_B$ and $c = c_A = c_B$. Fig. 4.1 shows the size of the giant components in networks A and B for several clustering coefficients. In each graph the simulation results agree well with the theoretical results obtained from Eqs. (4.12). Note that, for strong coupling, as we increase the clustering coefficient the two interdependent networks become less robust. When the coupling is weak, the weakening effect of the clustering on the robustness is smaller. This can be seen in Fig. 4.2, which shows p_c versus $q = q_A = q_B$ for different clustering coefficients for both $\langle k\rangle = 3$ and $\langle k\rangle = 4$. Note that, for the same coupling strength q, a larger clustering coefficient yields a larger p_c, making the networks less robust. In addition, the critical coupling strength q_c below which the first-order phase transition changes to a second-order

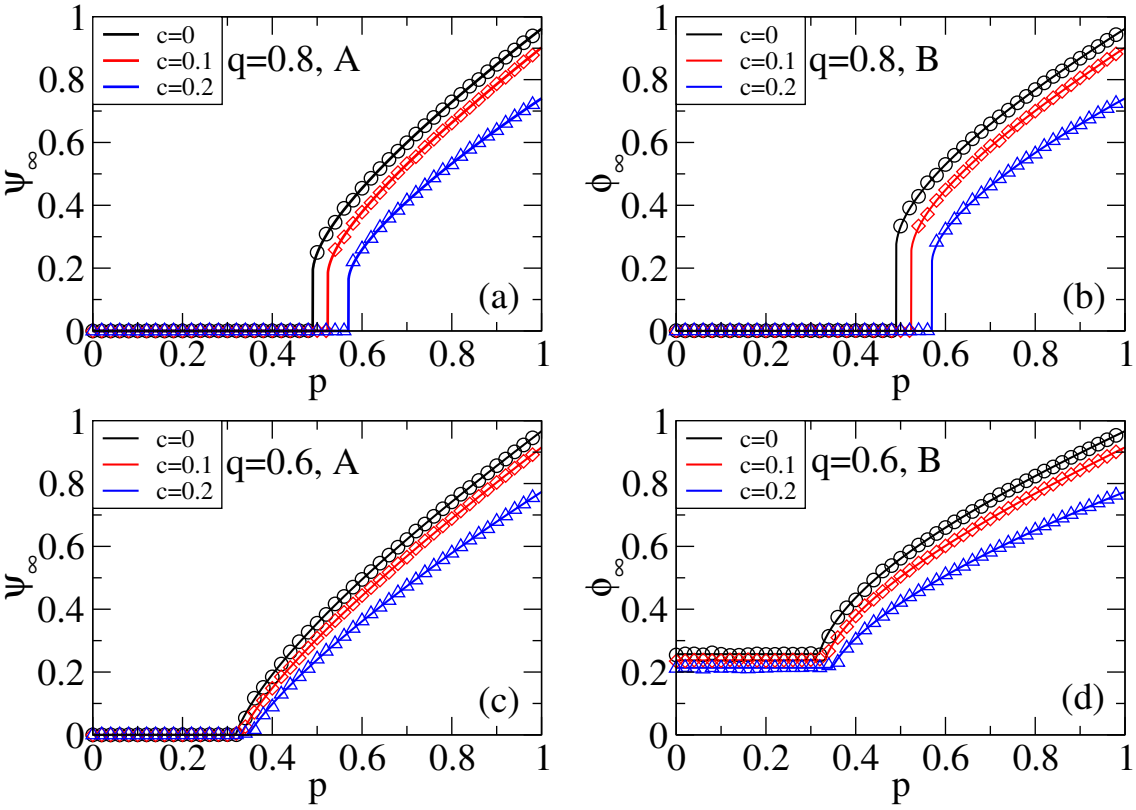


Figure 4.1: (Color online) Size of giant components as a function of p for $\langle k\rangle = \langle k\rangle_A = \langle k\rangle_B = 4$, where solid lines are from theoretical predictions, Eqs. (4.12), and symbols are from simulations with network size $N = 10^5$. (a) and (b) For strong coupling ($q = 0.8$), the sizes of giant components in (a) network A and (b) network B change abruptly at some critical threshold p_c, showing a first-order phase transition behavior. (c) and (d) For weak coupling ($q = 0.6$), on the contrary, the behavior is continuous, i.e., second-order. Note that while (c) network A collapses (d) network B does not collapse, since the initial failures are in A and q is relatively small to cause collapse of network B. Thus, the giant component of B is finite for all p values.

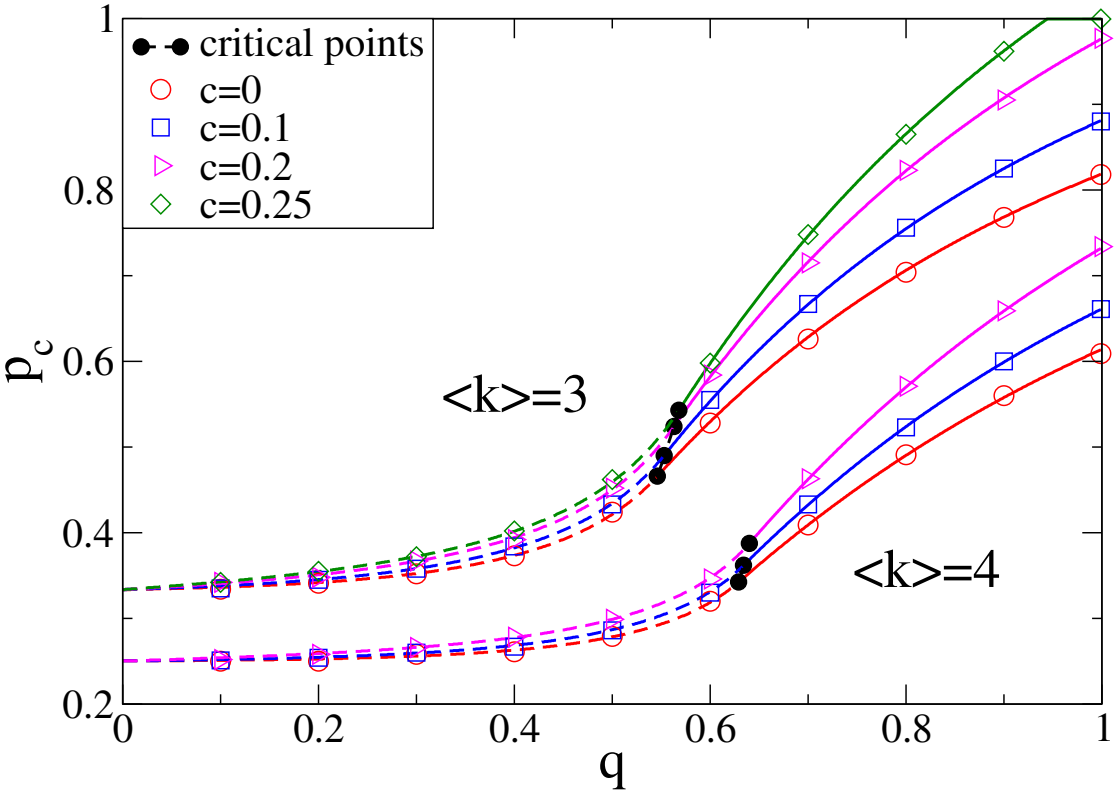


Figure 4.2: (Color online) Percolation shreshold, p_c, as a function of interdependency strength q ($q = q_A = q_B$) for $\langle k \rangle = \langle k \rangle_A = \langle k \rangle_B = 3$ and 4. Clustering coefficient c (c=c_A=c_B) ranges from 0 to 0.2 for $\langle k \rangle$=4 and from 0 to 0.25 for $\langle k \rangle$=3. For each $\langle k \rangle$ and c, there exists a critical point q_c (full circles). Above q_c, the system undergoes a first-order phase transition (solid lines) and below q_c, the system undergoes a second-order transtion (dashed lines). Symbols represent simulation results and are in good agreement with theoretical predictions (solid and dashed lines). Note that for the same average degree $\langle k \rangle$, increasing clustering coefficient c increases p_c and yields a larger critical coupling, q_c.

increases slightly as we increase clustering coefficient.

Fig. 4.3 shows the size of the giant component in network A after each cascading step around the critical threshold for the first-order phase transition case (Fig. 4.3(a)) and the second-order phase transition case (Fig. 4.3(b)). Note that the simulation results for the cascading failures agree well with analytical results (4.8) and (4.9). Different realizations give different results due to deviations from the mean field, rendering small fluctuations around the mean-field analytical results [61].

4.4 Network of Networks with Clustering

The framework discussed above can also be generalized to an interdependent system consisting of more than two clustered networks. Here we consider two cases of NON [49, 70, 71] composed of n interdependent clustered networks, (i) A star-like NON and (ii) a random regular NON (see Fig. 4.4). We assume that for each pair of interdependent networks i and j $(i, j = 1, 2, ..., n)$, there is a fraction q_{ji} of nodes in network i which depend on nodes in network j, i.e., they cannot function if the nodes upon which they depend fail. Similarly, q_{ij} denotes the fraction of nodes in network j which depend on nodes in network i. We also assume here that a node from one network may depend on no more than one node from the other network and, if a node i in network A depends on a node j in network B and j depends on a node l in network A, then $l = i$ (a no-feedback condition [49, 70, 71]). After an initial attack, only a fraction p_i $(i = 1, 2, ..., n)$ of nodes in each network will remain. After the period of cascading failures, a fraction $\psi_{\infty,i}$ of nodes in network i will remain functional. The final giant component of each network can be expressed as $\psi_{\infty,i} = x_i g_i(x_i)$ and the unknowns x_i can be found from a system of n equations [49, 70, 71],

$$x_i = p_i \prod_{j=1}^{K} [q_{ji} y_{ji} g_j(x_j) - q_{ji} + 1], \tag{4.14}$$

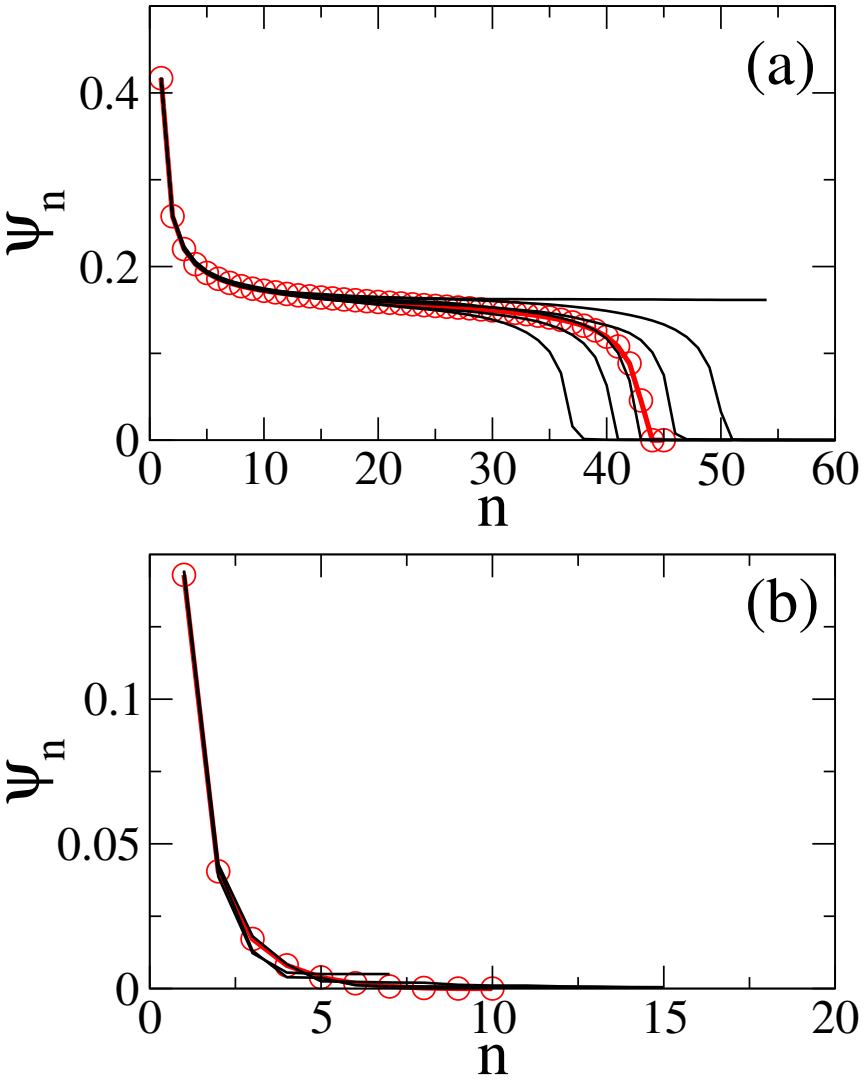


Figure 4.3: (Color online) Size of the giant component in network A (ψ_n) as a function of cascading failure steps n for $\langle k\rangle$=4, c=0.2 for (a) $q = 0.8$ (first-order transition) and (b) $q = 0.6$ (second-order transition). The symbols (circles) and their connecting line are from the theoretical prediction. The other lines are several random realizations from simulations ($N = 10^6$). The value of $p = 0.569$ for (a) the first-order phase transtion case and $p = 0.347$ for (b) the second-order phase transition case are both chosen to be just below critical thresholds obtained from theoretical predictions ($p_c = 0.57$ for the first-order case and $p_c = 0.3475$ for the second-order case). One can see that in both cases the agreement is very good. However, for first-order transition, after the plateau different realizations fluctuate.

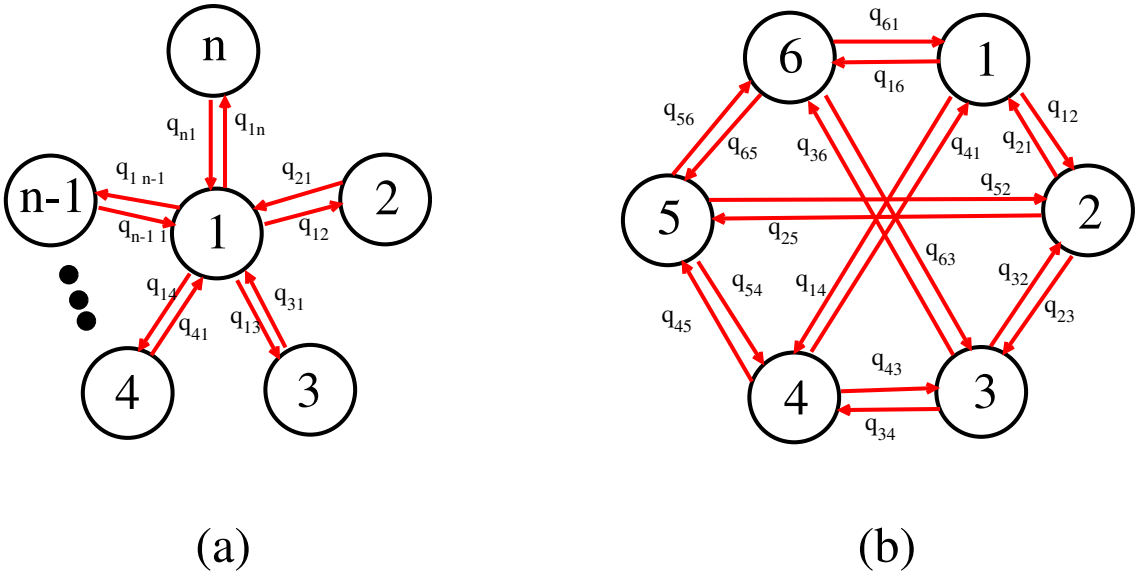


Figure 4.4: (Color online) Schematic representation of two types of NONs : (a) Star-like NON where one central network is interdependent with $(n-1)$ other networks. (b) Random regular NON where each network depends exactly on m (here, $m = 3$) other networks. Circles represent interdependent networks and arrows represent interdependency relations. For example, q_{12} represents a fraction q_{12} of nodes in network 2 depend on nodes in network 1.

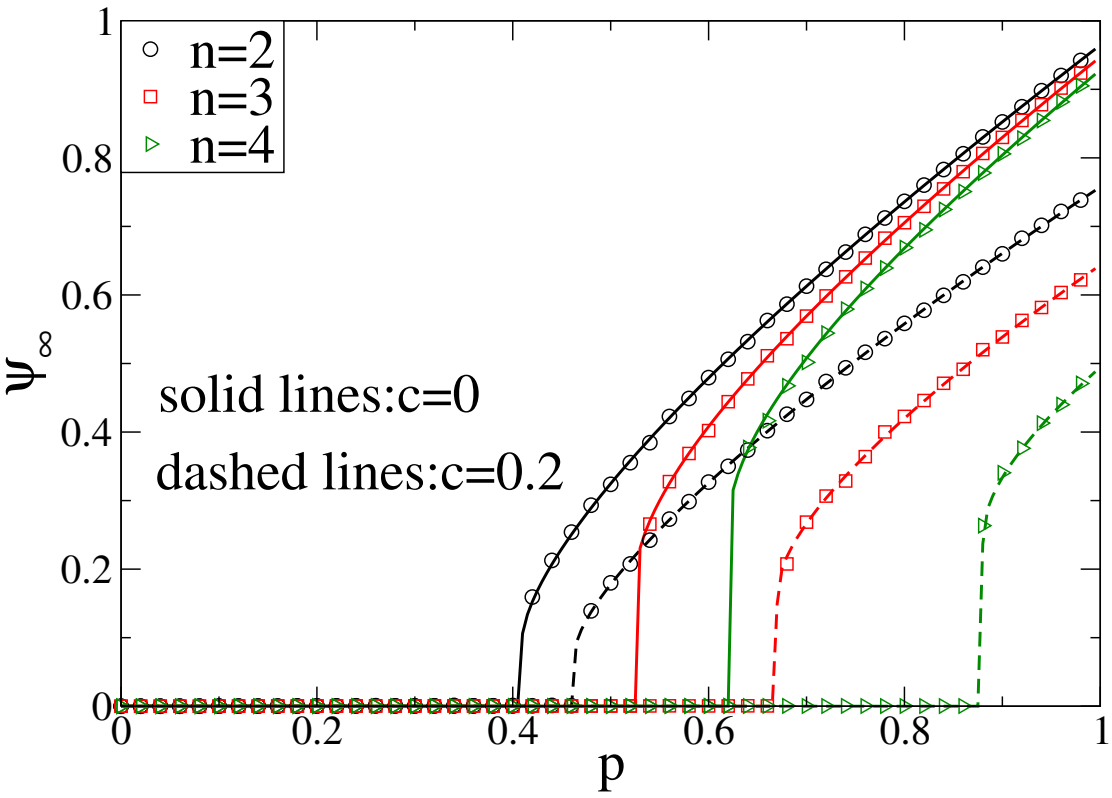


Figure 4.5: (Color online) Size of the giant component in the root network as a function of p for $n = 2, 3, 4$ and $c = 0, 0.2$ for star-like NON. Average degree of each network in the NON is $\langle k \rangle$=4. Symbols and lines represent simulations ($N = 10^5$) and theory, respectively.

where the product is taken over the K networks that are coupled with network i. Since we consider the no-feedback condition [49, 70, 71], we have

$$y_{ji} = \frac{x_j}{q_{ij} y_{ij} g_i(x_i) - q_{ij} + 1}, \tag{4.15}$$

where y_{ji} is the fraction of nodes left in network j after it has suffered damage from all networks other than network i. We next consider two analytically solvable examples of a NON, a star-like network of ER networks and a random regular (RR) network of ER networks, shown in Fig. 4.4.

4.4.1 Star-like NON with clustering

For a star-like NON (Fig. 4.4(a)), we have a root network which is interdependent with other $(n-1)$ networks. For simplicity, the initial attack is on the root network, and a fraction $(1-p)$ of its nodes is removed. This damage spreads to the other networks, and then returns to the root network, back and forth. Here we consider the case for n clustered ER networks with the same average degree $\langle k \rangle$ and same clustering coefficient c (thus the same average number of triangles $\langle t \rangle$). Assuming, again for simplicity, that for all i, $q_{i1} = q_{1i} = q$, Eq. (4.14) and Eq. (4.15) are simplified to two equations,

$$\begin{aligned} x_1 &= p[qg_2(x_2) - q + 1]^{n-1}, \\ x_2 &= pqg_1(x_1)[qg_2(x_2) - q + 1]^{n-2} - q + 1. \end{aligned} \tag{4.16}$$

For clustered ER networks, $f(x) = 1 - g(x)$ satisfies

$$f = exp[\langle t \rangle x^2(1-f)^2 - \langle k \rangle x(1-f)]. \tag{4.17}$$

By combining Eq. (4.16) and Eq. (4.17), we find x_1, x_2 and f_1, f_2, from which the sizes of the giant components in the root network (ψ_∞) and in the other networks (ϕ_∞) can be obtained.

Fig. 4.5 shows the size of the giant component in the root network for $n = 2$, 3, and 4 and compares two cases, $c = 0$ (no clustering) and $c = 0.2$ (high clustering). Note that the simulation results agree well with the theoretical predictions. Our results show that the NON becomes less robust with increasing n. For fixed n, the NON composed of networks with a larger clustering coefficient is less robust and the effect of clustering in reducing the robustness becomes larger as n increases. Similarly, the critical coupling q_c, where the behavior of phase transition changes from first-order to second-order decreases with n and increases slightly with the clustering coefficient (see Fig. 4.6).

4.4.2 Random regular (RR) NON of ER networks with clustering

We now consider the case in which each clustered ER network depends on exactly m other clustered ER networks, i.e., a random regular (RR) NON formed of clustered ER networks.

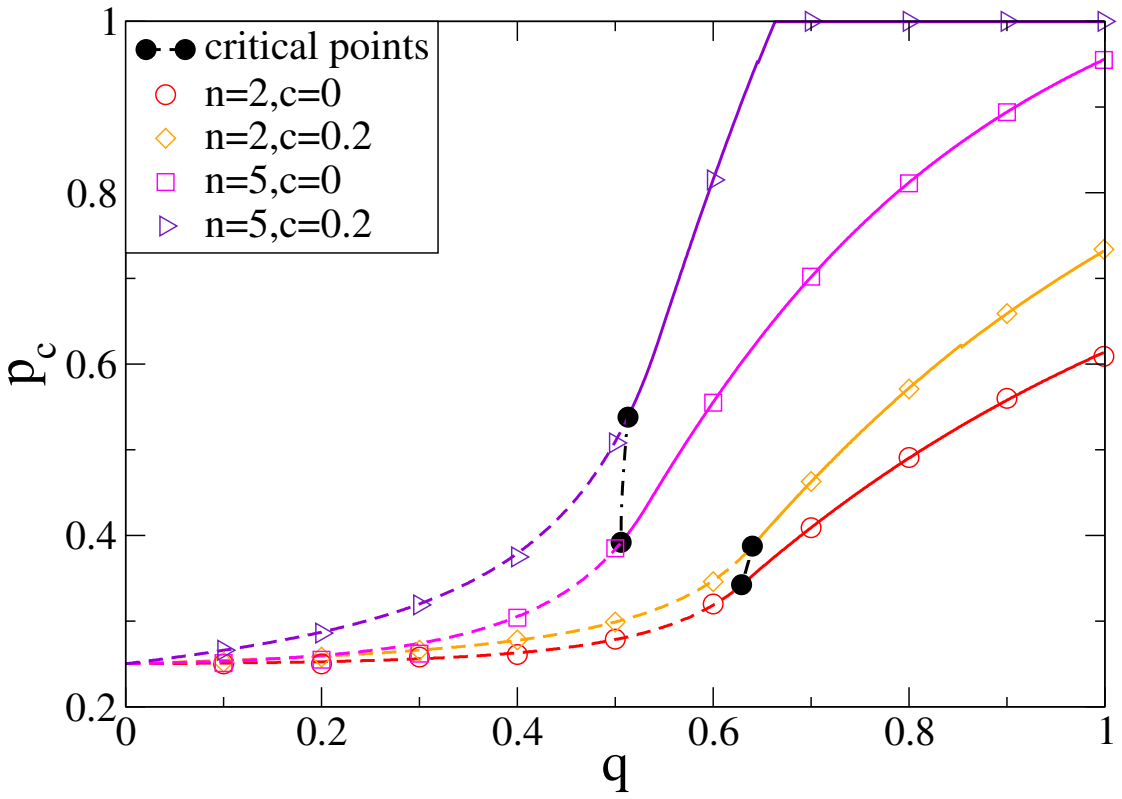


Figure 4.6: (Color online) Critical threshold p_c as a function of interdependency strength, q, for clustered star-like NON for $\langle k\rangle$=4, $n = 2, 5$ and $c = 0, 0.2$. For each n and c, there exsits a critical interdependency strength q_c (solid symbols) that separates the first-order (solid lines) and second-order (dashed lines) phase transitions.

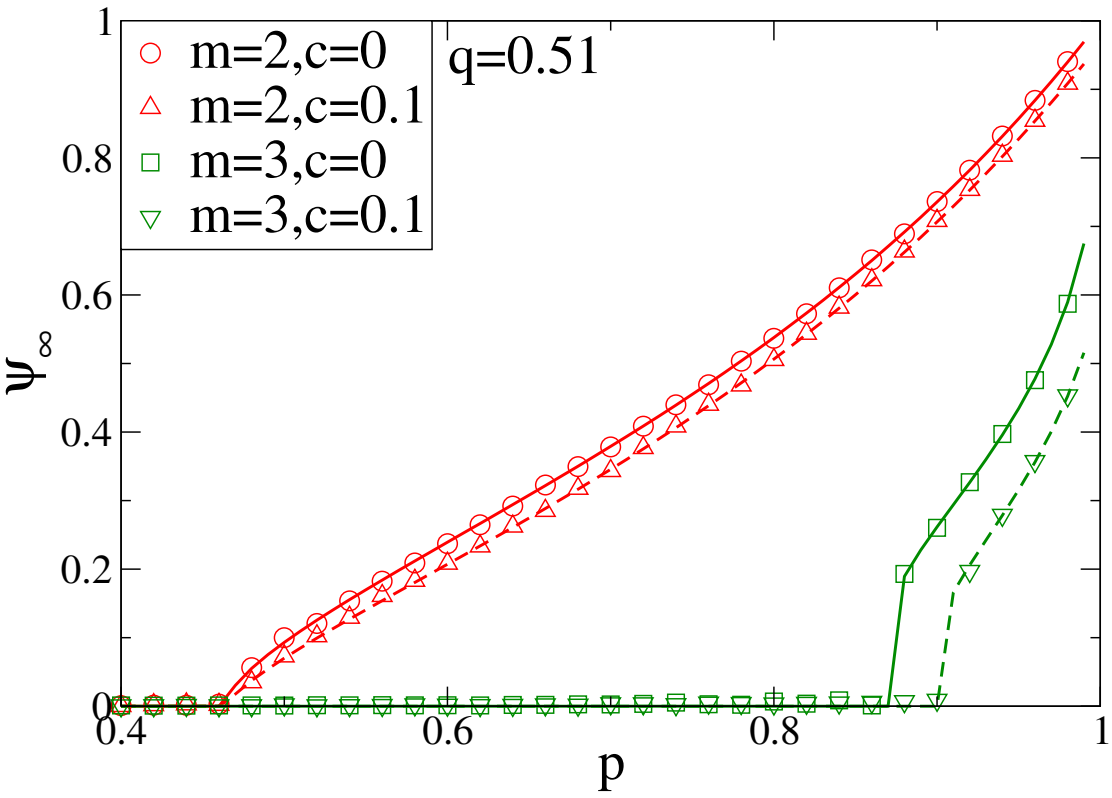


Figure 4.7: (Color online) Size of the giant component, ψ_∞, as a function of p for RR NON of clustered ER networks for fixed q (q=0.51). The average degree is $\langle k\rangle$=9, m=2, 3 and c=0, 0.1. For $m = 3$, the system shows a first-order percolation transition as we change the value of p. While for $m = 2$, the phase transition is second-order.

Assume that the initial attack is on each network and randomly removes a fraction $(1 - p)$ of nodes and that the interacting strengths are all equal to q. Assume also that all ER networks have the same average degree $\langle k\rangle$ and the same average number of triangles $\langle t\rangle$. Because of symmetry, all equations in Eqs. (4.14) and Eqs. (4.15) are reduced into a single equation and the size of the giant component in each network is

$$\psi_\infty = p(1 - e^{\langle t\rangle\psi_\infty^2 - \langle k\rangle\psi_\infty})\Big[\frac{1 - q + \sqrt{(1-q)^2 + 4q\psi_\infty}}{2}\Big]^m. \tag{4.18}$$

Fig. 4.7 and Fig. 4.8 show numerical solutions of Eq. (4.18) and simulation results. Note that the simulations agree well with theory. For a given $\langle k\rangle$, the size of the giant component ψ_∞ in each network displays a first-order or a second-order phase transition as a function of

p, depending on the values of q, m, and the clustering coefficient c. Fig. 4.7 shows that, for some fixed values of $\langle k\rangle$ and q, the behavior of the phase transition can be either first-order or second-order for different values of m. Similarly, as shown in Fig. 4.8, for fixed values of $\langle k\rangle$ and m, different values of q can cause the phase transition to be first-order or second-order. In each scenario, when the transition is first-order the clustering within networks reduces the resistance of the NON to random node failure, but when it is second-order the effect of clustering is similar but very small. This again is due to the smaller coupling value q in the second-order phase transition region. Note that for q=1 and m=1, the limit of two fully-interdependent networks, Eq. (4.18) reduces to an equation similar to Eq. (16) in Huang et al. [90]. The only difference is because here we initially attack all networks, not just network A as in [90]. For $\langle t\rangle = 0$ (the no-clustering case), Eq. (4.18) reduces to Eq. (23) in Gao et al. [71].

By adding the condition that the first derivative of both sides of Eq. (4.18) with respect to ψ_∞ are equal, we obtain the critical threshold of the first-order phase transition, p_I. The critical threshold of the second-order phase transition p_{II} is solved by adding the condition $\psi_\infty(p_{II}) \to 0$ to Eq. (4.18). If we equate p_I and p_{II}, the critical coupling q_c where the first-order phase transition changes to a second-order phase transition can be derived analytically,

$$(\langle k\rangle^2 + 2\langle t\rangle)(1-q_c)^2 = 2\langle k\rangle q_c m. \tag{4.19}$$

By substituting $c = \frac{2\langle t\rangle}{\langle k\rangle^2+2\langle t\rangle}$, we have

$$q_c = 1 + x - \sqrt{x(x+2)}, \tag{4.20}$$

where $x \equiv \frac{m}{\langle k\rangle}(1-c)$. Note that increasing the clustering coefficient c increases the critical dependency q_c. Note also that for $c = 0$, Eq. (4.20) reduces to Eq. (30) of Ref. [71].

4.5 The Fixed Degree Distribution

The double-Poisson distribution model can display the features of clustering and it is possible to solve it analytically. Although in this model the average degree does not change, the

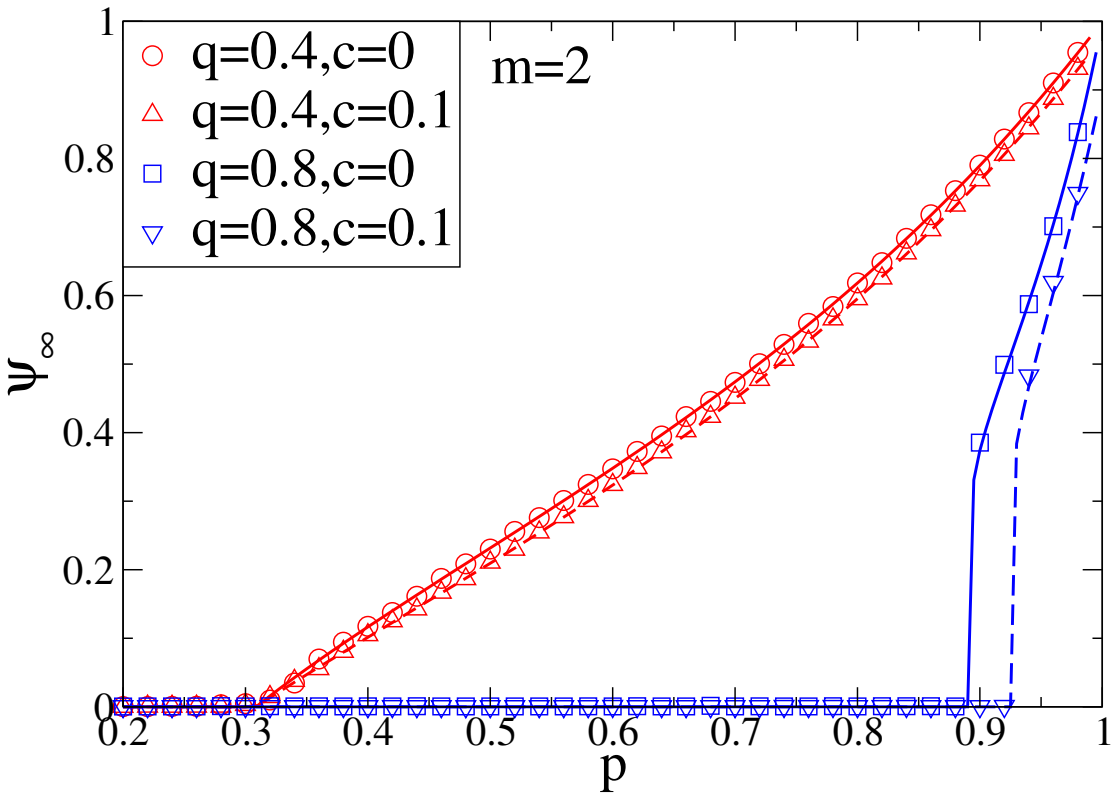


Figure 4.8: (Color online) Size of giant component, ψ_∞, as a function of p for RR NON composed of clustered ER networks for fixed m (m=2). The average degree is $\langle k\rangle$=9, q=0.4, 0.8 and c=0, 0.1. The behavior of the phase transition is first-order for $q = 0.8$ and second-order for $q = 0.4$.

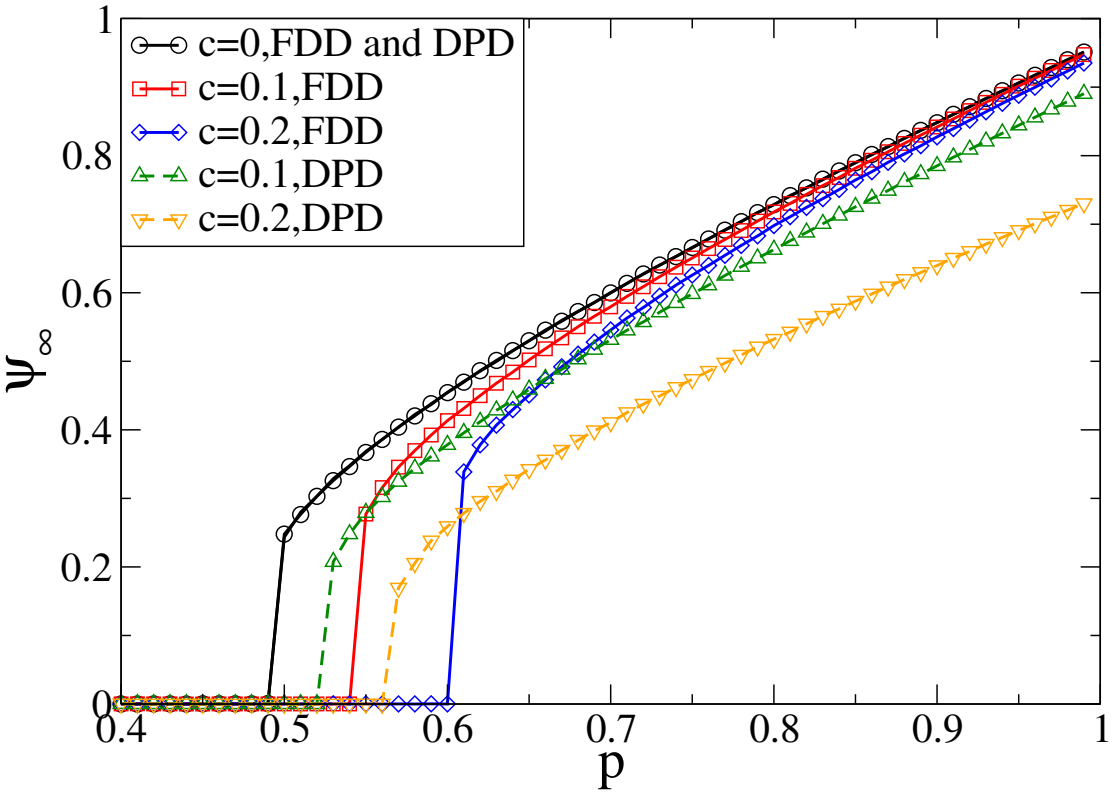


Figure 4.9: (Color online) Size of giant component in network A for two partially interdependent networks with clustering. Circles, squares and diamonds represent results for a joint degree distribution which fix the total degree distribution being Poisson as we change the clustering coefficient (FDD). Circles, up-triangles and down-triangles represent results for a double-Poisson distribution (DPD) with total average degree fixed. All resulted are from simulations with N=10^6, $\langle k\rangle$=4 and q=0.8> q_c. The behavior of the phase transition is first-order in both cases but p_c is larger for FDD.

degree distribution does change as the clustering coefficient changes. Here we consider another kind of joint distribution P_{st} proposed by Hackett et al. [91, 95], which also preserves the total degree distribution $P(k)$ for different clustering coefficients. We set

$$P_{st} = P(k)\delta_{k,s+2t}[(1-f)\delta_{t,0} + f\delta_{t,\lfloor (s+2t)/2 \rfloor}], \tag{4.21}$$

where $f \in [0,1]$ and $\lfloor . \rfloor$ is the floor function.

Eq. (4.21) allows us to construct P_{st} from a given degree distribution $P(k)$ by picking a fraction f of nodes being attached to a maximum possible number of triangles while the remaining $(1-f)$ nodes are attached to single edges only. From the definition of a clustering coefficient, we have

$$c = f\frac{\sum_k k(P(2k) + P(2k+1))}{\sum_k \binom{k}{2} P(k)}, \tag{4.22}$$

hence the clustering coefficient can be adjusted by tuning the parameter f.

We investigate the effect of the joint degree distribution, Eq. (4.21), on the robustness of partially interdependent networks by comparing the two joint degree distributions. One is the fixed degree distribution (FDD), which is defined by Eq. (4.21) with $P(k)$ obeying a Poisson distribution ($P(k)$=$\langle k \rangle^k e^{-\langle k \rangle}/k!$). The other is the double-Poisson distribution (DPD) discussed in Sec. III, with P_{st}=$e^{-\langle s \rangle}\frac{\langle s \rangle^s}{s!}e^{-\langle t \rangle}\frac{\langle t \rangle^t}{t!}$.

Fig. 4.9 plots the size of the giant component in network A for two partially-interdependent networks with clustering. The joint degree distribution in each network is fixed as either FDD or DPD. The interdependent strength q is fixed as first-order. Note that the critical threshold p_c in FDD is larger than that in DPD when the clustering coefficient is the same. This difference in p_c is caused by the broadening of $P(k)$ in the double-Poisson distribution. Note that for site percolation on a single clustered network, a larger clustering coefficient leads to a higher critical threshold for both distributions [90, 91]. For a system of two interdependent networks, the general trend is similar and, for both degree distributions, p_c increases as the clustering coefficient increases. Fig. 4.10 shows the size of the giant components in partially-interdependent networks with a second-order phase transition for both FDD and DPD. The influence of clustering on the robustness of partially-interdependent

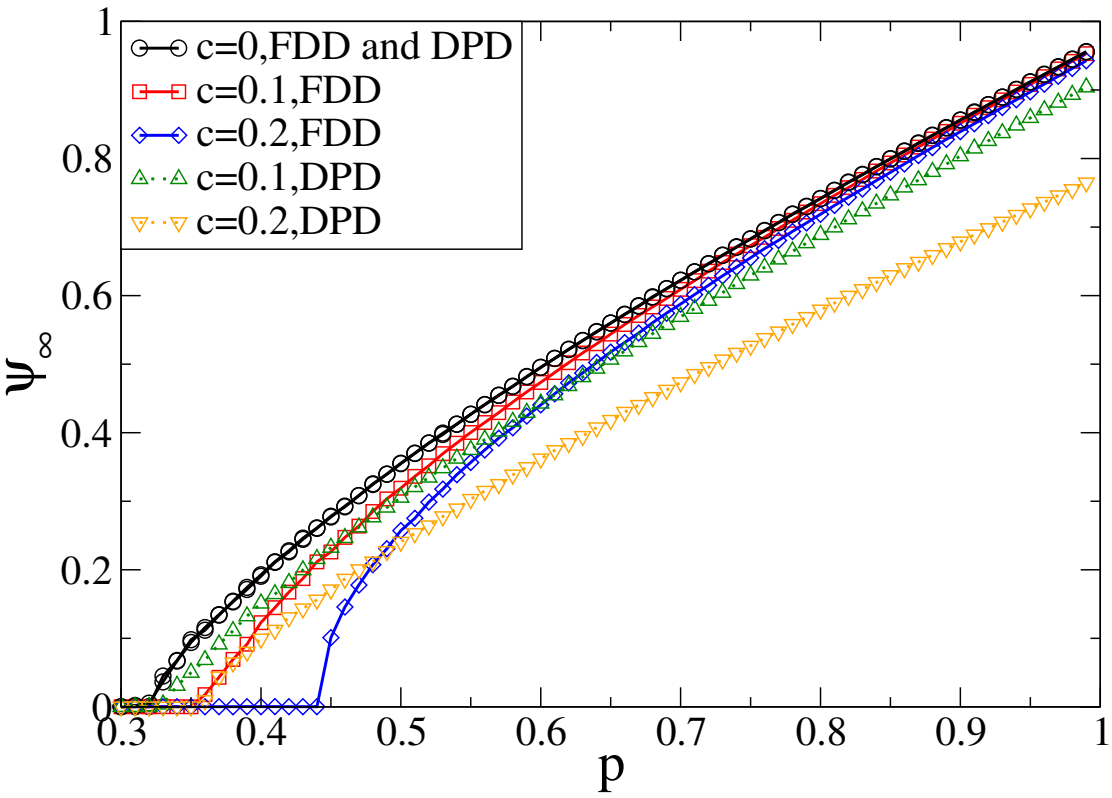


Figure 4.10: (Color online) Size of giant component in network A for two partially interdependent networks with clustering. Circles, squares and diamonds represent results for FDD, while circles, up-triangles and down-triangles represent results for DPD. All results are from simulations with N=10^6, $\langle k\rangle$=4 and q=0.6< q_c. The behavior of the phase transition is second-order.

networks is larger for FDD than for DPD, but the general trend is similar in both distributions.

4.6 Summary

We have developed a framework for studying percolation in a network formed of interdependent ER networks with clustering. For each clustering coefficient, the system shows a first-order to second-order transition as we decrease coupling strength q. As we increase the clustering coefficient of each network, the system becomes less robust. This influence of the clustering coefficient on network robustness decreases as we decrease the coupling strength, and the critical coupling strength q_c, at which the first-order phase transition changes to second-order, increases as we increase the clustering coefficient. We have also investigated the differences and similarities between two different joint degree distributions, FDD and DPD. We have found that, although the percolation threshold is different in the two cases, the general conclusion that an increase in the clustering coefficient causes interdependent networks to become less robust holds.

Chapter 5

Conclusion

This dissertation is a review of the original research of me and my collaborators during my PhD studies at Boston University. My research is focused on studying the robustness of complex networks under attacks from the perspective of statistical physics. Complex networks appear in every aspect of our daily life and are widely studied in Physics, Mathematics, Biology, and Computer Science. Understanding the robustness of complex networks under attacks is crucial for protecting complex systems and designing robust infrastructures. Networks' robustness depends crucially on the structure of the networks as well as the nature of the attacks. This dissertation coverers two major parts of my research on the robustness of complex networks: i) proposing a new type of attack – localized attack and modeling the robustness of complex networks under this type of attack; ii) discovering the clustering structure in complex networks, and investigating it's influence on the robustness of both fully and partially interdependent network of networks.

In Chapter 2, we model a new type of attack which we call the localized attack. Previous research of the robustness of complex networks has focused on two types of attacks: random attack and targeted attack. For random attack, each node in the network is attacked and removed with the same probability. While for targeted attack, the probability for each node to be attacked is dependent upon the degree of each node. However, these two types of attacks fail to describe many real-world scenarios where the damages or failures on the networks are localized. We propose a theoretical framework to study the robustness

of complex networks under localized attack based on percolation theory and generating function method from statistical physics. We investigate the percolation properties, such as size of the giant component in the network, critical threshold of the phase transition where the giant component disappears. We derive the analytical expression for the above properties and compare the result with that of the random attack. Specifically, we find that while RR networks are more robust against localized attack, ER networks are equally robust under both. As for scale-free networks, their robustness depends crucially on the degree exponent λ. We also run simulations on two real-world networks to test our model: a peer-to-peer computer network and an airline network. We find that the real-world networks are much more vulnerable to localized attack compared with random attack. These results can provide useful insights into the protection of networked systems and the design of resilient infrastructures.

In Chapter 3, we present a generating function formalism solution for site percolation on both single and fully interdependent networks with clustering. Clustering quantifies the property for two neighbors of the same node to also be neighbors of each other, forming triangle-shaped configurations in the networks. Unlike random networks in which there is very little or no clustering, real-world networks exhibit significant clustering. We present a mathematical framework for understanding how the robustness of a pair of fully interdependent networks is affected by clustering within the network components. We extended the percolation method for single clustered networks to interdependent clustered networks. We find that interdependent networks that exhibit significant clustering are more vulnerable to random attacks than networks without significant clustering. We also discuss the influence of a change of degree distribution and the degree-degree correlation associated with clustering in the model on the critical threshold of interdependent networks and conclude that p_c for interdependent networks increases when networks are more highly clustered. Our results help to better understand the effect of clustering on the percolation of fully interdependent networks.

In Chapter 4, we further extend our model to a partially interdependent network of

networks with clustering within each network. For each clustering coefficient, the system shows a first-order to second-order phase transition as we decrease coupling strength q between networks. We find that as we increase the clustering coefficient of each network, the system becomes less robust. The influence of the clustering coefficient on network robustness decreases as we decrease the coupling strength, and the critical strength q_c, at which the first-order transition changes to second-order, increases as we increase the clustering coefficient in each network. We also investigate two different joint degree distributions, FDD and DPD. We find that although the percolation threshold is different in two cases, the general conclusion that an increase in the clustering coefficient causes partially interdependent networks to become less robust holds.

Bibliography

[1] H. Albert, R. Jeong and A. L Barabási. Error and attack tolerance of complex networks. *Nature*, 406:6794, 2000.

[2] ben-Avraham D Cohen R, Erez K and Havlin S. Resilience of the internet to random breakdowns. *Physical Review Letters*, 85:4626, 2000.

[3] Newman M E J Strogatz S H Callaway, D. S. and Watts D J. Network robustness and fragility: Percolation on random graphs. *Physical Review Letters*, 85:5468, 2000.

[4] ben-Avraham D Cohen R, Erez K and Havlin S. Breakdown of the internet under intentional attack. *Physical Review Letters*, 86:3682, 2001.

[5] A. L. Barabási and R. Albert. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74:47, 2002.

[6] Derényi *etal*. Clique percolation in random networks. *Physical Review Letters*, 94:160202, 2005.

[7] L. Gallos *etal*. Stability and topology of scale-free networks under attack and defense strategies. *Physical Review Letters*, 94:188701, 2005.

[8] M. E. J. Newman. *Networks: An Introduction*. Oxford University Press, New York, 2010.

[9] A. Bashan, R. Parshani, and S. Havlin. Percolation in networks composed of connectivity and dependency links. *Physical Review E*, 83:051127, 2011.

[10] S.V. Buldyrev *etal*. Catastrophic cascade of failures in interdependent networks. *Nature*, 464:1025, 2010.

[11] R. Parshani, S. V. Buldyrev, and S. Havlin. Interdependent networks: Reducing the coupling strength leads to a change from a first to second order percolation transition. *Physical Review Letters*, 105:048701, 2010.

[12] X. Huang, J. Gao, S. V. Buldyrev, S. Havlin, and H. E Stanley. Robustness of interdependent networks under targeted attack. *Physical Review E*, 83:065101, 2011.

[13] A. Bashan, R. P. Bartsch, J. W. Kantelhardt, and S. Ivanov P C Havlin. Network physiology reveals relations between network topology and physiological function. *Nature Communications*, 3:702, 2012.

[14] J. Gao, S. V. Buldyrev, S. Havlin, and H. E Stanley. Robustness of a network of networks. *Physical Review Letters*, 107:195701, 2011.

[15] J. Gao, S. V. Buldyrev, H. E. Stanley, and S. Havlin. Networks formed from interdependent networks. *Nature Physics*, 8:40, 2012.

[16] J. Gao, S. V. Buldyrev, H. E. Stanley, X. Xu, and S. Havlin. Percolation of a general network of networks. *Physical Review E*, 88:062816, 2013.

[17] C. D. Brummitt and R. M. Leicht E A D'Souza. The extreme vulnerability of interdependent spatially embedded networks. *Proceedings of the National Academy of Sciences*, 109:680, 2012.

[18] G. J. Baxter, S. N. Dorogovtsev, A. V. Goltsev, and J. F F Mendes. Avalanche collapse of interdependent networks. *Physical Review Letters*, 109:248701, 2012.

[19] T. P. Peixoto and S. Bornholdt. Evolution of robust network topologies: Emergence of central backbones. *Physical Review Letters*, 109:118703, 2012.

[20] Y. Shang, W. Luo, and S. Xu. L-hop percolation on networks with arbitrary degree distributions and its applications. *Physical Review E*, 84:031113, 2011.

[21] R. Cohen and S. Havlin. *Complex Networks, Structure, Robustness and Function*. Cambridge University Press, Cambridge, 2010.

[22] A. Bunde and S. Havlin. *Fractals and Disordered Systems*. Springer, 1991.

[23] D. Stauffer and A. Aharony. *Introduction to Percolation Theory*. CRC Press, 1994.

[24] A. Coniglio. Cluster structure near the percolation threshold. *Journal of Physics A: Mathematical and General*, 15:3829, 1982.

[25] S. Neumayer, G. Zussman, and R. Modiano E Cohen. Assessing the vulnerability of the fiber infrastructure to disasters. *INFOCOM*, IEEE:1566–1574, 2009.

[26] Y. Berezin, A. Bashan, M. M. Danziger, D. Li, and S. Havlin. Spatially localized attacks on interdependent networks: the existence of a finite critical attack size. *arXiv:*, 1310.0996.

[27] B. Bollobás. *Random Graphs*. Academic Press, London, 1985.

[28] M. E. J. Strogatz S H Newman and Watts D J. Random graphs with arbitrary degree distributions and their applications. *Physical Review E*, 64:026118, 2001.

[29] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. *Random Structures and Algorithms*, 6:161, 1995.

[30] T. Kalisky, R. Cohen, O. Mokryn, D. Dolev, Y. Shavitt, and S. Havlin. Width of percolation transition in complex networks. *Physical Review E*, 74:066108, 2006.

[31] J. Shao, S. V. Buldyrev, L. A. Braunstein, S. Havlin, and H. E Stanley. Structure of shells in complex networks. *Physical Review E*, 80:036105, 2009.

[32] M.E.J. Newman. The spread of epidemic disease on networks. *Physical Review E*, 66:016128, 2002.

[33] Stanford Large Network Collection. Internet peer-to-peer network data.

[34] Openflight.org. Airport network data.

[35] V. et al Rosato. Modeling interdependent infrastructures using interacting dynamical models. *International Journal of Critical Infrastructures*, 4:63, 2008.

[36] Us-canada power system outage task force: Final report on the august 14th 2003 blackout in the united states and canada. *The Task Force*, 2004.

[37] J. Peerenboom, R. Fischer, and Whitefield R. *Mitigating the Vulnerability of critical Infrastructures to Catastrophic Failures*. Proc. CRIS/DRM/IIIT/NSF Workshop, 2001.

[38] S. Rinaldi, J. Peerenboomand, and Kelly T. Identifying, understanding, and analyzing critical infrastructure interdepedencies. *IEEE Control Systems Magazine*, 21:11, 2001.

[39] O. Yagan, D. Qian, J. Zhang, and D. Cochran. *Optimal allocation of interconnecting links in cyber-physical systems: Interdependence, cascading failures and robustness.* Special Issue of IEEE Transactions on Parallel and Distributed Systems (TPDS) on Cyber-Physical Systems, 2012.

[40] A. Vespignani. The fragility of interdependency. *Nature*, 464:984, 2010.

[41] R. Zimmerman. Decision-making and the vulnerability of interdependent critical infrastructure. *2004 IEEE International Conference on Systems, Man, and Cybernetics*, 5:4059, 2005.

[42] D. Mendonca and Wallace W. Impacts of the 2001 world trade center attack on new york city critical infrastructures. *Journal of Infrastructure Systems*, 12:260, 2006.

[43] B. Robert, L. Morabito, and Christie R. D. The operational tools for managing physical interdependencies among critical infrastructures. *International Journal of Critical Infrastructures*, 4:353, 2008.

[44] D. A. Reed, K. C. Kapur, and Christie R. D. Methodology for assessing the resilience of networked infrastructure. *IEEE Systems Journal*, 3:174, 2009.

[45] E. Bagheri and Ghorbani A. A. A reference model for profiling critical infrastructure systems. *Information Systems Frontiers*, 12:115, 2009.

[46] D. Mansson, R. Thottappillil, M. Backstrom, and Ludvika H. V. V. Methodology for classifying facilities with respect to intentional emi. *IEEE Transactions on electromagnetic compatibility*, 95:46, 2009.

[47] J. Shao, S. V. Buldyrev, S. Havlin, and Stanley H. E. Cascade of failures in coupled network systems with multiple support-dependence relations. *Physical Review E*, 83:036116, 2011.

[48] J. Gao, S. V. Buldyrev, H. E. Stanley, and S. Havlin. Networks formed from interdependent networks. *Nature Physics*, 8:40, 2011.

[49] J. Gao, S. V. Buldyrev, S. Havlin, and Stanley H. E. Robustness of a network of networks. *Physical Review Letters*, 107:195701, 2011.

[50] D. J. Watts and Strogatz S. H. Collective dynamics of 'small-world' networks. *Nature*, 393:440, 1998.

[51] M. A. Serrano and Boguñá M. Clustering in complex networks. i. general formalism. *Physical Review E*, 74:056114, 2006.

[52] M. E. J. Newman and Park J. Why social networks are different from other types of networks. *Physical Review E*, 68:036122, 2003.

[53] M. E. J. Newman. Random graphs with clustering. *Physical Review Letters*, 103:058701, 2009.

[54] J. C. Miller. Percolation and epidemics in random clustered networks. *Physical Review E*, 80:020901, 2009.

[55] J. P. Gleeson. Bond percolation on a class of clustered random networks. *Physical Review E*, 80:036107, 2009.

[56] J. P. Gleeson, S. Melnik, and A. Hackett. How clustering affects the bond percolation threshold in complex networks. *Physical Review E*, 81:066114, 2010.

[57] A. Hackett, S. Melnik, and Gleeson J. P. Cascades on a class of clustered random networks. *Physical Review E*, 83:056107, 2011.

[58] C. Molina and L. Stone. Modelling the spread of diseases in clustered networks. *Journal of Theoretical Biology*, 315:110–118, 2012.

[59] P. Van Mieghem, H. Wang, X. Ge, S. Tang, and Kuipers F. A. Influence of assortativity and degree-preserving rewiring on the spectra of networks. *The European Physical Journal B*, 76:643, 2010.

[60] R. et al Parshani. Critical effect of dependency groups on the function of networks. *Proceedings of the National Academy of Sciences*, 108:1007, 2011.

[61] D. Zhou, G. D'Agostino, A. Scala, and H. E. Stanley. Assortativity decreases the robustness of interdependent networks. *arXiv:*, 1203.0029v1, 2012.

[62] S. Wasserman and K. Faust. *Social Network Analysis: Methods and Applications*. Cambridge University Press, England, 1994.

[63] E. Ravasz and A. L. Barabasi. Hierarchical organization in complex networks. *Physical Review E*, 67:026112, 2003.

[64] E. M. Jin, M. Girvan, and M. E. J. Newman. Structure of growing social networks. *Physical Review E*, 64:046132, 2001.

[65] P. Holme and B. J. Kim. Growing scale-free networks with tunable clustering. *Physical Review E*, 65:026107, 2002.

[66] K. Klemm and V. M. Eguiluz. Highly clustered scale-free networks. *Physical Review E*, 65:036123, 2002.

[67] M. A. Serrano and M. Boguñá. Tuning clustering in random networks with arbitrary degree distributions. *Physical Review E*, 72:036133, 2005.

[68] S. Bansal, S. Khandelwal, and L. A. Meyers. Exploring biological network structure with clustered random networks. *BMC Bioinformatics*, 10:405, 2009.

[69] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin. Catastrophic cascade of failures in interdependent networks. *Nature*, 464:1025, 2010.

[70] J. Gao, S. V. Buldyrev, H. E. Stanley, and S. Havlin. Networks formed from interdependent networks. *Nature Physics*, 8:40–48, 2012.

[71] J. Gao, S. V. Buldyrev, H. E. Stanley, X. Xu, and S. Havlin. Percolation of a general network of networks. *arXiv:*, 1306.3416, 2013.

[72] G. Dong et. al. Percolation of partially interdependent networks under targeted attack. *Physical Review E*, 85:016112, 2012.

[73] J. Shao, S. V. Buldyrev, S. Havlin, and H. E. Stanley. Cascade of failures in coupled network systems with multiple support-dependence relations. *Physical Review E*, 83:036116, 2011.

[74] A. Vespignani. Complex networks: The fragility of interdependency. *Nature*, 464:984–985, 2010.

[75] E. A. Leicht and R. M. D'Souza. Percolation on interacting networks. *arXiv:cond-mat*, 0907.0894.

[76] R. G. Morris and M. Barthelemy. Transport on coupled spatial networks. *Physical Review Letters*, 109:128703, 2012.

[77] S.W. Son et al. Percolation theory on interdependent networks based on epidemic spreading. *Europhysics Letters*, 97:16006, 2012.

[78] S.M. Anna, S.M Ángeles, and M. Bogu ná. Epidemic spreading on interconnected networks. *Physical Review E*, 86:026106, 2012.

[79] S. Gómez et al. Diffusion dynamics on multiplex networks. *Physical Review Letters*, 110:028701, 2013.

[80] J. Aguirre, D. Papo, and J. M. Buldú. Successful strategies for competing networks. *Nature Physics*, 9:230–234, 2013.

[81] C. D Brummitt, R. M D'Souza, and E. A Leicht. Suppressing cascades of load in interdependent networks. *Proceedings of the National Academy of Sciences*, 109:680–689, 2012.

[82] C. M Schneider et al. Towards designing robust coupled networks. *Scientific Reports*, 3:1969, 2013.

[83] R. Parshani et al. Inter-similarity between coupled networks. *Europhysics Letters*, 92:68002, 2010.

[84] Y. Hu, B. Ksherim, R. Cohen, and S. Havlin. Percolation in interdependent and interconnected networks: Abrupt change from second- to first-order transitions. *Physical Review E*, 84:066116, 2011.

[85] S. V. Buldyrev et al. Interdependent networks with identical degrees of mutually dependent nodes. *Physical Review E*, 83:016112, 2011.

[86] A. Bashan, Y. Berezin, S. V. Buldyrev, and S. Havlin. The extreme vulnerability of interdependent spatially embedded networks. *Nature Physics*, 9:667, 2013.

[87] D. Cellai, E. López, J. Zhou, J. P. Gleeson, and G. Bianconi. Percolation in multiplex networks with overlap. *arXiv:*, 1307.6359.

[88] F. Radicchi and A. Arenas. Abrupt transition in the structural formation of interconnected networks. *arXiv:*, 1307.4544.

[89] M. De Domenico, A. Sole, S. Gomez, and A. Arenas. Random walks on multiplex networks. *arXiv:*, 1306.0519.

[90] X. Huang, S. Shao, H. Wang, S. V. Buldyrev, S. Havlin, and H. E. Stanley. The robustness of interdependent clustered networks. *Europhysics Letters*, 101:18002, 2013.

[91] A. Hackett, S. Melnik, and J. P. Gleeson. Cascades on a class of clustered random networks. *Physical Review E*, 83:056107, 2011.

[92] P. Erdős and A. Rényi. On random graphs. i. *Publicationes Mathematicae*, 6:209, 1959.

[93] P. Erdős and A. Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5:17, 1960.

[94] B. Bollobás. *Random Graphs.* Academic, London, 1985.

[95] J. P. Gleeson, S. Melnik, and A. Hackett. How clustering affects the bond percolation threshold in complex networks. *Physical Review E*, 81:066114, 2010.

Curriculum Vitae