Interdependent Cyber Physical Systems: Robustness and Cascading Failures

by

Zhen Huang

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements for the degree of

Doctorate in Philosophy in

Electrical and Computer Engineering

Ottawa-Carleton Institute for Electrical and Computer Engineering
School of Electrical Engineering and Computer Science
Faculty of Engineering
University of Ottawa

“What we cannot speak about we must pass over in silence.”

Ludwig Wittgenstein

Abstract

The cyber-physical systems (CPS), such as smart grid and intelligent transportation system, permeate into our modern societies recently. The infrastructures in such systems are closely interconnected and related, e.g., the intelligent transportation system is based on the reliable communication system, which requires the stable electricity provided by power grid for the proper function. We call such mutually related systems *interdependent networks*.

This thesis addresses the cascading failure issue in interdependent cyber physical system. We consider CPS as a system that consists of physical-resource and computational-resource networks. The failure in physical-resource network might cause the failures in computational-resource network, and vice versa. This failure may recursively occur and cause a sequence of failures in both networks.

In this thesis, we propose two novel interdependence models that better capture the interdependent networks. Then, we study the effect of cascading failures using percolation theory and present the detailed mathematical analysis on failure propagation in the system. By calculating the size of functioning parts in both networks, we analyze the robustness of our models against the random attacks and failures.

The cascading failures in smart grid is also investigated, where two types of cascading failures are mixed. We estimate how the node *tolerance parameter* T (ratio of capacity to initial workload) affect the system performance. This thesis also explores the small clusters. We give insightful views on small cluster in interdependent networks, under different interdependence models and network topologies.

Keywords: Cyber Physical Systems; Interdependent Networks; Smart Grid; Percolation Theory

Acknowledgements

I would give my sincerest gratitude to my thesis supervisors Dr. Amiya Nayak and Dr. Ivan Stojmenovic, for their brilliant scientific guidance and steady financial support through the whole doctoral program. They have helped me grow in the scientific field, and made me a strong researcher. I am proud to be a member in this group and have worked with them.

My Special thanks go to Dr. Cheng Wang from Tongji University, who I worked with along my doctoral program. We had so many wonderful discussions on both emerging scientific topics and our life philosophies. He gave me valuable advices, suggestions and encouragements on this thesis work.

I would also like to thank our former group members Dr. Sushmita Ruj, Dr. Dandan Liu and Dr.Tieyin Zhu. The discussions with them gave me an eye to find more interesting areas.

I wish to express my cheers to my friends whom I learned from. We will be friends as we were before.

Finally, my gratitude and huge hug go to my parents Mr. Yiping Huang and Mrs. Guiqin Zhang, who always understand and support my ideals. This work is dedicated to them.

Contents

List of Figures

List of Tables

Dedicated to my parents

Chapter 1

Introduction

A cyber-physical system (CPS) is a system that integrates pervasive sensing, computation, communication and actuation together to infuse engineered systems with intelligent capabilities. The research advances in CPS promise to transform our world with new relationships between computer-based control and communication systems, engineered systems and physical reality. The CPS will require a comprehensive effort from computer scientists, engineers, and physical domain experts to develop new theories, methods, tools, as well as new hardwares and software components, run-time substrates.

Based on the emerging technologies, we will be able to develop the new generation of intelligent infrastructures, smart highways, factories and robotic systems. Fig. 1.1 indicates the frame of the cyber physical system. The physical devices, such as cell phones, vehicles and sensors are viewed as the physical components. The embedded computers, servers and communication networks are considered as the cyber components [5]. The data gathered by the physical components is passed to the cyber components, which compute and handle the huge amount of information. The cyber components provide the control information to the physical components.

Imagine, for example, the fast moving vehicles, from small cars to huge trucks, can recognize and report the traffic conditions, automatically switch information

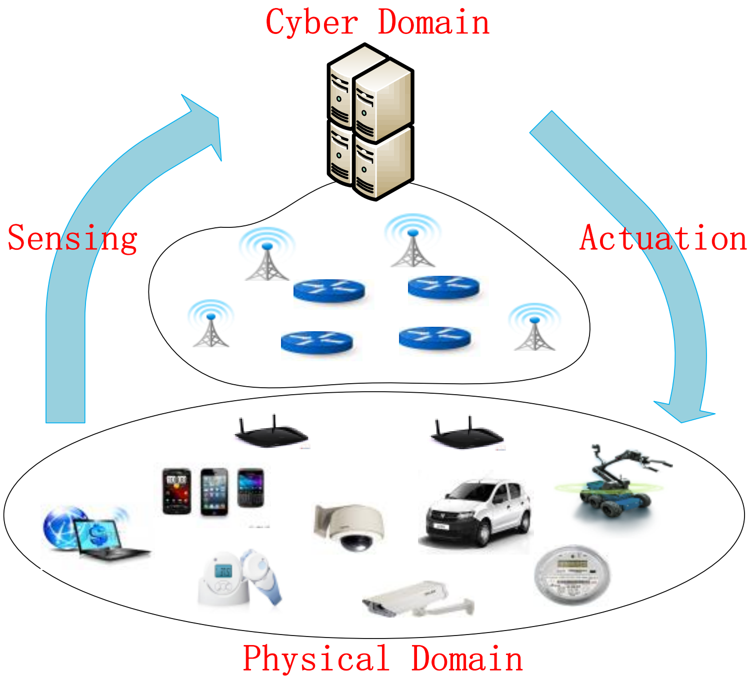


FIGURE 1.1: Cyber-physical system

with other peers. By introducing Android/IOS, the vehicles could cooperate to solve more complex issues online. One emerging application is the self-diagnosis system that could provide online car-monitor functionality to the drivers anytime and everywhere. More intelligently, the vehicles could call for their own site-specific maintenance, contribute to the cloud-based real time road condition map, dynamically choose the optimal paths to the destination. On the other hand, highways and roads are not only 'road'. They monitor the events and conditions, and thus could play the role as the managers of the traffic flow.

Another typical example of CPS is smart grid, which inspires the plenty of research topics. The entire electrical power system from generation to its use should incorporate sensing, computing, and communication capabilities.

1.1 Motivation

The cyber-physical system refers to the tight conjoining and coordination between *computational* and *physical* resources [6, 7]. The tight and complex interdependence between the cyber and physical components is one of key characteristics of CPS. The current research on CPS mainly aims to improve the ability to understand and exploit the interfaces between the cyber and physical worlds and to innovate new behaviors and capabilities from their seamless integration [8].

In a typical CPS system, embedded computers and communication networks are sort of computational resources, while the actuators and sensors are viewed as physical resources. Nowadays, computational resources and physical resources are strongly coupled. That is, computational resources schedule and control physical resources, on the other hand, the proper functioning of computational resources requires steady operation of physical resources. As one of the important applications of CPS, smart power grid is composed of power grid and communication/-control network. Communication network operates depending on power grid for the electricity. Power grid, from another point of view, has to be controlled by communication network. The two networks are mutually-close connected. Similar to smart power grid, interdependence also exists in the wireless sensor robot/actuator network. Robot/actuator receives or transmits messages through sensor network. While, sensor placement, maintenance and relocation are completed by robot/actuator.

Accordingly, CPS could be considered as two interdependent large-scale networks, i.e., *physical-resource* and *computational-resource* networks. These two networks are mutually connected and dependent.

In such interdependent networks, cascading failure is an important issue. It is an extensively studied topic in the individual large-scale network [9–14]. A cascading failure in a system occurs if the failure of a part can trigger the failure of remaining successive parts [9]. When it happens, the nearby parts have to take the responsibility of the failure nodes. The overload of the neighbors may lead to more failures. Such a failure could happen in many types of physical-resource and computational-resource networks, such as power grids, Internet, and transportation networks.

While in cyber-physical systems, the relation between physical-resource and computational-resource networks leads to another type of cascading failure that is caused due to the interdependence, which is more complex to be analyzed than the cascading failure in a single system [15]. The unexpected malfunction of nodes initiates the sequence of failures not only within the single network, but also between the networks. One good example is the huge blackout happened in Italy in 2003: the initial breakdown of a few power stations led to the overload of their neighbors, which consequently led to the cascading failures within the power grid. Since the communication network relies on power grid, a part of communication nodes was affected and thus got failed. As a result, the operation of the power grid was influenced and some stations were disordered. This process continued and recursively affected both networks.

In this case, two types of cascading failures happened simultaneously, where they could be roughly classified to: **interdependence cascading failure** and **load propagation cascading failure.** The former one is caused due to the interdependence relation, the latter one is caused due to the overload propagation within the single network.

The interdependence in cyber-physical system is growing, which leads to the system more fragile than the single network [3]. Our aim in this thesis is to reveal the

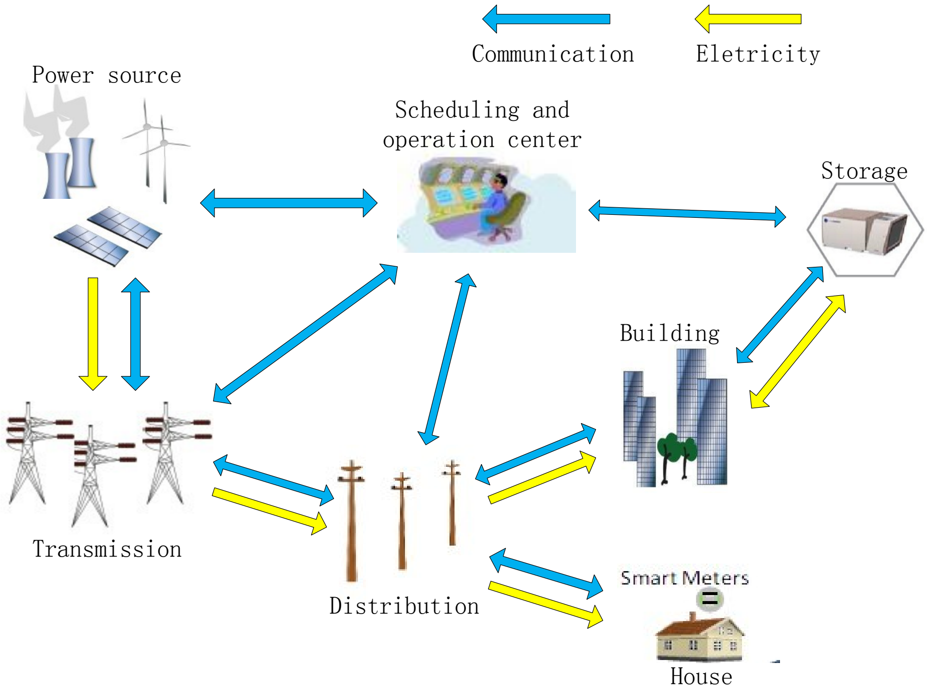


FIGURE 1.2: Frame of smart grid

failure propagation process in CPS and estimate how robust the system is under the random failure.

1.1.1 Important Case: Smart Grid

Conventional electrical grids utilize centralized command and control structures, e.g., SCADA (Supervisory Control And Data Acquisition) system relying on human monitors for identifying faults and decision making. Massive blackouts occurred in the past since the existing system lacks real-time control ability, e.g., the recent huge blackout happened in July 2012, affected more than 600 million people in India [16].

The smart grid concept addresses the real-time control and energy efficiency. It is an electrical grid that integrates information and communications technology and different sources of power generation e.g., fossil-fuel, solar and wind. It predicts the electricity demands in different regions, monitors the power consumption of customers using smart meters, and deals with system failure rapidly. Smart grid standardization depends on three technologies: power engineering (for interoperability), communications technologies (to provide communication grid side-by-side to power grid), and information technology (to resolve privacy and security, provide data integrity and interfaces, and intelligent control). Fig 1.2 gives the frame of smart grid.

As one of the important CPS applications, in this thesis, smart grid is studied and analyzed to demonstrate the general application of this research. However, we are not considering the physical properties, such like power transmission line limit and voltage sag.

1.2 Thesis Goals

Recently, there are some studies on interdependent networks to assess system vulnerability due to cascading failures, such as [15, 17–20]. The common assumptions of these works for a node to properly operate are: 1) the node has interdependence relation; 2) the node necessarily belongs to the *giant component*, which is defined as the largest connected subnetwork that fills most of the entire network [21]. By calculating the fractions of remaining giant component in the final steady state, they studied to what extent the initial failures could impact the whole interdependent networks. The models proposed in [15, 22] are 'one-on-one' interdependence, within which the node properly operates relying on one unique node in the other network. Such models do not correctly reflect the characteristics of cyber-physical systems since the physical-resource and computational-resource are not one-on-one. The 'multiple-to-multiple' interdependence models developed in [20] assumed that each node is assigned with the same number of inter relations. Each node functions as long as at least one of its supporting nodes is operating. [23] points out that the interdependence in real world network is *unidirectional* rather than *bidirectional*. To find the critical nodes that might cause fatal cascading failures, [17] indicated a near-optimal heuristic greedy algorithm. Cascading link failures was discussed in [24]. Huang *et al.* [25] introduced the targeted failure rather than the random malfunction. [3] pointed out the different roles of nodes, and indicated a multiple-to-multiple interdependence model. [26] studied the robustness of n interdependent networks, with partial interdependence relationship. The targeted attack was studied in [27].

On the other hand, current research on robustness in smart power grid are mainly focusing on load distribution and malicious attacks. An architecture for distributed generation, which can help prevent cascading failures, is described in [28]. To decrease the impact of cascading failure or even to prevent it, [29] analyzed the tripping of overloaded lines and proposed a model to control these lines. [30] proposed three mitigation strategies and simulated them on real-world network structures to find an effective way for reducing cascading failure. Software overlays and multiple routes to deliver critical data to prevent failures were studied in [31]. Load distribution attack was deeply investigated in [32] to provide effective prevention on false data injection. The work on load propagation cascading failures can be found in [33–38]. Hines *et al.* [33] introduced the backgrounds of power node overload cascading failures. Some load propagation schemes were proposed and discussed in [34]. [39] developed a new vulnerability assessment by combining attack methodology with DC flow.

The current works on interdependent networks are by topological methods and they only study the cascading failure caused due to interdependence between networks. On the other hand, the studies on load propagation cascading failure are within the single power grid, and thus, can not be applied to smart grid because the key feature of interdependence is missing.

In this thesis, we study the cascading failure in interdependent cyber

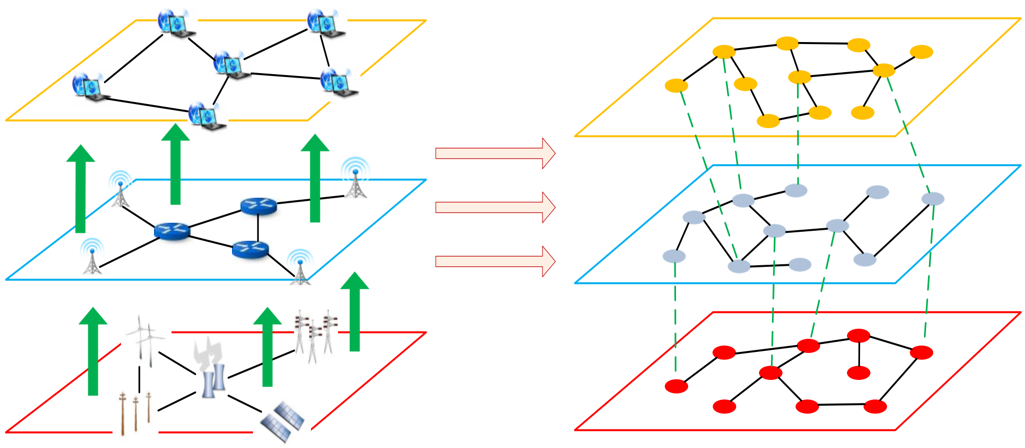


FIGURE 1.3: Modeling interdependent cyber-physical system

physical system. We limit the work into the scope on network topology and network theory. Taking communication network protocols and electrical features into account would be our future research directions.

1.2.1 Milestone 1: Modelling Interdependent Cyber-Physical System Using Complex Network Theory

As the aim is to study the CPS robustness, a practical interdependence model is thus necessary. While, an unique model of a generic cyber-physical system appears to be infeasible due to the diversities of actuation and physical world reaction. The challenge is to extract the common ingredients and components of CPS, model and investigate them, and combine and apply them to the certain categories of CPS.

The existing interdependence models can not represent and apply to all cyber-physical systems. The first aim of this thesis is to indicate the practical interdependence models by using complex network theory, as shown in Fig. 1.3. We are using the complex network models to feature the topologies of the real world networks, and indicating the inter relation between the nodes within different networks.

1.2.2 Milestone 2: Analyzing CPS Robustness Against Cascading Failures

Based on the models proposed, we are interested in how the entire system behaves once the cascading failures occur. The interdependent networks are more fragile than the single network, as the *first-order transition* has been found in such a system [22], where there exists a threshold for the proportion of initial failure, above which the entire system collapses. Different network topologies and interdependence models could lead to quite different system robustness.

In this thesis, the widely used *giant component* is introduced to show the fraction of survived nodes from cascading failures in the system. We aim to find an efficient mathematical tool to estimate the giant components, by given the initial failure. This tool could be extended and applied to some more general cases.

1.2.3 Milestone 3: Smart Grid Case Study

As mentioned above, smart grid is an important cyber-physical system in our daily life. The communications and control infrastructures need energy to properly operate, while the power stations and electricity transmission are controlled by *operation centers*. The reliability of such system is worthy to focus on. The failure in either of them may lead to the failure in another, and continues in a recursive manner.

However, the current work on smart grid cascading failure is by pure topological methods, which can not represent the real smart grid properly due to the lost of electricity characteristics [40]. On the other hand, the studies on load propagation cascading failure are within the single power grid, and thus, can not be applied to smart grid because the key feature of interdependence is missing.

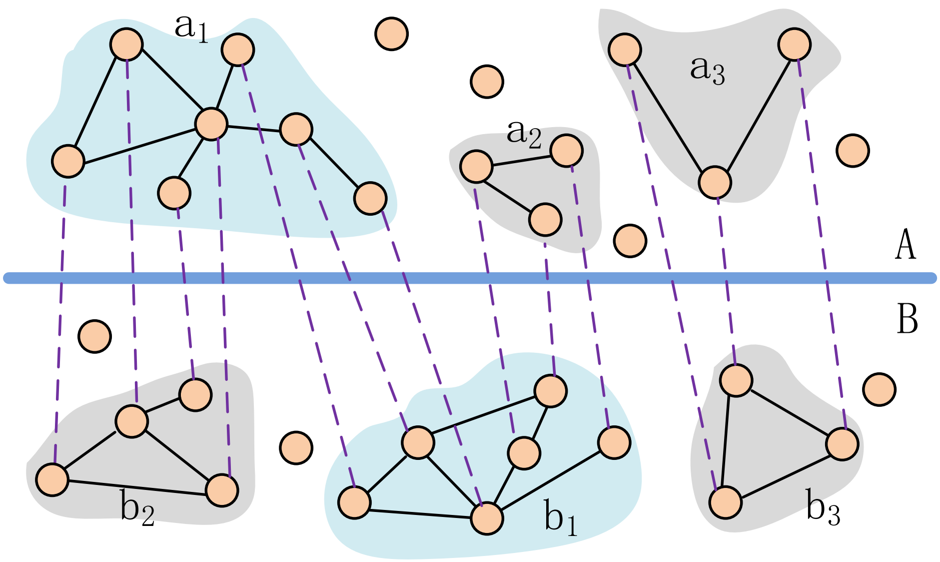


FIGURE 1.4: Two 'one-to-one' interdependent networks A and B fragment into disconnected clusters and pieces. Giant components a1 and b1 are mutually dependent, thus function. The small clusters a2, a3, b2 and b3 also operate because the nodes involved have inter links. Thus, not only the giant component, but also the small clusters work properly in this steady state.

The third aim of this thesis is to study the joint effects of interdependence cascading failure and load propagation cascading failure in smart power grid.

1.2.4 Milestone 4: Small Cluster

The common underlying assumptions of existing works are that a node can function only if: *1) it belongs to the giant component; 2) it has at least one inter link*. While, the first assumption is demanding and incompatible with many engineering reality. In many application scenarios, once the network gets fragmented due to the cascading failures, except the giant component, the network possibly has plenty of *autonomous* and *self-sufficient* disconnected clusters.

Fig. 1.4 gives an example of two networks A and B with 'one-to-one' [15] interdependence, where both of them fragment to clusters after the cascading failures. Network A has a giant component a_1, and two small clusters a_2, a_3. B is similar with A. Notice that except the two giant components, the small clusters a_2, a_3,

b_2 and b_3 do operate because the nodes involved have the supporting inter links. This example illustrates that in the interdependent networks, if we only consider the giant components, some small clusters that could function are neglected. Consequently, the results the mentioned studies obtained are not accurate enough to reflect the real coupled cyber physical systems. It is convincing that such ignored clusters could play important roles in the network of some real-life systems.

Therefore, the fourth aim of this thesis is to study the small clusters caused by cascading failures in interdependent networks.

1.3 Contributions

In this research, we indicate two practical interdependence models for cyber-physical system, study the effect of cascading failure using percolation theory and present detailed mathematical analysis of the failure propagation. We analyze the robustness of our model in terms of random attacks or failures by calculating the size of functioning parts after the cascading failures stop.

1.3.1 Contribution 1: One-Multiple Model

Chapter 3 indicates an 'one-multiple' interdependence model for cyber-physical system. We adopt *balls-bins* model to define the interdependence relation between physical-resource and computational-resource networks. Each computational station has exact one inter link from physical-resource network, assuming that the computational node need support from one specific physical-resource station. Each physical-resource station can provide support to multiple computational stations, then has possibly multiple outgoing inter links. We consider both physical-resource network and computational-resource network as scale-free networks [21], in which the node degrees follow power law distributions. We then study the effect of cascading failure using percolation theory and present detailed mathematical analysis of the failure propagation in the interdependent system. We analyze the robustness of our model in terms of random attacks or failures by calculating the size of functioning parts after the cascading failures stop.

Our mathematical analysis proves that for the interdependent CPS, given the power law exponents for physical-resource and computational-resource network, there always exists a critical value of the initial failure fraction of entire network, above which the entire system collapses, i.e., the giant components disappear; Below which the system can still function after the cascading failures stops, with loss of part of the system.

The simulations show that when the power law exponents for physical-resource and computational-resource networks are both 2.2, the critical value is approximately 49%. It is 15% when the power law exponents are 2.5 for both networks. Thus, critical value appears sensitive to the power law exponent. The size of physical-resource and computational-resource networks, and the ratio between their sizes do not affect the system robustness.

1.3.2 Contribution 2: Extended 'k-n' Interdependence Model

In Chapter 4, we propose a more general interdependence model for cyber physical systems. We use smart grid as the example to elaborate this model. In Section 4.6, we discussed the possible applications of this model.

For smart grid, we realize that each power station could be controlled by multiple distinct operation centers, and functions as long as at least one of its operation centers is working. We believe that the more control relations (links) and devices the system has, the higher cost is required. We define the *system controlling cost* as proportionally dependent on the number of links between operation centers and power stations. Our primary interest is to find out the relation between controlling cost and system robustness.

We propose an novel 'k-n' model, that each power station is operated by k distinct operation centers, and each operation center could monitor and control n power stations. The defined 'system controlling cost' is simply equal to k units, and the 'monitoring capability' for each operation center is n. We follow a three-step scheme to construct 'k-n' interdependent network. Both power grid and communication networks are type of scale-free networks [21], in which the degree distribution follows power law. One important parameter for discussing the smart grid robustness is the fraction of properly functioning nodes, i.e., survival ratio. Using percolation theory, we calculate the survival ratio of the system. We mathematically study the relation between system robustness, controlling cost and monitoring capability.

Our analysis shows that the system robustness experiences a sub-linear improvement with the increase of k, while n has no impact on system robustness when the size of network is large. The extensive simulation validates our analysis. The system robustness against random failure can be improved by increasing monitoring cost though they have nonlinear relation. Meanwhile, the critical value for a higher k is smaller, which means the system can tolerate a higher fraction of random failures. The robustness improves between $k = 1$ and $k = 2$ significantly, while the gap between $k = 10$ and $k = 15$ is small. Hence, for building a smart power grid, adding as many as possible control link is not a good strategy since the massive extra cost does not improve system reliability significantly.

Our experimental results illustrate that higher n improves system robustness slightly and at the same time decreases the critical value. But the disadvantage is that the transition becomes flat. Therefore, for building smart grid infrastructures, we need carefully choose the value of k and n depending on our demands.

Compared with the existing works [15, 20, 22], the proposed 'k-n' model is more practical because that we believe nodes in network are not equal as they play different roles. Therefore, we divide the nodes into different categories. This model enables us to understand how the system performs under different monitoring schemes, and the relation between robustness with monitoring cost.

1.3.3 Contribution 3: Joint Cascading Failures in Smart Grid

In Chapter 5, by integrating the topological method and load redistribution, we study the joint effect of interdependence cascading failure and load propagation cascading failure to be more closer to the real-world smart grid. We use the load propagation model in [34] where the whole workload of removed node is equally redistributed to its neighbors. For the interdependence relation, we apply our previous 'k-n' model. First of all, due to the capacity limit, the number of inter links each power node has is restricted. Thus, we propose a new modified *Balls and Bins* allocation scheme for 'k-n' model. Assuming that both power grid and communication network are scale-free [41], in terms of random initial failure, we develop a method to estimate the node *tolerance parameter* threshold T_c (the ratio of capacity to initial workload) in single power grid, below which the grid suffers from the load propagation cascading. A mathematical tool is proposed to calculate the fraction of survivals in power grid after the load propagation failure stops.

We indicate a percolation-based mathematical method to simulate the process of interdependence cascading failure and obtain a set of transcendental function for the fractions of survivals in both power grid and communication network. Similar to the single power grid, we find a method to estimate the *tolerance parameter* threshold T'_c for smart grid. Our analysis demonstrates that given the same network topologies and initial removals, $T'_c > T_c$. We also find the relation between T'_c and 'k-n' interdependence, that higher k indicates lower T'_c. The system robustness approaches to the upper bound once tolerance parameter $T \to \infty$. We proves that this upper bound is the performance when the system only suffers from the interdependence cascading failure.

The extensive simulations determine the values of T'_c and validate our mathematical analysis. Besides, we observe a super linear relation between the initial removals with T'_c in the specific system. Also, the system robustness of single power grid is better than smart grid, given the same parameters. The interesting finding is that, in smart grid, although the initial failure occurs in power grid, the fraction of survivals in power grid is always greater than that in communication network.

1.3.4 Contribution 4: Understanding Small Clusters

In Chapter 6, using real-world network data and extensive simulations, we first introduce the small cluster overviews, including the size distribution and the number of clusters, in single complex network. We show that the clusters in Erdös-Rényi random network are tiny, e.g., size is less than 5. While, the cluster size in scale-free network and small-world network could be much larger.

We indicate a percolation based mathematical analysis method to study how small clusters be affected by the interdependence between two coupled networks. We prove that both the fraction and the number of operating small clusters have upper bounds.

Without loss of generality, in experiment part, we use both synthetic network and real network data to study the impacts of two different interdependence models 'k-n' model [3] and 'one-to-one' model [15] on small clusters. We find that the fraction of small clusters do not monotonously increase as the growing of compactness of interdependence. The simulation validates our mathematical analysis that an upper bound for the fraction of small clusters exists. We also show that except the giant component and small clusters, the fragmentation of interdependent networks also bring massive tiny pieces, i.e., isolated single vertex.

Considering the initial failure could be either random or targeted, to this end, three different attack strategies are discussed: *Inter Degree Priority Attack*, *Intra Degree Priority Attack* and *Random Attack*. The interesting finding is that the proportion of functioning small clusters keeps stable and is independent from attack strategies, which, at the same time, leads to quite different sizes of the giant components. We observe that the Intra Degree Priority Attack destroys the whole system very effectively, compared with others.

Our main contributions in this part of work are:

- Show small cluster size distributions in various single networks
- Indicate a percolation based math tool to study the impacts of different interdependence models on small cluster in coupled networks
- Use extensive simulations to give insightful views on small cluster in interdependent networks, under different interdependence models and network topologies

1.4 Thesis Outline

The remainder of this thesis is organized as follows:

Chapter 2 navigates the readers the background of the recent advances on interdependent networks, as well the cascading failures in such a system.

Chapter 3 proposes an 'one-on-one' interdependence model and focuses on studying the system performance against the cascading failures.

Chapter 4 is dedicated to analyze the smart grid, where an specific 'k-n' model is proposed.

Chapter 5 indicates a method for us to understand the joint effect of interdependence cascading failures and load propagation cascading failures in smart grid.

Chapter 6 is focusing on studying the small clusters under different interdependence models and network topologies.

Chapter 7 wraps up the thesis and enunciates potential research directions.

Chapter 2

Literature Review

The major topics related with this thesis are discussed in this chapter. Cyber-physical system finds direct applicability in a wide range of areas, including:

- Actuator-sensor network: sensor monitors the area and gathers the information, which is manipulated and conducted by the actuators. Robot sensor network is a very general application [42, 43]. The communication is through the sensor network, while the robots cooperate to place the minimum number of sensors to cover the whole area.

- Smart grid: aim to integrate electricity resources (wind, power plant and water), facilitate the intelligent monitoring, schedule the electricity distribution and dynamically adjust the pricing schemes [44, 45].

- Intelligent transportation systems: to combine computing, communication, and storage capabilities with monitoring and control of vehicles to deal with the challenges of safe, green, and efficient transportation [46, 47].

- Smart phone network: the smart phones held by users consist of a huge computational system, where each person and device are the source of information. Social-based and cloud-based applications form a very emerging cyber-physical network.

The readers will first encounter the state-of-the-art research on the communication, computation and architecture applied in CPS. Then, the very recent work on interdependent networks and smart grid will be introduced. The specific algorithms used to deal with cascading failures and load propagation in power grid are discussed.

2.1 Cyber-Physical System: State-of-the-art

A part of existing research on CPS are focusing on designing analytical CPS model. A unique model of a generic CPS appears to be infeasible due to specifics of actuation and physical world reaction. The challenge is to identify common ingredients and components of CPS present in variety of scenarios, model and investigate them, and combine and apply them to certain categories of CPS and corresponding concrete systems. One example is modelling the interconnection and interdependency between cyber and physical systems, or CPS network elements. The solutions to the remaining objectives are model dependent. Existing work on modelling is primarily about extracting properties from physical systems and assumed associated cyber system and matching with some network families. For example, Wang et al. [48] proposed an algorithm that generates random topology power grids featuring the same topology and electrical characteristics derived from the real data. Derler et al. [49] focused on the challenges of modeling CPSs that arise from the intrinsic heterogeneity, concurrency, and sensitivity to timing.

Specific technologies applied in a particular CPS include hybrid system modeling and simulation, concurrent and heterogeneous models of computation, the use of domain-specific ontologies to enhance modularity, and the joint modeling of functionality and implementation architectures [49].

In cyber physical system, the physical device such as battery, sensors are viewed as physical components. The embedded computers and communication networks are considered as cyber components [5]. The interaction and relation between physical components and cyber components are essential to estimate the remaining battery life of the system. Zhang *et al.* [5] proposed an approach to study the characteristic behaviors of each component in battery-based CPS, and measured the robustness of scheduling algorithms as well as battery management schemes.

2.1.1 Machine to Machine Communication

Machine to Machine (M2M) communication allows both wired and wireless devices communicate with each other with little or no human intervention. M2M uses a monitoring or sensing device to capture the events or parameters, such as humidity and temperature, which is relayed through a wired or wireless network to the devices that translate the captured event into meaningful information, which can trigger an actuation [50]. Examples include telemetry, industrial, automation, and SCADA (in power grids) applications. Existing M2M concepts incorporate a *central point* for gathering information, making decisions, and acting. The information may be gathered via, for example, wireless sensor network by single-hop or multi-hop communication, to the central point. The real-time monitoring capability is one of crucial requirements. A good introduction of M2M, as well as its applications, major technology gaps and standards are studied in [51].

Lu *et al.* [52] explored the architecture and technologies for M2M from three perspectives: energy efficiency (green), reliability, and security (GRS). They applied the scheduling algorithm proposed in [53] to improve the energy efficiency, where the 'sleep-active' mode is used for each node. The time is divided to slots. In 'sleep' slot, the node listens to the neighbors' messages, based on which it decides to be 'active' or not. The reliability is achieved by using the hardware redundancy, where two servers are standing at the gateway to deal with the arrival packets. This scheme, apparently, leads to a high cost.

Niyato *et al.* [54] introduced a M2M network architecture for home area, where the energy information from a few neighbored houses is coupled together using a gateway, where they called *concentrator*. Then the information is collected and passed to the operation centers through the base stations. They proposed a dynamic programming algorithm to optimally allocate the concentrators in the whole network to minimize the communication cost. Their simulation results showed that the small cluster size indicates high concentrator installation fees, while large cluster suffers considerable packet delay and loss.

Fadlullah *et al.* [55] surveyed a number of existing communication technologies that can be applied on smart grid. They concluded that in the home area, Zigbee is a better candidate than Ultra-wide band (UWB), WiFi and Bluetooth, which either has a long communication latency or has a high power consumption.

The above work was on the centralized M2M system. Fu *et al.* [56] proposed the data gathering mechanisms to deal with the energy issue for mobile M2M nodes. They indicated both *energy-efficient centralized reporting scheme (ECR)* and *energy-efficient distributed reporting scheme (EDR)*. In ECR, the M2M gateway is reasonable to schedule the data gathering from nodes. In EDR, this transmission schedule is determined by each mobile node. The authors formulated the problems to a integer linear programming, and proved that the energy minimization problem is NP-hard. Therefore, they proposed a greedy algorithm to find the near-optimal solution, with low computational complexity. The simulation results

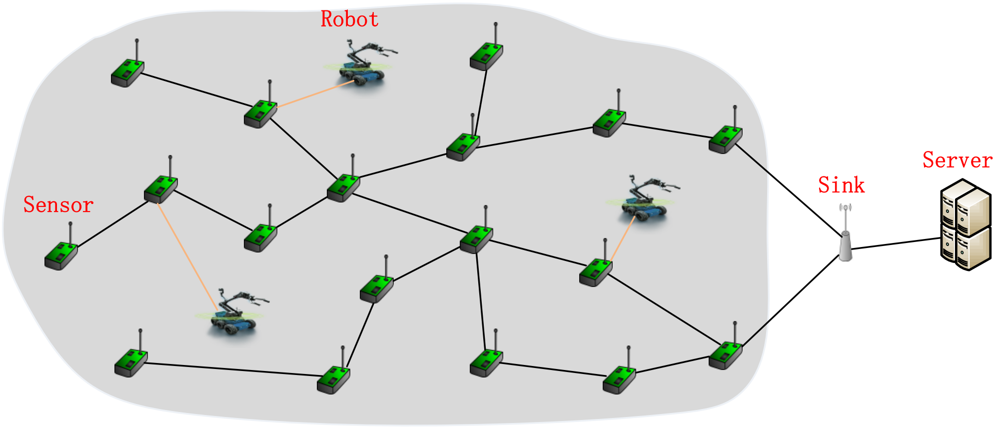


FIGURE 2.1: Robot sensor network

showed that both ECR and EDR are efficient on saving energy,while in the mobile case, EDR is much better.

The existing research was based on the small scale M2M models, and most of them were based on the centralized schemes [57]. Stojmenovic [57] surveyed the existing work on M2M communications architecture, data aggregation, intrusion detection, security and privacy.

2.1.2 Robot-Sensor Networks

Robot-sensor networks are widely used to monitor the area and gather the information. The data is passed to the server through the *sink node* on the edge of the area, as shown in Fig. 2.1. A team of robots is applied to deploy the sensors to fully cover the required area. There are some research objectives in this area [58]:

- Sensor placement and relocation [59–61] . The mobile sensors coordinate for self-deployment, or a team of robots carries the sensors to the target positions. The aim is to fully cover the area to satisfy both sensing and communication requirements. Sensors report failures of neighbouring sensors or lack of area coverage. Robots (with limited carrying capacity) search for spare sensors and move them to replace failed ones or to mitigate connectivity weakness.

- Robots coordination and task management [62–66]. To save the energy and time, a team of robots has to dynamically cooperate to select a required number of robots closest to a given event location to minimize event response delay, and find a optimal movement and carrying strategy. More importantly, once a sensor fails, the robots need to find a optimal solution, either place a new sensor or move a neighbored sensor, to replace it.

- Data aggregation and routing [67–70]. Information obtained by sensors needs to be transmitted to the sink nodes or actuators. Choosing a short and efficient path from source node to destination is essential for saving the energy and improving the lifetime of the network. The solution should handle node mobility, intermittent connectivity, presence of other traffic and collision issues at the medium access, and should provide QoS for different messages and levels of urgency.

- Energy consumption [71]. Energy efficiency is a crucial topic in wireless sensor network, as it defines the lifetime of the whole system. Data aggregation and routing protocols require energy for sending and receiving message from node to node. The node could run out of the energy, and thus stops functioning and generates a hole, which leads to the uncovered area and disconnected communication. A common method in existing algorithms use 'sleep-active' mode for the sensors to save the energy, where the sensor only transmits the message in 'active' state.

Wang *et al.* [60] proposed a sensor deployment scheme, where the mobile sensors could move from dense areas to sparse areas to improve the overall coverage. The uncovered area, called 'coverage hole', is detected by the static sensors, which give a bid depending on the size of the coverage hole. Each mobile node has a price for serving the hole, while the price is depending on the size of the new holes generated by the node movement. The mobile sensors choose the highest bid and move to cover this area. The proposed algorithm was centralized, as the bid and price information has to be disseminated to the all mobile sensors.

Abbasi *et al.* [61] studied the sensor relocation in a wireless sensor network when a critical node is failure. A centralized algorithm LeDiR was proposed, where each node is aware of the entire network topology and routing table. The sensors detect the failure node and find the smallest block cluster, which is the one with the least number of nodes. Then, the neighbored node of failure node, as the same time, belonging to the smallest block cluster, moves towards the hole to replace the faulty node. They called the moving node as 'parent'. The sensors connected with the 'parent', referred as 'child', are moving towards it to maintain the connectivity. This algorithm does not require to deploy new sensors in the hole, thus reduce the cost and overhead. On the other hand, it creates new holes because the 'child' sensors move to cover the hole generated due to the moving of 'parent' sensor.

Several robotic sensor self-deployment techniques were developed in the literature. [72] proposed an adaptive triangular deployment algorithm to maximize the coverage area. They divided the communication range to six sections, within which each node adjusts the relative distance to its one-hop neighbors. [73] indicated a new deployment scheme regardless the network density, sensor communication range and obstacles in the field. In [65, 74], the robot followed snake-like deterministic sensor placement which can leave uncovered areas behind. Fletcher *et al.* [75] resolved this problems by combining the snake-like algorithm with a novel back-tracking technique. The robot moved forward along the grid in open directions until a dead end is reached, where the four directions are obstructed. Then it moved back along the backward trace to the nearest uncovered vertex. Pei *et al.* [76] studied the efficient use of mobile robots to sense not only the region, but also deploy relays to build the networking infrastructure. Essentially, a centralized heuristic solution generates tours by clustering points, generating optimal single-traveller tours, and tour pruning and balance.

Robot task assignment is a complex combinatorial optimization problem involving single/multiple events, single/multiple robots, energy and time constraints, and any other application-specific requirements. [64] considered a single robot single task scenario and proposed auction aggregation protocols for task assignment in multi-hop wireless robot networks. An event, detected by a sensor, is reported to a robot collector using any sensor-robot coordination framework. This robot organizes an auction to select the best robot to respond to the event, so that the communication cost, and the cost of performing the task are minimized. [66] balanced the workload of each robot by partitioning the whole network using Voronoi diagram. After the initial deployment, the proposed algorithm computed the load in each partition and dynamically moved the robots to achieve the balanced workload distribution.

2.1.3 Vehicular Ad Hoc Networks

Vehicular ad hoc network (VANET) is a mobile ad hoc network that provides wireless inter-communication among nearby vehicles or the communication between vehicles and stationary roadside infrastructures. In a vehicular network, vehicles monitor and sense the surrounding parameters, such as event and road condition, and disseminate it to the neighbors, through either Vehicle-to-Vehicle (V2V) communication or Vehicle-to-Infrastructure (V2I) communication. The information is collected to the server through the roadside infrastructures. Fig. 2.2 introduces the architecture.

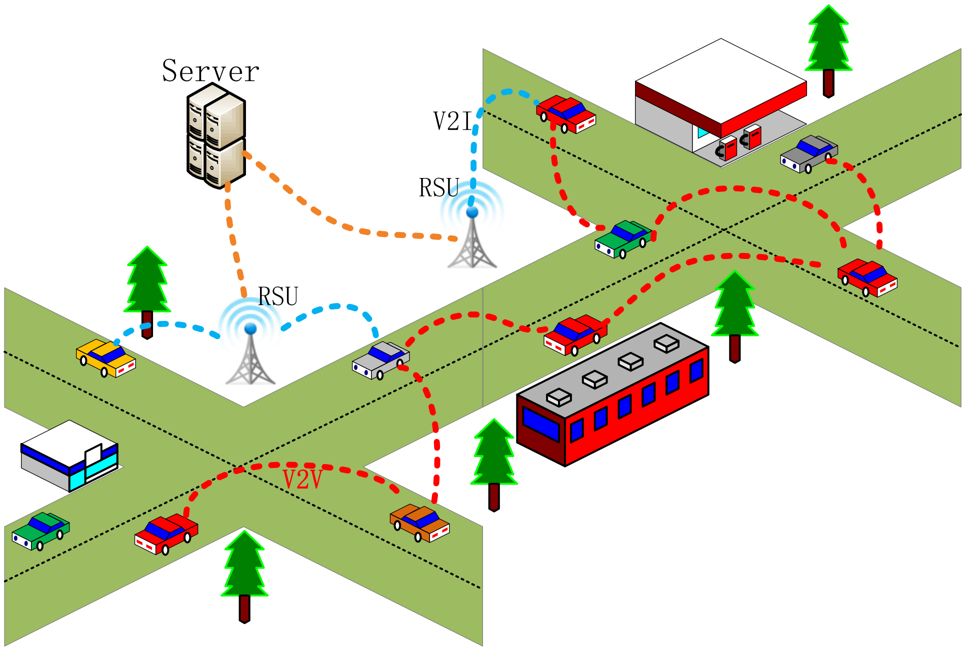


FIGURE 2.2: Vehicular ad hoc network

VANET is such a system that combines computing, communication, and storage capabilities with monitoring and control of the vehicles. [77] surveyed the technologies and research directions of vehicular networks. Compared with other ad hoc networks, intermittent connectivity, dynamic changes in network topology and low packets reception rate are the key characteristics of VANET [78]. The existing work is mainly on data dissemination [79–82], M2M communication architecture [83], security and privacy [84, 85], trust management [86, 87].

Booysen *et al.* [83] elaborated on a generic large scale M2M communication architecture for vehicular networks, where a five-layer M2M mechanism was introduced. The communication within the single vehicle could be achieved by wired Controller Area Network (CAN) and Media Oriented Systems Transport (MOST) ring. The V2V and V2I communications apply WiMAX or Cellular standards. The access points and gateways collect the data and transmit it to the server using wired Internet backbone.

Liu *et al.* [80, 81] proposed a receiver consensus algorithm for urgent and reliable warning message dissemination in vehicular networks. The algorithm exploits the geographical information of vehicles to autonomously choose the message forwarding strategies. Each forwarding candidate assigns the different priorities to the neighbored nodes by ranking itself and its neighbors by the distance to the centroid of neighbors in need of message. This scheme enables the best candidate to relay the packets without waiting.

Huang *et al.* [47] introduced the trust cascading failures in vehicular networks, where the decision of each vehicle depends on the decisions of neighbored vehicles. Therefore, an incorrect decision could cause a sequence of failure decisions. A novel voting scheme was proposed, in which each vehicle has different voting weight according to its distance from the event. The vehicle closer to the event has the higher voting weight.

2.2 A Brief Introduction to Complex Networks

To understand the structures of the real world networks, along with the aim to know how the human being activities spread on such networks, the graph theory was came out. Models and algorithms were extensively proposed in the past few decades. But the most social network, biological network and computer networks are not regularly purely random. Such networks have high clustering coefficients and short diameters [88, 89], which are not properly characterized by the original graph theories. Therefore, the complex network theory was indicated to reflect the real world essentials and help researchers understand them more profoundly. Fig. 2.3 is the face to face contacts in a primary school of a normal school day [90], with 236 students and teachers, and 5899 contacts. The nodes are colored for each classroom.

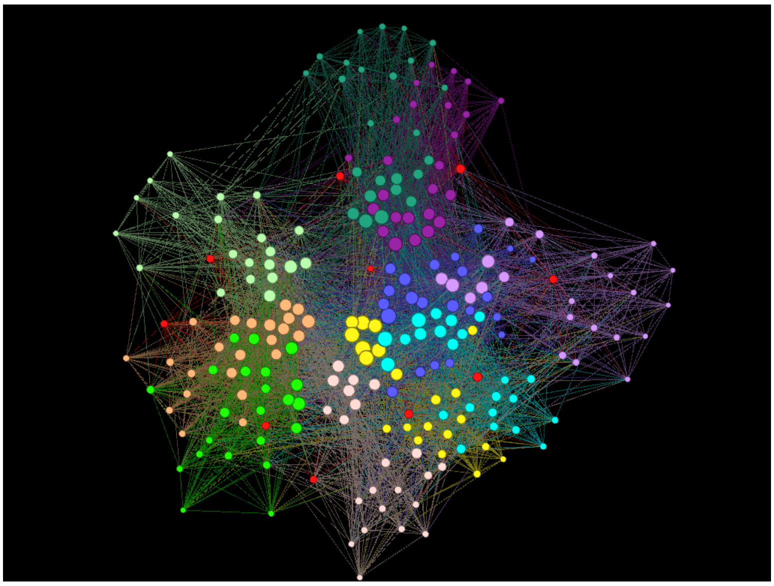


FIGURE 2.3: Contacts in a primary school

2.2.1 Network Metrics

To explore the networks, some measurements and metrics are widely used as follows:

- Degrees and degree distribution. The degree of a node is defined as the number of edges it connects with. In a large scale network, the degree distribution is one of the most fundamental metrics to study with. It is shown that power-law distribution is a common form in many real world networks [21], and thus is of particular interest to us.

- Clustering coefficient. To study the solidity of the network, or say the density of triangles, we apply the metric *clustering coefficient*. It measures the probability that two neighbors of a node are themselves neighbors. Experiments and empirical observations suggest that many real world networks tend to have a high clustering coefficient [88, 91].

- Path length. The average path length of the network is calculated by adding the shortest path length of all pairs of nodes. One of the most remarkable phenomena in real networks is *small-world effect*, where the average path between each paid of nodes is surprisingly small. This famous phenomena was found in Stanley Milgram's letter-passing experiment in 1967 [92], and called *Six Degree of Separation*. The experiment in [93] concludes the similar results that the average path length among Microsoft Messenger users is 6.6.

- Network diameter. The diameter of a network is defined as the longest length of the shortest path lengths of all pairs of vertices. It is shown that the diameter is usually scale logarithmically with the size of the network [21].

- Node centrality. Different nodes are of different importance contributing to the network. Node centrality is to evaluate the influence and importance of a vertex. The average distance of a node to all other vertices is provided by *Closeness centrality*. Another widely used metric is *betweenness centrality*, which measures the number of shortest paths go through that node.

- Component. In a network, normally, there is a large component fills most part of the network, along with a large number of small components that disconnect with each other. In most of the studies, the *giant component*, defined as the largest connected subnetwork that fills most of the entire network, is evaluated and analyzed. Fig. 2.4 sketches a case.

2.2.2 Three Network Models

Some models are widely used to depict the real world networks [88, 94, 95]:

- Erdös-Rényi model (ER). This model is used to generate a random graph, where the probability that a node has a link with another node is same for each node pair. It can be shown that the degree distribution of nodes follows Poisson distribution. The name of this model was for Paul Erdös and

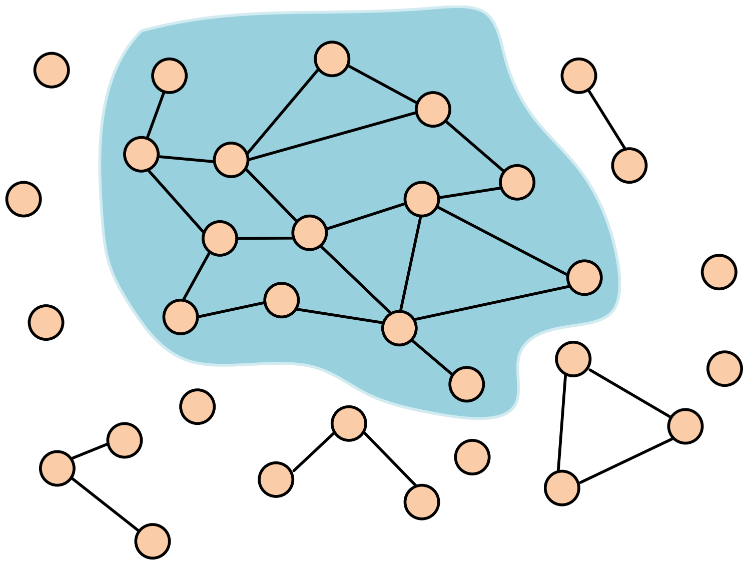


Figure 2.4: Components in a network

Alfréd Rényi [96, 97], where they used a probabilistic method to add the edge to each pair of nodes. The network generated by ER model normally has relatively low clustering coefficient and long network diameter. Therefore, this model can not properly describe many real world networks.

- Barabási-Albert model [98, 99] (BA). A common feature of random graphs is that the degree distribution of the network is inhomogeneous. The observation on large scale real world networks, including WWW, protein-protein interaction networks and interbank payment flow [100], is that they are scale-free and the degree distributions follow power law forms. In such networks, some nodes, which called 'hubs', have more connections than the average. The remaining nodes have few connections. Barabási and Albert proposed a preferential attachment algorithm to explain and reflect the growth of the real world network. The BA model suggests two steps to construct a scale-free network: a new node is adding into the network; it connects to the existing nodes with a probability proportional to the degrees the existing nodes already have. Numerical results demonstrate that BA model has a shorter diameter than random graph. There are some scale-free network model variations [21, 101].

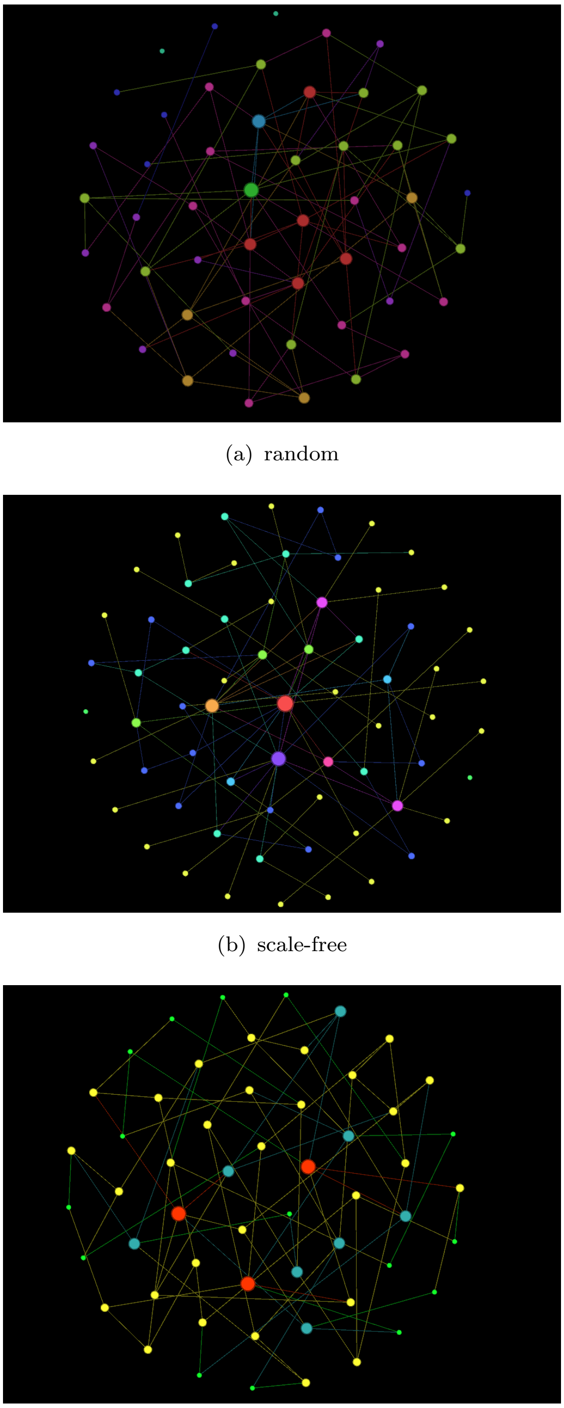
- Watts-Strogatz model (WS) [4]. In many real networks, a 'small-world' effect exists, popularly known as *six degree of separation* [102]. In other words, any two persons on the earth are connected by an average path length of six. Empirical results show that the small world networks have short average path length and high clustering coefficient. Watts and Strogatz proposed a small world model by interpolating a regular ring lattice to the random graph. The study on WS model metrics, including degree distribution, connectivity and clustering coefficient, was presented in [103].

Fig. 2.5 are the synthetic networks created using the above mentioned models in the generator Gephi [1]. The size of the node is proportional to its degree. It clearly shows that in scale-free network Fig. 2.5(b), a few nodes have more connections than the average. Fig. 2.6 illustrates the degree distribution forms of each network. The random graph of 2.6(a) has an Poisson distribution, a significant part of nodes is with the degree around the average. The power-law distribution is demonstrated in Fig. 2.6(b), most of nodes has 1 or 2 neighbors. A fat-tail effect exists. The small-world degree follows an exponential distribution with a cut-off on the left.

2.3 Interdependent Networks

Nowadays, the infrastructures such as water supply, power grid, transportation system, fuel stations are becoming increasingly interconnected. Studying the interactions and understanding how robustness is affected by the interdependency in such networks is one of the main challenges when designing reliable infrastructures.

Percolation theory [21] is an elegant method for estimating the network reliability. It studies the effect of random node removal on the network. It can be used to model a variety of real world network failures. One important characteristic



(a) random

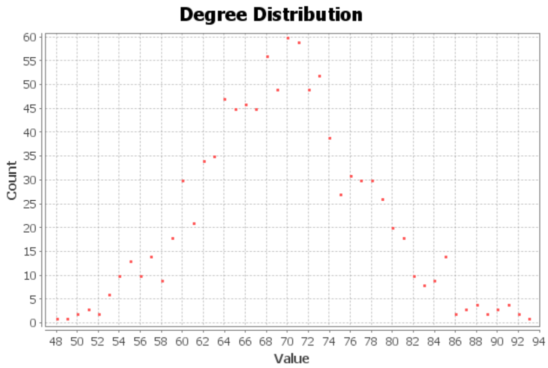
(b) scale-free

(c) small-world

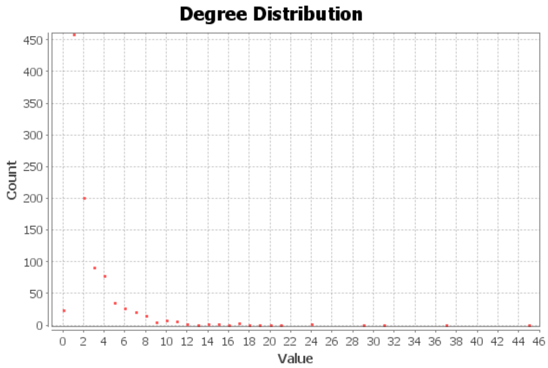
FIGURE 2.5: Using network generator Gephi [1] create synthetic networks

used to reflect the network connectivity is *giant component*, defined as the largest connected subnetwork that fills most of the entire network [21]. Percolation theory was used to measure the interdependent network characteristics in [15, 20].

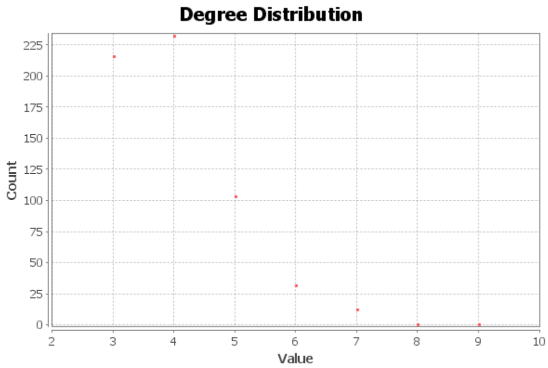
Buldyrev *et al.* [15] developed a framework for understanding the robustness of two



(a) random



(b) scale-free



(c) small-world

FIGURE 2.6: The degree distribution forms

interacting networks in terms of cascading failures. A 'one-to-one correspondence' model was proposed, where the two interdependent networks A and B have same number of nodes, each node in A functions depending on the unique node in B, and vice versa. The authors developed a percolation based method to study the robustness of this model by calculating the number of survived nodes in the final state. The cause of cascading is the random attack or system failure upon

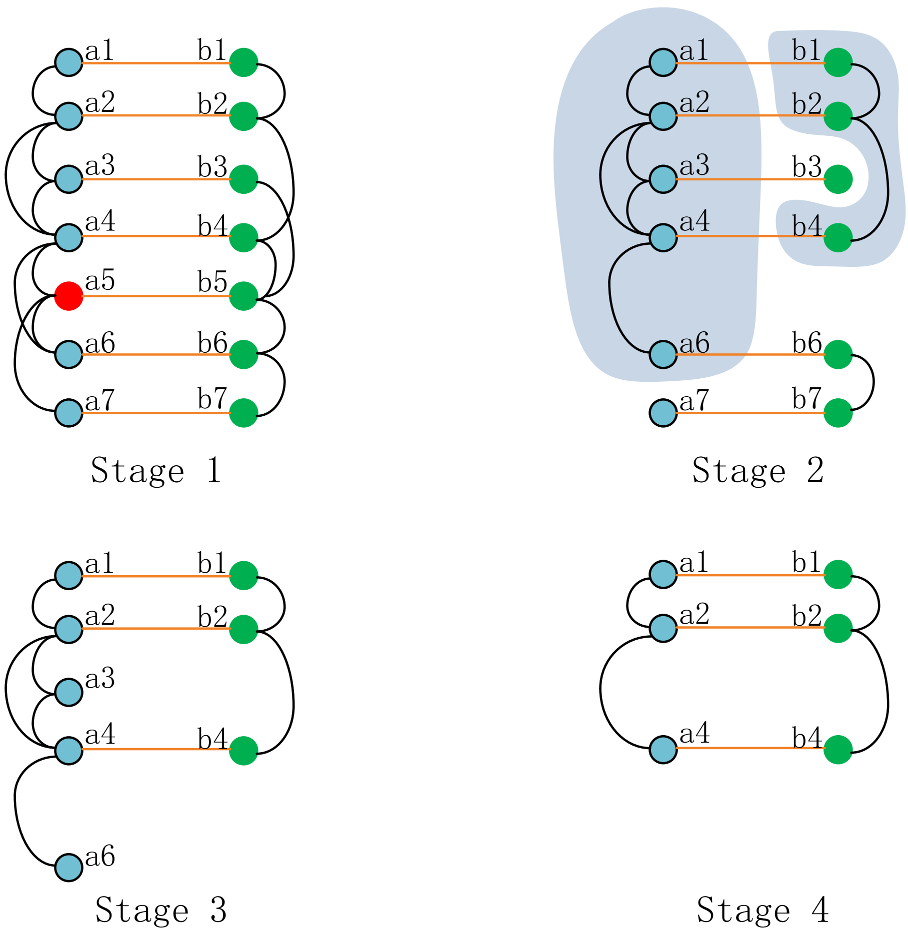


FIGURE 2.7: Network A and B are coupled using one to one model

one network. Their work estimated a critical value such that if the initial failure outnumbers it, the whole system breaks down. They discussed the coupled Erdös-Rényi networks. The number of intra-degree is randomly distributed.

Fig. 2.7 gives an example of the propose model, where the inter links are all *bidirectional* which means that every connected pair is mutually dependent. Initially, a random attack is on a_5, both its links and itself are removed, the node b_5 is also removed since it loses the dependent source. Similarly, the intra links of b_5 are wiped out. In Stage 2, b_1, b_2 and b_4 consist of the giant component of network B, the others are considered as faulty. As a result, in Stage 3, a_3, a_6 and a_7 lose the inter links, and thus fail. In Stage 4, the cascading failure stops, no more nodes or links are removed.

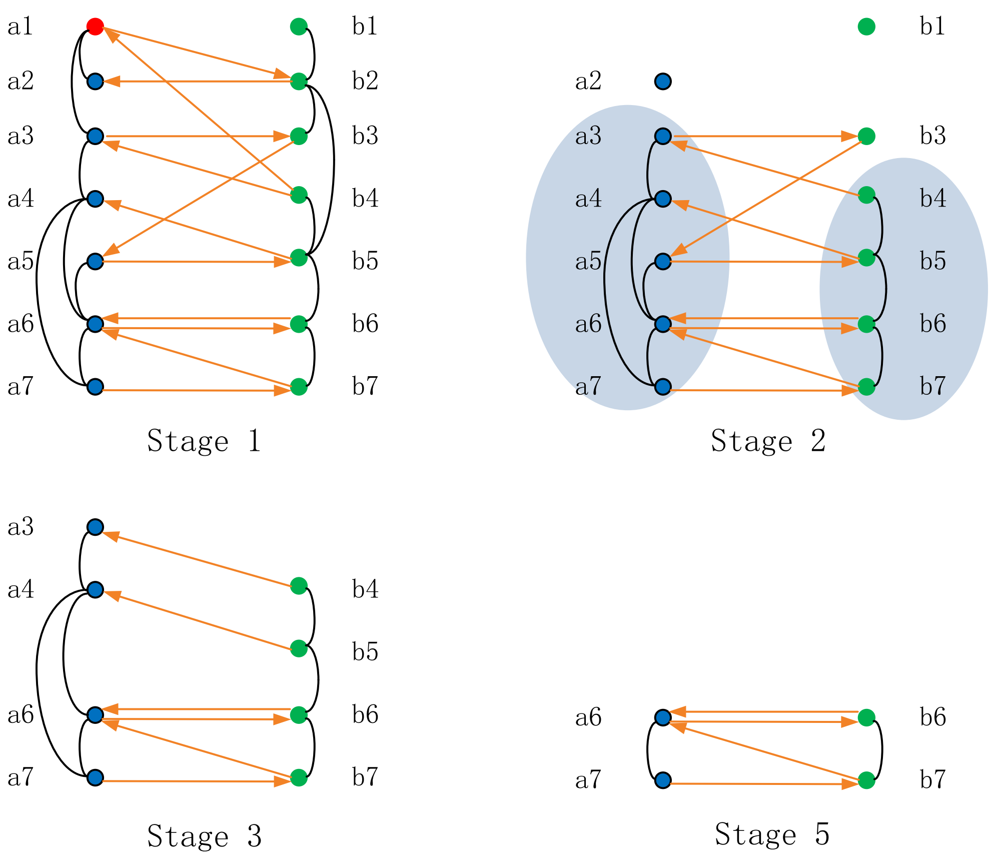


Figure 2.8: Unidirectional inter link with multiple-to-multiple correspondence

Buldyrev's work initiated the research on interdependent networks. [22] considered a new 'one-to-one' correspondence model, where the mutually dependent nodes have the same number of connectivity links, i.e., their degrees coincide. The model and assumption were not practical because we can not ensure that all dependent pairs of nodes have the same number of intra-degrees. However, the conclusion of [22] indicated a potential method to design the interdependent networks to improve the reliability: if each mutually dependent pair has the identical number of neighbors, for the same value of initial failure, the number of survivals is always greater than that in [15].

Shao *et al.* [23] pointed out that in real world mutually dependent networks, a single node in network A operates depending on more than one node in network B, and vice versa. The more realistic model should be 'multiple-to-multiple' correspondence, where each node functions as long as at least one of its supporting nodes is operating. They also assume that not all the pairs of nodes are mutually dependent. For instance, the on-line banking system is upon Internet, but Internet does not work depending on the banking system, so the interdependence is *unidirectional*. Fig. 2.8 is an example of proposed model. Each node has a random number of inter degrees, the mutually dependent counterpart is also chosen randomly. Initially, a_1 is attacked, all the related links are removed. In Stage 2, b_2 loses the support links from Network A and thus fail. a_2, b_1 and b_3 are disconnected from the giant components, hence fail. In Stage 3, a_5 is gone and b_5 will be considered as faulty due to the interdependency relation. Node b_4 is excluded from the giant component in Stage 4.

Parshani *et al.* [104] described two types of inter links: *dependency link* and *connectivity link*. The *dependency link* makes failure in one network cause failure in the other network, while *connectivity link* enables the nodes work cooperatively. High density of dependency link makes networks more vulnerable.

In the previous work, the links between two networks are randomly distributed ('random allocation'). The work of [15, 23] is extended in [20] by a so-called 'regular allocation', where every node in the system is assigned the same number of inter links. The regular allocation scheme is proven to be optimal when the topology of each individual network is unknown. The conclusions suggests the regular allocation of bidirectional inter links achieves higher system performance, rather than the random unidirectional allocation scheme. The example is demonstrated in Fig. 2.9.

Similarly, Dong *et al.* [26] studied the system of n interdependent networks with partial interdependence relationship. The random failure is applied to the above work, while in [27], the authors focused on the targeted attack on a network of networks(NON). They investigated the system robustness of both fully interdependent and partially interdependent networks. Nguyen *et al.* [17] proposed an

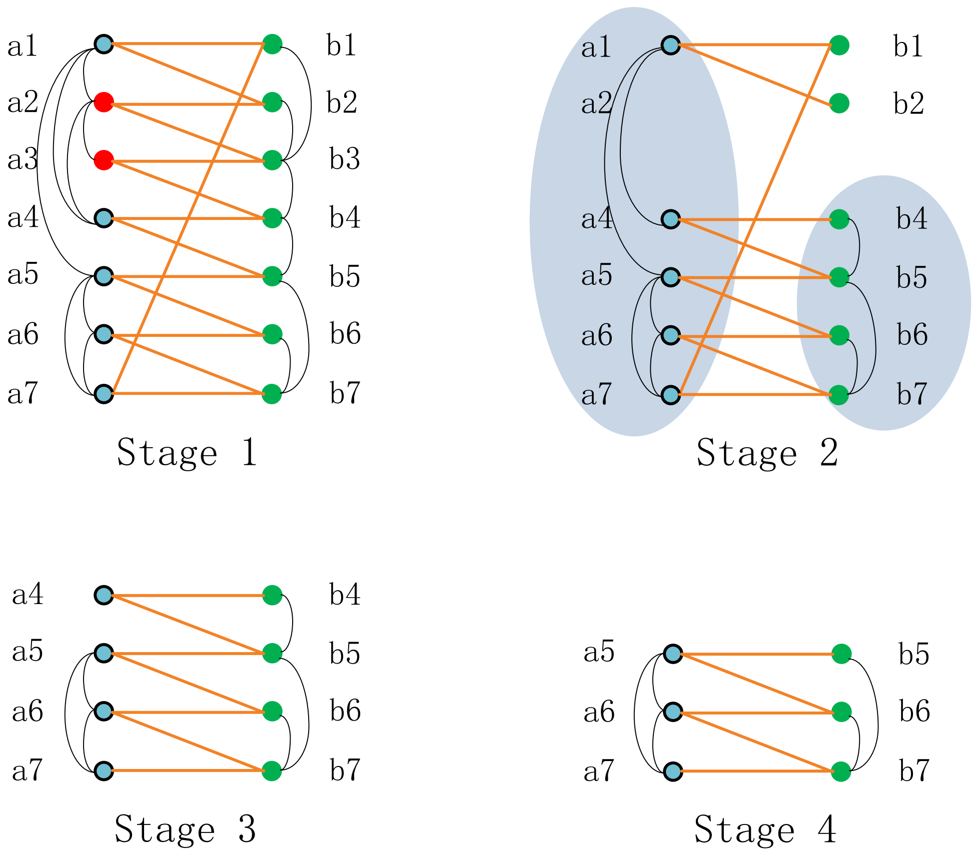


FIGURE 2.9: Regular allocation with bidirectional inter link

algorithm to identify the critical nodes whose removals could lead to the fatal destroy of the networks.

Zhou *et al.* [105] studied the percolation of partially interdependent scale-free networks under the 'one-to-one' interdependence [15]. They assumed that a fraction q of nodes in both networks are interdependent. The results showed that there are three types of behaviors with different values of q. For a large value, the *first-order discontinuous* transition appears. The *second-order continuous* transition occurs for a small q. A mixed first-second-order is for the intermediate value of q.

A targeted attack was discussed in [25], where a mapping algorithm was indicated to transfer target attacks to random failures. The authors pointed out that protecting the high degree nodes can improve system robustness significantly. Gao et al. [106] studied the interacting networks and presented an percolation law for a network of several interdependent networks.

BakTang-Wiesenfeld sandpile model was applied to study how interdependence affects cascading behaviors [107]. They found that the inter connectivity provides benefits which can balance the dangers at stable point. While, too much inter connectivity impacts the system. Huang et al. [108] developed a method for evaluating how clustering within the networks affects the system robustness. [109] proposed an epidemic spreading based theory, which is more straightforward than previous work, to formulate the interdependent networks. Ruj *et al.* [110] randomly assigned the inter links and studied the targeted attack in smart grids. High degree nodes are compromised with higher probability. They found that an adversary can disrupt significant part of network by launching a targeted attack.

A completed and detailed survey on interdependent networks could be found [19, 111]. The authors elaborate this new area on its every aspect.

2.4 Smart Grid

The demand for power electricity recently increases dramatically. Developing smart grid has become an urgent global priority, as the need to cope with future electricity demand and to provide a more efficient way of delivering such electricity has become more pronounced. Smart grid within a system should intelligently integrate and control the different elements of the grid so as to realize the full potential of the system.

Plenty of work has been done on the following aspects: electricity delivery, demand managing, real-time communication, intelligent control, pricing and system robustness. In this literature review, we are going to cover the work on power grid load propagation and power grid topology.

The *Task Force on Understanding, Prediction, Mitigation and Restoration of Cascading Failures* provided the completed background on dealing with cascading failure in power grid in [112, 113], including techniques, modelling and simulation approaches. They defined the cascading failure, discussed the challenges of cascading failure, and summarized a variety of state-of-the-art techniques that has been used in simulation and analysis. The authors also revealed the limitations of existing techniques and approaches, and indicated some future research directions.

2.4.1 Power Grid Topology

To study and predict the cascading failures, it is fundamental to have a large number of power grid information, such like the topology, size and capacity of each node. However, the real information of power system is difficult to obtain due to the privacy and security reasons [114]. Therefore, complex network theory is widely used to simulate and study the structure, properties and cascading failures of power grid. There are some work [114–118] on revealing the topologies of the power grids all around the world. Unfortunately, it is very hard to define a general model for all the power grids. The different power grids in different nations and regions yield different topologies [119].

A few papers reported the power grid is a kind of small-world network. Watts and Strogatz [4] measured some metrics of power grids, such as path length and clustering coefficient. They found that the small world characteristics appear in such systems. Similarly, Amaral *et al.* [120] concluded that the Southern California power grid shows an exponential degree distribution. Albert *et al.* [117] also analyzed the North America power grid and reported an exponential cumulative degree distribution. Rosato *et al.* [118] studied three European power transmission systems in Italy, France and Spain. It was found that despite the

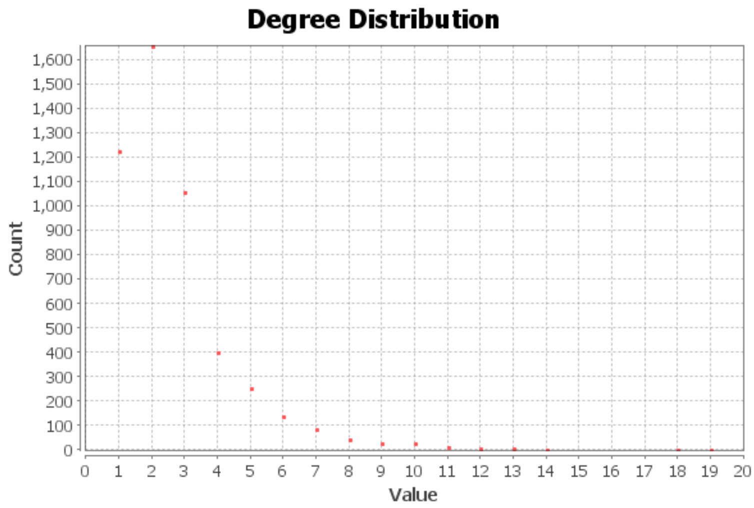


FIGURE 2.10: Western Power Grid Degree Distribution. 4941 nodes with 6594 edges [2]

locations, they all have high clustering coefficients, a large path length and display some characteristics of small world network.

While, Crucitti *et al.* [121] reported the Italic power grid with a power-law degree distribution. The interesting thing is in [122], Chassin *et al.* studied the same power grid with [117], but they concluded with the different results that the degree distribution is power-law. We use Gephi to calculate the degree distribution of the network that has been studied in [117, 122], as shown in Fig. 2.10. A detailed survey on topologies of power grid could be found in [41].

The mentioned papers were focusing on the physical topology of power grid, and neglected the electricity features. Hines *et al.* [119, 123] studied the structures of IEEE 300 bus and the Eastern United States power grid, and compared them with random graph, scale-free and small world network models. They demonstrated that the real power grid differs with these models in degree distribution, clustering coefficient and network diameter. Therefore, they proposed a novel algorithm to create the synthetic network with featuring the power grid properties. The analysis showed that the degree distributions of their sampled power grids, including Eastern and Texas Interconnects, follow exponential rather than a power-law. Further, the clustering coefficient is not as high as the small-world network. The conclusion was that the physical topologies of their samples are neither small-world or scale-free.

Wang *et al.* [48, 114] also aimed to generate a random topology model for power grid, with including the topology features and electricity features. Firstly, the authors studied both the topological and electrical characteristics of power grids based on a number of synthetic and real-world power systems. They founded that the power grid is sparsely connected with small-world properties. The degree distribution could be fitted by a mixture distribution coming from the sum of a truncated geometric random variable and an irregular discrete random variable. They also studied the line impedance distribution and the graph spectral density. Based on the discoveries, they indicated a random topology algorithm for power grid to feature the same characteristics derived from the real data.

2.4.2 Load Propagation Models

The cascading failures are caused due to the load propagation. Therefore, it is necessary for us to understand the load propagation mechanism and some widely used models that have been used to simplify the research.

2.4.2.1 Initial Load Allocation

There are mainly four ways to define the initial work load assigned to each node in power grid.

- For a node i, the initial load L_i is assigned uniformly between the predefined minimum value L_{min} and maximum value L_{max} [34]. Both L_{min} and L_{max} are set up with the same values for all the nodes. Therefore, the L_i follows a uniform distribution $U(L_{min},L_{max})$.

- The load L_i on node i be proportional to the number of shortest paths passing through it [39, 124]. If there is more than one shortest path between two nodes, the load is evenly divided on the critical nodes. Obviously, this scheme requires to calculate the betweenness of each node at every cascading step, which leads to a burden on computation and analysis.

- The load L_i is a function of its degree z, given by

$$L_i = \alpha \cdot z^\eta, \tag{2.1}$$

  where α and η are tunable parameters [35, 125, 126]. The advantage of this approach is that once given the degree distribution of the network could we start the analysis by using the mean field.

- L_i is a function of its degree z and the summation of the degrees of its adjacent neighbors [127–129], given by

$$L_i = (z \cdot \sum_{m \in \Gamma_i} z_m)^\eta, \tag{2.2}$$

  where Γ_i is the set of neighbors of node i, z_m is the degree of node m.

In a real-world power grid, the capacity of each station is limited due to the cost. It is reasonable for us to assume that the capacity of node is proportional to its initial load [35, 39, 125, 128]

$$C_i = T \cdot L_i, T \geq 1 \tag{2.3}$$

where constant T is the *Tolerance parameter* that controls the tolerance of the power grid.

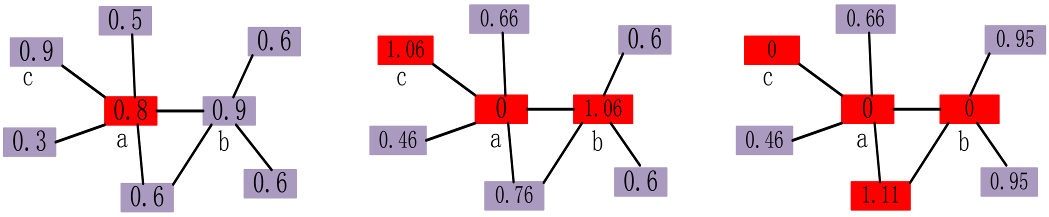


FIGURE 2.11: Load propagation

2.4.2.2 Load Redistribution

A common characteristic of blackouts is an initial power station breakdown leads to the load propagation, which could cause the overload of neighbored stations. The load propagation continues once the stations are faulty due to the overload. Therefore, the cascading failure occurs. Some widely used load redistribution models are summarized as follows:

- For each failed station, the load of adjacent neighbors increases by an pre-defined amount [34]. This scheme is simple to be applied and analyzed. However, it can not be able to reflect the true load redistribution.

- The load of failed station is evenly divided and redistributed to the neighbors [34]. Therefore, the neighbor station j would take a piece more load τ_{ij}, given by

$$\tau_{ij} = \frac{L_i}{z_i}, \tag{2.4}$$

  where z_i is the degree of i. Fig. 2.11 shows a simple example of load propagation. We assign each power station an initial workload, and assume they have the same capability of 1. The failure starts from station a, where the workload is transmitted to the neighbors equally. As a result, stations b and c run out of the limits, and thus become to being overloaded. Similarly, the workload of b, c is equally shifted to the neighbors, and this process causes more failures. The load redistribution scheme in this example is to evenly

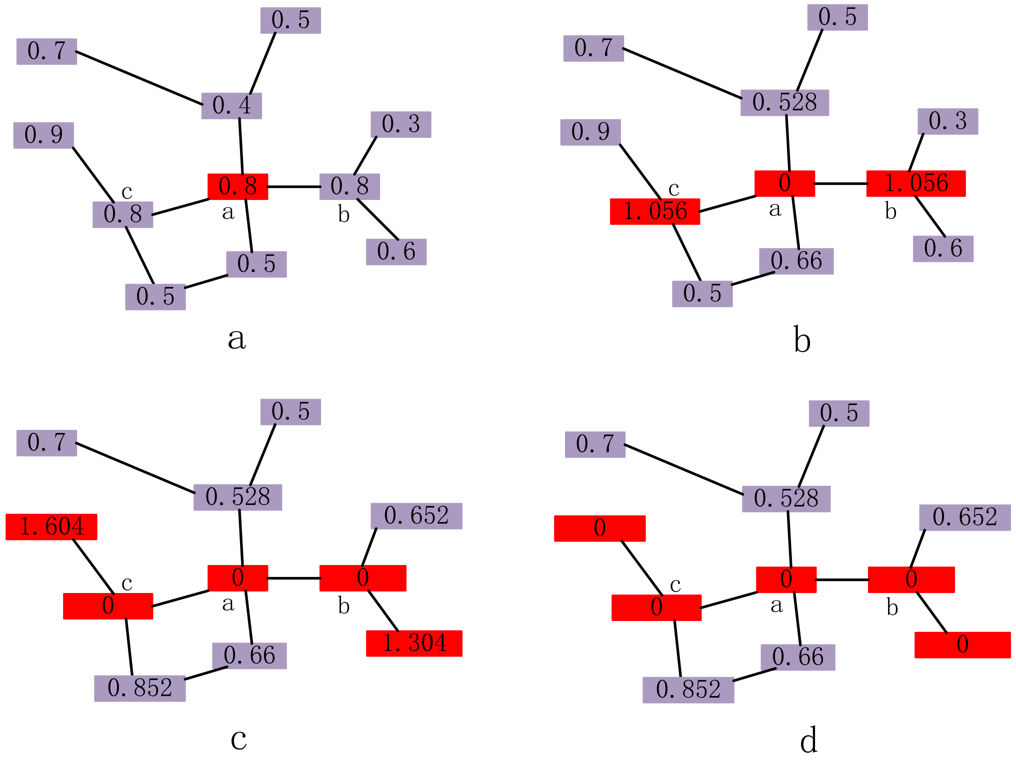


FIGURE 2.12: Load propagation with preferential probability scheme

divide the load to the neighbors. If there is no operating component in the neighborhood, e.g., station c, the failure spread stops in that direction.

- The load L_i of node i is redistributed to its neighbors according to the preferential probability [126]. The load shifted to the neighbor station j is proportional to its own initial load L_j [129]. Therefore, when node i is down, the redistributed load τ_{ij} to node j is given by

$$\tau_{ij} = L_i \cdot \frac{L_j}{\sum_{m \in \Gamma_i} L_m}. \tag{2.5}$$

Fig. 2.12 sketches an example. The load of node a shifts to the neighbor nodes based on Eq. (2.5), and so forth.

2.4.3 State of the art on Cascading Failure in Power Grid

Many papers are dedicated to understand the electricity load propagation to prevent the cascading failure. A detailed introduction of power grid cascading failure was provided in [33], where the authors also described two schemes to reduce the blackout size and the cost: *survivability* and *reciprocal altruism*. The former one suggests to build a more redundant electricity supply system for the vital services, such as traffic lights and hospitals. Reciprocal altruism is common in biological systems. The power grid could be designed as a decentralized, reciprocally altruistic system, where each power station shares its goals and information with neighbors. By negotiating, each station considers not only the local goals, but also the goals of the neighbors. The simulation illustrated that the cascading failure size is reduced by using reciprocal altruism.

There are two main lines in this area: *pure topological approach* and *hybrid approach*. The pure topological approach [10, 35, 126, 127, 130–132] focuses on studying the static failure in the network, primarily on how the connectivity changes. These work normally use graph theory, stochastic method and percolation theory to assess the power grid vulnerability. The analysis based on hybrid approach [36, 133–138] is combining the complex network theory with the features of the electricity, such like the transmission line flow limit and voltage sag. Compared with pure topological approach, hybrid approach can provide fundamental insights on power systems. While, on the other hand, the computation burden is heavy due to the large scale and complexity. A detailed power flow model is also computational expensive.

2.4.3.1 Pure Topological Approach

Zhao *et al.* [35] studied the cascading failure trigged by the attack on a single node on a scale-free network. By both theoretically and experimentally, they found a phase transition phenomenon in terms of the ratio of node initial load to the capacity (tolerance parameter). They proved that the breakdown due to cascading is only possible when the ratio is below the transition threshold. The simulations determined this threshold value. For small capacity, the attack on the node with high degree can trigger a cascade failure through the entire network. In [125], the authors tried to address the problem that how large of the capacity could make the network immune to the cascading breakdown in the scale-free network.

Wang *et al.* [127] discussed the cascade-based attacks on the US Western power grid. They used extensive simulations to study two attack schemes: attack on the nodes with highest loads and the lowest loads, respectively. For each attack, 50 nodes are initially chosen as the targets. The interesting finding was that the attack on the nodes with the lowest loads is a more effective way to destroy the whole system in some cases.

Kadloor and Santhi [37] modeled power grid as a graph and studied the system robustness. By extensive mathematical analysis, they estimated the disturbance levels the system can tolerate before a few overloaded nodes trigger a large blackout. Kinney et al. [38] used the real network structure of the North American power grid and modeled it as a weighted graph. Combining dynamical approach of the Crucitti-Latora-Marchiori model [10] and complex networks, they studied two types of node overload progression, and showed that the disruption of 40% transmission substations leads to the cascading failure, and a single node failure can cause up to 25% loss of transmission efficiency.

Fu *et al.* [131] analyzed the regional power grid in China by modelling it as a small-world network. They studied both targeted attack mode and random attack mode, and concluded that multiple failure occurring on the nodes with low degree can easily lead to the regional blackout. This conclusion is similar with the results from [127].

While, Wang *et al.* [126] held a controversial opinion that they believe if the removed node has a relatively small load, it would not cause major changes on the balance of load, and consequently would not lead to the cascading failure. Based on this assumption, the authors proposed a cascading model where the initial load of each node is assigned based on Eq. (2.1). The load redistribution to the neighbors follows the preferential probability scheme, where the accommodated load of each neighbor is proportional to its initial load. They studied the single node failure and developed a metric, called *avalanche size*, to calculate the total number of broken nodes. The simulation results showed that when $\eta = 1$, the single attack has no impact on the network; while for $\eta < 1$, the attack on node with low degree significantly affects the network. [126, 127, 131] indicated a point that we should take care of the nodes with low degree and load.

Pahwa *et al.* [30] found that as the load-carrying nodes getting disconnected, the total load on the network keeps reducing. In other words, a part of shifted load can not be afforded by the neighbors, thus has to be eliminated from the network. Node c in Fig. 2.11 is a good example. The authors indicated three load mitigation strategies to reduce the size of cascading failure. Homogeneous load reduction aims to reduce a certain percentage of load on each node after the initial failure occurs. Targeted range-based load reduction is to select a node connecting to the link that failed and create a tree, on which the nodes reduce the load. Islanding is to optimally select a cluster and separate it with the main grid to prevent the cascading failure.

Wang *et al.* [129] studied the power grid robustness from the attacker perspective. It suggested that it is not a good strategy to choose the node with the largest load for the attackers. The authors proposed three metrics to efficiently choose the critical victim to destroy the network as much as possible. Simulation results showed that a high value of the *Risk if Failure*, defined as the ratio between the initial load of the node and the total initial load of its neighbors, is extremely efficient to damage the power grid.

Bao *et al.* [124] proposed the concept of *load entropy* to study the load redistribution, as well as developed a new cascading model, where they divided the cascading failure into steps varying with time. The aim was to understand the dynamic load change. The simulation results indicated that the load entropy for a large scale cascading failure increases sharply than the small scale. Therefore, we can distinguish the size of the cascading failure in the very beginning.

The case of static overload failure was discussed in [139]. Optimization technique was used and a distance-to-failure algorithm was proposed to predict the weak points in power grid. They applied their algorithm to two real power grid examples and concluded that the failures due to overload are sufficiently sparse if the normal operational stations are healthy. Load Redistribution (LR) attack was developed and studied in [32] by analyzing their damage to power grid operation. It proposed an attack model describing the main goal of LR attack and then based on that, indicated the theory and criterion of protecting the system from LR attack.

Pfitzner et al. [29] proposed a model to focus on the analysis of tripping of already overloaded lines. By simulating on a real-world power grid structure, they showed that such controlled tripping leads to significant mitigation of cascading failure.

2.4.3.2 Hybrid Approach

The power grid is a special complex network along with many physical constraints, such as power flow and power line limit. Compared with the pure topological approach mentioned above, the hybrid approach takes the electricity features into account. The biggest advantage that hybrid models provide is the fundamental insights of cascading failure behavior. While on the other hand, for a large scale power grid, this approach is computationally expensive. An accurate power system fault model is still difficult to be found [136].

In [36], the authors pointed out that the pure topological methods disregard the electricity features. They proposed an extended topological model with taking into account the specific features of power systems, such as electrical distance and line flow limits. They also developed two metrics *extended betweenness* and *net-ability* to rank the criticality of network components. The simulations were used to demonstrate that the components found by their method are indeed critical. Arianos *et al.* [138] used *net-ability* to estimate the impact of the line outages in order to find the critical power transmission lines. They compared the efficiency on identifying the critical lines of three different methods, and suggested the one they called *computation of the line overloads by dc flow* as the reference method, as it takes into account the features and information of power grid. The *net-ability* approach, with less computation burden, however, could find part of the lines.

A new approach of vulnerability assessment for power grid was proposed in [39], where the authors combined a DC flow model with a traditional attack methodology. Random attack, degree-based attack and betweenness-based attack were studied.

Chen *et al.* [135] proposed a modelling algorithm for power system cascading failure, by utilizing sampling technique, to relieve the burden on simulation computation without affecting the modelling accuracy. A sequential Monte Carlo approach was applied to simulate the AC power flow, DC power flow and voltage sag. Then, they utilized stochastic methods *importance sampling* and *importance sampling and antithetic variates combined* to decrease the Monte Carlo trials. IEEE Reliability Test System was used to assess the performance of proposed methods. The simulation results found both dramatically reduce the computation burden.

Wang *et al.* [133] developed a Markov transition model for power grid, incorporating the line flows and the flow distribution. The proposed model could indicate which part in the system is under the stress and therefore likely to break down. Eppstein *et al.* [140] proposed a 'random chemistry' search scheme to identify the collection of initial failures that lead to the cascading failure. Based on the extended model of [36], Zhu *et al.* [136] indicated a new metric, called 'risk graph', was constructed based on the historical data, which records the top strongest nodes/links that lead to the cascading failure. Then the authors developed a new attack strategy. The simulation results showed their strategy is much more efficient to damage the network than the traditional load-based or degree-based attacks. However, this algorithm requires the network information to construct the graph, which is not practical in the real world.

Kim *et al.* [141] developed two branching-process-based methods to study the load shed in the cascading failure. The aim of the paper is to accurately predict the distribution of load shed by reducing the number of simulations, i.e., reduce the computation burden. The stochastic method is used. This work could help us estimate the load shed in the cascades by given a little knowledge that has been already observed.

Moslehi and Kumar [142] critically reviewed the reliability impacts of major components such as energy generator, demand response, communications and electricity transportation in smart power grid. They presented a grid-wide IT architectural framework to improve the robustness of system, and discussed the technical feasibility. A power generation and distribution architecture has been discussed in [28], arguing that a distributed generation enhances the robustness of the system. Using optimization techniques and simulations, the authors showed that fault tolerance increases with the number of generators.

Chapter 3

One to Multiple Interdependence Model

3.1 Model Description

We consider the cyber physical system consisting of two interdependent networks, i.e., the *physical-resource* and *computational-resource* networks, denoted by G_p and G_c, respectively. We refer to the edges that connect the nodes within the individual networks as *intra link* and those that connect the nodes from different networks as *inter link*. *Bidirectional* link between nodes u and v implies u and v depend on each other for proper operation. *Unidirectional* inter link from node u to node v means v function only if u is operating, but u does not depend on v.

The so-called physical and computational/cyber resources in CPS consist of a large number of components, e.g., embedded system, battery, control devices. Parts of these components make up the essential backbone of the CPS, while the others could be considered as terminals. In most situations the influence of these terminals to the system reliability is limited. For instance, in smart power grid, the breakdown of a smart meter in one house might not cause the failure of others, or a failure of a single computer or router might not impact the information exchange. Thus, we only consider the components which have important roles as the nodes in G_p and G_c. For instance, the computational stations in computational-resource networks, the data gateways in machine-to-machine controls play the key role. In smart grid, power plants, substations, transformers, new energy generators and operation centers are essential parts.

3.1.1 Intra Link Distribution

Complex networks have been discussed extensively. A significant part of networks in real life are *Erdös-Rényi* networks, *scale-free* networks and *small-world* networks [21]. While some networks may have more complex structures. To simplify our research, we consider both G_p and G_c are scale-free networks, within which the node degree distribution follows a power law. This implies $\mathbf{P}(z) \propto z^{-\gamma}$, where $\mathbf{P}(z)$ is the probability that a node has z links, γ is the *power law exponent*.

3.1.2 Inter Links

We now introduce the interdependence of our system model. One important type of interdependent networks is 'one-to-multiple' correspondence: each node in G_c has only one support link from G_p, while each node in G_p is connected to multiple nodes G_c. This 'one-to-multiple' structure could be applied on several cyber physical system, such as smart power grid, wireless sensing and computational-resource network to measure their robustness.

We use smart power grid to illustrate the importance and practice of 'one-to-multiple' structure. The power station is monitored and controlled by its operation\control center which consists of a large number of computers, sensors and routers. The energy required by each operation center to function is supported from the power station, the power grid and the control network are 'one-to-multiple' dependent. Each power station and substation has one operation center, each power station might provide energy to more than one operation centers.

Let $|\mathcal{N}|$ denote the size of $\mathcal{N}$ (the number of nodes in $\mathcal{N}$). To allocate the inter links, we consider each node in G_p is a bin and the nodes of G_c are balls. The allocation follows the well-known *Balls and Bins* problem, where $|\mathcal{C}|$ balls have to be independently and uniformly put into $|\mathcal{P}|$ bins, the probability that one ball is assigned to i-th bin is $\frac{1}{|\mathcal{P}|}$. For any $1 \leq i \leq |\mathcal{P}|$, denote the number of balls in the i-th bin by $l_{P,i}$. Then,

$$\Pr(l_{P,i} = k) = \binom{|\mathcal{C}|}{k} \cdot (\frac{1}{|\mathcal{P}|})^k \cdot (1 - \frac{1}{|\mathcal{P}|})^{|\mathcal{C}|-k}. \tag{3.1}$$

That is, the number of nodes in G_c supported by each node in G_p follows a binomial distribution $\mathbf{B}(|\mathcal{C}|, \frac{1}{|\mathcal{P}|})$. In our system model, the links from G_p to G_c are assumed to be unidirectional.

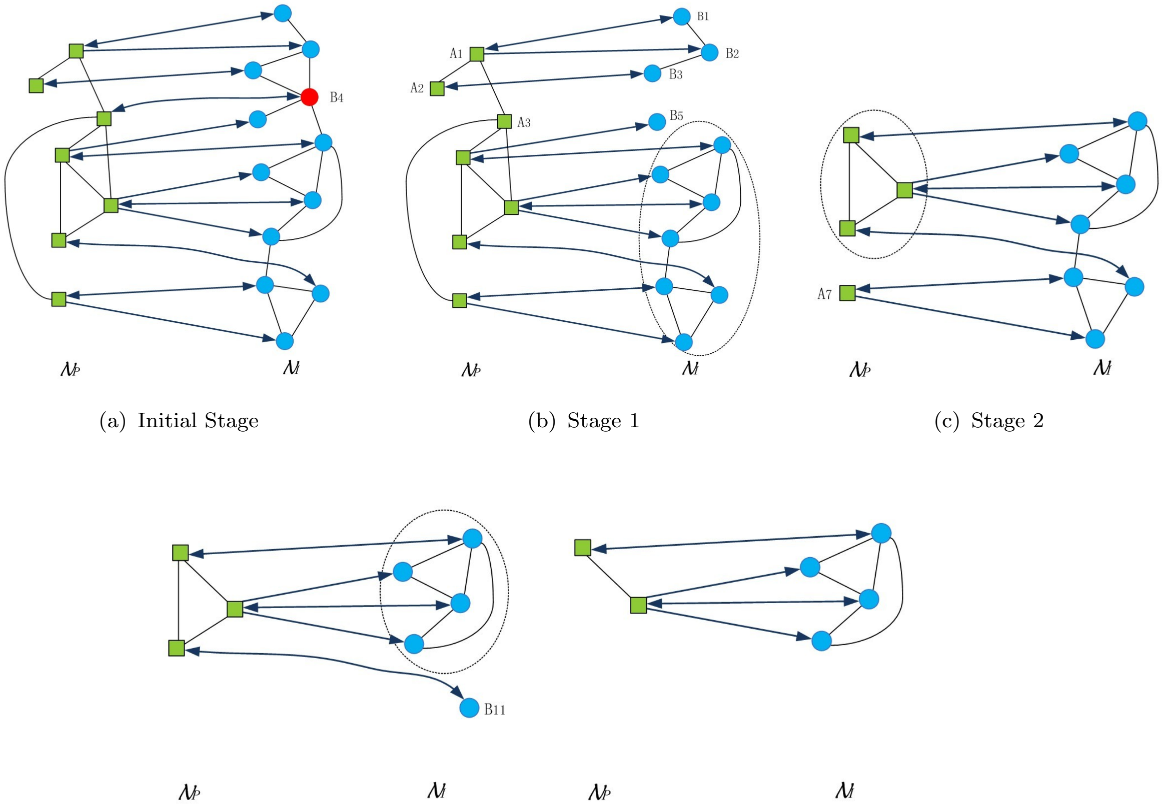
We assign the unidirectional inter link from G_c to G_p: We allocate one node in G_c for each node in G_p. For the i-th node in G_p, we randomly choose one inter link from its k links (value k is decided as per discussion above) as one bidirectional link. This means that the G_p node supports the G_c node and this G_c node controls the i-th G_p node. For any $1 \leq i \leq |\mathcal{P}|$ and $1 \leq j \leq |\mathcal{C}|$, define an event $\mathcal{E}_{i,j}$: the j-th G_c node controls the i-th G_p node; and define an event $\mathcal{E}_j$: the j-th AS is chosen as the operation center. Since $\mathcal{E}_{i,j} \cap \mathcal{E}_{h,j} = \varnothing$ when $i \neq h$, it holds that

$$\Pr(\mathcal{E}_j) = \sum_{i=1}^{|\mathcal{P}|} \Pr(\mathcal{E}_{i,j}) = \sum_{i=1}^{|\mathcal{P}|} \sum_{k=1}^{|\mathcal{C}|} \Pr(l_{P,i} = k) \cdot \frac{k}{|\mathcal{C}|} \cdot \frac{1}{k},$$

then,

$$\Pr(\mathcal{E}_j) = \frac{|\mathcal{P}|}{|\mathcal{C}|} \cdot \sum_{k=1}^{|\mathcal{C}|} \Pr(l_{P,i} = k) = \frac{|\mathcal{P}|}{|\mathcal{C}|}. \tag{3.2}$$

Since each node in G_c has only one inter link, the inter degree distribution in G_c follows a Bernoulli distribution with mean of $p = \frac{|\mathcal{P}|}{|\mathcal{C}|}$.



(a) Initial Stage (b) Stage 1 (c) Stage 2

(d) Stage 3 (e) Stage 4

FIGURE 3.1: A sketch of cascading failures. Initially, G_p has 7 nodes while G_c has 12 nodes. The random attack upon network G_c causes the failure of node B_4. In stage 1, B_4 is removed along with all the links. As a result, nodes B_1, B_2, B_3 and B_5 are disconnected from the giant component (subnetwork within the circle), thus fails. Consequently, nodes A_1, A_2 and A_3 are going to be faulty since they lose the supporting inter links from B_1, B_3 and B_4. In Stage 2, their related intra links and inter links are removed. As a result, 4 nodes left in G_p, while node A_7 is excluded from the giant component. Therefore, in Stage 3, A_7 and all its inter links are removed, network G_c fragments into components, while node B_{11} is disconnected from the giant component, thus fails. The corresponding node in G_p also fails. In the final stage, the remaining nodes of both networks are reachable to each other and mutually dependent, the cascading failure stops.

3.1.3 Assumption and Example

We are interested in the system robustness to a random failure or attack. In the dynamics of cascading failure, we estimate the fraction of nodes that can survive at the end (the number of nodes that remain functioning). In previous works [15, 20], it was assumed that the node has to belong into the giant component to operate. This assumption is reasonable and simplifies the mathematical analysis for two reasons: 1) it only considers the nodes within the largest connected cluster; 2) giant component has been widely studied and used in network theory. Therefore, in this work, we stay on the assumptions from previous work that a node can operate only if the following conditions are satisfied:

1. It has at least one inter link with a node that functions.
2. It belongs to the *giant component* of its own network.

The cascading failure is triggered by the failure of $1-\phi$ fraction of nodes in either G_p and G_c.

We begin by randomly removing $(1-\phi)\cdot|\mathcal{C}|$ of nodes in G_c. As a result, both intra and inter links of these deleted nodes are also removed. Owing to the interdependence, in the next stage, some nodes in G_p might disfunction since they lose their inter links. As nodes and links are removed, G_p begins to fragment into components. Due to condition 2, the nodes operate properly only if they belong to the giant component, thus more nodes and links are removed. The fragmentation in G_p might lead to further failures in G_c, because now some nodes in G_c have no inter links coming from G_p. This cascading failure continues recursively between both networks, and reaches one of the following two final states: *1)* all nodes are removed and the entire system collapses; *2)* the giant component is mutually inter connected between G_p and G_c. Figure 4.2 gives an example of failure propagation on our model. In the initial stage, we have 7 nodes and 12 nodes operating in G_p

and G_c. Then, one node (B_4) is attacked in G_c and triggers the cascading failure. When the cascade stops, only four and two nodes remain functioning in G_c and G_p respectively.

3.2 Mathematical Analysis of Cascading Failures

In this section, we give a mathematical analysis of cascading failures in our system model. As the failure could occur in either G_p or G_c, so we arbitrarily choose G_c as the start point.

Assuming that a fraction of $1-\phi$ nodes in G_c are attacked initially, we use percolation theory to model the cascading failures and calculate the fraction of surviving nodes. Then a set of transcendental functions are derived, whose roots represent the remaining fraction of nodes in both networks.

Recall percolation is a process of removing vertices and edges in network. To simplify our analysis, we divided the entire percolation into stages. For each stage, a portion of nodes and associated edges are removed if they do not satisfied the two conditions. In this study, to be specific, the failure begins from G_c, then goes to G_p. Then the failure returns back to G_c, and so on.

Besides the notations in Table 3.1, we introduce a notation $F_{\mathcal{N}}(\phi)$ from [19, 20] to represent the expected fraction of giant component in the subnetwork which occupies the fraction ϕ of the nodes in the entire network $\mathcal{N}$. In our work, we focus on a subclass of scale-free networks whose node degree strictly follows the power law degree distribution. Then, besides the fraction ϕ, $F_{\mathcal{N}}(\phi)$ only depends on the node degree distribution of $\mathcal{N}$, [15, 19]. Furthermore, since we only consider the networks (including the sequence of subnetworks and giant components) with

TABLE 3.1: Key notations in mathematical analysis

$\mathcal{C}'_i$, $\mathcal{P}'_j$	the sets of remaining nodes in G_c and G_p which has supporting interlink at Stage i and j
μ'_i, μ'_j	the fraction corresponding to $\lvert\mathcal{C}'_i\rvert$ and $\lvert\mathcal{P}'_j\rvert$, $\lvert\mathcal{C}'_i\rvert = \mu'_i \cdot \lvert\mathcal{C}\rvert$, $\lvert\mathcal{P}'_i\rvert = \mu'_j \cdot \lvert\mathcal{P}\rvert$
$\mathcal{C}_i$, $\mathcal{P}_j$	the functioning giant components in $\mathcal{C}'_i$ and $\mathcal{P}'_j$
μ_i, μ_j	the fraction corresponding to $\lvert\mathcal{C}_i\rvert$ and $\lvert\mathcal{P}_j\rvert$, i.e., $\lvert\mathcal{C}_i\rvert = \mu_i \cdot \lvert\mathcal{C}\rvert$, $\lvert\mathcal{P}_i\rvert = \mu_j \cdot \lvert\mathcal{P}\rvert$
$\lambda_{\mathcal{P}}$	Power law exponent of node degree distribution in G_p. ($\Pr(d_{\mathcal{P}} = k) \propto k^{-\lambda_{\mathcal{P}}}$)
$\lambda_{\mathcal{C}}$	Power law exponent of node degree distribution in G_c. ($\Pr(d_{\mathcal{C}} = k) \propto k^{-\lambda_{\mathcal{C}}}$)

infinite size, the power law distribution is completely determined by the power law exponent. Thus, in this chapter, we let $F(\phi, \lambda)$ represent the expected fraction of giant component in the subnetwork which occupies the fraction ϕ of the nodes in the entire scale-free network with a power law degree distribution, where λ is the power law exponent. Accordingly, $F(\phi, \lambda_{\mathcal{P}})$ and $F(\phi, \lambda_{\mathcal{C}})$ represent the fractions of giant components for network G_p and G_c whose power law exponents are $\lambda_{\mathcal{P}}$ and $\lambda_{\mathcal{C}}$, respectively.

3.2.1 Stage 1: Random Removal in G_c

We begin our estimation with random removal of a fraction $1 - \phi$ of nodes in G_c. The remaining network $\mathcal{C}'_1$ has $|\mathcal{C}| \cdot \phi$ nodes, this value is a positive integer when $|\mathcal{C}|$ is sufficiently large. Then,

$$|\mathcal{C}'_1| = |\mathcal{C}| \cdot \phi = \mu'_1 \cdot |\mathcal{C}|,$$

where μ'_1 denotes the proportion of $\mathcal{C}'_1$ to the entire network G_c, $\mu'_1 = \phi$. Since we assume that only the nodes contained in the giant component can function, how to calculate the giant component size is the next step. Denote the giant component in $\mathcal{C}_1'$ by $\mathcal{C}_1$, we have

$$|\mathcal{C}_1| = |\mathcal{C}_1'| \cdot F(\mu_1', \lambda_{\mathcal{C}}) = \mu_1' \cdot |\mathcal{C}| \cdot F(\mu_1', \lambda_{\mathcal{C}}) = \mu_1 \cdot |\mathcal{C}| \tag{3.3}$$

where μ_1 is analogous with μ_1', and implies the fraction of $\mathcal{C}_1$ on G_c. We notice that at the end of this stage, only the nodes in subnetwork $\mathcal{C}_1$ can properly operate.

3.2.2 Stage 2: Fragmentation in G_p

The removal of nodes and inter links in Stage 1 affects G_p, i.e., some nodes in G_p would be removed due to losing inter links from G_c. We observe that the links removed in Stage 1 include unidirectional links as well bidirectional links, we are focusing on the bidirectional links since the nodes in G_p function only depending on them. Given by (3.2), the probability that each node in G_c has one bidirectional link is $\frac{|\mathcal{P}|}{|\mathcal{C}|}$, hence the expected number of bidirectional links that have been removed is $(|\mathcal{C}| - |\mathcal{C}_1|) \cdot \frac{|\mathcal{P}|}{|\mathcal{C}|}$, which is also the number of nodes in G_p that has to be removed. Let $\mathcal{P}_2'$ denote the set of nodes in G_p which still has the bidirectional inter link, we have

$$|\mathcal{P}_2'| = |\mathcal{P}| - (|\mathcal{C}| - |\mathcal{C}_1|) \cdot \frac{|\mathcal{P}|}{|\mathcal{C}|} = \mu_1' \cdot F(\mu_1', \lambda_{\mathcal{C}}) \cdot |\mathcal{P}| \tag{3.4}$$

$$\mu_2' = \mu_1' \cdot F(\mu_1', \lambda_{\mathcal{C}}) \tag{3.5}$$

The giant component in $\mathcal{P}_2'$, denoted by $\mathcal{P}_2$, follows that:

$$|\mathcal{P}_2| = |\mathcal{P}_2'| \cdot F(\mu_2', \lambda_{\mathcal{P}}) = \mu_2' \cdot |\mathcal{P}| \cdot F(\mu_2', \lambda_{\mathcal{P}}), \tag{3.6}$$

$$\mu_2 = \mu_2' \cdot F(\mu_2', \lambda_{\mathcal{P}}), \tag{3.7}$$

Note that $F(\mu_2', \lambda_{\mathcal{P}})$ represents the fraction of giant component in $\mathcal{P}_2'$.

3.2.3 Stage 3: Further Failures in G_c

Due to the fragmentation of functional nodes in network G_p, some of nodes in network G_c might accordingly lose the inter link and stop operating as a result. To compute such number of nodes, we first observe that the expected number of inter links in $\mathcal{P}_2$ is $|\mathcal{P}_2| \cdot \sum_{k=0}^{|\mathcal{C}|} \Pr(d_{\mathcal{P}} = k) \cdot k$. We recall the two conditions for a node to properly function *1) it belongs to giant component* and *2) it has at least one inter link that support it to work*, observe the expected number of nodes in $\mathcal{C}_1$ that has an inter link from $\mathcal{P}_2$ is given by

$$|\mathcal{C}_3'| = \sum_{k=0}^{|\mathcal{C}|} \Pr(d_{\mathcal{P}} = k) \cdot k \cdot |\mathcal{P}_2| \cdot \frac{|\mathcal{C}_1|}{|\mathcal{C}|} = \mu_2 \cdot \mu_1 \cdot |\mathcal{C}|.$$

That is, passing from $\mathcal{C}_1$ to $\mathcal{C}_3'$, a fraction $1 - \frac{|\mathcal{C}_3'|}{|\mathcal{C}_1|} = 1 - \mu_2$ of nodes are broken. As we did in the previous stages, the next step is to calculate the size of giant component of $\mathcal{C}_3'$, but it is indeed a difficult task, which has been discussed in [20]. We now borrow an idea from [15] that in terms of the size of giant component in $\mathcal{C}_3'$, we consider the effect of removing the fraction of $1 - \mu_2$ of nodes in $\mathcal{C}_1$ has the same effect as taking out the same fraction sizes from $\mathcal{C}_1'$. Therefore, with the respect of giant component size in $\mathcal{C}_3'$, now we can model the process of failure in network G_c from the initial state $|\mathcal{C}|$ to $\mathcal{C}_3'$ by removing a fraction of

$$1 - \phi + \phi \cdot (1 - \mu_2) = 1 - \phi \cdot \mu_2$$

nodes. So, $\mu_3' = \phi \cdot \mu_2$, by which the size of giant component $\mathcal{C}_3$ in subnetwork $\mathcal{C}_3'$ is

$$|\mathcal{C}_3| = \mu_3' \cdot |\mathcal{C}| \cdot F(\mu_3', \lambda_{\mathcal{C}}) = \mu_3 \cdot |\mathcal{C}|. \tag{3.8}$$

3.2.4 Stage 4: Cascading Failures in G_p Once Again

Owing to the removal of nodes and inter links in Stage 3, some relevant bidirectional links are removed, consequently $\mathcal{P}_2$ will fragment into components. Notice that since each node in $\mathcal{C}_3'$ has only one inter link, the removal of $1 - \frac{|\mathcal{C}_3|}{|\mathcal{C}_3'|}$ nodes results in the removal of $1 - \frac{|\mathcal{C}_3|}{|\mathcal{C}_3'|}$ inter links. During this process, the number of removed bidirectional links is $(1 - \frac{|\mathcal{C}_3|}{|\mathcal{C}_3'|}) \cdot \mathcal{P}_2$. Subsequently, the number of nodes which still has the bidirectional link (means still operate) in $\mathcal{P}_2$ is

$$\begin{aligned}
|\mathcal{P}_4'| &= |\mathcal{P}_2| - (1 - |\mathcal{C}_3|/|\mathcal{C}_3'|) \cdot |\mathcal{P}_2| \\
&= \frac{\phi \cdot \mu_2 \cdot F(\mu_3', \lambda_{\mathcal{C}})}{\mu_2 \cdot \mu_1} \cdot |\mathcal{P}_2| \\
&= \frac{F(\mu_3', \lambda_{\mathcal{C}})}{F(\mu_1', \lambda_{\mathcal{C}})} \cdot |\mathcal{P}_2|
\end{aligned}$$

Passing from $\mathcal{P}_2$ to $\mathcal{P}_4'$, the fraction $1 - |\mathcal{P}_4'|/|\mathcal{P}_2| = 1 - F(\mu_3', \lambda_{\mathcal{C}})/F(\mu_1', \lambda_{\mathcal{C}})$ of nodes are removed, as we do in Stage 3, in terms of the size of giant component in $\mathcal{P}_4'$, that is equivalent to remove the same fraction of nodes from $\mathcal{P}_2'$. Therefore, if we model the failure process from initial state $|\mathcal{P}|$ to $\mathcal{P}_4'$, the proportion of nodes that has to be removed is

$$1 - \mu_2' + \mu_2' \cdot \left(1 - \frac{F(\mu_3', \lambda_{\mathcal{C}})}{F(\mu_1', \lambda_{\mathcal{C}})}\right) = 1 - \mu_1' \cdot F(\mu_3', \lambda_{\mathcal{C}}).$$

As before, $\mu_4' = \mu_1' \cdot F(\mu_3', \lambda_{\mathcal{C}})$, and with this value, the corresponding giant component $\mathcal{P}_4$ is given to be

$$|\mathcal{P}_4| = \mu_4' \cdot F(\mu_4', \lambda_{\mathcal{P}}) \cdot |\mathcal{P}|. \tag{3.9}$$

3.2.5 Transcendental Equations for Cascading Failures

Following the previous steps, one can obtain the size of giant component $|\mathcal{C}_1|, |\mathcal{P}_2|, |\mathcal{C}_3| \cdots$ in a certain stage, but no one knows in which stage the cascading failure stops, and the entire system reaches the final steady state. Our main aim is to estimate the size of survivals in the final stage. When we repeat the above calculations we can obtain the recursive equations to represent the size of giant component in both networks G_c and G_p as follows:

	Network G_c	**Network** G_p
Stage 1	$\mu_1' = \phi$ $\mu_1 = \mu_1' \cdot F(\mu_1', \lambda_{\mathcal{C}})$	
Stage 2		$\mu_2' = \mu_1' \cdot F(\mu_1', \lambda_{\mathcal{C}})$ $\mu_2 = \mu_2' \cdot F(\mu_2', \lambda_{\mathcal{P}})$
Stage 3	$\mu_3' = \phi \cdot \mu_2$ $\mu_3 = \mu_3' \cdot F(\mu_3', \lambda_{\mathcal{C}})$	
Stage 4		$\mu_4' = \mu_1' \cdot F(\mu_3', \lambda_{\mathcal{C}})$ $\mu_4 = \mu_4' \cdot F(\mu_4', \lambda_{\mathcal{P}})$
$\cdots$	$\cdots$	$\cdots$
Stage $2j$		$\mu_{2j}' = \mu_1' \cdot F(\mu_{2j-1}', \lambda_{\mathcal{C}})$ $\mu_{2j} = \mu_{2j}' \cdot F(\mu_{2j}', \lambda_{\mathcal{P}})$
Stage $2j+1$	$\mu_{2j+1}' = \phi \cdot \mu_{2j}$ $\mu_{2j+1} = \mu_{2j+1}' \cdot F(\mu_{2j+1}', \lambda_{\mathcal{C}})$	

When the cascading failure stops, the following equations hold,

$$\mu_{2j+1}' = \mu_{2j+3}' = \mu_{2j-1}'$$
$$\mu_{2j}' = \mu_{2j+2}' = \mu_{2j-2}'$$

because the giant components in both networks have no further fragments. Let us denote $x = \mu'_{2j+1} = \mu'_{2j+3} = \mu'_{2j-1}$ and $y = \mu'_{2j} = \mu'_{2j+2} = \mu'_{2j-2}$, we have

$$\begin{cases} x = \phi \cdot y \cdot F(y, \lambda_{\mathcal{P}}) \\ y = \phi \cdot F(x, \lambda_{\mathcal{C}}) \end{cases} \tag{3.10}$$

where $0 \leq x, y \leq 1$. There may be two sets of giant component sizes. In this case, the solution is the point that is closer to the initial status, because the system fragmentation stops at this point and never goes to the small one.

The fraction of nodes that still functions in the final steady state in both networks can be calculated by

$$\begin{cases} \lim_{j\to\infty} \mu_{2j} = y \cdot F(y, \lambda_{\mathcal{P}}) \\ \lim_{j\to\infty} \mu_{2j+1} = x \cdot F(x, \lambda_{\mathcal{C}}) \end{cases} \tag{3.11}$$

This gives us a complete solution of the fraction of survivals in G_p and G_c. If we can find a non-trivial solution for x and y, then we can compute the remaining number of survivals.

3.3 Generating Function for Calculating the Survivals

In this section, we discuss how to calculate the functioning parts μ_{2j} and μ_{2j+1}. That is equivalent to finding the solution of x and y. We observe that Eq. (3.10) has a trivial solution $x = y = 0$; however, we are indeed interested in the non-trivial solution. Table 3.2 shows the notations to be used in this section.

Table 3.2: Notations in the analysis of generating functions

$d_{\mathcal{P}}$	Intra node degree for network G_p
$d_{\mathcal{C}}$	Intra degree for network G_c
$G_{0,\mathcal{P}}$, $G_{1,\mathcal{P}}$	Generating functions for the intra node degree distribution in G_p
$G_{0,\mathcal{C}}$, $G_{1,\mathcal{C}}$	Generating functions for the intra node degree distribution in G_c

We firstly introduce the generating function [21] of degree distribution for network $\mathcal{N}$, i.e.,

$$G_{0,\mathcal{N}}(u) = \sum_{k=0}^{\infty} \Pr(d_{\mathcal{N}} = k) \cdot u^k, \tag{3.12}$$

where $d_{\mathcal{N}}$ follows the intra degree distribution for network $\mathcal{N}$, and $\Pr(d_{\mathcal{N}} = k)$ represents the probability that a node has k intra links. Recall that we assume the degree distribution of physical-resource network follows power-law, i.e., $\Pr(d = k) = \kappa \cdot k^{-\lambda}$, where κ is a constant, power-law exponent λ varies for different network structures.

The *excess degree distribution* [15, 21] is the number of edges attached to a vertex other than the edge we arrived along, and given by

$$G_{1,\mathcal{N}}(u) = \sum_{k=0}^{\infty} Q_k \cdot u^k \tag{3.13}$$

$$Q_{k,\mathcal{N}} = \frac{(k+1) \cdot \Pr(d_{\mathcal{N}} = k+1)}{\bar{K}} \tag{3.14}$$

where $\bar{K}$ is the first moment of $d_{\mathcal{N}}$. The following two lemmas are used to calculate the functioning giant component.

Lemma 3.1 ([21]). *For the network $\mathcal{N}$, the expected proportion of giant component after removing a fraction $1-\phi$ of nodes from the initial network, denoted by $F_{\mathcal{N}}$, holds that*

$$\begin{cases} F_{\mathcal{N}}(\phi) = 1 - G_{0,\mathcal{N}}(u) \\ u = 1 - \phi + \phi \cdot G_{1,\mathcal{N}}(u) \end{cases} \tag{3.15}$$

where $F_{\mathcal{N}}(\phi) \leq 1$.

Lemma 3.2 ([143]). *For a single infinite scale-free network $\mathcal{N}$ with a power law exponent λ, it holds that $F(\phi, \lambda) \propto \kappa \cdot \phi^{1/(3-\lambda)}$ for $2 < \lambda < 3$, where κ is a constant.*

Theorem 3.3. *Eq. (3.10) always has two solutions, one of which is a trivial one,* i.e., $x = y = 0$. *For the other solution, there exists a critical value ϕ_c, such that: for $\phi < \phi_c$, $x, y > 1$, which makes no sense due to the condition of Eq. (3.10) that $0 \leq x, y \leq 1$; otherwise, one efficient solution can be obtained,* i.e., $x, y \leq 1$, *which leads to a solution of Eq. (3.11), representing the remaining part in both networks.*

Proof. Since we consider both G_p and G_c are scale-free networks, the generating functions are in the same form of Eqs. (3.12) and (3.13). We assume the degree distribution of network G_p and G_c are $\Pr(d_P = k) = \kappa_a \cdot k^{-\lambda_{\mathcal{P}}}$, $\Pr(d_I = k) = \kappa_b \cdot k^{-\lambda_{\mathcal{C}}}$ respectively, where κ_a and κ_b are constants. Now, combining Eqs. (3.10) and (4.8), we obtain a set of equations as follows:

$$\begin{cases} x = \phi \cdot y \cdot F(y, \lambda_{\mathcal{P}}) \\ y = \phi \cdot F(x, \lambda_{\mathcal{C}}) \\ F(y, \lambda_{\mathcal{P}}) = 1 - G_{0,\mathcal{P}}(u_1) \\ u_1 = 1 - y + y \cdot G_{1,\mathcal{P}}(u_1) \\ F(x, \lambda_{\mathcal{C}}) = 1 - G_{0,\mathcal{C}}(u_2) \\ u_2 = 1 - x + x \cdot G_{1,\mathcal{C}}(u_2) \end{cases} \tag{3.16}$$

This can give us a complete solution for the value of x and y, although in practice it is not possible to solve Eq. (4.8) in closed form [21].

According to Lemma 3.2, let $F(y, \lambda_{\mathcal{P}}) = \kappa_1 \cdot y^{1/(3-\lambda_{\mathcal{P}})}$ and $F(x, \lambda_{\mathcal{C}}) = \kappa_2 \cdot x^{1/(3-\lambda_{\mathcal{C}})}$, where κ_1 and κ_2 are predefined constants based on $\lambda_{\mathcal{P}}$ and $\lambda_{\mathcal{C}}$ respectively. Now

Eq. (3.16) becomes

$$\begin{cases} x = \phi \cdot y \cdot \kappa_1 \cdot y^{1/(3-\lambda_{\mathcal{P}})} \\ y = \phi \cdot \kappa_2 \cdot x^{1/(3-\lambda_{\mathcal{C}})} \end{cases} \tag{3.17}$$

Excluding y, we get

$$x = \kappa_1 \cdot \kappa_2^{1+1/(3-\lambda_{\mathcal{P}})} \cdot \phi^{2+1/(3-\lambda_{\mathcal{P}})} \cdot x^{(1+1/(3-\lambda_{\mathcal{P}}))\cdot(1/(3-\lambda_{\mathcal{C}}))} \tag{3.18}$$

The right side of Eq. (3.18) can be simplified as the form of $C \cdot x^{\eta}$, where

$$\eta = (1 + 1/(3 - \lambda_{\mathcal{P}})) \cdot (1/(3 - \lambda_{\mathcal{C}})), \tag{3.19}$$

since $2 < \lambda_{\mathcal{P}}, \lambda_{\mathcal{C}} < 3$, η is greater than 1. We notice that Eq. (3.18) has a trivial solution $x = 0$, which indicates that there is no node left in the giant component. Also, the right side curve always goes through the line $y = x$ once for $x > 0$, as shown in Fig. 3.2. We recall the condition of Eq. (3.10) that $0 \leq x, y \leq 1$, thus only the intersection between [0,1] is an efficient solution for x.

The case of Fig 3.2(b) shows the borderline between case (a) and (c). The efficient solution for x exists in case (a), but not in (c). Mathematically case (b) is the point of the threshold; in other words, the threshold occurs when

$$\kappa_1 \cdot \kappa_2^{1+1/(3-\lambda_{\mathcal{P}})} \cdot \phi_c^{2+1/(3-\lambda_{\mathcal{P}})} = 1. \tag{3.20}$$

So, the critical value for the solution of x is when

$$\phi_x = \sqrt[2+1/(3-\lambda_{\mathcal{P}})]{\kappa_1^{-1} \cdot \kappa_2^{1/(\lambda_{\mathcal{P}}-3)-1}}. \tag{3.21}$$

We repeat the above process for calculating the critical point of y. The threshold occurs when

$$\phi_y = \sqrt[1+1/(3-\lambda_{\mathcal{C}})]{\kappa_1^{1/(\lambda_{\mathcal{C}}-3)} \cdot \kappa_2^{-1}}. \tag{3.22}$$

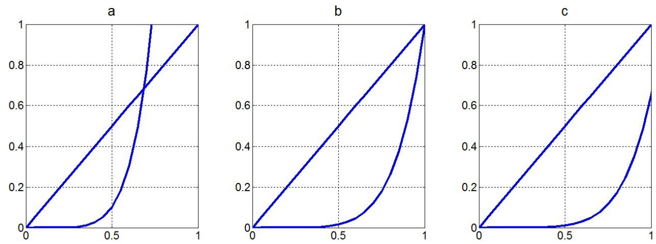


FIGURE 3.2: The left side and right side curves of Eq. (3.18). For a given $\kappa_1 = \kappa_2 = 2$, $\lambda_{\mathcal{P}} = \lambda_{\mathcal{C}} = 2.5$, $\phi = 0.8, 0.6, 0.4$ respectively. Two intersections exist for cases (a), (b).

To satisfy the condition $x, y \subset$ [0,1], the system threshold ϕ_c is Max$[\phi_x, \phi_y]$.

Therefore, Eq. (3.10) always have two solutions: one is $x = y = 0$ which is a trivial solution, and the other one depends on the value of ϕ. If $\phi \geq \phi_c$, an efficient solution exists, which means functioning part remains in both networks; otherwise, the entire system collapses. $\square$

However, finding the non-trivial solution is challenging since the exact value of κ is still unknown [143]. To the best of our knowledge, deriving the closed-form solution of Eq. (4.8) is still open. Therefore, in this work, we resort to the standard graphical method for simulations.

3.4 Simulation Results and Analysis

In this section, we simulate the system and the cascading failures to obtain the fractions of survivals μ_{2j} and μ_{2j+1} in the final steady stage and show how the system behaves. To generate the scale-free network with different power-law distributions, we adopt the *generalized Barabási-Albert model* [21], whose power law exponent can vary in the range $(2, +\infty)$. In the case of a CPS such as smart grid, for example, the power grid and Internet are growing every day, which causes the change of degree distribution. As pointed out in [143], the power-law exponent for Internet on *Autonomous System* level is 2.2 in 1999, but becomes to 2.1 in 2000, while that data from [21] is 2.5. Therefore, we intentionally construct the models with various parameters to carter the dynamic of a CPS. We also use real world Western States Power Grid data [4] in our simulation to study the smart power grid.

We apply a random removal to represent the initial failure upon network G_c, and obtain an accurate result by implementing extensive simulations.

3.4.1 Simulation Setup

We write a specific program using Java to simulate the entire percolation process in interdependent networks. In all the experiments the same approach is employed:

- first, we construct two scale-free networks representing G_p and G_c, by using generalized Barabási-Albert model.
- then, we couple the two networks using our proposed interdependence model. The Balls and Bins scheme is applied.
- we randomly choose the fraction of $1 - \phi$ nodes in network G_c as random failure or attack.
- the failure propagates between the two networks. Our program simulates each stage, stores and updates the network information.
- finally, our program calculates the number of survived nodes in both networks.

3.4.2 System Robustness

First of all, we compare the different values of μ_{2j} and μ_{2j+1} by changing ϕ, with the same values of $\lambda_{\mathcal{P}}$ and $\lambda_{\mathcal{C}}$. Fig. 3.3 shows the comparison. Generally, the

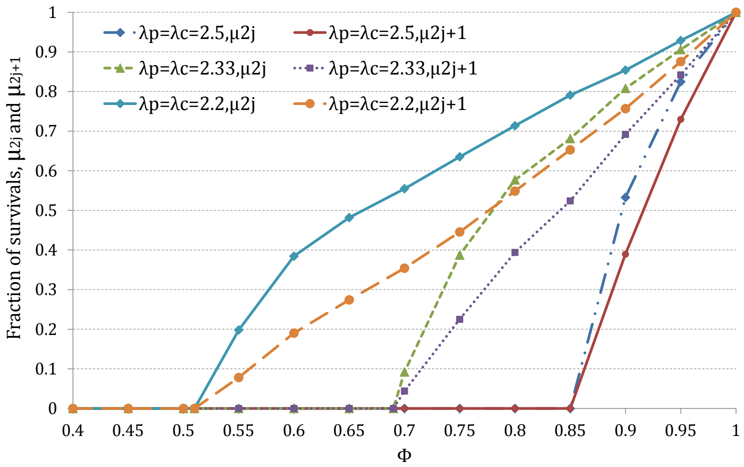


FIGURE 3.3: The fraction of survivals in both networks. For our simulation, the initial network size of G_p is 1000 with $10,000$ nodes in G_c. We use generalized Barabási-Albert model to construct both networks. The initial random failure $1-\phi$ occurs in network G_c.

fraction of nodes which remains functional after the cascading failure stops in both networks grows as $\lambda_{\mathcal{P}}$ and $\lambda_{\mathcal{C}}$ become smaller, which means the system is more reliable. The reason is that the network with small power-law exponent has *fat-tail effect*, the number of nodes with high degree is larger, these 'hub's have more neighbors, so the connectivity of network is better. Another finding is that the value of ϕ_c increases by the reducing power-law exponent. For $\lambda_{\mathcal{P}} = \lambda_{\mathcal{C}} = 2.2$, ϕ_c is approximate to 0.51, while this value is 0.85 for $\lambda_{\mathcal{P}} = \lambda_{\mathcal{C}} = 2.5$, the former system outperforms greatly the latter one.

We observe that the values of μ_{2j} and μ_{2j+1} can *never* be in the situation where one is zero but another is not. This means that collapse of one network results in the failure of the other dependent network.

3.4.3 Second-order Transition of Giant Component Size

The proposed *first-order discontinuous transition* [15] in interdependent networks says the size of giant component meets a sharp transition when ϕ approaches ϕ_c. To be specific, if the network size is infinite, the transition becomes a *step function*.

TABLE 3.3: The sharp transition of giant component size when $\lambda_{\mathcal{P}} = \lambda_{\mathcal{C}} = 2.33, \phi_c = 0.69$, network size $|\mathcal{P}| = 1000, |\mathcal{C}| = 10000$. The values in the table are the number of nodes remain functioning in the end. Each row is one sample for implementation.

$\phi = 0.73$		$\phi = 0.71$	
Nodes in G_p	Nodes in G_c	Nodes in G_p	Nodes in G_c
483	2725	1	1
249	1035	303	1397
1	1	337	1580
425	2251	1	1
389	2079	1	1
1	1	2	2
375	1898	423	2200
1	1	1	1
1	1	1	1
331	1640	1	1
438	2377	400	2146
400	2139	385	2054
385	2047	328	1541
460	2613	1	1
1	1	1	1

Since the network sizes in simulation are all finite, this transition becomes much more flat. Thus, we say there is an transition interval.

Therefore, there is a small transition interval including the value of ϕ_c. Fig. 3.4 shows an example of transition interval. The giant components always exist in both networks for $\phi > 0.76$, while the system collapses when $\phi < 0.69$. Thus, the transition interval is [0.69, 0.76], during which the system might either collapses or not.

Table 3.3 shows part of simulation results. Each row in the table is a sample of survivals in two networks. In some cases, the size of giant component scales in proportion to that of network, e.g., the remaining nodes in G_p is 385, corresponding to 2047 nodes in G_c; while in other cases, there is no giant component remaining, i.e., both are only 1 node left.

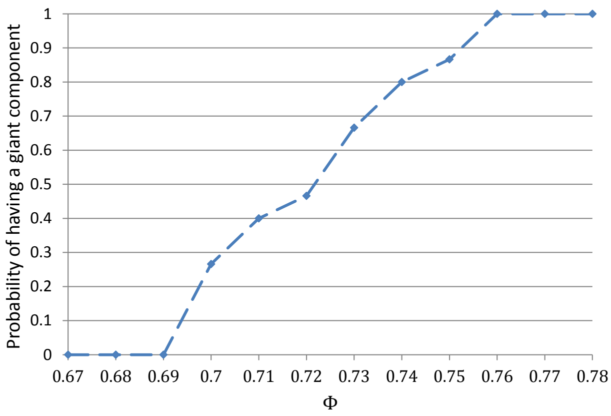


FIGURE 3.4: The probability for having a giant component in the final steady state. $\lambda_{\mathcal{P}} = \lambda_{\mathcal{C}} = 2.33, \phi_c = 0.69$, network size $|\mathcal{P}| = 1000, |\mathcal{C}| = 10000$. The entire system totally collapses when $\phi < 0.69$. For $\phi \geq 0.76$, giant components always exist in both G_p and G_c.

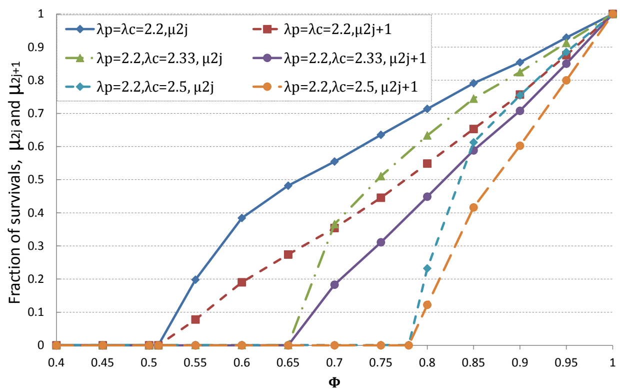
TABLE 3.4: Comparison of network size

$\lambda_{\mathcal{P}} = \lambda_{\mathcal{C}} = 2.2$	$\lvert\mathcal{P}\rvert = 500,\ \lvert\mathcal{C}\rvert = 10000$		$\lvert\mathcal{P}\rvert = 1000,\ \lvert\mathcal{C}\rvert = 10000$		$\lvert\mathcal{P}\rvert = 1200,\ \lvert\mathcal{C}\rvert = 10000$		$\lvert\mathcal{P}\rvert = 500,\ \lvert\mathcal{C}\rvert = 5000$	
	μ_{2j}	μ_{2j+1}	μ_{2j}	μ_{2j+1}	μ_{2j}	μ_{2j+1}	μ_{2j}	μ_{2j+1}
$\phi = 0.9$	0.86	0.76	0.854	0.756	0.861	0.768	0.859	0.76
$\phi = 0.8$	0.711	0.532	0.713	0.548	0.712	0.549	0.715	0.548
$\phi = 0.7$	0.554	0.34	0.554	0.354	0.562	0.365	0.546	0.347

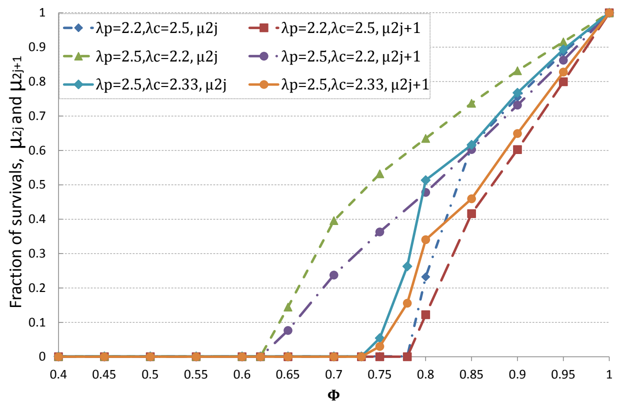
3.4.4 Comparison between Different $\lambda_{\mathcal{P}}$ and $\lambda_{\mathcal{C}}$

We now show how the system robustness depends on the different values of $\lambda_{\mathcal{P}}$ and $\lambda_{\mathcal{C}}$. Fig. 3.5(a) assuming $\lambda_{\mathcal{P}} = 2.2$, illustrates that the CPS robustness goes down when $\lambda_{\mathcal{C}}$ increases. Comparing Fig. 3.3 with Fig. 3.5(a), we notice that ϕ_c for $\lambda_{\mathcal{P}} = \lambda_{\mathcal{C}} = 2.5$ is 0.85, while it is 0.78 if $\lambda_{\mathcal{P}}$ is 2.2. Thus, reducing one network's degree distribution exponent will improve the reliability of the entire system. We notice that the curves for $\lambda_{\mathcal{P}} = 2.2, \lambda_{\mathcal{C}} = 2.5$ and $\lambda_{\mathcal{P}} = 2.5, \lambda_{\mathcal{C}} = 2.2$ in Fig. 3.5(b) have a huge gap. The ϕ_c for the former one is approximately 0.78, but for the latter one it is 0.62. This implies that the initial attack upon different networks leads to very different results.

As a typical application of CPS, smart grid is an important interdependent networks. We couple the real Western States Power Grid data [4] with an synthetic scale-free network in Fig. 3.6(a) and Fig. 3.6(b). Western States Power Grid has



(a)



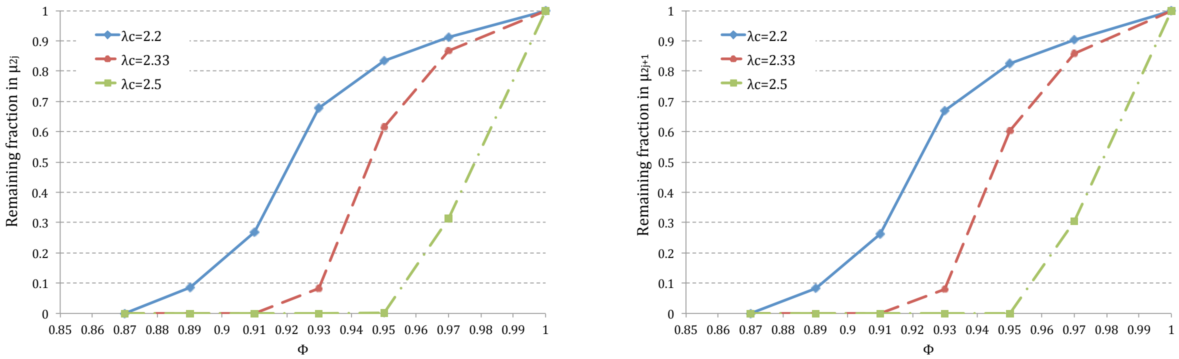
(b)

FIGURE 3.5: The fraction of survivals μ_{2j} and μ_{2j+1}. For the simulation, the network size $|\mathcal{P}|$=1000, $|\mathcal{C}|$=10000, the initial random failure $1-\phi$ is upon network G_c.

4941 nodes with the mean degree of 2.67. The system robustness and ϕ_c are compared under various $\lambda_{\mathcal{C}}$. The reason why the values of ϕ_c of Fig. 3.6(a) and Fig. 3.6(b) are higher than Fig. 3.5(a) and Fig. 3.5(b) is that the degree distribution in Western States Power Grid does not strictly follow the power law.

3.4.5 Impact of Network Size

Our mathematical Eqs. (3.10) and (4.19) demonstrate that μ_{2j} and μ_{2j+1} do not depend on $|\mathcal{P}|$ and $|\mathcal{C}|$, as well as the ratio between them. In other words, the fraction of remaining functioning giant component is determined by power law exponent of the network.



(a) (b)

FIGURE 3.6: The fraction of survivals μ_{2j} and μ_{2j+1}. We couple Western States Power Grid (G_p) and an synthetic scale-free network (G_c) for (c) and (d). The initial network sizes are 4941 and 8000 for G_p and G_c respectively.

Our simulation tells the same story. According to Table 3.4 that fixing $\lambda_{\mathcal{P}} = \lambda_{\mathcal{C}} = 2.2$, giving the μ_{2j} and μ_{2j+1} with different network size and ratio, we observe that no mater what the network sizes are or what the ratio between $|\mathcal{P}|$ and $|\mathcal{C}|$ is, the remaining fraction of nodes in the end is actually the same.

So for a real-life CPS, such as smart grid system, our model shows that whether it grows or not, the system robustness in terms of random failure remains in a specific level.

Chapter 4

Extended ‘k-n’ Interdependence Model

This chapter is going to propose an interdependence model, called ‘k-n’ model, where the nodes in network have different roles. To better explain this model, we use smart grid as example. However, to simplify our analysis, we limit the scope into the network topologies and inter relations, and ignore the network physical features, such like network protocols and electricity characteristics.

This is a general interdependence model, and thus could be applied to many other cyber physical systems, such as vehicular ad hoc networks.

4.1 System Model

4.1.1 Assumptions and Definitions

We model the smart power grid as two interconnected networks $\mathcal{N}_{\mathrm{P}}$ and $\mathcal{N}_{\mathrm{I}}$ (power grid and communication/control network, respectively). $\mathcal{N}_{\mathrm{I}}$ is envisioned as a part of the Internet backbone, extended with some wireless links. Both networks consist of a large number of components; some of them are considered as terminals, with limited influence to the system reliability. For instance, the breakdown of a smart meter in one house is not likely to cause the failure of other components. We only consider the components that have some dependencies or are providing connectivity. Nodes in $\mathcal{N}_{\text{P}}$ represent the power plants, substations, transformers and new energy generators. Nodes in $\mathcal{N}_{\text{I}}$ are *Autonomous Systems (AS)*. There are two types of ASs: *Operation Centers*, which monitor and control power stations and exchange information with other communication devices, and *Relaying ASs*, for relaying messages in smart grid system, with no direct impact on power stations. S_{P} and S_{I} are sizes of $\mathcal{N}_{\text{P}}$ and $\mathcal{N}_{\text{I}}$, respectively.

The distribution of power stations and edges in power grid was studied and modeled in [122, 144], which considered power grid as a scale-free network. A scale-free network is a network whose degree distribution follows a power law, $\mathbf{P}(z) \propto z^{-\gamma}$, where $\mathbf{P}(z)$ is the probability that the degree of a node is z, γ is power law exponent. Extensive data show that Internet is also a scale-free network [21, 143].

4.1.2 Interdependence Model

We refer to edges that connect the nodes from different networks as *Inter Links*. In smart grid, the operation centers actually have limited capability to monitor and control the power grid [145]. More importantly, the centers in different regions need to share the data and see a big picture of the entire grid. Therefore, dozens of operations centers are required to meet the demand to control and manage the smart grid [146, 147].

To simplify our model, we are not considering which power station is managed by which specific operation center. The operation centers as well as the controlled stations are randomly chosen.

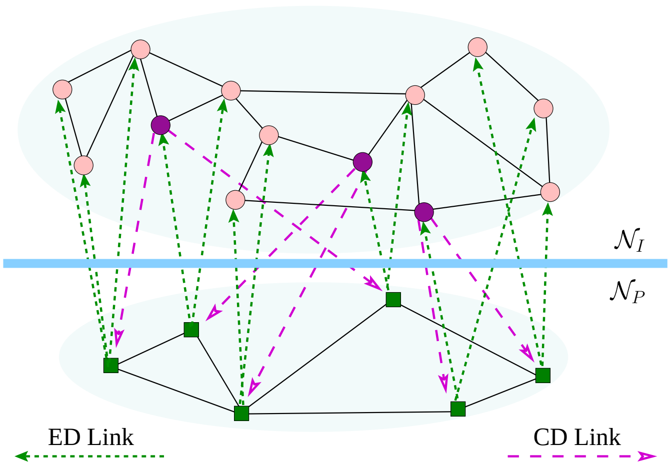


FIGURE 4.1: Each node in $\mathcal{N}_\mathrm{I}$ has one energy inter link from $\mathcal{N}_\mathrm{P}$, and each node in $\mathcal{N}_\mathrm{P}$ is controlled by $k = 1$ operation centers. Three dark nodes ($m = 3$) are operation centers and each controls $n = 2$ nodes from $\mathcal{N}_\mathrm{P}$.

We assume that each power station in $\mathcal{N}_\mathrm{P}$ is operated by k operation centers, and functions properly as long as at least one of them works. The *monitoring capability* (limited capability) of each operation center is defined as the number n of power stations it can control. Thus, each node $\mathcal{N}_\mathrm{P}$ has k CD (Control-Dependency) inter links and each operation center has n CD links. Meanwhile, each power station has multiple ED (Energy-Dependency) links to ASs in $\mathcal{N}_\mathrm{I}$, while each node in $\mathcal{N}_\mathrm{I}$ has only one ED link.

We define the cost including hardware, software and labour as *Controlling Cost*, measured based on two parameters: k and the number of operation centers m. k determines $k \cdot S_\mathrm{P}$, the total number of CD inter links required. To minimize the cost with a fixed k, we minimize m by fully utilizing the control capability of each operation center. In other words, the minimum controlling cost is achieved for $m = \frac{k \cdot S_\mathrm{P}}{n}$. Thus the minimum controlling cost is proportional to k. In the rest of paper, we simply consider the controlling cost to be k. For each k, we automatically calculate the minimum m in the cost formula. Figure 4.1 gives a sketch of our model, with $k = 1, m = 3, n = 2$.

We construct inter links by applying a three step procedure:

4.1.2.1 Allocating ED Link

We consider each node in $\mathcal{N}_{\mathrm{P}}$ as a bin and nodes in $\mathcal{N}_{\mathrm{I}}$ are balls. The allocation follows the well-known *Balls and Bins* problem, where S_{I} balls have to be independently and uniformly put into S_{P} bins, the probability that one ball is assigned into i-th bin is $\frac{1}{S_{\mathrm{P}}}$. For each bin, the probability it has t balls is given by:

$$\mathbf{P}(t) = \binom{S_{\mathrm{I}}}{t} \cdot (\frac{1}{S_{\mathrm{P}}})^{t} \cdot (1 - \frac{1}{S_{\mathrm{P}}})^{S_{\mathrm{I}}-t}. \tag{4.1}$$

Hence, the number of *AS* t that is supported by each power station follows a binomial distribution with $\mathbf{B}(S_{\mathrm{I}}, \frac{1}{S_{\mathrm{P}}})$.

4.1.2.2 Choosing Minimum Number of Operation Center

We choose uniformly m nodes at random as operation centers from $\mathcal{N}_{\mathrm{I}}$. The relation between m, n and k is given by

$$m = \frac{k \cdot S_{\mathrm{P}}}{n}, \tag{4.2}$$

where each node in $\mathcal{N}_{\mathrm{P}}$ is monitored by k operation centers, thus totally $k \cdot S_{\mathrm{P}}$ links are required, and each operation center can control n power stations.

4.1.2.3 Allocating CD Link

Subsequently, we match m operation centers and S_{P} power stations so that each operation center is allocated n power stations, and each power station is controlled by k operating centers. This can be done in a sequence, by operating centers selecting power stations at random but only among stations not selected k times already.

4.2 Mathematical Analysis

In this section, we show how cascading failure propagates in our model. Then we introduce the math tools that be used in a single complex network. Before starting, we list the definitions and notations in Table 5.1. We introduce a notation $F_{\mathcal{N}}(\phi)$ from [19, 20] to represent the expected fraction of giant component in the subnetwork which occupies the fraction ϕ of the nodes in the entire network $\mathcal{N}$. In our work, we focus on a subclass of scale-free networks whose node degree strictly follows the power law degree distribution. Then, besides the fraction ϕ, $F_{\mathcal{N}}(\phi)$ only depends on the node degree distribution of $\mathcal{N}$, [15, 19]. Furthermore, since we only consider the networks (including the sequence of subnetworks and giant components) with infinite size, the power law distribution is completely determined by the power law exponent. Thus, in this paper, we let $F(\phi, \lambda)$ represent the expected fraction of giant component in the subnetwork which occupies the fraction ϕ of the nodes in the entire scale-free network with a power law degree distribution, where λ is the power law exponent. Accordingly, $F(\phi, \lambda_{\mathrm{P}})$ and $F(\phi, \lambda_{\mathrm{I}})$ represent the fractions of giant components for network $\mathcal{N}_{\mathrm{P}}$ and $\mathcal{N}_{\mathrm{I}}$ whose power law exponents are λ_{P} and λ_{I}, respectively.

4.2.1 Failure Cascading Process

We define system robustness in our model as the fraction of survivals after the cascading failure stops, i.e., the nodes that still can operate. We focus on how this failure cascading propagates and then estimate the survivals. Considering smart grid as two complex networks, we assume two conditions should be satisfied if a node is in work:

1. The node belongs to the *giant component*.

TABLE 4.1: Notations for the analysis

k	number of operation centers that control a power station, it represents the system controlling cost
n	monitoring capability of operation center
m	number of operation centers
$\mathbf{P}_\mathrm{P}(z), \mathbf{P}_\mathrm{I}(z)$	intra degree distribution for network $\mathcal{N}_\mathrm{P}, \mathcal{N}_\mathrm{I}$
$\lambda_\mathrm{P}, \lambda_\mathrm{I}$	power law exponent of $\mathbf{P}_\mathrm{P}(z), \mathbf{P}_\mathrm{I}(z)$.
$F(\phi, \lambda_\mathrm{P}), F(\phi, \lambda_\mathrm{I})$	the fraction of giant component in the subnetwork which occupies the fraction ϕ of the nodes in the entire network $\mathcal{N}_\mathrm{P}$ and $\mathcal{N}_\mathrm{I}$

2. At least one inter link is connected to this node, where this link comes from a functioning node in the other network.

Again, we arbitrarily choose $\mathcal{N}_\mathrm{I}$ to start. We begin with a random removal of $(1-\phi)\cdot S_\mathrm{I}$ nodes in $\mathcal{N}_\mathrm{I}$ as the simulation of initial failures or attacks. After this removal, the related intra links and inter links of deleted nodes are removed. As a result, $\mathcal{N}_\mathrm{I}$ begins to fragment into disconnected components. Due to our Condition 1, only the nodes belong to giant component can operate properly. Therefore, the nodes in small components are considered as failure. Now owning to the interdependency, a part of nodes in $\mathcal{N}_\mathrm{P}$ lost inter links so they are unsatisfied with Condition 2. Then these nodes and related links are removed. The fragmentation in $\mathcal{N}_\mathrm{P}$ might lead to further failures in $\mathcal{N}_\mathrm{I}$, because now some nodes in $\mathcal{N}_\mathrm{I}$ have no inter links. This cascading failure continues recursively between two networks, and reaches one of the following two final status: *1)* all nodes are faulty and the giant component disappears; *2)* the two giant components in $\mathcal{N}_\mathrm{P}$ and $\mathcal{N}_\mathrm{I}$ are mutually connected. One complete cascading failure example is given by Fig. 4.2. Initially, ten nodes and six nodes are in $\mathcal{N}_\mathrm{I}$ and $\mathcal{N}_\mathrm{P}$ respectively. After one node is attacked in $\mathcal{N}_\mathrm{I}$, the entire system collapses.

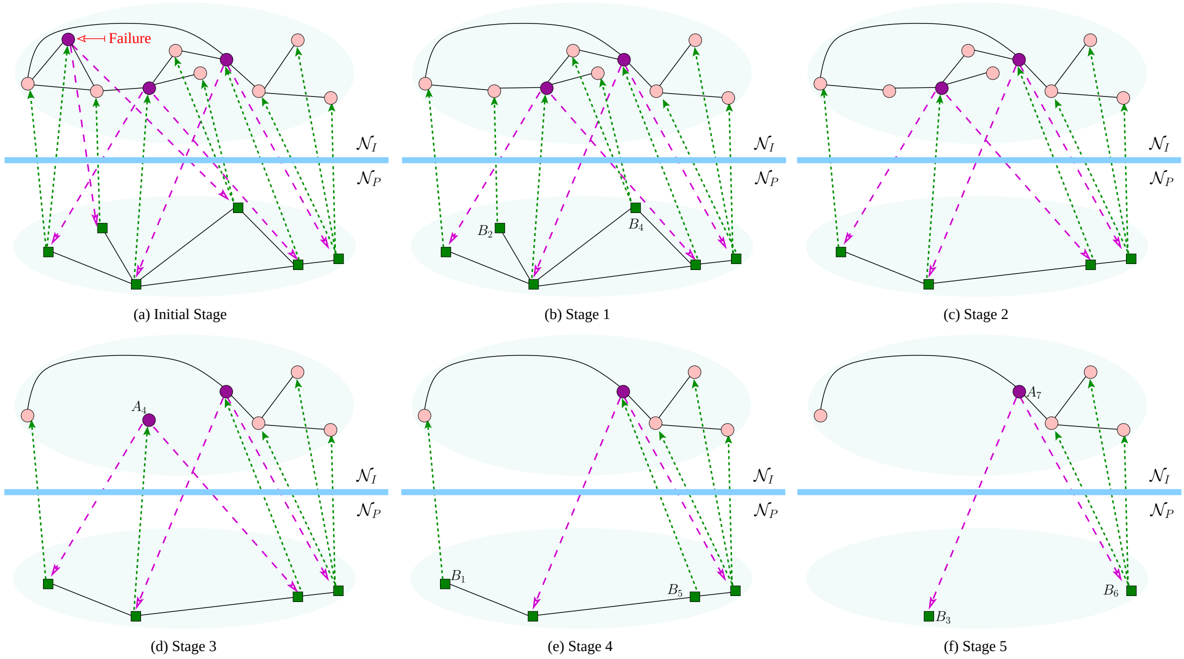


FIGURE 4.2: A sketch of cascading failure in smart grid. Initially, power grid $\mathcal{N}_{\mathrm{P}}$ has 6 nodes while Internet has 10 nodes. Each node in $\mathcal{N}_{\mathrm{P}}$ is operated by one operation center, which implies $k = 1$. Each operation center in $\mathcal{N}_{\mathrm{I}}$ can operate 2 power stations, $n = 2$. Three dark nodes are randomly chosen and assigned as operation centers. The random attack causes the failure of one operation center, thus all its related links are removed. In stage 1, the remaining nodes in $\mathcal{N}_{\mathrm{I}}$ are mutually reachable, therefore consist of a giant component. B_2 and B_4 lose control, thus in Stage 2 they are out of work. In Stage 3, due to the loss of energy inter links, some nodes in $\mathcal{N}_{\mathrm{I}}$ are removed. As a result, nodes A_4 is disconnected from the giant component, thus fails. Consequently, B_1, B_5 are faulty since they lose the control inter links from A_4. In Stage 5, the remaining nodes in $\mathcal{N}_{\mathrm{P}}$ are disconnected, therefore no giant component exists. Due to our Condition 1, the entire $\mathcal{N}_{\mathrm{P}}$ collapses. As a result, the nodes in $\mathcal{N}_{\mathrm{I}}$ lost energy support, thus fail.

4.2.2 Generating Function for An Individual Network

Generating function and percolation theory are widely used to solve the problems in complex network. We describe the generating function for a single network that will also be used in studying interdependent networks. We introduce generating function into our model. Let us assume the nodes in $\mathcal{N}_{\mathrm{P}}$ are assigned a degree z with the same probability $\mathbf{P}_{\mathrm{P}}(z)$, which follows power law in scale-free network.

Thus $\mathbf{P}_{\mathrm{P}}(z) \propto z^{-\lambda_{\mathrm{P}}}$. The generating function is defined as

$$G_{\mathrm{P}}(u) = \sum_{z=0}^{\infty} \mathbf{P}_{\mathrm{P}}(z) \cdot u^z, \tag{4.3}$$

where u is an arbitrary variable. The *excess degree distribution* [21] is the number of edges attached to a vertex other than the edge we arrived along, and given by

$$H_{\mathrm{P}}(u) = \sum_{z=0}^{\infty} Q_{\mathrm{P}}(z) \cdot u^z \tag{4.4}$$

$$Q_{\mathrm{P}}(z) = \frac{(z+1) \cdot \mathbf{P}_{\mathrm{P}}(z+1)}{\bar{z}}, \tag{4.5}$$

where $\bar{z}$ is the average degree of network $\mathcal{N}_{\mathrm{P}}$. Using Eq. (4.3), $\bar{z}$ is calculated

$$\bar{z} = \sum_{z=0}^{\infty} z \cdot \mathbf{P}_{\mathrm{P}}(z) = \frac{\partial G_{\mathrm{P}}}{\partial u}|_{u \to 1} = G'_{\mathrm{P}}(1) \tag{4.6}$$

Then $H_{\mathrm{P}}(u)$ can be written as

$$\begin{aligned} H_{\mathrm{P}}(u) &= \bar{z} \cdot \sum_{z=0}^{\infty} (z+1) \cdot \mathbf{P}_{\mathrm{P}}(z+1) u^z \\ &= \bar{z} \cdot \sum_{z=0}^{\infty} z \mathbf{P}_{\mathrm{P}}(z) u^{z-1} \\ &= \bar{z} \frac{\partial G_{\mathrm{P}}}{\partial u} = \frac{G'_{\mathrm{P}}(u)}{G'_{\mathrm{P}}(1)} \end{aligned} \tag{4.7}$$

Once removing a fraction $1-\phi$ of nodes from network $\mathcal{N}_{\mathrm{P}}$, the remaining fraction ϕ of the network will have different degree distribution [23], with a new argument $1-\phi+\phi u$ [19]. According to the results of a single network, the proportion of giant component $F(\phi, \lambda_{\mathrm{P}})$ of subnetwork ϕ is given by

$$\begin{cases} F(\phi, \lambda_{\mathrm{P}}) = 1 - G_{\mathrm{P}}[1-\phi+\phi \cdot u] \\ u = H_{\mathrm{P}}[1-\phi+\phi \cdot u] \end{cases} \tag{4.8}$$

TABLE 4.2: Notations for math approximation

P'_i, C'_i	the remaining subnetwork in $\mathcal{N}_{\mathrm{P}}$ and $\mathcal{N}_{\mathrm{I}}$ with at least one support inter links in stage i
P_i, C_i	the functioning giant component in P'_i, C'_i
$\widetilde{P'_i}, \widetilde{C'_i}, \widetilde{P_i}, \widetilde{C_i}$	the number of nodes in P'_i, C'_i, P_i, C_i
$\mu'_{p_i}, \mu'_{c_i}, \mu_{p_i}, \mu_{c_i}$	the fractions to P'_i, C'_i, P_i, C_i. i.e., $\widetilde{P'_i} = \mu'_{p_i} \cdot S_{\mathrm{P}}$

The functional forms of $G_{\mathrm{P}}(u)$ and $H_{\mathrm{P}}(u)$ are complicated, deriving the closed form of Eq. (4.8) is still a challenge [15, 21, 22]. However, an approximation expression for $F(\phi, \lambda_{\mathrm{P}})$ was introduced:

Lemma 4.1 ([143]). *For a single scale-free network $\mathcal{N}_{\mathrm{P}}$, for $2 < \lambda < 3$, with some approximation and simplification, it holds that $F(\phi, \lambda_{\mathrm{P}}) \propto \epsilon \cdot \phi^{1/(3-\lambda_{\mathrm{P}})}$, where ϵ is a predefined constant.*

We can analogously define $G_{\mathrm{I}}(u), H_{\mathrm{I}}(u), F(\phi, \lambda_{\mathrm{I}})$ for network $\mathcal{N}_{\mathrm{I}}$ as counterparts to $G_{\mathrm{P}}(u), H_{\mathrm{P}}(u), F(\phi, \lambda_{\mathrm{P}})$ for $\mathcal{N}_{\mathrm{P}}$.

4.3 Math Approximation for Failure Cascading

We analyze the dynamics of cascading failure using percolation theory in this section. The objective of this study is to quantify the system robustness with different k and n, by means of estimating the functioning giant component size in both $\mathcal{N}_{\mathrm{P}}$ and $\mathcal{N}_{\mathrm{I}}$. The notations needed in this section are listed in Table 6.2.

4.3.1 Stage 1: Random Removal in $\mathcal{N}_{\mathrm{I}}$

We begin our estimation with random removal of a fraction $1 - \phi$ of nodes in $\mathcal{N}_{\mathrm{I}}$. The size of remaining network C'_1 is $S_{\mathrm{I}} \cdot \phi$, therefore $\mu'_{c_1} = \phi$. We assume this value is a positive integer if S_{I} is sufficiently large. According to our Condition 1, only the nodes belonging to the giant component can operate properly, next step is to calculate the size of giant component. The size of giant component C_1 is

$$\begin{aligned} \widetilde{C_1} = \widetilde{C_1'} \cdot F(\mu'_{c_1}, \lambda_{\mathrm{I}}) = \mu'_{c_1} \cdot S_{\mathrm{I}} \cdot F(\mu'_{c_1}, \lambda_{\mathrm{I}}) = \mu_{c_1} \cdot S_{\mathrm{I}} \\ \mu_{c_1} = \mu'_{c_1} \cdot F(\mu'_{c_1}, \lambda_{\mathrm{I}}), \end{aligned} \tag{4.9}$$

where μ_{c_1} implies the fraction of C_1 to $\mathcal{N}_{\mathrm{I}}$. We notice only the nodes in C_1 are still in work at the end of this stage.

4.3.2 Stage 2: Fragmentation on $\mathcal{N}_{\mathrm{P}}$

As the network fragmentation from $\mathcal{N}_{\mathrm{I}}$ to C_1, a part of inter links are removed. Thus, some nodes in $\mathcal{N}_{\mathrm{P}}$ would be faulty, because they lost the control inter links. As a result of Stage 1, the fraction $1 - \mu_{c_1}$ of nodes are gone. If the network size is large enough, then the fraction $1 - \mu_{c_1}$ of operation centers are removed since the initial failures are random. Because each operation center has the same control capability n, the probability for each control link to be removed can be approximated by $1 - \mu_{c_1}$. With this respective, a node in $\mathcal{N}_{\mathrm{P}}$ loses all its k control links (totally out of control) with the probability of $(1 - \mu_{c_1})^k$. Denoting P'_2 is the subnetwork belong which the nodes retain at least one control link, we have

$$\begin{aligned} \widetilde{P'_2} = (1 - (1 - \mu_{c_1})^k) \cdot S_{\mathrm{P}} \\ \mu_{p'_2} = 1 - (1 - \mu_{c_1})^k \end{aligned} \tag{4.10}$$

Within P_2', the probability for a node loses $k-i$ of its control links follows the binomial distribution $\mathbf{B}(k, \mu_{c_1})$. The size of giant component in P_2' is

$$\widetilde{P_2} = \widetilde{P_2'} \cdot F(\mu_{p_2}', \lambda_{\mathrm{P}}) = \mu_{p_2}' \cdot S_{\mathrm{P}} \cdot F(\mu_{p_2}', \lambda_{\mathrm{P}}), \tag{4.11}$$

$$\mu_{p2} = \mu_{p_2}' \cdot F(\mu_{p_2}', \lambda_{\mathrm{P}}). \tag{4.12}$$

Each node in P_2 can survive from P_2' with a probability $\widetilde{P_2}/\widetilde{P_2'} = F(\mu_{p_2}', \lambda_{\mathrm{P}})$. As a result, the expected number of nodes with i control inter links in P_2 is given by

$$\widetilde{P_2}|_i = \binom{k}{i} \mu_{c_1}^i (1-\mu_{c_1})^{k-i} \cdot F(\mu_{p_2}', \lambda_{\mathrm{P}}) \cdot S_{\mathrm{P}}, i <= k. \tag{4.13}$$

As a consequence of this fragmentation, the control inter links which operate the nodes belong to P_2' but not to P_2 become *ineffectiveness* even they are still alive. They will have no impact on the cascading failure in the following stages. Thus, we consider these links as faulty and remove them from the system. From P_2' to P_2, the probability for the control link to be removed is approximate to $1-F(\mu_{p_2}', \lambda_{\mathrm{P}})$, then an operation center in C_1 has i control links with the probability of

$$\mathbf{P}_{\mathrm{o}}(i) = \binom{n}{i} F(\mu_{p_2}', \lambda_{\mathrm{P}})^i \cdot (1-F(\mu_{p_2}', \lambda_{\mathrm{P}}))^{n-i}, i <= n. \tag{4.14}$$

4.3.3 Stage 3: Recursive Failure to $\mathcal{N}_{\mathrm{I}}$

The removal of nodes and inter links in Stage 2 affects $\mathcal{N}_{\mathrm{I}}$. The nodes in network $\mathcal{N}_{\mathrm{I}}$ may lose energy inter link thus stops operating. We observe the number of energy inter links subnetwork P_2 provides is approximately $\widetilde{P_2} \cdot \sum_{t=0}^{S_{\mathrm{I}}} P_t \cdot t$, when $\widetilde{P_2}$ is large enough (law of large numbers is satisfied). Initially the total number of energy inter links is S_{I} since each node in S_{I} depends on only one power stations.

With this respective, the probability for one energy inter link removal is

$$1 - \frac{\widetilde{P_2} \cdot \sum_{t=0}^{S_{\mathrm{I}}} P_t \cdot t}{S_{\mathrm{I}}} = 1 - \mu_{p2}.$$

The number of nodes in C_1 that has an energy inter link from P_2 is given by

$$\widetilde{C_3'} = \widetilde{C_1} - \widetilde{C_1} \cdot (1 - \mu_{p_2}) = \mu_{p_2} \cdot \mu_{c_1} \cdot S_{\mathrm{I}}.$$

That is, passing from C_1 to C_3', a fraction $1 - \mu_{p_2}$ of nodes are broken. As we did in the previous stages, the next step is to calculate the size of giant component of C_3'. As mentioned in [15], it is indeed not an easy task. Instead, we consider the joint effect of the node removal in Stage 1 and Stage 3 are equivalent, i.e., the effect of removing the fraction of $1 - \mu_{p_2}$ of nodes in C_1 has the same effect as taking out the same fraction size from C_1' in terms of calculating the giant component size of C_3'. We find that the fragmentation from $\mathcal{N}_{\mathrm{I}}$ to C_3' can be modeled by removing node of a fraction of

$$1 - \phi + \phi \cdot (1 - \mu_{p_2}) = 1 - \phi \cdot \mu_{p_2}$$

in Stage 1. Thus, the equivalent $\mu_{c_3}' = \phi \cdot \mu_{p_2}$, by which the size of giant component C_3 in subnetwork C_3' is

$$\widetilde{C_3} = \mu_{c_3}' \cdot S_{\mathrm{I}} \cdot F(\mu_{c_3}', \lambda_{\mathrm{I}}) = \mu_{c_3} \cdot S_{\mathrm{I}}. \tag{4.15}$$

4.3.4 Stage 4: Further Fragmentation in $\mathcal{N}_{\mathrm{P}}$

As network $\mathcal{N}_{\mathrm{I}}$ splits to C_3 in previous stage, more control inter links are removed and so forth more nodes in $\mathcal{N}_{\mathrm{P}}$ is going to be faulty. Notice the probability each operation center survives from C_1 to C_3 is $\frac{\mu_{c3}}{\mu_{c_1}}$. The number of control inter links the network C_1 has is approximated to $\sum_{i=1}^{n} i \cdot \mathbf{P}_{\mathrm{o}}(i) \cdot \mu_{c_1} \cdot m$. Since each operation center survives with same probability, the number of existing control links in C_3 can be given by $\sum_{i=1}^{n} i \cdot \mathbf{P}_{\mathrm{o}}(i) \cdot \mu_{c_3} \cdot m$. Thus, the probability of a control inter link to be removed is $1 - \frac{\mu_{c_3}}{\mu_{c_1}}$.

Consequently, the probability for a node in P_2 with i control inter links stopping function is $(1 - \frac{\mu_{c_3}}{\mu_{c_1}})^i$. Combining with Eq. (4.13), the number of nodes that will be removed in P_2 is

$$
\begin{aligned}
R &= \sum_{i=1}^{k} \widetilde{P_2}|_i \cdot (1 - \frac{\mu_{c_3}}{\mu_{c_1}})^i \\
&= \sum_{i=1}^{k} \binom{k}{i} \mu_{c_1}^i (1-\mu_{c_1})^{k-i} \cdot F(\mu'_{p_2}, \lambda_{\mathrm{P}}) \cdot S_{\mathrm{P}} \cdot (1 - \frac{\mu_{c_3}}{\mu_{c_1}})^i \\
&= \sum_{i=1}^{k} S_{\mathrm{P}} \cdot F(\mu'_{p_2}, \lambda_{\mathrm{P}}) \cdot \binom{k}{i} \cdot (1-\mu_{c_1})^{k-i} \cdot (\mu_{c_1} - \mu_{c_3})^i \\
&= S_{\mathrm{P}} \cdot F(\mu'_{p_2}, \lambda_{\mathrm{P}}) \cdot ((1-\mu_{c_3})^k - (1-\mu_{c_1})^k)
\end{aligned}
\tag{4.16}
$$

So, the size of P'_4 is

$$
\begin{aligned}
\widetilde{P'_4} &= \widetilde{P_2} - R \\
&= S_{\mathrm{P}} \cdot F(\mu'_{p_2}, \lambda_{\mathrm{P}}) \cdot (1 - (1-\mu_{c_3})^k)
\end{aligned}
\tag{4.17}
$$

Passing from P_2 to P'_4, the fraction $1 - \widetilde{P'_4}/\widetilde{P_2} = 1 - (1 - (1-\mu_{c_3})^k / \mu'_{p_2}$ of nodes are removed. As we do in Stage 3, in terms of the size of giant component in P'_4, that is equivalent to remove the same fraction of nodes from P'_2. The proportion of nodes that has to be removed from S_{P} to P'_4 is

$$
1 - \mu'_{p_2} + \mu'_{p_2} \cdot (1 - \frac{1 - (1-\mu_{c_3})^k}{\mu'_{p_2}}) = (1-\mu_{c_3})^k
$$

Thus, the equivalent $\mu'_{p_4} = 1 - (1-\mu_{c_3})^k$. The corresponding fraction of giant component $\mu_{p_4} = \mu'_{p_4} \cdot F(\mu'_{p_4}, \lambda_{\mathrm{P}})$.

4.3.5 Transcendental Equations for Failure Cascading

Following the previous steps, we can obtain the size of giant component $\widetilde{C_1}, \widetilde{P_2}, \widetilde{C_3} \cdots$ in a certain stage, but no one knows in which stage the cascading failure stops. Our main aim is to estimate the size of functioning parts in the final stage. When repeating the above calculations, we can observe the pattern of equations:

$$\begin{gathered}
\mu'_{c_1} = \phi, \;\; \mu_{c_1} = \mu'_{c_1} \cdot F(\mu'_{c_1}, \lambda_{\mathrm{I}}) \\
\mu'_{c_3} = \phi \cdot \mu_{p_2}, \;\; \mu_{c_3} = \mu'_{c_3} \cdot F(\mu'_{c_3}, \lambda_{\mathrm{I}}) \\
\cdots, \;\; \cdots \\
\mu'_{c_{2j+1}} = \phi \cdot \mu_{p_{2j}}, \;\; \mu_{c_{2j+1}} = \mu'_{c_{2j+1}} \cdot F(\mu'_{c_{2j+1}}, \lambda_{\mathrm{I}})
\end{gathered}$$

and

$$\begin{gathered}
\mu'_{p_2} = 1 - (1 - \mu_{c_1})^k, \;\; \mu_{p_2} = \mu'_{p_2} \cdot F(\mu'_{p_2}, \lambda_{\mathrm{P}}) \\
\mu'_{p_4} = 1 - (1 - \mu_{c_3})^k, \;\; \mu_{p_4} = \mu'_{p_4} \cdot F(\mu'_{p_4}, \lambda_{\mathrm{P}}) \\
\cdots, \;\; \cdots \\
\mu'_{p_{2j}} = 1 - (1 - \mu_{c_{2j-1}})^k, \;\; \mu_{p_{2j}} = \mu'_{p_{2j}} \cdot F(\mu'_{p_{2j}}, \lambda_{\mathrm{P}})
\end{gathered}$$

To determine the state of the system in the end of cascading failure, we look at $j \to \infty$. The networks stop fragmenting and the functioning giant components are fixed. Thus, the following equations hold

$$\begin{gathered}
\mu'_{c_{2j+1}} = \mu'_{c_{2j+3}} = \mu'_{c_{2j-1}} \\
\mu'_{p_{2j}} = \mu'_{p_{2j+2}} = \mu'_{p_{2j-2}}
\end{gathered}$$

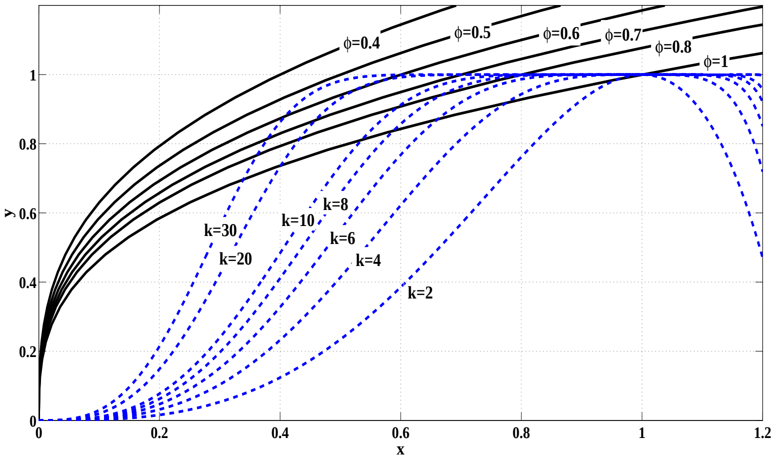


FIGURE 4.3: Solutions for Eq. (4.20) for $\lambda_{\mathrm{P}} = \lambda_{\mathrm{I}} = 2.5$, $\epsilon_1 = \epsilon_2 = 1$. The solutions are the corner points of two lines. It is clearly shown for some cases there is no intersection ($\phi = 0.8, k = 4$). For some cases one intersection exists ($\phi = 0.6, k = 20$). For the others, two non-trivial solution exist.

Let $x = \mu'_{c_{2j+1}} = \mu'_{c_{2j+3}} = \mu'_{c_{2j-1}}$ and $y = \mu'_{p_{2j}} = \mu'_{p_{2j+2}} = \mu'_{p_{2j-2}}$, then we obtain a set of transcendental equations

$$\begin{cases} x = \phi \cdot y \cdot F(y, \lambda_{\mathrm{P}}) \\ y = 1 - (1 - x \cdot F(x, \lambda_{\mathrm{I}}))^k \end{cases} \tag{4.18}$$

where $F(\cdot, \lambda_{\mathrm{P}}), F(\cdot, \lambda_{\mathrm{I}})$ are according to Eq. (4.8).

The fraction of nodes that still function in the final steady state in both networks can be calculated by

$$\begin{cases} \lim_{j\to\infty} \mu_{p_j} = \mu_{p_\infty} = y \cdot F(y, \lambda_{\mathrm{P}}) \\ \lim_{j\to\infty} \mu_{c_j} = \mu_{c_\infty} = x \cdot F(x, \lambda_{\mathrm{I}}) \end{cases} \tag{4.19}$$

This analysis can be applied to any type of networks. This gives us a complete solution for the fraction of survivals in $\mathcal{N}_{\mathrm{P}}$ and $\mathcal{N}_{\mathrm{I}}$. If we can find a non-trivial solution for x and y, then we can compute the remaining number of survivals.

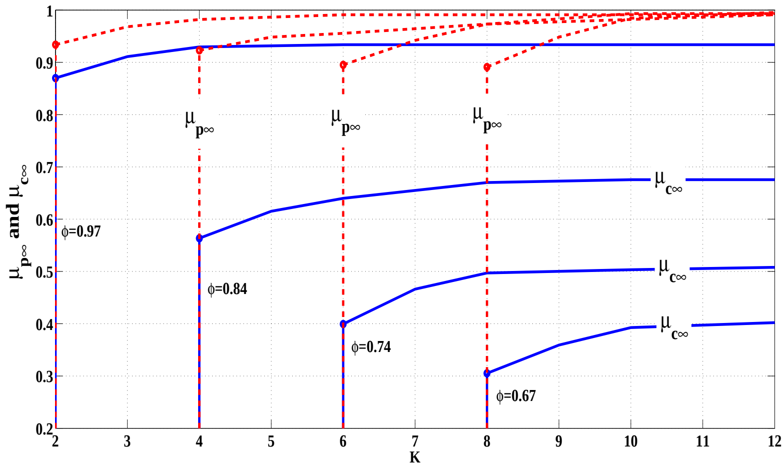


FIGURE 4.4: A solution for μ_{p_∞} and μ_{c_∞}, given $\lambda_{\mathrm{P}} = \lambda_{\mathrm{I}} = 2.2$, $\epsilon_1 = \epsilon_2 = 1$. For $k = 2, 4, 6, 8$, the threshold $\phi_c = 0.97, 0.84, 0.74, 0.67$ respectively. The first-order discontinues transition occurs at each ϕ_c.

4.3.6 Graphic Solution

According to Lemma (4.1), let $F(y, \lambda_{\mathrm{P}}) = \epsilon_1 \cdot y^{1/(3-\lambda_{\mathrm{P}})}$ and $F(x, \lambda_{\mathrm{I}}) = \epsilon_2 \cdot x^{1/(3-\lambda_{\mathrm{I}})}$, where ϵ_1 and ϵ_2 are predefined constants. Eq. (6.4) comes to

$$\begin{cases} x = \phi \cdot y \cdot \epsilon_1 \cdot y^{1/(3-\lambda_{\mathrm{P}})} \\ y = 1 - (1 - x \cdot \epsilon_2 \cdot x^{1/(3-\lambda_{\mathrm{I}})})^k \end{cases} \tag{4.20}$$

In general, it is difficult to derive an expression for x, y depending on k, ϕ. Instead, we can solve Eq. (4.20) with graphic method [15] for a given set of $\lambda_{\mathrm{P}}, \lambda_{\mathrm{I}}, \epsilon_1, \epsilon_2$.

We draw two function curves of Eq. (4.20) and the intersections are the solutions of x, y. Fig. 4.3 and Fig. 4.4 give complete graphic solutions, based on which we obtain some insightful findings as following:

• A critical threshold ϕ_c exists, beyond which the system collapses. For the case of $k = 20$, we find there is no intersection if $\phi < 0.6$, i.e., no solution exists. While for $\phi > 0.6$, the two curves have two intersection points. So ϕ_c is approximate to 0.6 for the case of $k = 20$. For $\phi < \phi_c$, both networks go into complete fragmentation in the end. If $\phi > \phi_c$, the two non-trivial intersections are corresponding to two sets of giant component sizes. In this case, the solution is the point that is closer to the initial status, because the system fragmentation stops at this point and never goes to the small one.

- The ϕ_c varies for different values of k. For instance, it is about 0.73 for $k = 10$ and 0.83 for $k = 6$ ($\lambda_{\mathrm{P}} = \lambda_{\mathrm{I}} = 2.5$). We find ϕ_c becomes lower for a higher k, which matches with the intuition that the smart power grid is more reliable if each power station has more control inter links.

- The survival ratios of both networks experience *first-order transition* at ϕ_c. Fig. (4.4) gives an example solution for μ_{p_∞} and μ_{c_∞}. For the case of $k = 4$, ϕ_c equals 0.84, there is no giant component existing for $k < 4$. While for $k = 4$, both μ_{p_∞} and μ_{c_∞} have a step function.

- The improvements of μ_{p_∞} and μ_{c_∞} are sublinear with the increasing of k, and reach upper bounds. Note that for all the cases in Fig. (4.4), μ_{p_∞} approaches to 1. This gives us a meaningful guide that adding control inter links improves system robustness significantly when k is small. While when k is sufficiently large, increasing k does not affect the robustness except ϕ_c.

However, finding the non-trivial solution for scale-free network is challenging since the exactly value of ϵ is still unknown, [15, 21, 106, 143]. To the best of our knowledge, deriving the closed-form solution of Eq. (4.8) is still open. Therefore, in this work, we resort to the standard graphical method for simulations.

4.4 Experimental Validation

In this section, we generate an interdependent network and simulate cascading failure to obtain the fractions of functioning parts (survival ratio) μ_{p_∞} and μ_{c_∞}. To generate the scale-free network with different power-law distributions, we adopt

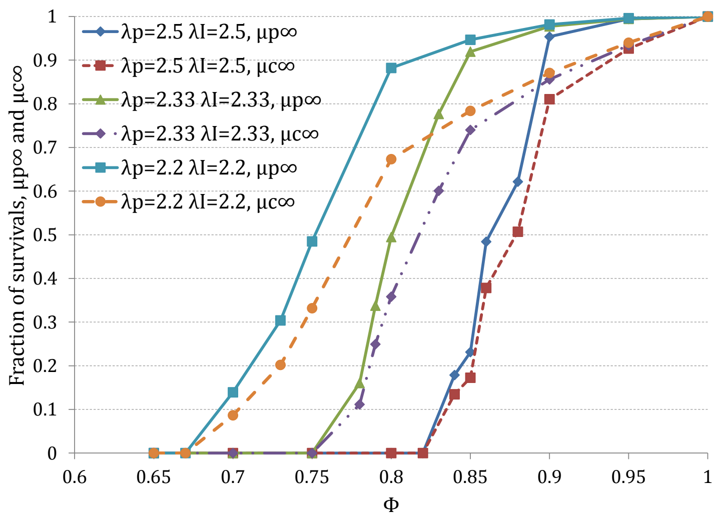


FIGURE 4.5: The fraction of survivals in both networks with different λ_{P} and λ_{I}. The initial networks size $S_{\mathrm{P}} = 1000, S_{\mathrm{I}} = 10000$, k=2, n=5. The generalized Barabási-Albert model is used to construct the two scale-free networks. The initial random failure $1 - \phi$ is occurs in network $\mathcal{N}_{\mathrm{I}}$.

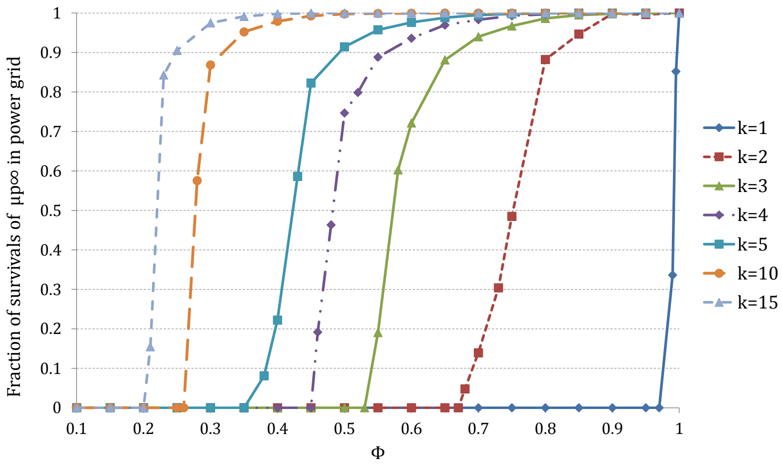
the *generalized Barabási-Albert model* [21], whose power law exponent can vary in the range $(2, +\infty)$. In all the experiments the same approach is employed: first, we construct two scale-free networks representing $\mathcal{N}_{\mathrm{P}}$ and $\mathcal{N}_{\mathrm{I}}$; then we remove the fraction of $1 - \phi$ nodes in $\mathcal{N}_{\mathrm{I}}$ as random failure.

4.4.1 ϕ_c and System Robustness

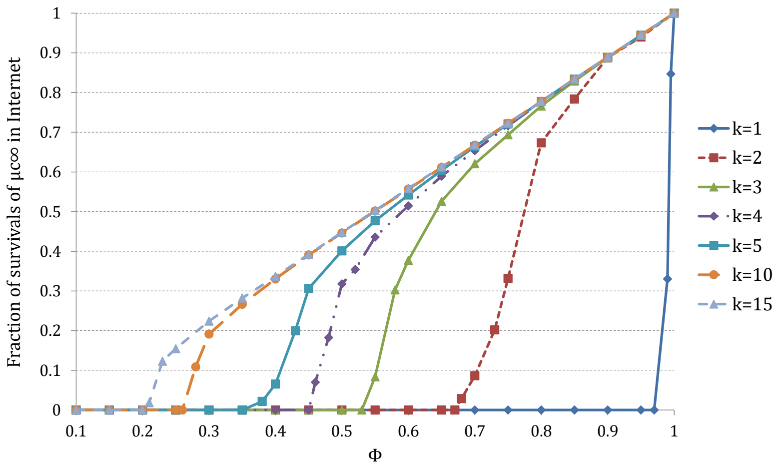
We firstly discuss the values of μ_{p_∞} and μ_{c_∞} with different exponents of λ_{P} and λ_{I}.

Fig. 4.5 clearly shows with the decreasing of λ_{P} and λ_{I}, which means that the more nodes with high degree, the better robustness of entire system. Only 50% of nodes survives for $\lambda_{\mathrm{P}} = \lambda_{\mathrm{I}} = 2.33$ when $\phi = 0.5$, but it is approximate 0.9 for $\lambda_{\mathrm{P}} = \lambda_{\mathrm{I}} = 2.2$. ϕ_c is also highly influenced by λ. It is 0.67 for $\lambda_{\mathrm{P}} = \lambda_{\mathrm{I}} = 2.2$, while this increases to 0.82 for $\lambda_{\mathrm{P}} = \lambda_{\mathrm{I}} = 2.5$.

We notice when the initial failure or attack upon Internet is small, i.e., less than 5%, the entire power grid is extremely reliable for all three systems. The curves of



(a) μ_{p_∞} varies with different value of k



(b) μ_{c_∞} varies with different value of k

FIGURE 4.6: System Robustness vs. k. The initial networks size $S_{\mathrm{P}} = 1000, S_{\mathrm{I}} = 10000$, $\lambda_{\mathrm{P}} = \lambda_{\mathrm{I}} = 2.2$, $n = 5$.

μ_{p_∞} remain stable on the right of corner points (0.8 for $\lambda_{\mathrm{P}} = \lambda_{\mathrm{I}} = 2.2$), and drop rapidly on the left. One more observation is that μ_{p_∞} and μ_{c_∞} are either zero or nonzero simultaneously which indicates that the power grid and communication network are either totally broken or not.

4.4.2 Cost and Robustness

Our main aim is to find the relation between system cost and system robustness. Fig. 4.6(a) and 4.6(b) give detailed curves on various values of k.

It is shown that as the increasing of k, $\mu_{p\infty}$ and $\mu_{c\infty}$ increase. This implies a practical meaning that if we let more operation centers control each power station, the system would has a much higher robustness. In the case of $k = 1$, the system could not bear even 2% failure, the threshold $\phi_c = 0.98$. If we add one more control link to each node, i.e., $k = 2$, the system robustness experiences a huge improvement. Not only the ϕ_c dramatically decreases to 0.67, but also the system can tolerate small scale of failure: there is no impact to power grid even 10% of communication nodes are initially faulty.

In the case of $k = 15$, the power grid remains totally function, even if 60% of communication network is destroyed. The ϕ_c equals 0.2, approximate five times promotion compared to 0.98 for the case $k = 1$. Noticing in Fig. 4.6(b), the decrease of $\mu_{c\infty}$ is quite flat for $k = 15$. The loss of Internet is almost all due to the initial failure.

The important finding according to Fig. 4.6(a) and 4.6(b) is that the relation between robustness and cost is sublinear. The improvement between $k = 2$ and $k = 1$ is significant, while the gap between $k = 15$ and $k = 10$ is tiny. Hence, for building a smart power grid, adding as many as possible control link is not our choice since the massive extra cost does not improve system reliability remarkably.

The proposed *first-order discontinuous transition* [15] in interdependent networks says the size of giant component meets a sharp transition when ϕ is approaching to ϕ_c. Since the network size in our simulation is non-infinite, there is a small transition interval including the value of ϕ_c. Fig. 4.7 shows how our model transits around ϕ_c for different k.

In case of $k = 5$, the giant components always exist in both $\mu_{p\infty}$ and $\mu_{c\infty}$ for $\phi > 0.42$, while the system collapses when $\phi < 0.36$. Thus, we say the transition interval is $[0.36, 0.42]$, during which the system might either collapses or not. Note that for larger k, the transition interval is smaller, i.e., the transition is more sharp.

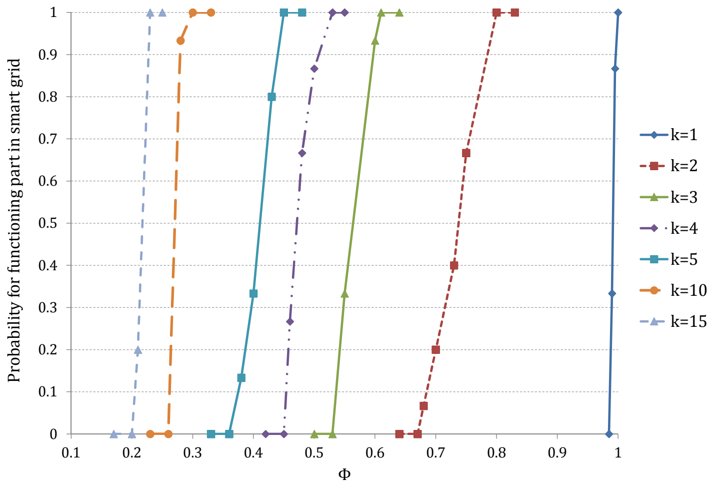


FIGURE 4.7: The different probabilities varies with k for the entire system to have a functioning part after cascading failure stops. $\lambda_{\mathrm{P}} = \lambda_{\mathrm{I}} = 2.2$, $n = 5$. For the case of $k = 3$, the entire smart grid system collapses if $\phi < 0.53$; for $\phi > 0.6$, giant components always exist in both μ_{p_∞} and μ_{c_∞}. During the interval, there is a probability for giant component existing.

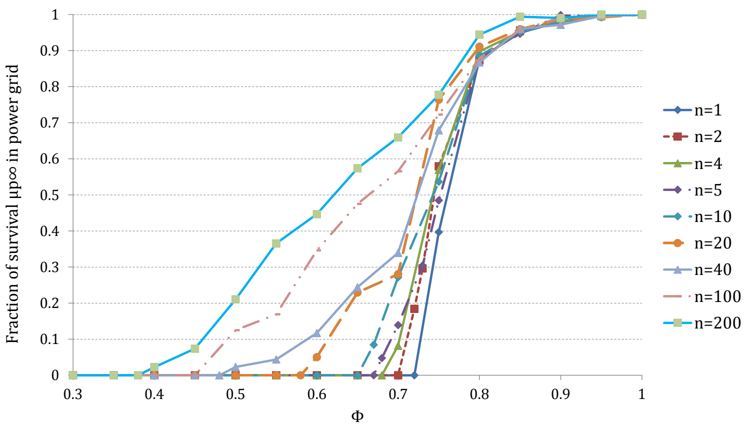
Our finding validates the conclusion proposed by [148], that is, when the coupling between the networks is reduced, at a critical coupling strength the transition becomes a second-order continuous phase transition.

We believe a sharp transition is preferred for the realistic smart grid, since the smaller transition interval makes the system easy to be predicted and controlled.

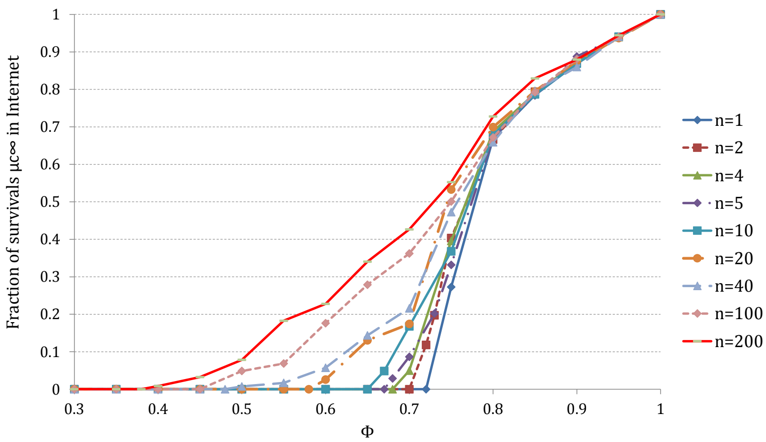
4.4.3 Various Allocation Strategies with System Robustness

We now discuss the impact of different operation center capability n. In this simulation series, the value of k is set to 2. We change the value of n to explore the different system robustness. Fig. 4.8(a) and Fig. 4.8(b) reveal the relations between robustness and n.

Generally, the system performance improves with the increasing n. As shown in Fig. 4.8(a), in the case of $n = 2$, μ_{p_∞} is equal to 0.95 when the initial failure is



(a) μ_{p_∞} varies with different value of n



(b) μ_{c_∞} varies with different value of n

FIGURE 4.8: System Robustness vs. n. The initial networks size $S_{\mathrm{P}} = 1000, S_{\mathrm{I}} = 10000$, $\lambda_{\mathrm{P}} = \lambda_{\mathrm{I}} = 2.2$, $k = 2$.

$1 - \phi = 0.15$. While it almost equals 1 for the case of $n = 200$ under the same situation. We observe promoting n decreases the threshold value ϕ_c in all cases.

We notice that for all the cases of n, on the right side of the corner point $\phi = 0.8$, both μ_{p_∞} and μ_{c_∞} approach to 1 steadily. While on the left side, the curves drop to zero more rapidly. Note that for higher n, the slope is smaller. Comparing Fig. 4.6(a) and Fig 4.8(a), it is clearly shown that the corner point position is determined by k rather than n. A distinct k completely has a different corner point, while there is only a sightly change for different n with a specific k.

Now we consider the transition intervals of our simulation. Fig. 4.9 gives the

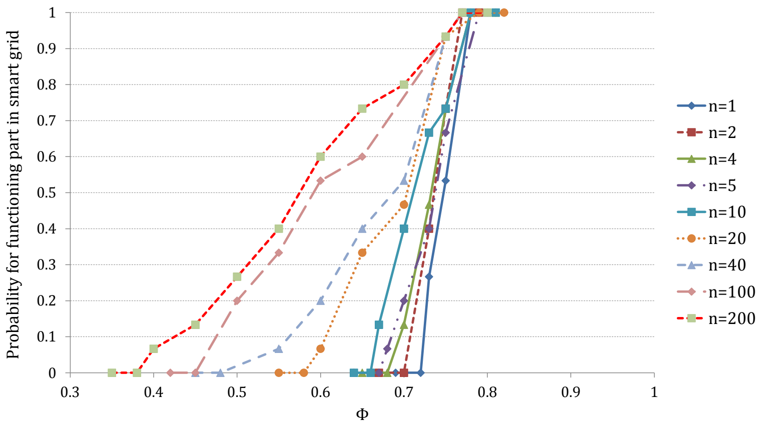


FIGURE 4.9: Transition intervals for various allocation strategies. $S_{\text{P}} = 1000, S_{\text{I}} = 10000,\ \lambda_{\text{P}} = \lambda_{\text{I}} = 2.2,\ k = 2.$

detailed curves for different values of n. When n is equal to 2, the interval is [0.71, 0.77]. It extends dramatically for the case of $n = 200$, with the approximate value between [0.37, 0.77]. Hence, increasing n extends the transition internal, i.e., the transition is much more flat.

4.5 Validation of Theoretical Analysis

Our extensive simulations validate the mathematical analysis in Section 4.3.6. Some conclusions are listed as follows:

• The system robustness against random failure can be improved by increasing controlling cost, i.e., adding more control links for each target (power station). The robustness improvement is nonlinear with cost promotion, which has been demonstrated by our figures. Both mathematical and experimental results show that the power grid is almost intact against random failure when k is large enough.

• The critical threshold ϕ_c is inverse proportional to the value of k. The transition at ϕ_c is first-order in mathematical analysis, while it is more flat in simulation. The reason is that the network size is non-infinite. Increasing controlling cost k shortens the transition interval in smart grid.

• The monitoring capability n of each operation center has no impact on system robustness according to the transcendental Eq. (6.4). While our simulations give a little bit different story. The reason is that the network size for simulation is finite, while our math calculation is based on the assumption that the number of nodes is large enough so that *Law of Large Numbers* is satisfied.

• Based on our analysis, we would conclude that it is important to find a trade-off between expenditure and performance for building smart grid infrastructures. While if the controlling cost is fixed and we still want to promote the system robustness. Then one possible way is: increasing monitoring capability n. Consequently, the total number of operation center required decreases. By this way, the reliability experiences a slight improvement. But the disadvantage of this method is obvious: the transition interval is extended.

4.6 Application of Proposed Model

We use smart grid to explain our proposed 'k-n' model. However, this model is general and could be applied in many cyber physical cases.

- Intelligent Vehicular Ad Hoc Networks. The mutually dependent two networks are vehicular network and Roadside Unit network (RSU). The traffic congestion and road condition censored by the sensors attached on vehicles are passed to data centers through roadside units. On the other side, the traffic control instructions disseminate to each vehicle through the vehicular network. Each data center and roadside unit is in charge of a specific area as well as the cars within this area. Each vehicle could get message from multiple roadside units nearby through vehicular network. The interdependence between these two networks is the correctness of information/message. If any of vehicle, data center or roadside unit got failed, the data/message is going to be incorrect, and thus, affect the function of the entire system. However, the devices may still keep operating, but remains in disorder.

- Actuator-sensor networks. Sensor monitors the area and gathers the information, which is manipulated and conducted by the actuators. On the other hand, robots cooperate to repair, replace battery or place new sensors if the old one got failed [149]. The cooperation and job assignment between multiple robots are achieved by the communication through the sensor network [150]. Each robot is in charge of an area and each sensor could be recovered by anyone among robot team. Similar to vehicular ad hoc networks, the interdependence here is the correctness of information. Any incorrect message/data could lead to inefficient cooperation between robots, and thus possibly cause the unsuccessful sensor recovery.

As we see in these two applications, the interdependence between networks could be any forms.

Chapter 5

Case Study: Smart Grid

In this chapter, we integrate the interdependence model with load propagation model to understand how cascading failure propagates in smart power grid. We limit our analysis into the scope on network theory. The electricity features, such as power transmission flow limit and voltage sag, are not considered in this work.

5.1 Load Propagation Model

We denote the single power grid by $G_p = (V_p, E_p)$, where V_p is the set of vertices and E_p is the set of edges. According to the previous studies [39, 125, 126], G_p is assumed to follow the scale-free network characteristics. Specifically, we denote the probability distribution of node degree by $\mathbf{P}_{\mathrm{d}}(z)$, and assume that $\mathbf{P}_{\mathrm{d}}(z) \propto z^{-\lambda}$. The electricity is transported from nodes to nodes through the power lines. When all nodes are on, the entire network operates properly. But, the breakdown of few nodes may cause the load dynamic redistribution, and consequently leads to the breakdown of the remaining nodes. To study this cascading process, we introduce a load propagation model, as follows.

For a node i in scale-free network, assuming its initial workload is the function of its degree z, we have

$$L_i(z) = \alpha \cdot z^{\eta}, \tag{5.1}$$

where α and η are the pre-defined constants. It is reasonable to assume that each node has the limit of load, called *capacity*, which is proportional to the initial load

$$C_i(z) = T \cdot L_i(z), T \geq 1 \tag{5.2}$$

where constant T is the *Tolerance parameter* [35, 125] that controls the tolerance of the power grid.

The breakdown or removal of a single node i would transfer its initial load to the remaining network. To simplify our study, these load is *uniformly* distributed to the neighbors. Therefore, the load of neighbor j of removed node i equals to the sum of its initial load plus the incremental load transferred from the removed node. The work load for node j is

$$L_j(z) = L_j(z) + \frac{L_i(z)}{z}. \tag{5.3}$$

Node j would be overloaded and out of work when $L_j(z) > C_j(z)$. Then this load redistribution process might continue to the neighbors of j. The overload cascading stops if one of the following situations is satisfied,

- The workload is less than the capacity for all of the remaining nodes,
- The entire nodes in V_p are removed.

5.2 Interdependent Networks

We denote the communication network by $G_c = (V_c, E_c)$. In smart power grid, G_p and G_c are mutually dependent and connected. Several interdependence models have been studied in previous work. In this article, to study how the interdependence affect the network robustness, we apply our 'k-n' model [3], where we briefly introduce as follows:

- The nodes in G_c are roughly divided into two roles: *information relay node* and *control node*. The control node, or called *operation center*, is responsible for monitoring and operating the nodes in G_p. While the control messages are transmitted through the information relay nodes. On the other hand, nodes in V_p provide electricity to the nodes in G_c. To improve the system robustness, we let each node in V_p be under the control of k control nodes in V_c, which at the same time, could totally monitor n nodes in V_p.
- *Control-Dependency* links (CD Link) and *Energy-Dependency* links (ED link) consist of the inter links. Each control node has n Control-Dependency links to operate n nodes in G_p. While each node of G_p is under the control of k operation centers, thus has k Control-Dependency links. A *Balls and Bins* allocation procedure is applied to assign the Energy-Dependency links so that each node in V_c could obtain the energy from G_p.

Fig. 5.1 gives a sketch of 'k-n' interdependence model.

5.2.1 Modified Balls and Bins Allocation

After taking the initial workload and capacity into account, we believe there is a limit on the number of V_c nodes that a power node can support. Thus, the

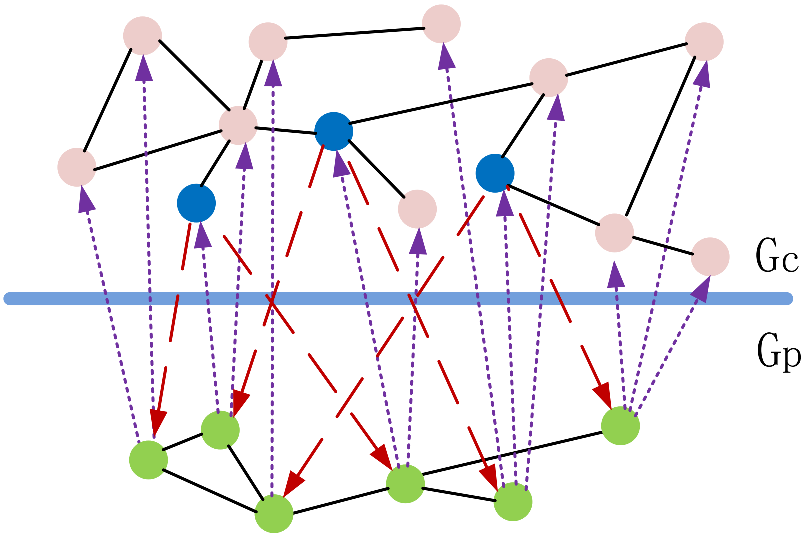


FIGURE 5.1: Each power node provides electricity to multiple communication nodes. 3 communication nodes are chosen as the operation centers (deep blue), where each of them operates 2 power nodes, thus $n = 2$. Each power station is controlled by 1 operation center, thus $k = 1$. ED link is from G_p to G_c, CD link is opposite.

previous *Balls and Bins* allocation is no longer accurate, since it is based on the assumption that all power nodes are equal and have unlimited capacities.

For a power node i, we assume that the maximum number of communication nodes it can support is proportional to its initial load, thus we have

$$N_i(z) = \beta \cdot \alpha \cdot z^{\eta}. \tag{5.4}$$

The mean maximum number is given by

$$\langle N_i(z) \rangle = \sum_{z=0}^{\infty} \mathbf{P}_{\mathrm{d}}(z) \cdot N_i(z) = \beta \cdot \alpha \sum_{z=0}^{\infty} \mathbf{P}_{\mathrm{d}}(z) \cdot z^{\eta}. \tag{5.5}$$

We consider the nodes in G_p are bins and the nodes in G_c are balls. Assuming the sizes of G_p and G_c are S_p and S_c respectively, the total positions a ball can choose is $S_p \cdot \langle N_i(z) \rangle$. Therefore, the probability for a ball to be allocated into the bins with initial load of $\alpha \cdot z^{\eta}$ equals to

$$p = \frac{\mathbf{P}_{\mathrm{d}}(z) \cdot z}{S_p \cdot \langle N_i(z) \rangle}. \tag{5.6}$$

Define the number of balls in a bin with initial load of $\alpha \cdot z^{\eta}$ as a random variable ξ_z. Therefore, it holds that

$$\mathbf{P}(\xi_z = t) = \binom{z}{t}\binom{S_c}{t} \cdot p^t \cdot (1-p)^{S_c - t}, t \leq z \tag{5.7}$$

Furthermore, define the number of balls in a bin as a random variable ξ. It follows that

$$\mathbf{P}(\xi = t) = \sum_{z=0}^{\infty} \mathbf{P}(\xi_z = t) \cdot \mathbf{P}_{\mathrm{d}}(z). \tag{5.8}$$

5.3 Mathematical Analysis

In this section, we first study the load propagation in single power grid G_p. We apply a random failure of fraction of $1-\phi$ on G_p. Then we estimate the robustness of interdependent smart power grid by calculating the size of functioning giant component, and discuss how the 'k-n' interdependence model and load propagation affect the system. The definitions and notations are listed in Table 5.1.

5.3.1 Calculation of Giant Component

We introduce a notation $F(\phi, G_p)$ [3] to denote the expected fraction of giant component in the subnetwork which occupies the proportion ϕ of nodes in network G_p. Accordingly, $F(\phi, G_c)$ is for network G_c. $F(\phi, G_p)$ is estimated by [21]

$$\begin{cases} F(\phi, G_p) = 1 - \sum_0^{\infty} \mathbf{P}_{\mathrm{d}}(z) \cdot u^z, \\ u = \sum_0^{\infty} \mathbf{Q}(z) \cdot (1 - \phi + \phi \cdot u^z), \end{cases} \tag{5.9}$$

TABLE 5.1: Notations for the analysis

k	number of operation centers that control a specific power station
n	number of power stations each operation center operates
T	tolerance parameter
u	the probability a node can not reach to the giant component through a specific neighbor
$1-\phi_c$	the initial failure threshold above which the entire system collapse [3, 15]
$\lambda_{\mathrm{P}}, \lambda_{\mathrm{c}}$	power law exponents of networks G_p and G_c.
$F(\phi, G_p), F(\phi, G_c)$	the fraction of giant component in the subnetwork which occupies the fraction ϕ of the nodes in the entire network G_p and G_c

where $\mathbf{Q}(z)$ is the number of edges connected with the node other than the edge we arrived with, called *excess degree distribution* [3, 21], and given by

$$\mathbf{Q}(z) = \frac{(z+1)\cdot \mathbf{P}_{\mathrm{d}}(z+1)}{\langle z \rangle}, \tag{5.10}$$

where $\langle z \rangle = \sum_0^\infty \mathbf{P}_{\mathrm{d}}(z)\cdot z$.

5.3.2 Threshold T_c in Single Power grid G_p

Once $1-\phi$ of G_p is removed, the proportion of remaining functioning network is approximated by $\phi \cdot F(\phi, G_p)$ [3]. After the removal, the load of removed nodes is distributed to the remaining part of the network. The total load that need to be shifted R is the summation of initial load of the removed nodes, which is given by $\sum_{i=1}^{S_p(1-\phi\cdot F(\phi,G_p))} L_i(z)$. As far as $S_p \to \infty$, we simplify it to

$$R = S_p \cdot (1-\phi\cdot F(\phi, G_p)) \cdot \sum_{z=0}^{\infty} \mathbf{P}_{\mathrm{d}}(z)\cdot \alpha \cdot z^{\eta}. \tag{5.11}$$

To avoid the load cascading failure, the shifted load R can not exceed the remaining volume of the network. Therefore,

$$R \leq \sum_{i=S_p(1-\phi \cdot F(\phi,G_p))}^{S_p} (C_i(z) - L_i(z))$$

should be satisfied. The threshold T_c is obtained when both sides are exactly equal, that is

$$R = S_p \cdot \phi \cdot F(\phi, G_p) \cdot (T_c - 1) \cdot \sum_{0}^{\infty} \mathbf{P}_{\mathrm{d}}(z) \cdot \alpha \cdot z^{\eta}. \tag{5.12}$$

Applying Eq. (5.11), we have

$$\begin{aligned} 1 - \phi \cdot F(\phi, G_p) &= \phi \cdot F(\phi, G_p) \cdot (T_c - 1), \\ 1 &= T_c \cdot \phi \cdot F(\phi, G_p), \\ T_c &= \frac{1}{\phi \cdot F(\phi, G_p)}. \end{aligned} \tag{5.13}$$

Smaller T would more likely cause load cascading failures in the entire network.

5.3.3 Failure Due to Load Redistribution

For an arbitrary node i, we define two random variables and an event as follows.

- Random Variable Z_i: the number of node i's neighbors.
- Random Variable X_i: the number of neighbors that could destroy node i by shifting their load.
- Event $\mathbf{O}_i$: node i is destroyed due to the overload shifted from its neighbors.

From the homogeneity of the distribution of node degrees, we denote the sequences of Z_i, X_i, and $\mathbf{O}_i$ by Z, X, and $\mathbf{O}$, without the loss of generality.

For a node, to compute $\mathbf{P}(\mathbf{O}|Z=z)$, we need to know $\mathbf{P}(\mathbf{O}|X=x)$, which is the probability that the shifted load from x neighbors is greater than the remaining volume of the node. Thus, applying Eqs. (5.1) and (5.3), we have

$$\mathbf{P}(\mathbf{O}|X=x) = \mathbf{P}\left\{x \cdot \sum_{z=1}^{\infty} \frac{\alpha \cdot z^{\eta}}{z} + \alpha \cdot z^{\eta} > T \cdot \alpha \cdot z^{\eta}\right\} \tag{5.14}$$

As x, z, α, η and T are constants here, the distribution of $\mathbf{P}(\mathbf{O}|X=x)$ is

$$\mathbf{P}(\mathbf{O}|X=x) = \begin{cases} 1, x > \frac{(T-1)\cdot\alpha\cdot z^{\eta}}{\sum_0^{\infty} \frac{\alpha\cdot z^{\eta}}{z}}, \\ 0, \text{Otherwise.} \end{cases} \tag{5.15}$$

The probability to choose x neighbors among total z neighbors $\mathbf{P}(X=x|Z=z)$ is given by

$$\mathbf{P}(X=x|Z=z) = \binom{z}{x} / 2^z \; .$$

Therefore, $\mathbf{P}(\mathbf{O}|Z=z)$ could be calculated by

$$\begin{aligned} \mathbf{P}(\mathbf{O}|Z=z) &= \sum_{x=0}^{z} \mathbf{P}(\mathbf{O}|X=x) \cdot \mathbf{P}(X=x|Z=z), \\ &= \sum_{\frac{(T-1)\cdot\alpha\cdot z^{\eta}}{\sum_0^{\infty} \frac{\alpha\cdot z^{\eta}}{z}}}^{z} \frac{\binom{z}{x}}{2^z}. \end{aligned} \tag{5.16}$$

5.3.4 Load Propagation in G_p

We denote u is the probability that a node can not reach to the giant component through a specific neighbor. After the initial removal of fraction $1-\phi$, for a node i with degree of z, there are three possible ways not belonging to the giant component

- It is removed

- Or it is present but not a member of the giant component
- Or the shifted load from neighbors destroys itself, i.e., i is overloaded.

The first case happens with probability of $1 - \phi$. The second case happens if i can not reach to giant component through all its neighbors. The probability that none of the neighbors connects to giant component is $\phi \cdot u^z$. $\mathbf{P}(\mathbf{O}|Z = z)$ is the probability of the third case, given the number of neighbors is z. Thus,

$$\mathbf{P}(\mathbf{E}(\phi)) = 1 - \phi + \phi \cdot (u^z + (1 - u^z) \cdot \mathbf{P}(\mathbf{O}|Z = z)), \tag{5.17}$$

where $\mathbf{E}(\phi)$ is the event that the considered node disconnects with the giant component.

The probability that the vertex does belong to the giant component after the removal of fraction $1 - \phi$ is

$$\sum_0^\infty \mathbf{P}_\mathrm{d}(z) \cdot (1 - \mathbf{P}(\mathbf{E}(\phi))).$$

In other words, this is the expected fraction of remaining giant component. We define a new notation $\mathbf{J}(\phi, G_p)$ which represents the expected fraction of giant component of the subnetwork which occupies the fraction ϕ of the original entire network G_p, given by

$$\begin{aligned}\mathbf{J}(\phi, G_p) &= \frac{\sum_0^\infty \mathbf{P}_\mathrm{d}(z) \cdot (1 - \mathbf{P}(\mathbf{E}(\phi)))}{\phi}, \\ &= \sum_0^\infty \mathbf{P}_\mathrm{d}(z) \cdot (1 - u^z)(1 - \mathbf{P}(\mathbf{O}|Z = z)). \end{aligned} \tag{5.18}$$

The probability that i does not connect to giant component through a specific neighbor j equals to the probability that j does not connect to the giant component through its remaining neighbors, i.e., except i. Therefore, we obtain the following equation

$$u = \sum_{z=0}^{\infty} \mathbf{Q}(z) \cdot (1 - \phi + \phi \cdot (u^z + (1 - u^z) \cdot \mathbf{P}(\mathbf{O}|Z = z))), \tag{5.19}$$

recall $\mathbf{Q}(z)$ is the excess degree distribution.

Combining Eqs. (5.16) and (5.19), we can obtain the value of u. Then applying it to Eq. (5.17) and (5.18), finally we can calculate the remaining fraction of network after the random removal $1-\phi$, with considering the load cascading into account.

5.3.5 Interdependence Cascading Failures

First of all, we would like to estimate the threshold T'_c below which the interdependent networks would suffer a cascading failure due to load redistribution. After the random removal of a fraction $1 - \phi$ of nodes in G_p, the proportion of remaining network equals to $\mu'_{p_1} = \phi$. We assume that the nodes belong to giant component could operate properly. Thus, the remaining functioning fraction μ_{p_1} in G_p is given by

$$\mu_{p_1} = \mu'_{p_1} \cdot F(\mu'_{p_1}, G_p). \tag{5.20}$$

After the fragmentation in network G_p, the Energy-Dependency inter links connected with these removed nodes are removed. The corresponding nodes in network G_c lose the inter links and thus stop operate. The number of removed Energy-Dependency inter links is $\sum_0^{(1-\mu_{p_1})\cdot S_p} t_i$, where t_i represents the number of Energy-Dependency inter links the node i connects with. As far as S_p goes infinite, it could be approximated to

$$(1 - \mu_{p_1}) \cdot S_p \cdot \langle t \rangle,$$

where $\langle t \rangle$ is the mean value, given by

$$\langle t \rangle = \sum_{0}^{\infty} \mathbf{P}(\xi = t) \cdot t,$$

recall $\mathbf{P}(\xi = t)$ is given by Eq. (5.8).

As the total number of ED inter links is $\langle t \rangle \cdot S_p$, the probability for one energy inter link to be removed is $1 - \mu_{p_1}$. Consequently, the fraction of nodes in G_c that has the ED inter link is approximate to μ_{p_1}. Thus, we have $\mu'_{c_2} = \mu_{p_1}$. The operating giant component in G_c is calculated by

$$\mu_{c_2} = \mu'_{c_2} \cdot F(\mu'_{c_2}, G_c). \tag{5.21}$$

Roughly, the fraction $1 - \mu_{c_2}$ of control nodes are removed, given the network size S_c is large enough. As a result, this fragmentation may lead to further failures in G_p. The probability for each Control-Dependency inter link to be removed is approximated to $1 - \mu_{c_2}$, because each control node has the same ability to operate n nodes. With this in mind, the probability for a node in G_p loses all its k control links is $(1 - \mu_{c_2})^k$. The remaining fraction of nodes in G_p with at least one CD inter link is

$$\mu'_{p_3} = (1 - (1 - \mu_{c_2})^k) \cdot \mu_{p_1}.$$

By the transformation idea in [3, 15], to calculate the giant component, we consider the effect of removing the fraction of $1 - \mu_{p_3}$ of nodes in μ_{p_1} has the same effect as taking out the same fraction size from μ'_{p_1}. Thereby, from initial network G_p to μ'_{p_3}, we have the following equivalent removing process:

$$1 - \phi + \phi \cdot (1 - \mu_{c_2})^k = 1 - (\phi - \phi \cdot (1 - \mu_{c_2})^k).$$

Thus, the equivalent $\mu'_{p_3} = \phi \cdot (1 - (1 - \mu_{c_2})^k)$. The fraction of functioning giant component is

$$\mu_{p_3} = \mu'_{p_3} \cdot F(\mu'_{p_3}, G_p). \tag{5.22}$$

Once follow the above steps and continue, we can observe the pattern of these equations as follows:

$$\begin{aligned}
&\mu'_{p_1} = \phi,\\
&\mu_{p_1} = \mu'_{p_1} \cdot F(\mu'_{p_1}, G_p),\\
&\mu'_{p_3} = \phi \cdot (1 - (1 - \mu_{c_2})^k),\\
&\mu_{p_3} = \mu'_{p_3} \cdot F(\mu'_{p_3}, G_p),\\
&\cdots, \quad \cdots\\
&\mu'_{p_{2j+1}} = \phi \cdot (1 - (1 - \mu_{c_{2j}})^k),\\
&\mu_{p_{2j+1}} = \mu'_{p_{2j+1}} \cdot F(\mu'_{p_{2j+1}}, G_p),
\end{aligned}$$

and

$$\begin{aligned}
&\mu'_{c_2} = \mu_{p_1},\\
&\mu_{c_2} = \mu'_{c_2} \cdot F(\mu'_{c_2}, G_c),\\
&\mu'_{c_4} = \mu_{p_3},\\
&\mu_{c_4} = \mu'_{c_4} \cdot F(\mu'_{c_4}, G_c),\\
&\cdots, \quad \cdots\\
&\mu'_{c_{2j}} = \mu_{p_{2j-1}},\\
&\mu_{c_{2j}} = \mu'_{c_{2j}} \cdot F(\mu'_{c_{2j}}, G_c).
\end{aligned}$$

Once the cascading failures stop, at the final steady state, the following equations hold

$$\begin{aligned}
&\mu'_{p_{2j+1}} = \mu'_{p_{2j+3}} = \mu'_{p_{2j-1}},\\
&\mu'_{c_{2j}} = \mu'_{c_{2j+2}} = \mu'_{c_{2j-2}}.
\end{aligned}$$

We let $\gamma = \mu'_{p_{2j+1}} = \mu'_{p_{2j+3}} = \mu'_{p_{2j-1}}$ and $\delta = \mu'_{c_{2j}} = \mu'_{c_{2j+2}} = \mu'_{c_{2j-2}}$, then we obtain the transcendental equations:

$$\begin{cases} \gamma = \phi \cdot (1 - (1 - \delta \cdot F(\delta, G_c))^k), \\ \delta = \gamma \cdot F(\gamma, G_p). \end{cases} \tag{5.23}$$

Thus, the fractions of operating giant components in the final steady state are given by

$$\begin{cases} \lim_{j\to\infty} \mu_{p_j} = \mu_{p_\infty} = \gamma \cdot F(\gamma, G_p), \\ \lim_{j\to\infty} \mu_{c_j} = \mu_{c_\infty} = \delta \cdot F(\delta, G_c). \end{cases} \tag{5.24}$$

5.3.6 Load Propagation in Interdependent Networks

Eqs. (5.24) gives the estimated result on the size of functioning factions of networks. It only consider the effect of interdependence cascading failures. While, to avoid the cascading failure of overload, the redistributed load R can not exceed the remaining volume of the network. The threshold T'_c is obtained when both sides are exactly equal, where we get

$$\begin{aligned} 1 - \gamma \cdot F(\gamma, G_p) &= \gamma \cdot F(\gamma, G_p) \cdot (T'_c - 1), \\ T'_c =& \frac{1}{\gamma \cdot F(\gamma, G_p)} = \frac{1}{\delta}. \end{aligned} \tag{5.25}$$

That is the threshold below which the entire system would suffer from the cascading failure caused by load redistribution.

Lemma 5.1. *For a given power law exponent λ_{p} and initial removal $1-\phi$, $T'_c > T_c$.*

Proof. Combining Eqs. (5.13) and (5.25), we have

$$\frac{T'_c}{T_c} = \frac{\phi \cdot F(\phi, G_p)}{\gamma \cdot F(\gamma, G_p)}. \tag{5.26}$$

According to Eq. (5.23), if γ and δ are nonzero, $F(\gamma, G_p)$ and $F(\delta, G_p)$ are nonzero too. Consequently, $(1-(1-\delta \cdot F(\delta, G_c))^k) < 1$. Thus, we have $\phi > \gamma$.

It is shown that the $F(\phi, G_p)$ is a monotone increasing function [21], thus $F(\gamma, G_p) < F(\phi, G_p)$. We have

$$\frac{T_c'}{T_c} = \frac{\phi \cdot F(\phi, G_p)}{\gamma \cdot F(\gamma, G_p)} > 1,$$

which completes the proof. □

Lemma 5.2. *Recall a node in G_p is under the control of k nodes in G_c. For a given ϕ, for two different values of $k = k_1$ and $k = k_2$, where $k_1 < k_2$, the threshold $T_c'|_{k=k_1} > T_c'|_{k=k_2}$.*

Proof. By excluding δ in Eq. (5.23), we obtain

$$\phi = \frac{\gamma}{1-(1-\gamma \cdot F(\gamma, G_p) \cdot F(\gamma \cdot F(\gamma, G_p), G_c))^k}.$$

The growth of k from k_1 to k_2 leads to the increase of denominator of the right side. As ϕ is fixed, γ is required to be improved to hold the equation. Once ϕ increasing, $F(\gamma, G_p)$ grows since it is the monotone increasing function. Applying Eq. (5.25), we have $T_c'|_{k=k_1} > T_c'|_{k=k_2}$. Notice n is not involved in the equations, thus has no effect on T_c'. □

5.3.7 Joint Effect of Interdependence and Load Cascading

Now we estimate the system robustness with taking into account the joint effect of interdependence and load propagation cascading failures. In power grid G_p, $\mathbf{J}(\phi, G_p)$ should be applied to replace $F(\phi, G_p)$, while in the communication network G_c, $F(\phi, G_c)$ is kept. Therefore, the Eqs. (6.4) and (6.5) come to

$$\begin{cases} \gamma = \phi \cdot (1 - (1 - \delta \cdot F(\delta, G_c))^k), \\ \delta = \gamma \cdot \mathbf{J}(\gamma, G_p), \end{cases} \tag{5.27}$$

and the final fractions of functioning sizes are

$$\begin{cases} \mu_{p_\infty} = \gamma \cdot \mathbf{J}(\gamma, G_p), \\ \mu_{c_\infty} = \delta \cdot F(\delta, G_c). \end{cases} \tag{5.28}$$

Theoretically, once solving γ and δ in Eq. (5.27), we can obtain the results. While, no analytical theory is yet available to give the closed-forms of $F(\cdot)$ and $\mathbf{J}(\cdot)$.

Lemma 5.3. *Although the initial removal is upon G_p, the remaining fraction of survivals in G_p is still greater than that in G_c. That is, $\mu_{p_\infty} \geq \mu_{c_\infty}$.*

Proof. Applying $\delta = \gamma \cdot \mathbf{J}(\gamma, G_p)$ in Eq. (5.27) to Eq. (5.28), we get

$$\begin{cases} \mu_{p_\infty} = \delta, \\ \mu_{c_\infty} = \mu_{p_\infty} \cdot F(\delta, G_c). \end{cases} \tag{5.29}$$

Since $F(\delta, G_c)$ is less than or equal to 1, we have $\mu_{p_\infty} \geq \mu_{c_\infty}$.

□

5.3.8 Infinite Tolerance Parameter

Empirically, infinite value of tolerance parameter T means each power station has the infinite capacity, thus never suffer from overload propagation failure. For neither a single power grid or interdependent smart grid, whether a node is to be removed or not only depends on topology effect. In other words, infinite T is equivalent to studying the cascading failures caused by the topology and interdependence.

Lemma 5.4. *As the tolerance parameter increasing to infinite, i.e., $T \to \infty$, the fractions of survivals in both G_p and G_c meet the upper bounds, which are the results of pure 'k-n' interdependence model effect.*

Proof. Recall that $\mathbf{P}(\mathbf{O}|Z = z)$ is the probability that a node i with degree z is destroyed by its neighbors the load shift. As the growing of T, we notice Eq. (5.16) is monotonously decreasing, which leads to the reduce of the right side of Eq. (5.19). A smaller u is required to hold this equation. Therefore, Eq. (5.18) is *monotonously increasing.*

Once tolerance parameter goes to infinite, Eq. (5.15) is always equal to 0, which consequently let Eq. (5.16) be 0. Thus Eq. (5.17) comes to

$$\mathbf{P}(\mathbf{E}(\phi)) = 1 - \phi + \phi \cdot u^z. \tag{5.30}$$

The probability that the vertex does belong to the giant component $\mathbf{J}(\phi, G_p)$ is

$$\begin{aligned}\mathbf{J}(\phi, G_p) &= \sum_{0}^{\infty} \mathbf{P}_{\mathrm{d}}(z) \cdot (1 - u^z), \\ &= 1 - \sum_{0}^{\infty} \mathbf{P}_{\mathrm{d}}(z) \cdot u^z. \end{aligned} \tag{5.31}$$

Eq. (5.19) now can be simplified to

$$u = \sum_{0}^{\infty} \mathbf{Q}(z) \cdot (1 - \phi + \phi \cdot u^z). \tag{5.32}$$

Combining Eqs. (5.31) and (5.32) and comparing with Eq. (5.9), we obtain

$$\mathbf{J}(\phi, G_p) = F(\phi, G_p). \tag{5.33}$$

Now, Eq. (6.5) and Eq. (5.28) are exactly the same. The upper bound exists for $\mathbf{J}(\phi, G_p)$, when the tolerance parameter T approaches to infinite.

$\square$

5.4 Simulation and Analysis

In this section, we couple two synthetic networks using 'k-n' interdependence model and apply the overload cascading model upon G_p. We evaluate how the system robustness is affected by these models and parameters. In all experiments the fraction $1-\phi$ of nodes in G_p is removed randomly.

5.4.1 Joint Effect of Interdependence and Overload Cascading

First of all, we study the smart grid system on different network topologies by giving different λ_{p} and λ_{c}. Fig. 5.2(a) and 5.2(b) illustrate the joint effect of interdependence and overload cascading. The system is fragile to random failures, where approximate 11% initial failure leads to the entire collapse when $\lambda_{\mathrm{p}} = \lambda_{\mathrm{c}} = 2.5$. As the number of removals increasing, i.e., increasing of $1-\phi$, the fractions of survivals $\mu_{p\infty}$ and $\mu_{c\infty}$ in both networks slightly decrease, then followed by the rapid falls to zero. This performance drop is theoretically proved to be first-order discontinuous transition [106]. While in [3], we pointed out that in real life network and simulation, there exists a small transition interval due to the non-infinite sizes of networks. The system performance is heavily depending on the network topologies, where about 19% initial failure can destroy the system, given the power low exponents are 2.2 for both. Therefore, on the attack perspective, higher $1-\phi$ is required to knock down the system for lower λ_{p} and λ_{c}.

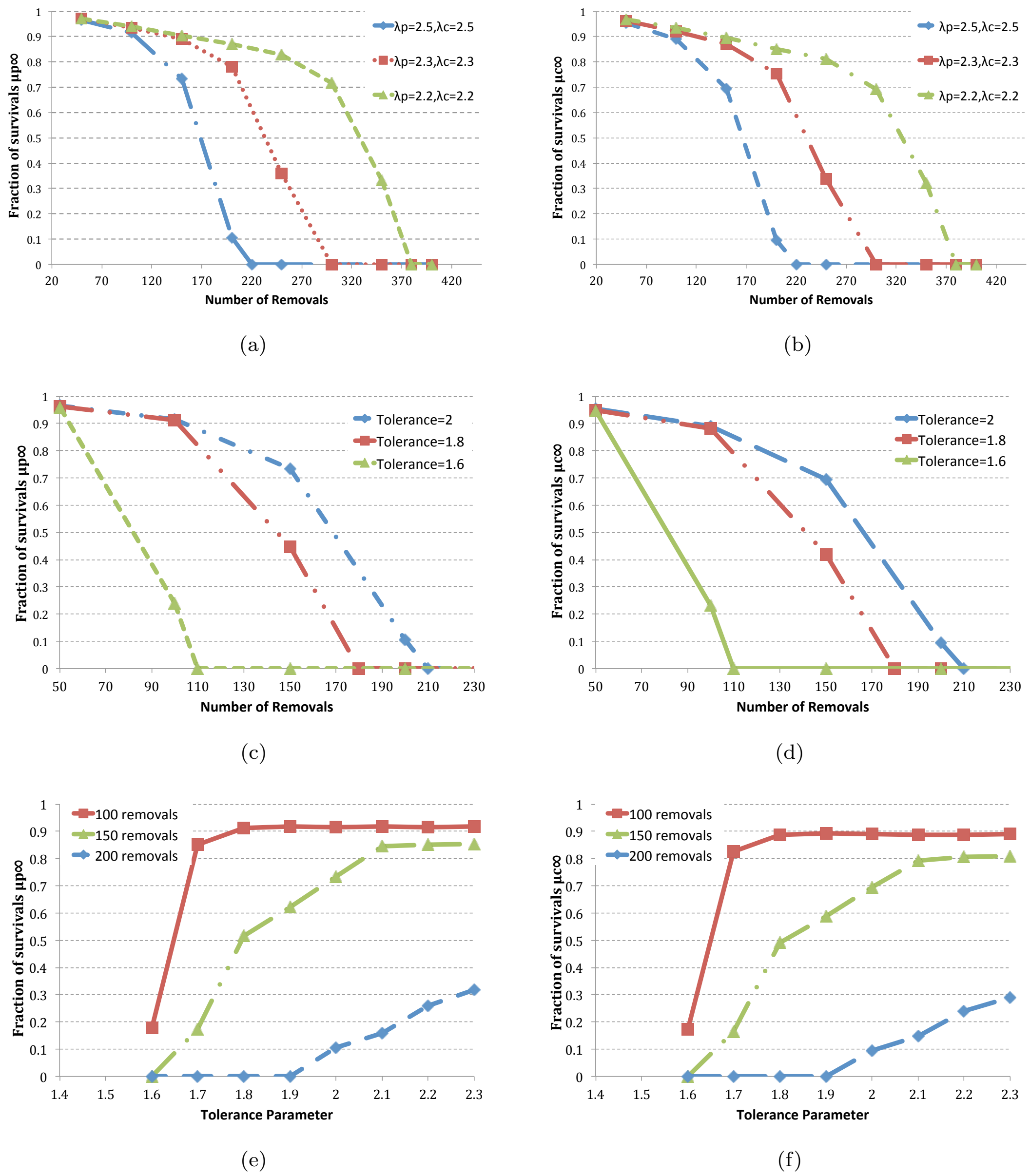
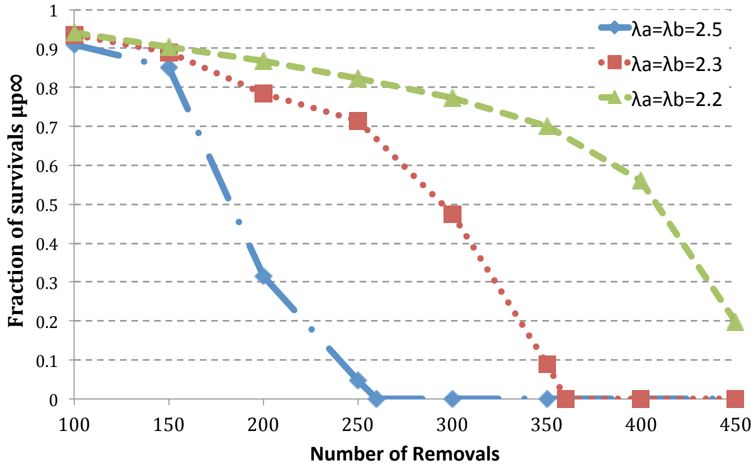
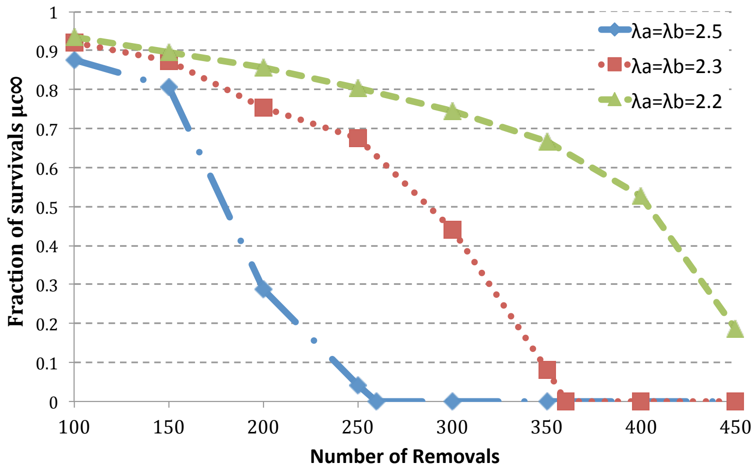


FIGURE 5.2: The initial size of G_p and G_c are 2000 and 7000 respectively. Random failure is applied on G_p. Both load propagation model and 'k-n' interdependence model are applied in the simulation. $k = 2, n = 5$. (a) and (b) compare the number of nodes left in different network topologies, with $T = 2$. In (c), (d), (e) and (f), we set up $\lambda_P = \lambda_c = 2.5$ and show how the system be affected by different tolerance parameter values and initial failures. Generally, a higher tolerance parameter T gives system a better performance, which reaches the upper bound once T keeps growing.

Fig. 5.2(c) and 5.2(d) show how system performs for various values of tolerance parameter. Accordingly, given the fixed removal, μ_{p_∞} and μ_{c_∞} grow once the tolerance parameter is increasing. The reason behind this phenomena is that the



(a)



(b)

FIGURE 5.3: The number of survivals in the system which only considers 'k-n' interdependence, with $k = 2, n = 5$. G_p and G_c have 2000 and 7000 nodes. Compared with Fig. 5.2(e) and 5.2(f), we observe the performance of only considering 'k-n' interdependence effect is the upper bound.

higher tolerance ability makes each power node has more vacant volume to adopt the shifted load from neighbors, thus is stronger in terms of overload propagation.

Comparing Fig. 5.2(a), 5.2(b), 5.2(c) and 5.2(d) , we observe $\mu_{p_\infty} \geq \mu_{c_\infty}$, implying that the fraction of survivals in G_p is greater than that in G_c. This validates our Lemma 5.3.

Another finding from Fig. 5.2(e) and 5.2(f) is that the system robustness finally reaches to the upper bound as the growing of tolerance parameter. Given the

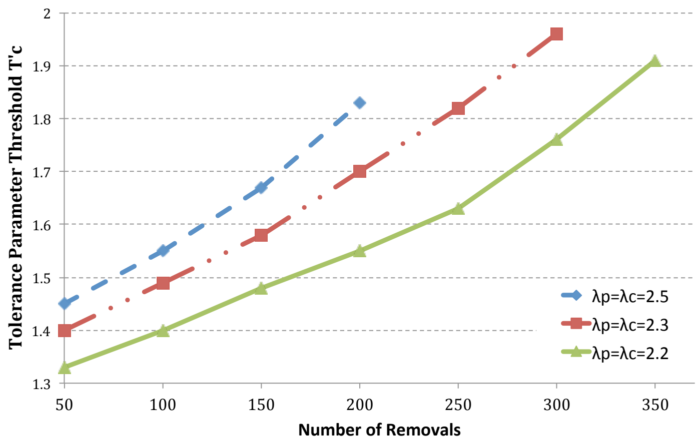
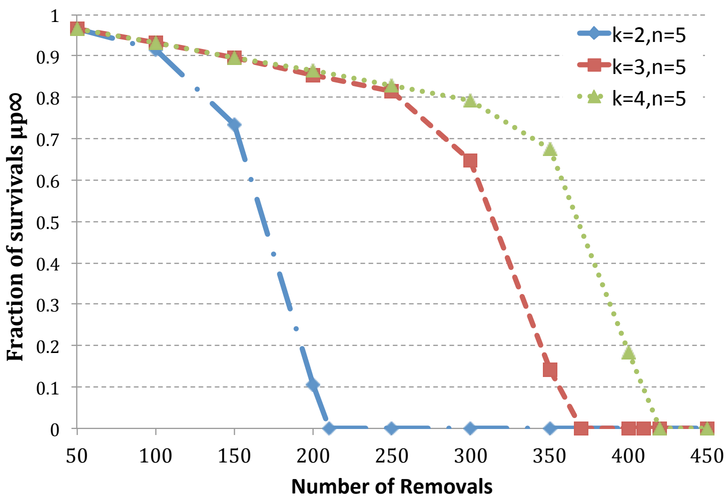


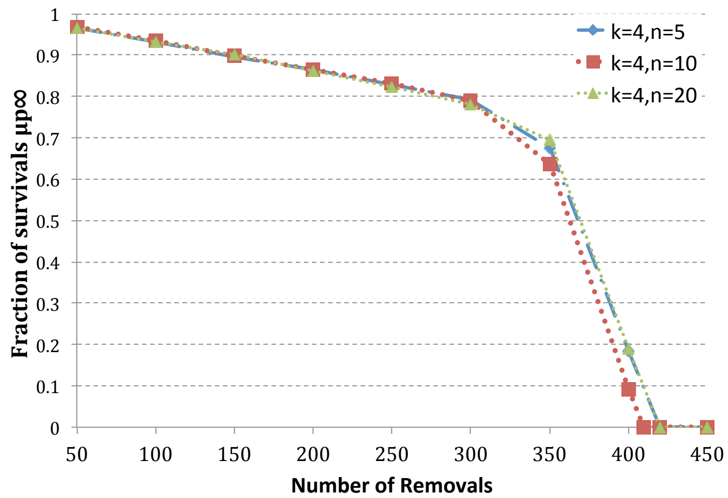
FIGURE 5.4: Tolerance threshold T'_c. $G_p = 2000$ and $G_c = 7000$, $k = 2, n = 5$. The super linear relation between the initial removal and T'_c is demonstrated.

initial failure of 0.05 (100 nodes out of 2000), $\mu_{p\infty}$ keeps stable on 0.91 when tolerance parameter exceeds 1.8. Same thing happens for the curve '150 removals'. We believe this upper bound is the system performance of single interdependence effect, where the load cascading effect is eliminated due to the huge vacant volume in power grid. Fig. 5.3 proves our finding, where it is the experiment that sets up T to infinite and only considers interdependence effect. For $\lambda_\mathrm{p} = \lambda_\mathrm{c} = 2.5$ and initial removal equals 100, $\mu_{p\infty}$ and $\mu_{c\infty}$ are 0.91 and 0.89 respectively, where that are exactly the same values of the upper bounds shown in Fig. 5.2(e) and 5.2(f). This validates our mathematical proof of Lemma 5.4.

We compare the tolerance threshold T'_c of various systems, as shown in Fig. 5.4. The value of T'_c for the system with lower λ_p and λ_c is also lower. To elaborate this, we firstly know that the scale-free network with power law of 2.2 has a longer fat tail than the network with power law of 2.5. The former one possesses more hubs, which, according to Eqs. (5.1) and (5.2), have more vacant space to absorb the shifted load from neighbors. Therefore, the network is firm to the load propagation. We also observe that the curve of $\lambda_\mathrm{p} = \lambda_\mathrm{c} = 2.5$ is cut out at the point that the number of removals is 200, where the curve is approaching to the interdependence threshold $1-\phi_c$, as illustrated in Fig. 5.2(a). No matter how big the T'_c is, the system keeps falling down when approaches to $1-\phi_c$. Therefore,



(a)



(b)

FIGURE 5.5: System robustness versus k and n. $G_p = 2000, G_c = 7000, T = 2$. Higher k indicates higher robustness, while n only has limited effects.

we conclude that the smart grid robustness depends on both T and $1 - \phi$. The system is totally knocked down whenever either of T'_c or $1 - \phi_c$ is reached first.

Another observation from Fig. 5.4 is that the relation between the number of initial removal and tolerance parameter threshold is super linear.

5.4.2 Comparison of k and n

The sub-linear relation between k and the fraction of survivals was demonstrated in our previous paper [3]. However, we would like to evaluate how the system robustness is affected by k and n in this new model with taking the load propagation into account. According to Fig. 5.5(a), both μ_{p_∞} and $1-\phi_c$ grow for the increasing k. With fixed $T=2$, for $k=4, n=5$, the threshold $1-\phi_c$ is 0.21, which is almost two times of curve $k=2, n=5$. The performance incremental quantity from $k=2$ to $k=3$ is much greater than that from $k=3$ to $k=4$. This validates our previous conclusion that the relation between k and robustness is sub-linear. Fig. 5.5(b) tells us another story that n has no effect on system. It validates our Eqs. (5.27) and (5.28), where n is not involved. In other words, the system performance is independent from n. However, we notice that the tiny fluctuations occur around threshold $1-\phi_c$. Recall n is the number of power stations that each operation center could operate. Given fixed size of S_p and k, the number of operation centers is inverse proportional to n [3]. Therefore, for each simulation, whether the initial random removal is upon these operation centers or not leads to quite different results. To weaken the fluctuation, we believe it is reasonable to repeat the simulation for massive times.

Fig. 5.6 shows that T_c' is also heavily depending on k rather than on n. For a given $1-\phi$, the tolerance threshold grows with the decreasing of k. For the case of $n=5, 1-\phi=0.1$, T_c' equals 1.83, 1.66, 1.63 respectively for $k=2,3,4$. This validates our Lemma 5.2. The reason could be explained as the following: higher k indicates higher system robustness in terms of topology failure. Thus, the system has more nodes survived, consequently has more vacant space to deal with the load propagation. The pressure of adopting shifted load on each individual node is therefore reduced.

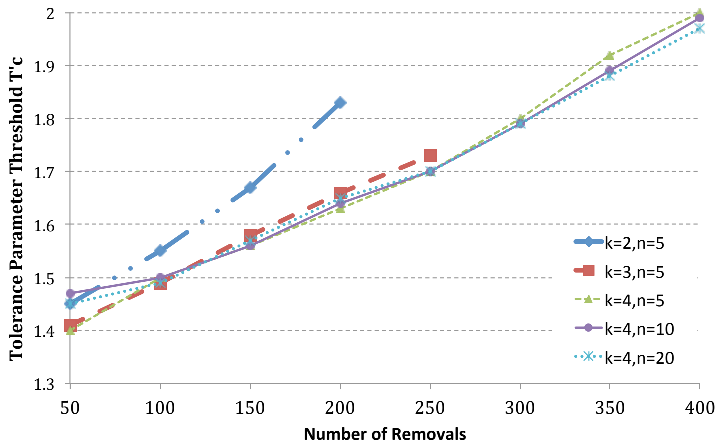
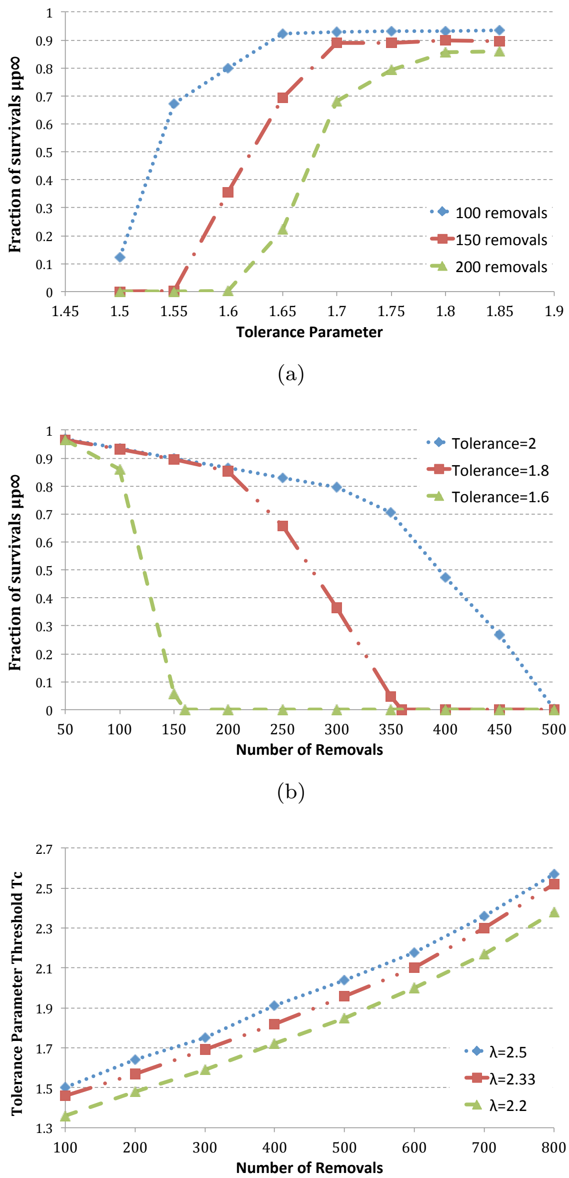


FIGURE 5.6: Tolerance threshold T_c' versus k, n. The fluctuations occur around the start points of $k = 4, n = 5, 10, 20$. The reason is that higher n indicates fewer operation centers [3]. For each single simulation, whether the initial random removal is upon these operation centers or not leads to quite different results.

For the case of fixed k, our mathematical Eq. (5.25) indicates that the threshold is not affected by n. While in our simulation, the change of n does affect T_c' limitedly. Generally, the three curves in Fig. 5.6 for $n = 5, n = 10$ and $n = 20$ are very close to each other. While, given the initial removal is 50, the values of T_c' for $n = 10$ and 20 are obviously greater than $n = 5$. We believe the reason is similar with Fig. 5.5(b), that whether the initial removal upon the operation centers or not causes quite different consequences. The mean results of simulation of these three curves would be almost the same once we enlarge the number of samples.

5.4.3 Single Power Grid versus Interdependent Networks

We compare the systems of single power grid and interdependent smart grid. Our comparison does not consider any physical properties on power grid. Fig. 5.7 indicates how different tolerance parameter affects the single power grid. Comparing Fig. 5.7(a) with Fig. 5.2(e), we observe that given the same $1 - \phi$ and network topology, the tolerance threshold for interdependent smart grid is higher than that for single power grid, e.g., $T_c' = 1.9$ and $T_c = 1.6$ if the initial removal is 200 nodes.



(a)

(b)

(c)

FIGURE 5.7: The system robustness and tolerance parameter in single power grid, with $G_p = 2000$. Similar with interdependent networks, growing k also leads to a upper bound performance, as shown in (a). Comparing (b) with Fig. 5.2(c), the number of survivals in single power grid is greater. (c) illustrates a super linear relation between initial removal and T_c.

Meanwhile, μ_{p_∞} in Fig. 5.7(a) is higher than that in Fig. 5.2(e), implying that the single power grid is stronger against random failures. Therefore, we conclude that the interdependence between the two networks decreases the system robustness dramatically. The comparison between Fig. 5.7(b) and Fig. 5.2(c) tells the same story that the number of survivals in single power grid is greater than that in coupled smart grid, thus the former one is much more reliable. We also notice that as the growing of T, the system performance is approaching to the upper bound in Fig. 5.7(a).

Fig. 5.7(c) draws the values of T_c of different network topologies. Generally, T_c grows with the increase of initial failures. We notice the growth speed is accelerating, e.g., for the curve of $\lambda = 2.2$, T_c increases from 1.35 to 1.6 when initial removal goes from 100 to 300, but expands from 2.0 to 2.37 when the removal grows from 600 to 800. This illustrates a super linear relation between initial removal and T_c.

According to Fig. 5.4, the T_c' for interdependent smart grid is 1.76 when $\lambda_{\mathrm{p}} = \lambda_{\mathrm{c}} = 2.2$ and initial removal is 300. Comparing with Fig. 5.7(c), $T_c = 1.6$ in the single power grid with same parameters. This validates our Lemma 5.1 that the required tolerance parameter for interdependent network is higher than for single power grid.

Chapter 6

An Exploration on Small Cluster

6.1 Small Disconnected Clusters

Recent studies explore network characteristics by using the metric of *giant component*, or specifically called *largest connected cluster* [17], which is defined as the largest connected cluster that fills most of the network. It is reasonable to study *giant component* of a single network, because each network could be approximately represented by its giant component. While, there is only one giant component in network and in most situations it does not fill the entire network [21]. There must also be some non-giant components in the network. We call these components *small components* or *small clusters*.

In an interdependent cyber physical system, except the giant component, the small clusters play an important role in the cascading failure process. In this section, we explore the role of small clusters in different single complex networks. In the remaining of this article ,we use *giant component* and *largest cluster* interchangeably.

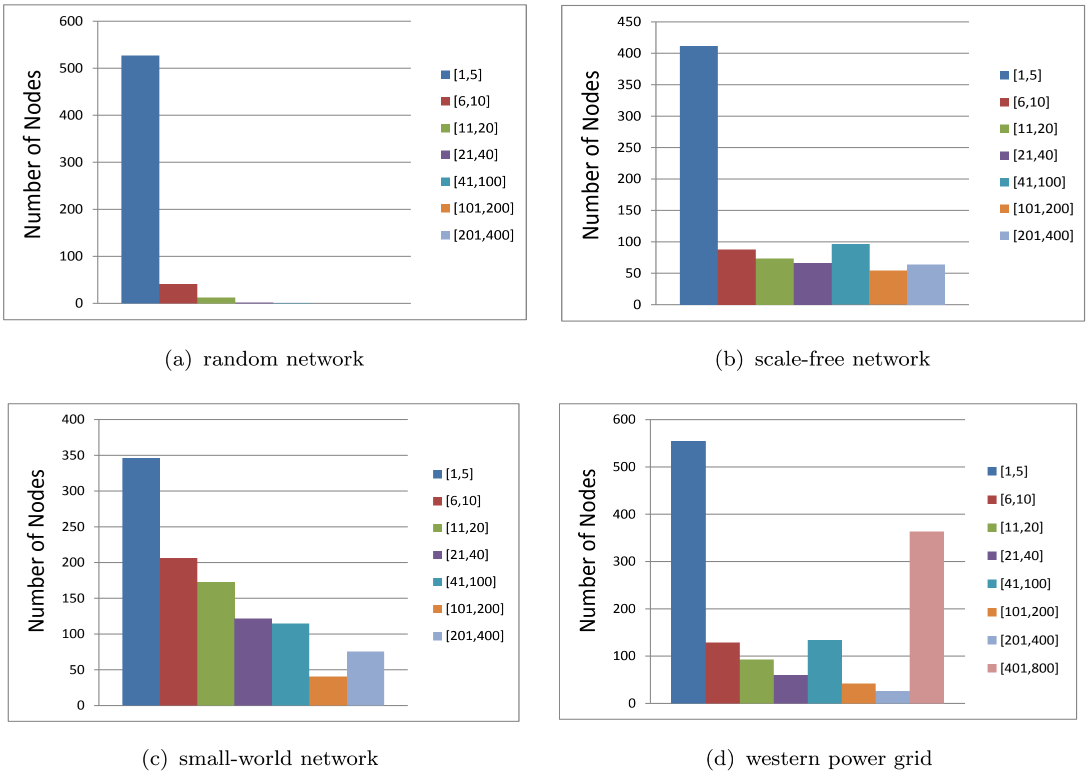
Considering a single network, we denote it as $G = (V, E)$, where V is the set of vertices of G, E is the set of edges between each pair in V. Once a set of vertices $V_f \subseteq V$ is failure or malfunctioning, the whole network G fragments to a set of disconnected sub networks, which denoted by $S(V_f) = \{G_1, G_2, G_3......G_j\}$, among which there exists one giant component whose size is proportional to the size of G [21]. Without loss of generality, we let G_1 represent the unique giant component, and call others *small clusters*. Previous studies were mostly focusing on exploring the characteristics of G_1. While, we give our argument that the remaining elements in $S(V_f)$ are also important, thus *cannot* be simply ignored.

6.1.1 Small Clusters in Single Network

Problem Definition : Given a single network $G = (V, E)$ and initial random failure $V_f \subseteq V$, let $|G|$ be the size of G. After V_f is applied to G, what are the sizes of $\{G_2, G_3......G_j\}$? How many small clusters exist in $S(V_f)$, i.e., the value of $j - 1$?

We use the extensive simulation to illustrate to what extent the small clusters could be important in a single network. We build synthetic random, scale-free and small-world networks using Erdös-Rényi model, Barabási-Albert model and Watts and Strogatz model [21] respectively. For each of them, the network size $|G| = 2000$, with random initial failure $|V_f| = 400$. To obtain the mean value of the small clusters, this construction-failure process was repeated for 60 times. Moreover, we use real Western States Power Grid data [4] and do the same process, as shown in Fig. 6.1.

With eliminating the largest cluster, one can clearly see the mean sizes of small clusters for each type of network. Most nodes in random network belong to clusters $\{G_i, G_{i+1}...\}$ whose sizes are less than 5 $(1 \leq \{|G_i|, |G_{i+1}|...\} \leq 5)$. In other words, the random network fragments to very tiny pieces. While for the other



(a) random network

(b) scale-free network

(c) small-world network

(d) western power grid

FIGURE 6.1: Distribution of small cluster sizes in various networks after random failure. The initial network sizes of (a), (b), (c) are 2000, with the mean degrees of 1.999, 1.999 and 2.19 respectively. (d) uses the real Western States Power Grid data compiled by [4], with 4941 nodes and the mean degree of 2.67. The 400 random failure nodes lead to the fragmentation upon each network. Each bar indicates the number of nodes locate into a cluster whose size is within the bar interval. For example, approximately 170 nodes belong to the small clusters whose sizes range from [11,20] in (c).

three, things are different. Significant part of nodes in scale-free and small-world networks locates into the clusters $\{G_i, G_{i+1}...\}$, where $11 \leq \{|G_i|, |G_{i+1}|...\} \leq 400$. Furthermore, the clusters with size over 200 appear sometimes in the networks. This gives us an important finding that the scale-free and small-world networks break into a largest cluster G_1, some relatively smaller clusters $G_2, G_3...$ where $|G_i| \geq 11$ for $i \geq 2$, and hundreds of small clusters whose sizes are less than 10.

For Western States Power Grid, the great observation is that the small clusters are not so 'small'. A large proportion of nodes are within the clusters whose sizes may be over than 380. In this experiment, we run the simulation and take the average. In some runs, the cluster does not form.

Table 6.1: Number of Small Clusters

	Random	Scale-free	Small World	Western Power
$\lvert G_i \rvert \leq 5$	403.8	290.7	148.7	361.0
$6 \leq \lvert G_i \rvert \leq 10$	5.8	11.8	27.4	17.7
$11 \leq \lvert G_i \rvert \leq 20$	0.9	5.1	12.3	6.6
$21 \leq \lvert G_i \rvert \leq 40$	0.1	2.4	4.5	2.1
$41 \leq \lvert G_i \rvert \leq 100$	0.02	1.6	2.0	2.1
$101 \leq \lvert G_i \rvert \leq 200$	0	0.4	0.3	0.4
$201 \leq \lvert G_i \rvert \leq 400$	0	0.3	0.3	0.1
$401 \leq \lvert G_i \rvert \leq 800$	0	0	0	0.7
Total	410.6	312.3	195.3	390.0

In terms of the number of small clusters $j-1$ after V_f, we repeat the simulation and calculate the mean value of the cluster size distribution, as shown in Table. 6.1. One can see that the random network has the highest number of small clusters. In other words, the average cluster size is the lowest, which could be proven by Fig. 6.1. We believe the reason is that the random network constructed by Erdös-Rényi model does not has the *small world effect*, which is also known as *six degree of separation*. Once the random failure occurs, the network fragments to massive tiny disconnect local clusters. This reason, at the same time, can explain why the small-world network has the less number of clusters than the others: a few long distance edges connect different local clusters, which consequently merge to the bigger clusters. Fig. 6.1(c) illustrates that the number of nodes belong to $|G_i| \leq 5$ are much less than others, proves that a part of tiny clusters merges.

We also observe that the number of clusters G_i where $|G_i| \leq 10$ dominates the total number of small clusters. Over 90% small clusters are quite small. Only a few clusters occupy a significant part of the whole network.

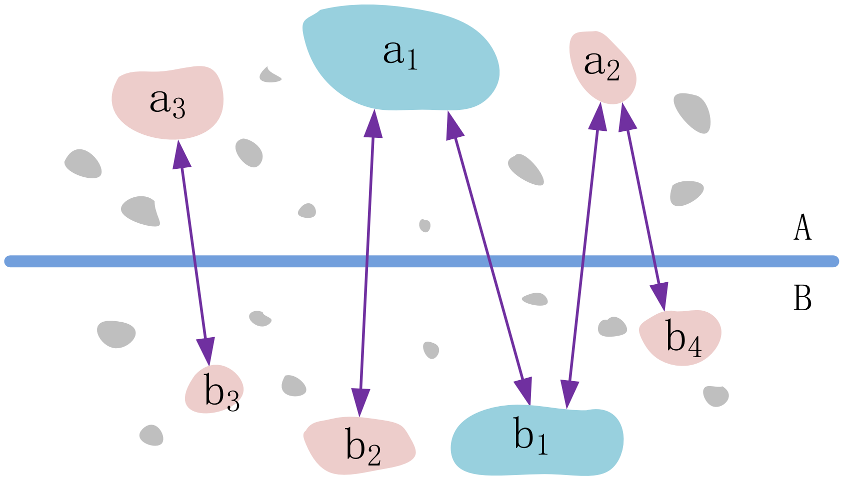


FIGURE 6.2: Two interdependent networks fragment into clusters. The largest clusters are partially dependent. Small clusters may fully interdependent or partially interdependent on the giant components to operate. Some small clusters (a_3, b_3) are fully mutually interdependent. Not only the giant component, but also the small clusters work properly in this steady state.

6.1.2 Small Clusters in Interdependent Networks

The findings according to Fig. 6.1 and Table. 6.1 bring us an overview on how the network behaves once got fragmented. Knowing this behavior directs us understand how to deal with modern systems. In a multiple sink wireless sensor actuator network, if the failure of nodes causes the disconnected of clusters, for each cluster it may has its own sink node. In this case, the original whole system is considered as several independently-working sub-systems. Not only the largest cluster, but also the small clusters are operating.

Similarly, the interdependent networks, such as smart power grid, once be fragmented, are working in the unit of isolated clusters. We believe the more reasonable assumptions that the node should satisfy to operate are:

- it has at least one inter link
- it belongs to a cluster with size greater than a pre-defined threshold Δ

Fig. 6.2 is a typical working pattern for the interdependent networks: the giant components a_1 and b_1 are partially dependent. Small clusters may fully interdependent (b_2, a_1) or partially interdependent (a_2, b_1) on the giant components. It is possible that two small clusters may mutually interdependent and thus operate properly, such as a_3 and b_3.

The way that the recent studies [3, 15, 20] used that only considering largest cluster is inappropriate, because significant part of network is ignored deliberately. In the remaining parts of this chapter, we will study the sizes of small clusters in the interdependent networks, and reveal how the network structure and initial failure impact small clusters.

Problem Definition : Given $G_p = (V_p, E_p)$ and $G_c = (V_c, E_c)$ representing physical resource network and computational resource network respectively, e.g., power grid and communication network, a set of failure nodes $V_f \subseteq V_p$ is applied to the network G_p. Due to the interdependence, G_p and G_c fragment to two sets of clusters $S_p(V_f) = \{G_{p1}, G_{p2}, G_{p3}......\}$ and $S_c(V_f) = \{G_{c1}, G_{c2}, G_{c3}......\}$ respectively. Besides the largest clusters G_{p1} and G_{c1}, we are interested in the fractions of operating small clusters $\{G_{p2}, ...G_{pi}\}$ and $\{G_{c2}, ...G_{cj}\}$, where the two assumptions mentioned above are satisfied that: *1)* $\{|G_{p2}|, |G_{c2}|, ...|G_{pi}|, |G_{cj}|\} > \Delta$; *2)* each node involved has at least one inter link. Moreover, as different interdependence models indicating quite different system vulnerabilities, we will compare our 'k-n' model [3] with 'one-to-one' [15, 17] to discuss how interdependence affects the small clusters.

6.2 Mathematical Estimation

In this section, we indicate a method to estimate the functioning giant components and small clusters in interdependent networks, under two interdependence models

'k-n' and 'one-to-one'.

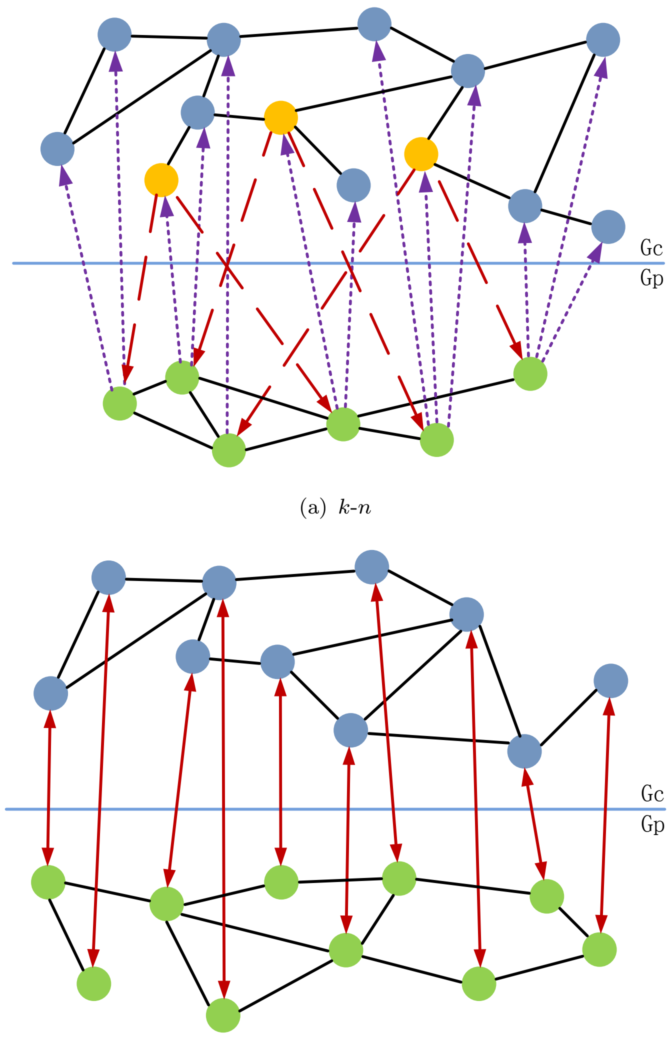
Although given the same network model parameters, e.g., power law exponent or mean degree, we may obtain different structures. Thus, in this section, we use mean filed to study the ensembles of G_p and G_c. Definitions and notations are listed in Table 6.2. We introduce a notation $F(\phi, G_p)$ from [20] to represent the expected fraction of giant component in the subnetwork which occupies the fraction ϕ of the nodes in the entire network G_p. Similarly, $F(\phi, G_c)$ is the fraction of giant component for network G_c. We begin random failure set V_f with the size of $(1 - \phi) \cdot |G_p|$ on network G_p.

We briefly introduce 'k-n' and 'one-to-one' models to help us understand this section more profoundly.

► *'k-n' interdependence model* was proposed for smart power grid, where the nodes in G_c are divided into *information relay nodes* and *control nodes*. Each node in G_c is supported by only one node in G_p, while each node in G_p is monitored and operated by k control nodes. On the other hand, each control node has the same ability to operate n nodes in G_p. Two types of unidirectional inter links exist in the model: *energy-dependency* and *control-dependency*. An example is given in Fig. 6.3(a).

► *'one-to-one' model* is a simple and widely used model, where each node in G_p depends on exactly one unique node in G_c, and vice versa. The nodes in each pair connected by one inter link are mutually dependent. Therefore, $|G_p| = |G_c|$ is satisfied in this model. Fig. 6.3(b) indicates an example of 'one-to-one' model.

Theorem 6.1. *Recall Δ is the threshold size above which we consider the small cluster could operate independently. Then, it follows that a smaller value of Δ leads to a higher robust interdependent system.*



(a) k-n

(b) one-to-one

FIGURE 6.3: (a) k-n interdependence model with k=1, n=2. The inter link is unidirectional. (b) each node in G_p depends on exactly one unique node in G_c.

Proof. Consider the 'k-n' interdependence model. Let $S(\phi, G_p, \Delta)$ represent the fraction of operating small clusters whose sizes are greater than Δ in network G_p, with $1-\phi$ of network failing. We begin the random removal of a fraction $1-\phi$ of nodes in G_p. The fraction of remaining network $\mu'_{p_1} = \phi$. The nodes belong to giant component and small clusters with sizes greater than Δ could operate. Thus, the remaining functioning fraction μ_{p_1} is given by

$$\mu_{p_1} = \mu'_{p_1} \cdot (F(\mu'_{p_1}, G_p) + S(\mu'_{p_1}, G_p, \Delta)) \tag{6.1}$$

Table 6.2: Notations for math approximation

G_p	physical resource network, e.g., power grid
G_c	computational resource network, e.g., communication network
μ'_{p_i}	the fraction of nodes retaining at least one inter link in G_p
μ'_{c_i}	the fraction of nodes retaining at least one inter link in G_c
μ_{p_i}	the fraction of giant component and operating small clusters in μ'_{p_i}
μ_{c_i}	the fraction of giant component and operating small clusters in μ'_{c_i}
Δ	the small cluster threshold size above which we consider it could operate
δ	the number of operating small clusters with sizes over Δ

As the network G_p fragmentation, a part of inter links is removed. Some nodes in network G_c lose inter link and thus stop function. the probability for one energy inter link removal is $1 - \mu_{p_1}$ (See details in [3]). The fraction of nodes in G_c that has an inter link is approximate to μ_{p_1}. So, we have $\mu'_{c_2} = \mu_{p_1}$. Then, the fraction of operating giant component and small clusters in G_c is given by

$$\mu_{c_2} = \mu'_{c_2} \cdot (F(\mu'_{c_2}, G_c) + S(\mu'_{c_2}, G_c, \Delta)) \tag{6.2}$$

The failure in G_c causes further failures in G_p. As a result, the fraction $1 - \mu_{c_2}$ of control centers are removed, given the network size $|G_p|$ is large enough. The probability for each control inter link to be removed is approximated by $1 - \mu_{c_2}$ because each control center has the same ability to control n nodes. With this respective, a node in G_p loses all its k control links with the probability of $(1-\mu_{c_2})^k$. The fraction μ'_{p_3} of nodes with at least one inter control link is $(1-(1-\mu_{c_2})^k)\cdot\mu_{p_1}$. By the transformation idea in [3], we consider the effect of removing the fraction of $1 - \mu'_{p_3}$ of nodes in μ_{p_1} has the same effect as taking out the same fraction size from μ'_{p_1}. Thereby, from initial network G_p to μ'_{p_3}, we have the following equivalent removing process:

$$1 - \phi + \phi \cdot (1 - \mu_{c_2})^k = 1 - (\phi - \phi \cdot (1 - \mu_{c_2})^k).$$

Thus, the equivalent $\mu'_{p_3} = \phi \cdot (1 - (1 - \mu_{c_2})^k)$. The fraction of operating clusters is

$$\mu_{p_3} = \mu'_{p_3} \cdot (F(\mu'_{p_3}, G_p) + S(\mu'_{p_3}, G_p, \Delta)) \tag{6.3}$$

Once repeat the above steps, we can observe the pattern of these equations as follows:

$$\begin{aligned}
&\mu'_{p_1} = \phi \\
&\mu_{p_1} = \mu'_{p_1} \cdot (F(\mu'_{p_1}, G_p) + S(\mu'_{p_1}, G_p, \Delta)) \\
&\mu'_{p_3} = \phi \cdot (1 - (1 - \mu_{c_2})^k) \\
&\mu_{p_3} = \mu'_{p_3} \cdot (F(\mu'_{p_3}, G_p) + S(\mu'_{p_3}, G_p, \Delta)) \\
&\cdots, \quad \cdots \\
&\mu'_{p_{2j+1}} = \phi \cdot (1 - (1 - \mu_{c_{2j}})^k) \\
&\mu_{p_{2j+1}} = \mu'_{p_{2j+1}} \cdot (F(\mu'_{p_{2j+1}}, G_p) + S(\mu'_{p_{2j+1}}, G_p, \Delta))
\end{aligned}$$

and

$$\begin{aligned}
&\mu'_{c_2} = \mu_{p_1} \\
&\mu_{c_2} = \mu'_{c_2} \cdot (F(\mu'_{c_2}, G_c) + S(\mu'_{c_2}, G_c, \Delta)) \\
&\mu'_{c_4} = \mu_{p_3} \\
&\mu_{c_4} = \mu'_{c_4} \cdot (F(\mu'_{c_4}, G_c) + S(\mu'_{c_4}, G_c, \Delta)) \\
&\cdots, \quad \cdots \\
&\mu'_{c_{2j}} = \mu_{p_{2j-1}} \\
&\mu_{c_{2j}} = \mu'_{c_{2j}} \cdot (F(\mu'_{c_{2j}}, G_c) + S(\mu'_{c_{2j}}, G_c, \Delta))
\end{aligned}$$

Once the cascading failures stop, at the final steady state, the following equations hold

$$\mu'_{p_{2j+1}} = \mu'_{p_{2j+3}} = \mu'_{p_{2j-1}}$$
$$\mu'_{c_{2j}} = \mu'_{c_{2j+2}} = \mu'_{c_{2j-2}}$$

because both G_p and G_c stop fragmenting. Let $x = \mu'_{p_{2j+1}} = \mu'_{p_{2j+3}} = \mu'_{p_{2j-1}}$ and $y = \mu'_{c_{2j}} = \mu'_{c_{2j+2}} = \mu'_{c_{2j-2}}$, then we have:

$$\begin{cases} x = \phi \cdot (1 - (1 - y \cdot (F(y, G_c) + S(y, G_c, \Delta)))^k) \\ y = x \cdot (F(x, G_p) + S(x, G_p, \Delta)) \end{cases} \tag{6.4}$$

The fractions of operating giant component and small clusters in the final steady state are given by

$$\begin{cases} \lim_{j\to\infty} \mu_{p_j} = \mu_{p_\infty} = x \cdot (F(x, G_p) + S(x, G_p, \Delta)) \\ \lim_{j\to\infty} \mu_{c_j} = \mu_{c_\infty} = y \cdot (F(y, G_c) + S(y, G_c, \Delta)) \end{cases} \tag{6.5}$$

Theoretically, once solving x and y in Eq. (6.4), we reach to the final results. While, no analytical theory is yet available to give the closed-forms of $F(\cdot)$ and $S(\cdot)$. But we still can derive some conclusions from these equations.

We use π_s to denote the probability that a random chosen vertex belongs to a small cluster with the size *exactly* s in total. The generating function of π_s is

$$H(z) = \pi_1 \cdot z + \pi_2 \cdot z^2 + \pi_3 \cdot z^3 + \pi_4 \cdot z^4 + \pi_5 \cdot z^5 + \cdots \tag{6.6}$$

Combining the threshold Δ, we have

$$H(z)|_{s\geq\Delta} = \pi_\Delta \cdot z^\Delta + \pi_{\Delta+1} \cdot z^{\Delta+1} + \cdots \tag{6.7}$$

Once reducing the size threshold Δ, the value of left of Eq. (6.7) increases. In other words, the fraction of operating small clusters increases. Therefore, we can conclude $S(y, G_c, \Delta)$ and $S(x, G_p, \Delta)$ on the right side of Eq. (6.4) would grow. Consequently, the left side x and y would increase. Given $F(\phi, G_p)$ and $F(\phi, G_c)$ are monotone increasing functions [21], Eq. (6.5) deterministically grows. Thus, a smaller value of Δ leads to a higher robust interdependent system. □

Lemma 6.2. *In 'k-n' interdependence model, there exists an upper bound on the fraction of operating small clusters.*

Proof. First of all, we have,

$$\begin{cases} 0 \leq F(x, G_p) + S(x, G_p, \Delta) \leq x \\ 0 \leq F(y, G_c) + S(y, G_c, \Delta) \leq y \end{cases} \tag{6.8}$$

From the result in [3], we get that for a large value of k, the giant components $F(x, G_p)$ and $F(y, G_c)$ approach to x and y respectively. In other words, with a strong interdependence, the remaining parts in both networks are almost fully connected. In this case, the fractions of small clusters hold that

$$\begin{cases} 0 \leq S(x, G_p, \Delta) \leq x - F(x, G_p) = 0 \\ 0 \leq S(y, G_c, \Delta) \leq y - F(y, G_c) = 0 \end{cases} \tag{6.9}$$

On the other hand, there are no survivals (both giant component and small cluster) if $k = 0$, according to Eq. (6.4). Thus, as k is increasing from 0 to infinite, both $S(\phi, G_p, \Delta)$ and $S(\phi, G_c, \Delta)$ start from zero and vanish to zero, which completes the proof. □

Theorem 6.3. *In 'k-n' interdependence model, an upper bound exists for the number of operating small clusters δ.*

Proof. Consider $S(x, G_p, \Delta)$, a loose bound follows that:

$$0 \leq \delta \leq S(x, G_p, \Delta)/\Delta \tag{6.10}$$

While, the mean size of the small cluster containing a randomly chosen operating node is given by: $\bar{s} = \sum_{s\geq\Delta}^{|G_{p1}|-1} s \cdot \pi_s / \sum_{s\geq\Delta}^{|G_{p1}|-1} \pi_s$, where $|G_{p1}|$ is the giant component size of G_p. A tight bound of δ could be obtained by

$$0 \leq \delta \leq \frac{S(x, G_p, \Delta)}{\bar{s}} = \frac{S(x, G_p, \Delta) \cdot \sum_{s\geq\Delta}^{|G_{p1}|-1} \pi_s}{\sum_{s\geq\Delta}^{|G_{p1}|-1} s \cdot \pi_s} \tag{6.11}$$

As proven in Lemma 6.2, $S(x, G_p, \Delta)$ has an upper bound, thus, we can conclude that δ also has an upper bound. □

We also study the widely used 'one-to-one' interdependence model [15]. Once repeat the steps above, we obtain new transcendental equations:

$$\begin{cases} x = \phi \cdot (F(y, G_c) + S(y, G_c, \Delta)) \\ y = x \cdot (F(x, G_p) + S(x, G_p, \Delta)) \end{cases} \tag{6.12}$$

In general, it is difficult to derive the expression for x and y. Deriving the closed-form of $F(y, G_c)$ and $S(y, G_c, \Delta)$ is a tough task and still an open problem.

6.3 Experiments

In this section, we use the extensive simulations to reveal how different network topologies and interdependence affect the small clusters.

6.3.1 Simulation Setup

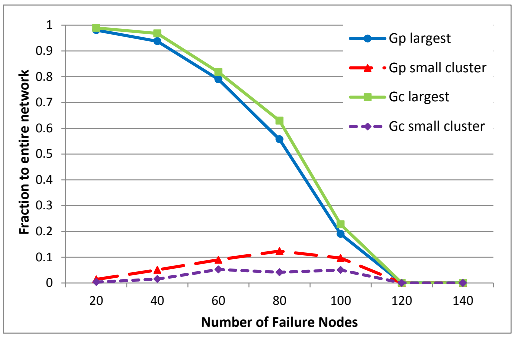
We build synthetic random, scale-free and small-world networks using Erdös-Rényi model, Barabási-Albert model and Watts and Strogatz model [21] respectively. We also use real Western States Power Grid data [4], where the data set could be found at [151].

We write a specific program using Java to simulate the cascading failure process in interdependent networks. In all the experiments the following approach is employed:

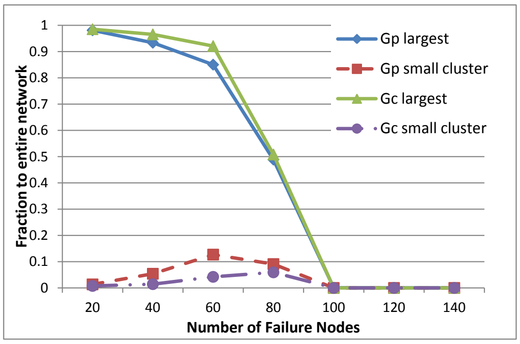
- we construct one or two synthetic networks
- then, we couple two synthetic networks or one synthetic network with Western States Power Grid using 'k-n' or 'one-to-one' interdependence models.
- we choose nodes in network G_p as failure or attack randomly or depending on the degrees.
- the failure propagates between the two networks. Our program simulates each stage, stores and updates the network information.
- finally, our program calculates the number of survived nodes and the number of clusters in both networks.

6.3.2 Interdependence Impacts

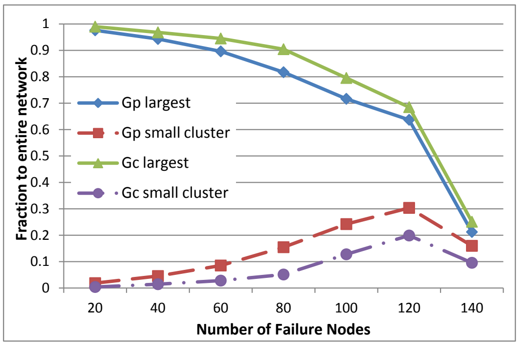
To study how small clusters be affected by the interdependence between coupled networks, we compare the 'k-n' interdependence model [3] with the widely used 'one-to-one' model [15, 17]. We couple two synthetic small-world networks using both models. The results are given in Fig. 6.4. Some observations are as follows:



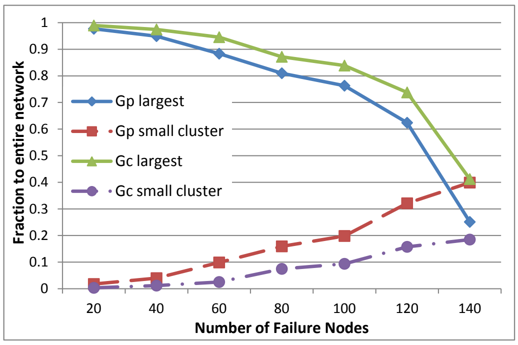
(a) k=2, n=3



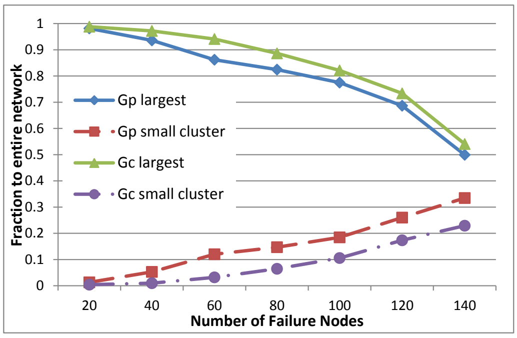
(b) k=2, n=5



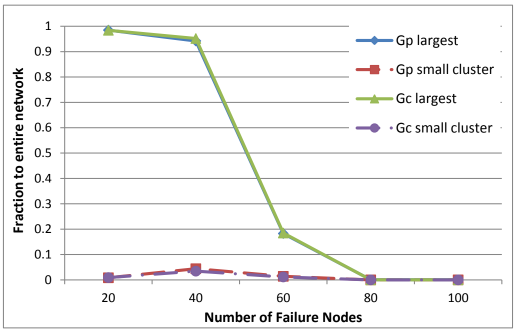
(c) k=3, n=3



(d) k=3, n=5



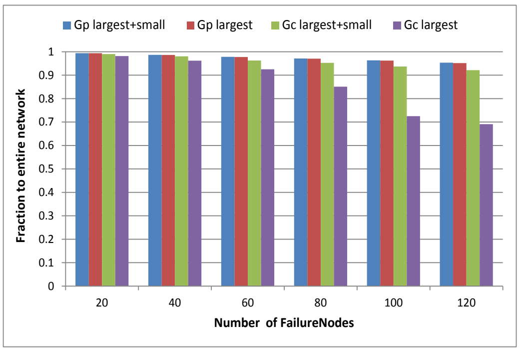
(e) k=4, n=5



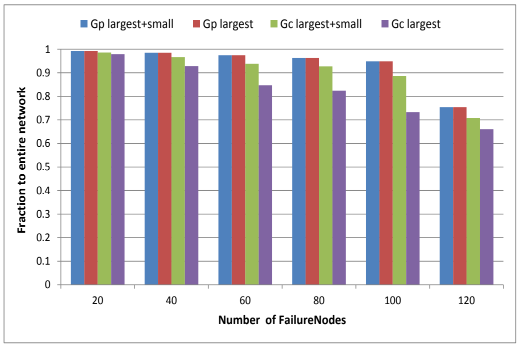
(f) One to One

FIGURE 6.4: The impacts of interdependencies on smaller clusters. We use coupled small-world networks for the implementations. For (a), (b), (c), (d) and (e), the sizes of G_p and G_c are 4000 and 6000. For 'one-to-one' model (f), both networks have the same size of 5000. We set the small cluster size threshold $\Delta = 20$.

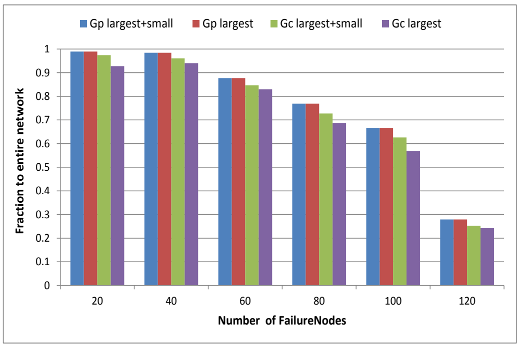
► Generally speaking, we observe that unless the interdependent networks collapse, the small clusters surviving from the cascading failures occupy a big part of the system. In the case of (c), (d) and (e), the fractions of operating small clusters could be greater than 40% of the whole system. In Fig. 6.4(d), when $|V_f| = 120$, the size of small clusters (41%) even exceeds the size of largest cluster (27%) in



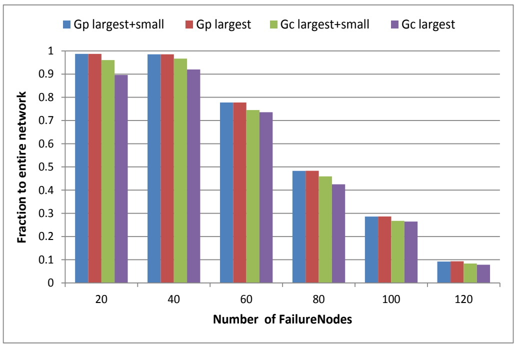
(a) $\Delta = 2$



(b) $\Delta = 5$



(c) $\Delta = 10$



(d) $\Delta = 20$

FIGURE 6.5: Largest cluster and small clusters for various Δ. Recall that Δ is the critical size beyond which the component is considered as the *small cluster*. We use 'k-n' interdependence model to couple two synthetic scale-free networks G_p and G_c with the power law exponents 2.5 and 3 respectively. Initially, G_p and G_c have 4000 and 6000 nodes, respectively, random failure set V_f is applied to G_p.

G_p.

► As shown in (a), (b) and (c), the proportions of small clusters first increase with the growing of $|V_f|$, then decrease to zero once the entire system fails. The sum of largest cluster and small clusters keeps stable (approximate 95%) for $|V_f| < 120$ in (c). This value for case (b) is also 95% if $|V_f| < 60$, after which the size of functioning clusters drops quickly.

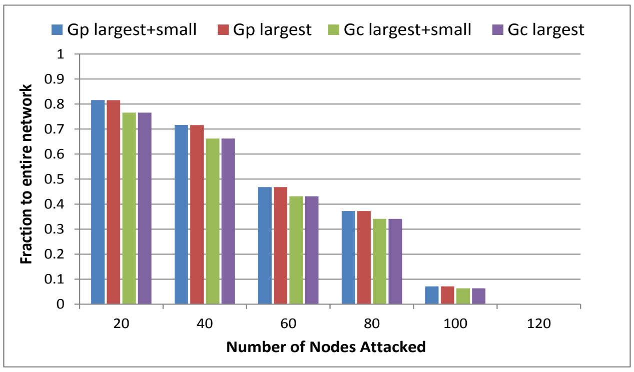
► In 'k-n' interdependence model, comparing (a), (c) and (b), (d), we find increasing k could increase both largest cluster and small cluster sizes. The reason is that higher value of k means the node in G_p is operated by more control nodes. The hardware redundancy improves the system robustness [3]. Notice that the proportion of small clusters does not grow unlimitedly. As shown in (d) and (e), the small cluster size experiences a slightly reduction even we raise k from 3 to 4. But, at the same time, the largest cluster keeps growing. The reason of this phenomenon is that once G_p and G_c are tightly coupled, the nodes are more likely to survive and stay in the largest cluster, which potentially affects the sizes of small clusters. Therefore, for the small values of k, the small cluster proportion is monotonously increasing as the raise of k. Then it decreases because the increasing largest cluster contests the nodes. The finding here proves our Lemma 6.2 that an upper bound exists for the fraction of operating small clusters.

▶ Recall different values of n do not affect the total number of inter links [3], thus do not improve the system robustness. The findings in (a), (b), (c) and (d) have the similar results. Comparing (a) and (b), one can observe that raising n leads to a flatter transition [3], and thus causes a different distribution of small clusters when $60 < |V_f| < 120$. Same thing occurs in (c) and (d). Therefore, according to our simulations, the value of n has the impacts on small clusters rather than largest cluster.

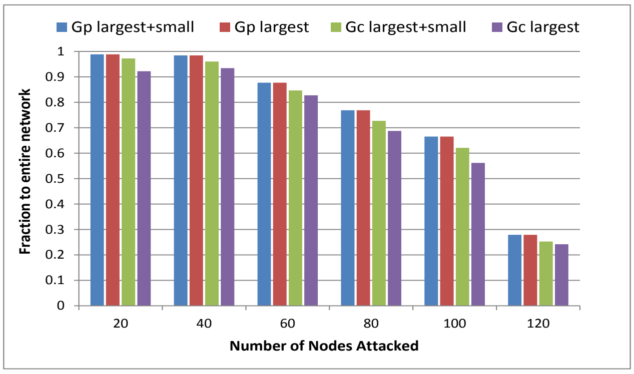
▶ (f) reports how 'one-to-one' interdependence model affects the small clusters, which only occupy a small fraction of the entire system. One interesting observation is that the sums of largest cluster and small cluster for G_p and G_c are equal. In other words, these two coupled networks have the same number of survivals. As required by 'one-to-one' model, each node in G_p depends on exactly one unique node in G_c. Each survival node has the unique interdependent node for operating. Thereby, the numbers of survival nodes in G_p and G_c are equal.

6.3.3 Tiny Pieces in Interdependent Networks

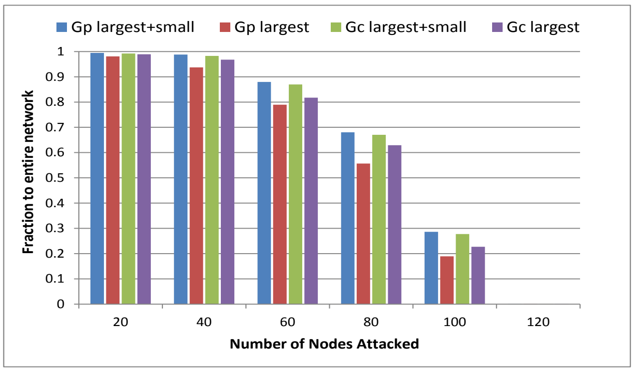
We are also interested in revealing how the interdependent networks look like once they get fragmented, in terms of clusters distributions. As percolation theory



(a) coupled random networks



(b) coupled scale-free networks



(c) coupled small world networks

FIGURE 6.6: Impacts of network topologies on small clusters. For coupled scale-free networks, G_p and G_c have the power law exponents of 2.5 and 3 respectively. The mean degrees of nodes in coupled random networks and small-world networks are 2.45 and 2.4. 'k-n' interdependence model is applied, with k=2, n=3. Initially, G_p has 4000 nodes, G_c has 6000 nodes. $\Delta = 20$.

pointed out, there exists a constant threshold for the proportion of faulty nodes, beyond which the system collapses [19]. Our study [3] found once the whole system collapse, even the small clusters *cannot* survive, i.e., the networks break to disconnected tiny pieces.

While, if the system does not collapse, the small clusters occupy a significant proportion according to Fig. 6.5. We notice with the increasing of Δ, the fraction of small clusters and the size of largest cluster dramatically reduce at the same time. Both G_p and G_c keep the stable sizes of survivals if Δ equals to 2, but are close to vanish as Δ growing to 20. This phenomenon indicates that considering small clusters into account or not highly affects the final results. The findings in Fig. 6.5 also prove our Theorem 6.1 that a smaller value of Δ leads to a higher robust interdependent system. Thereby, only using giant component to study the interdependent networks is not enough to draw the accurate conclusions.

The more important observation here is that the threshold Δ causes different proportions of small clusters in the final steady state. As shown in Fig. 6.5(a) and Fig. 6.5(b), the fraction of small clusters occupies at most 20% of the network. But for Fig. 6.5(c) and Fig. 6.5(d), 2% to 5% of the network is consisted by small clusters. This gives us an overview on how $\{|G_{p1}|, |G_{p2}|, ...|G_{pi}|\}$ and $\{|G_{c1}|, |G_{c2}|, ...|G_{cj}|\}$ distribute. In our simulation case, except the largest cluster, about 3% of networks are the small cluster whose sizes are over 20 in G_c. A significant fraction of G_c is consisted of the tiny clusters with the sizes approximate to 2. While, the remaining parts of the network are the isolated single vertex.

6.3.4 Various Network Topologies

To study how different network topologies affect the characteristics of small clusters, we implement three scenarios of coupled network topologies, as shown in Fig. 6.6. To make the results convincing, we let the nodes in each scenario have close mean degrees. According to Fig. 6.6(a), the small clusters is non-existent in coupled random networks. In other words, only the largest cluster could survive in the final steady state. While, the fraction of small clusters in coupled scale-free networks is more or less stable for various sizes of $|V_f|$. The interesting thing for coupled small-world networks is that small clusters occupy a large proportion of G_p, up to 10%. In the meanwhile, only 5% of G_c is small clusters.

Comparing Fig. 6.6(b) and Fig. 6.6(c), one can observe that the small cluster sizes of coupled small-world are greater than coupled scale-free, for both G_p and G_c. To explain the reason, we introduce an fundamental concept of scale-free network: that most part of nodes which have few edges are connected by the 'hub' whose degree is relatively high. The breakdown of 'hub' nodes may lead to disconnected tiny clusters, sometimes even to plenty of isolated vertex.

We notice that the coupled random networks are of the most vulnerable among all three implementations. Its counterpart, coupled small-world networks, has the fastest degradation speed. 120 failure nodes could devastate both of them.

6.3.5 Number of Functioning Clusters

One problem we would like to ask is as we know there indeed exists some small clusters surviving from the cascading failures, how many of them it could have. For real world interdependent networks, this problem is translated to how many isolated clusters could operate inter-dependently.

Fig. 6.7 reports the numbers of clusters that still function in interdependent networks. Coupled random networks only consist the giant components if it is still alive. None of working small clusters could be found. This is also proven by Fig. 6.6(a) that the proportion of small clusters is zero.

Similarly, coupled small-world networks have the largest number of functioning clusters, which are up to 15 in G_p and 12 in G_c, i.e., at most 27 different clusters could work together in this system. Coupled scale-free networks are in the middle, having approximate 9 clusters in working in G_c. These simulations results prove our Theorem 6.3.

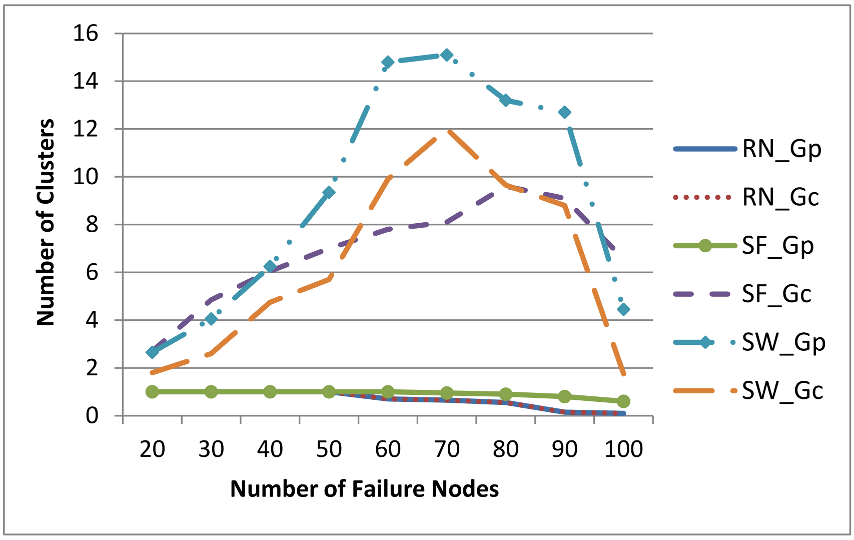


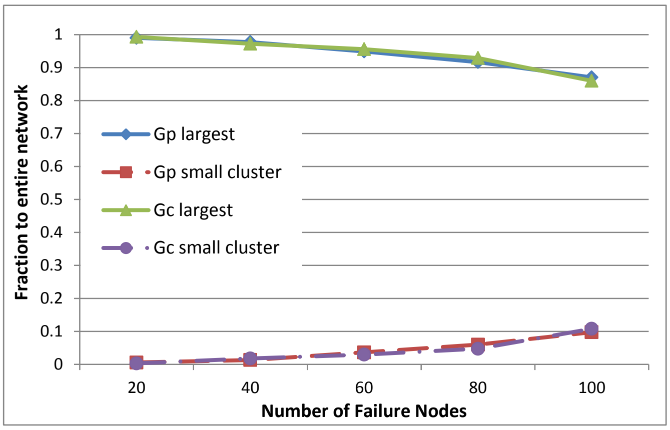
FIGURE 6.7: The number of clusters that still operate after cascading failures terminate. Coupled random networks (RN), scale-free networks (SF) and small-world networks (SW) are configured using same parameters as in Fig. 6.6.

The principle from the observations is that both the number and the size of small clusters are arrestive and have positive impacts on studying the interdependent networks.

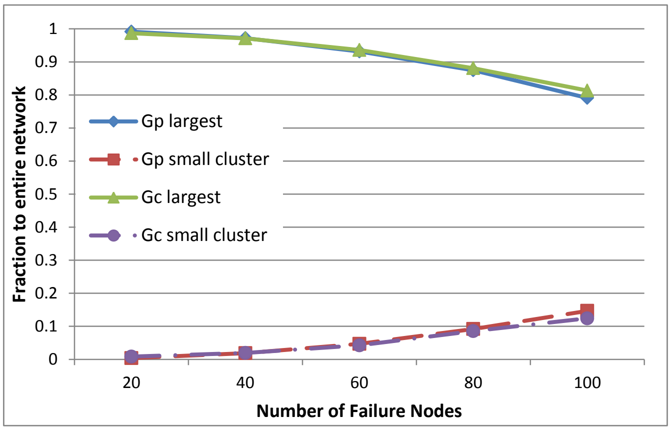
6.3.6 Various Coupling Schemes

The random interdependency usually does not occur in real interdependent systems. It is much more common that a central computational station depends on a central physical node, and vice versa [152]. Moreover, the coupled networks display some similarity in structure. An overpopulated area has many physical resources as well as many computational resources.

Fig. 6.8 shows different coupling methods in terms of 'one-to-one' model. We couple two small world networks by connecting high degree node in G_p with high degree node in G_c in Fig. 6.8(a). Compared with Fig. 6.4(f), degree-based coupling scheme is significantly more robust to random failure.



(a) Degree-based

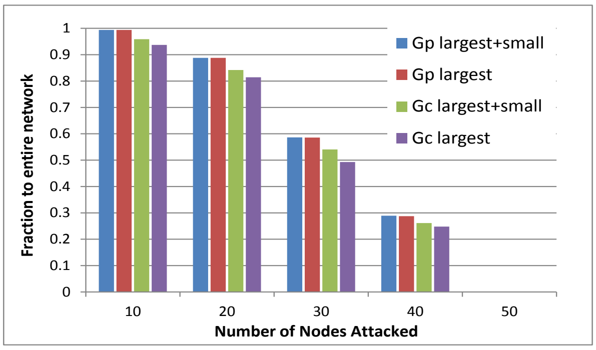


(b) Preferential-based

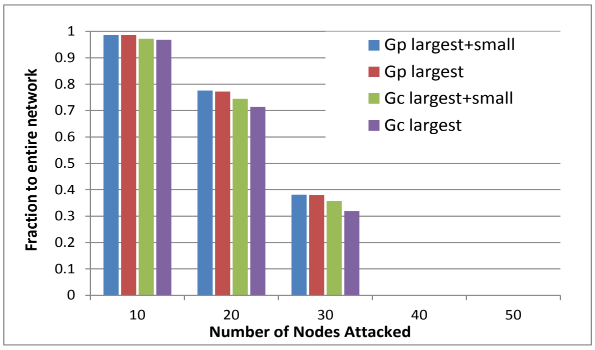
FIGURE 6.8: Different coupling methods. Both networks are small world topology, with the same size of 5000. We set the small cluster size threshold $\Delta = 10$.

We also compare preferential coupling scheme in Fig. 6.8(b), where the neighbors of coupled pair are tending to be coupled. This scheme is also much more robustness than random coupling scheme.

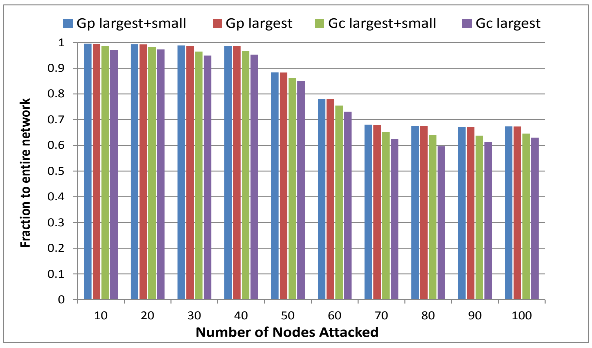
The fractions of operating small clusters in these two schemes are also significantly noticeable.



(a) Inter Degree Priority Attack



(b) Intra Degree Priority Attack



(c) Random Attack

FIGURE 6.9: The sizes of small clusters and largest cluster caused by initial failure V_f. We couple Western States Power Grid (G_p) and an synthetic scale-free network (G_c) using our 'k-n' interdependence model, with k=2, n=3. Barabási-Albert model is used to generate scale-free network with power law exponent 3. The initial network sizes are 4941 and 6000 for G_p and G_c respectively. The initial failures V_f start from G_p. We show the average results of 20 instances for each.

6.3.7 Three Attack Schemes

Previous sections study the small clusters caused by the random failure V_f. While, the targeted attacks to some critical nodes possibly happen in the real-life networks. We evaluate the sizes of small clusters of three different attack strategies:

- *Inter Degree Priority Attack*: the nodes in V_f have the highest degrees of inter links. Those nodes are targeted to be attacked in the beginning.
- *Intra Degree Priority Attack*: the nodes with highest intra degrees are prior to be attacked.
- *Random Attack*: failure nodes are randomly chosen.

We use the real Western States Power Grid and an synthetic scale-free network to construct the interdependent networks. The Western Power has 4941 nodes with 6594 edges, the synthetic scale-free network generated has the power law exponent of 3, with the initial size of 6000. Our previous 'k-n' model [3] is applied to build the interdependence. Recall $V_f \subseteq V_p$, we should first find the appropriate candidates for V_f among the Western States.

Fig. 6.9 reports the sizes of largest clusters and small clusters caused by different attack strategies. We observe no matter which attack we use, the size of functioning small clusters is almost equal to zero in G_p. In other words, if the system survives from the cascading failures, G_p only consists a largest cluster. In the meanwhile, the remaining parts fragment into massive tiny pieces. while, things are different in its counterpart G_c, where the considerable fraction of small clusters still operate. According to Fig. 6.9, approximate 3% to 5% proportion of the initial G_c survives. The interesting finding is that this proportion keeps stable for various $|V_f|$, which, at the same time, leads to quite different sizes of largest clusters.

One also can observe that the Intra Degree Priority Attack destroys the whole system very effectively. The first 40 nodes with the highest intra degrees could devastate both networks. It needs 50 nodes for Inter Degree Priority Attack strategy to complete the same mission. By contrast, the interdependent networks perform much better under Random Attack, where 40 failure nodes only affect a small part on G_p and G_c.

The principle given by these results is that different attack strategies do not affect the proportion of operating small clusters in the interdependent network. On the other hand, they indeed lead to the different vulnerabilities. Targeted attack makes the interdependent networks more fragile, especially when the attacker chooses the candidates with high intra degrees.

6.3.8 Discussion of Observations

As we found in Section 6.1, the small-world network has the highest average size of small clusters, compared with random network and scale-free network. What we observed in this section is quite similar. The coupled small-world networks generally have the highest fractions of operating small clusters. One difference is that Fig. 6.7 tells us the number of operating clusters for coupled small-world networks is higher than the counterparts. While, in the single network, this number should be the lowest, according to Table 6.1. The reason is that although the number of clusters is not as large as the others, the small-world network has significant part of clusters whose sizes are quite bigger, which are more likely to survive from cascading failures. We also find in interdependent random networks, small cluster is almost non existent. Thus, only studying giant component is enough in such networks.

6.4 Conclusions

In interdependent networks, except giant component, the small clusters play important roles. We indicate a percolation based mathematical tool to study the small clusters under different interdependence models and network topologies. Both mathematical and experimental results demonstrate that an upper bound exists for the fraction of operating small clusters. We also discover that the interdependent small world networks generally have the highest fractions of operating small clusters.

This work is helpful to find insightful characteristics of small clusters in interdependent networks. While, due to the lack of math tools, no analytical theory is yet available. This could be a possible direction we can work on.

Chapter 7

Summary and Conclusion

This chapter summarizes the research this thesis endeavours so far, and describes the plan for future work.

7.1 Conclusions on the Present Work

As described in Section 1.2, this thesis aims to study the cascading failures in interdependent cyber physical systems. The inter relation between the physical resource network and computational resource network is a key factor that affects the system robustness. We survey the existing research on CPS and interdependent networks in Chapter 2, and find the existing interdependence models are too simple to feature the real world CPS. For this reason, two novel models are proposed in Chapter 3 and 4.

The 'One-to-Multiple' model in Chapter 3 considers the fact that a physical resource node provides service for multiple computational nodes, e.g., a power station supplies hundreds of thousands computers and routers. A Ball-Bin allocation is applied to construct the inter relations between the nodes of two networks. Starting from a tiny fraction of random failure, the whole cascading failures process is shown by using the percolation theory. The results clearly demonstrate that the proposed model is fragile. Theoretically, the transition at the critical point is *first-order discontinuous*, while this only happens once the sizes of networks are infinite.

Now, we have a view on such interdependent system that a tiny failure could lead to the entire collapse. To improve the robustness, we extend to a new 'k-n' model in Chapter 4, where the inter relation is multiple to multiple. A few nodes in computational resource network are selected as the operation centers, which are used to deal with information and make decisions. Each physical resource node is operated by multiple operation centers. This model utilizes the hardware redundancy to improve the robustness. Apparently, redundancy means the increase of the cost. We study how this redundancy affects the system performance.

In Chapter 5, the smart grid is specifically analyzed in terms of two types of cascading failures that simultaneously occur. We use the load propagation model described in Section 2.4.2, and combine it with the 'k-n' model. As discussed, the tolerance parameter T, defined as the ratio of capacity to initial load, is the key factor in the analyzing of load propagation. A mathematical tool is developed to estimate the threshold T_c', below which the entire system may experience both the load propagation failures and interdependent cascading failures once suffers from the random attack. The results demonstrate that the upper bound performance exists once T goes to infinite. The interesting thing is that the fraction of survivals in power grid is always greater than that in communication network.

The previous chapters are based on the assumption that the operating node necessarily belongs to the giant component. However, we believe in interdependent networks, except the giant component, the small clusters play an important role in the cascading failure process. In Chapter 6, we explore the small cluster not only in interdependent networks, but also in single network. We first introduce an overview on how small clusters distribute in various single networks. Then we propose a percolation theory based mathematical method to study how small clusters be affected by the interdependence between two coupled networks. We prove that the upper bounds exist for both the fraction and the number of operating small clusters. Our simulation draw very insightful conclusions in this chapter.

The contributions in this work are helpful to understand the structures of cyber physical system, and understand how cascading failure affects this system. Our aim is to motivate a wide assortment of further research works on this topic.

7.2 Future Plans

In alignment with the overall thesis goal, we feel that the following research directions could be undertaken in the future.

7.2.1 Targeted Attack on Critical Nodes

This thesis work is focusing on random attack, where each node gets failed with same probability. We show that due to the coupling between two networks, the entire system is extremely vulnerable to random failure. However, initial failure in most real scenarios is not random. It may be due to the targeted attack on hub nodes with high degrees or high betweenness. It also could be attack on low degree nodes because hubs are purposely defended [25].

In Section 6.3.7, we simply compare the Inter Degree Attack, Intra Degree Attack and Random Attack under proposed 'k-n' model. It is shown that the targeted attack on high degree nodes has a dramatic effect on system robustness.

The research goal would then be to derive a general approach that analyzes the targeted attack upon different interdependence models.

7.2.2 Coupling Scenarios

The above studies assume that the pairs of interdependent nodes from two networks are randomly connected. However, in real world scenarios, it is much more common that a central computational station depends on a central physical node, and vice versa [152]. Moreover, the coupled networks display some similarity in structure. An overpopulated area has many physical resources as well as many computational resources.

The different coupling methods themselves will lead to very different results as well. We compare degree-based coupling scheme and preferential-based coupling scheme in Section 6.3.6, and find both these two are more robustness than random coupling against random failure. Therefore, future research on studying different coupling scenarios is quite necessary.

7.2.3 Electrical Features of Power Grid

In Section 2.4.3, we introduce two approaches to study the cascading failure in power grid: pure topological approach and hybrid approach. The pure topological approach focuses on studying the static failure in the network, primarily on how the connectivity changes. These work normally use graph theory, stochastic method and percolation theory to assess the power grid vulnerability.

The analysis based on hybrid approach is combining the complex network theory with the features of the electricity, such like the transmission line flow limit and voltage sag. Hybrid approach could provide fundamental insights on power systems, but on the other hand, the computation burden is heavy due to the large scale and complexity [136].

However, most of the existing work on interdependent networks is trying to indicate a general analysis tool, with only considering the network topological features, e.g., node degree and betweenness. Therefore, the physical features of networks are neglected. In future work, we would take more real physical features into account, such like power flow and node capacity. This would give more insightful results for interdependent cyber physical system.

The publications related to this thesis are listed:

- Zhen Huang, Cheng Wang, Amiya Nayak and Ivan Stojmenovic. Small Cluster in Cyber Physical Systems: Network Topology, Interdependence and Cascading Failures. Accepted by IEEE Transaction on Parallel and Distributed Systems (TPDS). 2014.

- Zhen Huang, Cheng Wang, Milos Stojmenovic and Amiya Nayak. Characterization of Cascading Failures in Interdependent Cyber-Physical Systems. Accepted by IEEE Transactions on Computers (TC). 2014.

- Zhen Huang, Cheng Wang, Sushmita Ruj, Milos Stojmenovic and Amiya Nayak. Modeling cascading failures in smart power grid using interdependent complex networks and percolation theory. IEEE Conference on Industrial Electronics and Applications, pp:1023-1028. 2013.

- Zhen Huang, Cheng Wang, Milos Stojmenovic and Amiya Nayak. Balancing System Survivability and Cost of Smart Grid Via Modeling Cascading Failures. IEEE Transactions on Emerging Topics in Computing, 1(1):45-56. 2013.

- Zhen Huang, Cheng Wang, Amiya Nayak and Ivan Stojmenovic. Cascading Failures in Smart Grid: Joint Effect of Load Propagation and Interdependence. Submitted to IEEE Transactions on Industrial Informatics. 2014.

Bibliography

[1] . URL https://gephi.org/.

[2] . URL http://wiki.gephi.org/index.php/Datasets.

[3] Zhen Huang, Cheng Wang, Milos Stojmenovic, and Amiya Nayak. Balancing system survivability and cost of smart grid via modeling cascading failures. *IEEE Transactions on Emerging Topics in Computing*, 1(1):45–56, 2013.

[4] Duncan J. Watts and Steven H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, 1998.

[5] Fumin Zhang, Zhenwu Shi, and Shayok Mukhopadhyay. Robustness analysis for battery-supported cyber-physical systems. *ACM Transactions on Embedded Computing Systems (TECS)*, 12(3):69, 2013.

[6] Marius C Bujorianu and Howard Barringer. An integrated specification logic for cyber-physical systems. In *14th IEEE International Conference on Engineering of Complex Computer Systems*, pages 291–300. IEEE, 2009.

[7] Jonathan Fink, Alejandro Ribeiro, and Vijay Kumar. Robust control for mobility and wireless communication in cyber–physical systems with application to robot teams. *Proceedings of the IEEE*, 100(1):164–178, 2012.

[8] Kyoung-Dae Kim and PR Kumar. Cyber–physical systems: A perspective at the centennial. *Proceedings of the IEEE*, 100(13):1287–1308, 2012.

[9] Adilson E Motter and Ying-Cheng Lai. Cascade-based attacks on complex networks. *Physical Review E*, 66(6):065102, 2002.

[10] Paolo Crucitti, Vito Latora, and Massimo Marchiori. Model for cascading failures in complex networks. *Physical Review E*, 69(4), 2004.

[11] Ryan Kinney, Paolo Crucitti, Reka Albert, and Vito Latora. Modeling cascading failures in the north american power grid. *The European Physical Journal B-Condensed Matter and Complex Systems*, 46(1):101–107, 2005.

[12] Wendi Rabiner Heinzelman, Anantha Chandrakasan, and Hari Balakrishnan. Energy-efficient communication protocol for wireless microsensor networks. In *Proc. the 33rd Annual Hawaii International Conference on System Sciences*, pages 1–10. IEEE, 2000.

[13] Wen-Xu Wang and Guanrong Chen. Universal robustness characteristic of weighted networks against cascading failure. *Physical Review E*, 77(2): 026101, 2008.

[14] Ian Dobson, Benjamin A Carreras, and David E Newman. A loading-dependent model of probabilistic cascading failure. *Probability in the Engineering and Informational Sciences*, 19(01):15–32, 2005.

[15] Sergey V. Buldyrev, Roni Parshani, Gerald Paul, H. Eugene Stanley, and Shlomo Havlin. Catastrophic cascade of failures in interdependent networks. *Nature*, 464:1025–1028, 2010.

[16] URL http://spectrum.ieee.org/energywise/energy/the-smarter-grid/a-postmortem-on-indias-blackout.

[17] Dung T Nguyen, Yilin Shen, and My T Thai. Detecting critical nodes in interdependent power networks for vulnerability assessment. *IEEE Trans. Smart Grid*, 4(1):151–159, 2013.

[18] Zhen Huang, Cheng Wang, Sushmita Ruj, Milos Stojmenovic, and Amiya Nayak. Modeling cascading failures in smart power grid using interdependent complex networks and percolation theory. *IEEE ICIEA*, pages 1023 – 1028, 2013.

[19] Jianxi Gao, Sergey V. Buldyrev, H. Eugene Stanley, and Shlomo Havlin. Networks formed from interdependent networks. *Nature Physics*, 8:40–48, 2012.

[20] Osman Yagan, Dajun Qian, Junshan Zhang, and Douglas Cochran. Optimal allocation of interconnecting links in cyber-physical systems: Interdependence, cascading failures and robustness. *IEEE Transactions on Parallel and Distributed Systems*, 23(9):1708–1720, 2012.

[21] Mark.E.J. Newman. *Networks: An Introduction*. Oxford University Press, 2010.

[22] Sergey V. Buldyrev, Nathaniel W. Shere, and Gabriel A. Cwilich. Interdependent networks with identical degrees of mutually dependent nodes. *Phys. Rev. E*, 83:016112, Jan 2011. doi: 10.1103/PhysRevE.83.016112. URL http://link.aps.org/doi/10.1103/PhysRevE.83.016112.

[23] Jia Shao, Sergey V. Buldyrev, Shlomo Havlin, and H. Eugene Stanley. Cascade of failures in coupled network systems with multiple support-dependent relations. *CoRR*, 2010.

[24] Hongda Xiao and Edmund M Yeh. Cascading link failure in the power grid: A percolation-based analysis. In *Proc. IEEE ICC 2011*, pages 1–6. IEEE, 2011.

[25] Xuqing Huang, Jianxi Gao, Sergey V Buldyrev, Shlomo Havlin, and H Eugene Stanley. Robustness of interdependent networks under targeted attack. *Physical Review E*, 83(6):065101, 2011.

[26] Gaogao Dong, Lixin Tian, Di Zhou, Ruijin Du, Jiang Xiao, and H. E. Stanley. Robustness of n interdependent networks with partial support-dependence relationship. *Europhysics Letters*, 102(6), 2013.

[27] Gaogao Dong, Jianxi Gao, Ruijin Du, Lixin Tian, H. Eugene Stanley, and Shlomo Havlin. Robustness of network of networks under targeted attack. *Phys. Rev. E*, 87(5), 2013.

[28] Xian Chen, Hieu Dinh, and Bing Wang. Cascading failures in smart grid - benefits of distributed generation. In *SmartGridComm*, pages 73–78, 2010.

[29] R. Pfitzner, K. S. Turitsyn, and M. Chertkov. Controlled tripping of overheated lines mitigates power outages. *CoRR*, abs/1104.4558, 2011.

[30] Sakshi Pahwa, Amelia Hodges, Caterina Scoglio, and Sean Wood. Topological analysis of the power grid and mitigation strategies against cascading failures. *4th IEEE Systems Conference*, 2010.

[31] Christopher Zimmer and Frank Mueller. Fault tolerant network routing through software overlays for intelligent power grids. In *ICPADS*, pages 542–549. IEEE, 2010.

[32] Y. Yuan, Z. Li, and K. Ren. Modeling load redistribution attacks in power systems. *IEEE Trans. Smart Grid*, 2(2):382–390, 2011.

[33] P. Hines, K. Balasubramaniam, and E. Cotilla-Sanchez. Cascading failures in power grids. *IEEE Potentials*, 28(5):24–30, 2009.

[34] Enrico Zio and Giovanni Sansavini. Component criticality in failure cascade processes of network systems. *Risk Analysis*, 31(8):1196–1210, 2011.

[35] Liang Zhao, Kwangho Park, and Ying-Cheng Lai. Attack vulnerability of scale-free networks due to cascading breakdown. *Phys. Rev. E*, 70(3), 2004.

[36] Ettore Bompard, Di Wu, and Fei Xue. Structural vulnerability of power systems: A topological approach. *Electric Power Systems Research*, 81(7): 1334–1340, 2011.

[37] S. Kadloor and N. Santhi. Understanding cascading failures in power grids. *CoRR*, abs/1011.4098, 2010.

[38] R. Kinney, P. Crucitti, R. Albert, and V. Latora. Modeling cascading failures in the north american power grid. *Eur. Phys. J. B*, 46:101–107, 2005.

[39] Guo Chen, Zhao Yang Dong, David J. Hill, Guo Hua Zhang, and Ke Qian Hua. Attack structural vulnerability of power grids: A hybrid approach based on complex networks. *Physica A: Statistical Mechanics and its Applications*, 389(3):595–603, 2010.

[40] P. Hines, E. Cotilla-Sanchez, and S.Blumsack. Do topological models provide good information about vulnerability in electric power networks? *Chaos: An interdisciplinary journal of non-linear science*, 20(3):24–30, 2010.

[41] Giuliano Andrea Pagani and Marco Aiello. The power grid as a complex network: A survey. *Physica A: Statistical Mechanics and its Applications*, 392(11):2688–2700, 2013.

[42] Rafael Jesus Falcon Martinez. Towards fault reactiveness in wireless sensor networks with mobile carrier robots. *Doctoral Thesis, University of Ottawa*, 2012.

[43] X. Li, G. Fletcher, A. Nayak, and I. Stojmenovic. Placing sensors for area coverage in a complex environment by a team of robots. *ACM Transactions on Sensor Networks. To appear*, 2013.

[44] Zubair Md. Fadlullah, Duong Minh Quan, Nei Kato, and Ivan Stojmenovic. Gtes: An optimized game-theoretic demand side management scheme for smart grid. *IEEE Systems Journal. To appear*, 2013.

[45] Xi Fang, Satyajayant Misra, Guoliang Xue, and Dejun Yang. Smart grid-the new and improved power grid: A survey. *IEEE Communications Surveys and Tutorials*, 14(4):944–980, 2012.

[46] Zhen Huang. On reputation and data-centric misbehavior detection mechanisms for vanet. *Master Thesis, University of Ottawa*, 2011.

[47] Zhen Huang, Sushmita Ruj, Marcos A. Cavenaghi, Milos Stojmenovic, and Amiya Nayak. A social network approach to trust management in vanets. *Journal of Peer-to-Peer Networking and Applications*, pages 1–14, 2012.

[48] Zhifang Wang, Anna Scaglione, and Robert J Thomas. Generating statistically correct random topologies for testing smart grid communication and control networks. *IEEE Transactions on Smart Grid*, 1(1):28–39, 2010.

[49] Patricia Derler, Edward A Lee, and A Sangiovanni Vincentelli. Modeling cyber–physical systems. *Proceedings of the IEEE*, 100(1):13–28, 2012.

[50] Bob Emmerson. M2m: the internet of 50 billion devices. *WinWin Magazine*, 2010.

[51] Geng Wu, S. Talwar, K. Johnsson, N. Himayat, and K.D. Johnson. M2m: From mobile to embedded internet. *IEEE Communications Magazine*, 49 (4):36–43, 2011.

[52] Rongxing Lu, Xu Li, Xiaohui Liang, Xuemin Shen, and Xiaodong Lin. Grs: The green, reliability, and security of emerging machine to machine communications. *IEEE Communications Magazine*, 49(4):28–35, 2011.

[53] A. Gallais, J. Carle, D. Simplot-Ryl, and I. Stojmenovic. Localized sensor area coverage with low communication overhead. *IEEE Transactions on Mobile Computing*, 7(5):661–672, 2008.

[54] Dusit Niyato, Lu Xiao, and Ping Wang. Machine-to-machine communications for home energy management system in smart grid. *IEEE Communications Magazine*, 49(4):53–59, 2011.

[55] Zubair Md. Fadlullah, Mostafa M. Fouda, Nei Kato, Akira Takeuchi, Noboru Iwasaki, and Yousuke Nozaki. Toward intelligent machine-to-machine communications in smart grid. *IEEE Communications Magazine*, 49(4):60–65, 2011.

[56] Huai-Lei Fu, Hou-Chun Chen, Phone Lin, and Yuguang Fang. Energy-efficient reporting mechanisms for multi-type real-time monitoring in machine-to-machine communications networks. *IEEE INFOCOM*, pages 136–144, 2012.

[57] Ivan Stojmenovic. Large scale cyber-physical systems: Distributed actuation, in-network processing and machine-to-machine communications. *EUROMICRO/IEEE Workshop on Embedded and Cyber Physical Systems*, pages 21–24, 2013.

[58] Rafael Jesus Falcon Martinez. Towards fault reactiveness in wireless sensor networks with mobile carrier robots. *Doctoral Thesis, University of Ottawa*, 2012.

[59] Xu Li, H. Frey, N. Santoro, and I. Stojmenovic. Strictly localized sensor self-deployment for optimal focused coverage. *IEEE Trans. on Mobile Computing*, 10(11):1520–1533, 2011.

[60] Guiling Wang, Guohong Cao, P. Berman, and T.F. La Porta. Bidding protocols for deploying mobile sensors. *IEEE Transactions on Mobile Computing*, 6(5):563–576, 2007.

[61] Ameer A. Abbasi, Mohamed F. Younis, and Uthman A. Baroudi. Recovering from a node failure in wireless sensor-actor networks with minimal topology changes. *IEEE Trans. on Vehicular Technology*, 62(1):256–271, 2013.

[62] Martín López-Nores et al. Application-level assessment of approaches to coordinate node mobility in wireless sensor and actor networks. *Computer Communications*, 33(7):860–867, 2010.

[63] Melodia T., Pompili D., Gungor V.C., and Akyildiz I.F. Communication and coordination in wireless sensor and actor networks. *IEEE Tran. on Mobile Computing*, 6(10):1116–1129, 2007.

[64] Mezei I., Malbaša V., and Stojmenovic I. Communication aspects of robot-robot coordination for single task single robot wireless multi-hop networks. *IEEE Robotics and Automation Magazine*, 17(4):63–69, 2010.

[65] Chih-Yung Chang, Jang-Ping Sheu, Yu-Chieh Chen, and Sheng-Wen Chang. An obstacle-free and power-efficient deployment algorithm for wireless sensor networks. *IEEE Tran. on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 39(4):795–806, 2009.

[66] Wu C., Tewolde G.S., Sheng W., Xu B., and Y. Wang. Distributed multi-actuator control for workload balancing in wireless sensor and actuator networks. *IEEE Trans. on Automatic Control*, 56(10):2462–2467, 2011.

[67] Zhenjiang Li, Yunhao Liu, Mo Li, Jiliang Wang, and Zhichao Cao. Exploiting ubiquitous data collection for mobile users in wireless sensor networks. *IEEE Transactions on Parallel and Distributed Systems*, 24(2):312–326, 2013.

[68] R. Falcon, Hai Liu, A. Nayak, and I. Stojmenovic. Controlled straight mobility and energy-aware routing in robotic wireless sensor networks. *IEEE 8th International Conference on Distributed Computing in Sensor Systems (DCOSS)*, pages 150–157, 2012.

[69] R.C. Shah and J.M. Rabaey. Energy aware routing for low energy ad hoc sensor networks. *IEEE Wireless Communications and Networking Conference*, pages 350–355, 2002.

[70] I. Stojmenovic and Xu Lin. Power-aware localized routing in wireless networks. *IEEE Transactions on Parallel and Distributed Systems*, 12(11):1122–1133, 2001.

[71] Xu Li, Shuo Yan, Chendong Xu, Amiya Nayak, and Ivan Stojmenovic. Localized delay-bounded and energy-efficient data aggregation in wireless sensor and actor networks. *Wireless Communications and Mobile Computing*, 11 (12):1063–1617, 2011.

[72] Ming Ma and Yuanyuan Yang. Adaptive triangular deployment algorithm for unattended mobile sensor networks. *IEEE Transactions on Computers*, 56(7):946–958, 2007.

[73] Guang Tan, S.A. Jarvis, and A.M. Kermarrec. Connectivity-guaranteed and obstacle-adaptive deployment schemes for mobile sensor networks. *The 28th International Conference on Distributed Computing Systems*, pages 429–437, 2008.

[74] Chih-Yung Chang, Chih-Yung Chang, Yu-Chieh Chen, and Hsu-Ruey Chang. Obstacle-resistant deployment algorithms for wireless sensor networks. *IEEE Trans. on Vehicular Technology*, 58(6):2925–2941, 2009.

[75] Fletcher G., Li X., Nayak A., and Stojmenovic I. Back-tracking based sensor deployment by a robot team. *IEEE Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, pages 385–393, 2010.

[76] Yuanteng Pei and M.W. Mutka. Stars: Static relays for remote sensing in multirobot real-time search and monitoring. *IEEE Transactions on Parallel and Distributed Systems*, 24(10):2079–2089, 2013.

[77] Zhen Huang. On reputation and data-centric misbehavior detection mechanisms for vanet. *Master Thesis, University of Ottawa*, 2011.

[78] Yousef-Awwad Daraghmi, Ivan Stojmenovic, and Chi-Wei Yi. Forwarding methods in data dissemination and routing protocols for vehicular ad hoc networks. *IEEE Network*, 27(6):74–79, 2013.

[79] Ivan Stojmenovic. A general framework for broadcasting in static to highly mobile wireless ad hoc, sensor, robot and vehicular networks. *IEEE 18th International Conference on Parallel and Distributed Systems*, pages 8–13, 2012.

[80] Junliang Liu, Zheng Yang, and Ivan Stojmenovic. Receiver consensus: On-time warning delivery for vehicular ad-hoc networks. *IEEE 32nd International Conference on Distributed Computing Systems (ICDCS)*, pages 386–395, 2012.

[81] Junliang Liu, Zheng Yang, and Ivan Stojmenovic. Receiver consensus: On-time warning delivery for vehicular ad-hoc networks. *IEEE Trans. on Emerging Topics in Computing*, 1(1):57–68, 2013.

[82] T. Nadeem, P. Shankar, and L. Iftode. A comparative study of data dissemination models for vanets. *International Conference on Mobile and Ubiquitous Systems: Networking and Services*, pages 1–10, 2006.

[83] Booysen M.J., Gilmore J.S., Zeadally S., and van Rooyen G.J. Machine-to-machine (m2m) communications in vehicular networks. *KSII Transactions on Internet and Information Systems*, 6(2):529–546, 2012.

[84] Giorgio Callandriello, Panos Papadimitratos, Jean-Pierre Hubaux, and Antonio Lioy. Efficient and robust pseudonymous authentication in vanet. pages 19–28, 2007.

[85] Yixin Jiang, Minghui Shi, Xuemin Shen, and Chuang Lin. Bat: A robust signature scheme for vehicular networks using binary authentication tree. *IEEE Transactions on Wireless Communications*, 8(4):1974–1983, 2009.

[86] M. Raya, P. Papadimitratos, V.D. Gligor, and J. P. Hubaux. On data-centric trust establishment in ephemeral ad hoc networks. *IEEE 27th Conference on Computer Communications*, 2008.

[87] Zhen Huang, S. Ruj, M. Cavenaghi, and A. Nayak. Limitations of trust management schemes in vanet and countermeasures. *IEEE 22nd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, pages 1228–1232, 2011.

[88] Xiao Fan Wang and Guanrong Chen. Complex networks: small-world, scale-free and beyond. *IEEE Circuits and Systems Magazine*, 3(1):6–20, 2003.

[89] Luciano da F. Costa, Francisco A. Rodrigues, Gonzalo Travieso, and P. R. Villas Boas. Characterization of complex networks: A survey of measurements. *Advences in Physics*, 56(1):167–242, 2007.

[90] Juliette Stehlé, Nicolas Voirin, Alain Barrat, Ciro Cattuto, Lorenzo Isella, Jean-Francois Pinton, Marco Quaggiotto, Wouter Van den Broeck, Corinne Régis, Bruno Lina, and Philippe Vanhems. High-resolution measurements of face-to-face contact patterns in a primary school. *PLoS ONE*, 6(8), 2011.

[91] Steven H. Strogatz. Exploring complex networks. *Nature*, 410:268–276, 2001.

[92] Stanley Milgram. The small-world problem. *Psychology Today*, 1(1):61–67, 1967.

[93] Jure Leskovec and Eric Horvitz. Planetary-scale views on an instant-messaging network. *arXiv:0803.0939*, 2008.

[94] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45:167–256, 2003.

[95] R van der Hofstad. Random graphs and complex networks.

[96] Reuven Cohen and Shlomo Havlin. *Complex Networks: Structure, Robustness and Function*. Cambridge University Press, 2010.

[97] Paul Erdös and Alfréd Rényi. On the strength of connectedness of a random graph. *Acta Mathematica Hungarica*, 12(1-2):261–267, 1961.

[98] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[99] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74:47–97, 2002.

[100] Kimmo Soramäkia, Morten L. Bechb, Jeffrey Arnoldb, Robert J. Glassc, and Walter E. Beyelerc. The topology of interbank payment flows. *Physica A: Statistical Mechanics and its Applications*, 379(1):317–333, 2007.

[101] P. L. Krapivsky, S. Redner, and F. Leyvraz. Connectivity of growing random networks. *Phys. Rev. Lett*, 85(21):4629–4632, 2000.

[102] Seth A. Marvel, Travis Martin, Charles R. Doering, David Lusseau, , and M. E. J. Newman. The small-world effect is a modern phenomenon. *arXiv:1310.2636*, 2013.

[103] A. Barrat and M. Weigt. On the properties of small-world network models. *European Physical Journal B*, 13(3):547–560, 2000.

[104] R. Parshani, V. Buldyrev, and S. Havlin. The critical effect of dependency groups on the function of networks. *Proc. National Academy of Sciences*, 108, 2011.

[105] Di Zhou, Jianxi Gao, H. Eugene Stanley, and Shlomo Havlin. Percolation of partially interdependent scale-free networks. *Physical Review E*, 87(5), 2013. doi: http://dx.doi.org/10.1103/PhysRevE.87.052812.

[106] S. Havlin Shlomo Havlin J. Gao, S. V. Buldyrev and H. E. Stanle. Robustness of a network of networks. *Phys. Rev. Lett.*, 107(19), 2011.

[107] Charles D. Brummitt, Raissa M. D'Souza, , and E. A. Leicht. Suppressing cascades of load in interdependent networks. *PNAS*, 109(12):680–689, 2012.

[108] Xuqing Huang, Shuai Shao, Huijuan Wang, Sergey V. Buldyrev, H. Eugene Stanley, and Shlomo Havlin. The robustness of interdependent clustered networks. *Europhysics Letters*, 101(1), 2013.

[109] Seung-Woo Son, Golnoosh Bizhani, Claire Christensen, Peter Grassberger, and Maya Paczuski. Percolation theory on interdependent networks based on epidemic spreading. *Europhysics Letters*, 97(1), 2012.

[110] Sushmita Ruj and Arindam Pal. Analyzing cascading failures in smart grids under random and targeted attacks. 2014.

[111] Dror Y. Kenett, Jianxi Gao, Xuqing Huang, Shuai Shao, Irena Vodenska, Sergey V. Buldyrev, Gerald Paul, H. Eugene Stanley, and Shlomo Havlin. Network of interdependent networks: Overview of theory and applications. *Networks of Networks: The Last Frontier of Complexity Understanding Complex Systems*, pages 3–36, 2014.

[112] R. Baldick, B. Chowdhury, I. Dobson, Zhaoyang Dong, Bei Gou, D. Hawkins, Zhenyu Huang, Manho Joung, Janghoon Kim, D. Kirschen, S. Lee, Fangxing Li, Juan Li, Zuyi Li, Chen-Ching Liu, Xiaochuan Luo, L. Mili, S. Miller, M. Nakayama, M. Papic, R. Podmore, J. Rossmaier, K. Schneider, Hongbin Sun, Kai Sun, D. Wang, Zhigang Wu, Liangzhong Yao, Pei Zhang, Wenjie

Zhang, and Xiaoping Zhang. Vulnerability assessment for cascading failures in electric power systems. *IEEE/PES Power Systems Conference and Exposition*, pages 1–9, 2009.

[113] M. Vaiman, K. Bell, Y. Chen, B. Chowdhury, I. Dobson, P. Hines, M. Papic, S. Miller, and P. Zhang. Risk assessment of cascading outages: Methodologies and challenges. *IEEE Trans. on Power Systems*, 27(2):631–641, 2012.

[114] Zhifang Wang, R.J. Thomas, and A. Scaglione. Generating random topology power grids. *41st Hawaii International Conference on System Sciences*, page 183, 2008.

[115] M. Rosas-Casals. Power grids as complex networks: Topology and fragility. *Complexity in Engineering*, pages 21–26, 2010.

[116] Pingping Han and Ming Ding. Analysis of cascading failures in small-world power grid. *International Journal of Energy Science*, 1(2):99–104, 2011.

[117] Réka Albert, István Albert, and Gary L. Nakarado. Structual vulnerability of the north american power grid. *Phys. Rev. E*, 69(2), 2004.

[118] V. Rosato, S. Bologna, and F. Tiriticco. Topological properties of high-voltage electrical transmission networks. *Electric Power Systems Research*, 77(2):99–105, 2007.

[119] P. Hines, S. Blumsack, E. Cotilla Sanchez, and C. Barrows. The topological and electrical structure of power grids. *43rd Hawaii International Conference on System Sciences*, pages 1–10, 2010.

[120] L. A. N. Amaral, A. Scala, M. Barthélémy, and H. E. Stanley. Classes of small-world networks. *PNAS*, 97(21):11149–11152, 2000.

[121] Paolo Crucitti, Vito Latora, and Massimo Marchiori. A topological analysis of the italian electric power grid. *Physica A: Statistical Mechanics and its Applications*, 338(1-2):92–97, 2004.

[122] David P. Chassin and Christian Posse. Evaluating north american electric grid reliability using the barabási–albert network model. *Physica A: Statistical Mechanics and its Applications*, 355(2-4):667–677, September 2005.

[123] E. Cotilla-Sanchez, P.D.H. Hines, C. Barrows, and S. Blumsack. Comparing the topological and electrical structure of the north american electric power infrastructure. *IEEE Systems Journal*, 6(4):616–626, 2012.

[124] Z.J. Bao, Y.J. Cao, L.J. Ding, Z.X. Han, and G.Z. Wang. Dynamics of load entropy during cascading failure propagation in scale-free networks. *Physics Letters A*, 372(36):5778–5782, 2008.

[125] Liang Zhao, Kwangho Park, Ying-Cheng Lai, and Nong Yek. Tolerance of scale-free networks against attack-induced cascades. *Phys. Rev. E*, 72(2), 2005.

[126] Jianwei Wang, Lili Rong, Liang Zhang, and Zhongzhi Zhang. Attack vulnerability of scale-free networks due to cascading failures. *Elsevier Physica A: Statistical Mechanics and its Applications*, 387(26):6671–6678, 2008.

[127] Jianwei Wang and Lili Rong. Cascade-based attack vulnerability on the us power grid. *Safety Science*, 47(10):1332–1336, 2009.

[128] Wenkai Wang, Yan Sun, and Haibo He. Topological analysis of cascading failures in bay area power grid. *IEEE Power and Energy Society General Meeting*, pages 1–9, 2012.

[129] Wenkai Wang, Qiao Cai, Yan Sun, and Haibo He. Risk-aware attacks and catastrophic cascading failures in u.s. power grid. *2011 IEEE Global Telecommunications Conference*, pages 1–6, 2011.

[130] M. Rosas-Casals, S. Valverde, and R. Solé. Topological vulnerability of the european power grid under errors and attacks. *International Journal of Bifurcation and Chaos*, 17(7):2465–2475, 2007.

[131] Li Fu, Wenjie Huang, Sheng Xiao, and Yuan Li. Vulnerability assessment for power grid based on small-world topological model. *IEEE Power and Energy Engineering Conference (APPEEC)*, pages 1–4, 2010.

[132] Lu Zongxiang, Meng Zhongwei, and Zhou Shuangxi. Cascading failure analysis of bulk power system using small-world network model. *IEEE Conference on Probabilistic Methods Applied to Power Systems*, pages 635–640, 2004.

[133] Zhifang Wang, Anna Scaglione, and Robert J. Thomas. A markov-transition model for cascading failures in power grids. *45th Hawaii International Conference on System Science*, pages 2115–2124, 2012.

[134] E. Bompard, R. Napoli, and F. Xue. Extended topological approach for the assessment of structural vulnerability in transmission networks. *IET Generation, Transmission and Distribution*, 4(6):716–724, 2010.

[135] Quan Chen and Lamine Mili. Composite power system vulnerability evaluation to cascading failures using importance sampling and antithetic variates. *IEEE Trans. on Power Systems*, 28(3):2321–2330, 2013.

[136] Yihai Zhu, Jun Yan, Yan Sun, and Haibo He. Revealing cascading failure vulnerability in power grids using risk-graph. *IEEE Trans. on Parallel and Distributed Systems*, 2014. doi: 10.1109/TPDS.2013.2295814.

[137] Z.J. Bao, Y.J. Cao, G.Z. Wang, and L.J. Ding. Analysis of cascading failure in electric grid based on power flow entropy. *Physics Letters A*, 373(34): 3032–3040, 2009.

[138] S. Arianos, E. Bompard, A. Carbone, and F. Xue. Power grid vulnerability: A complex network approach. *Chaos*, 19(1), 2009.

[139] Michael Chertkov, Feng Pan, and Mikhail G. Stepanov. Predicting failures in power grids: The case of static overloads. *IEEE Trans. Smart Grid*, 2(1): 162–172, 2011.

[140] Margaret J. Eppstein and Paul D. H. Hines. A "random chemistry" algorithm for identifying collections of multiple contingencies that initiate cascading failure. *IEEE Trans. on Power Systems*, 27(3):1698–1705, 2012.

[141] Janghoon Kim, Kevin R. Wierzbicki, Ian Dobson, and Rodney C. Hardiman. Estimating propagation and distribution of load shed in simulations of cascading blackouts. *IEEE Systems Journal*, 6(3):548–557, 2012.

[142] K. Moslehi and R. Kumar. A reliability perspective of the smart grid. *IEEE Transactions on Smart Grid*, 1:57–64, June 2010. doi: 10.1109/TSG.2010.2046346.

[143] S. N. Dorogovtsev and J. F. F. Mendes. *Evolution of Networks: From Biological Nets to the Internet and WWW.* Oxford University Press, 2003.

[144] Yang Nan, Liu Wenying, and Guo Wei. Study on scale-free characteristic on propagation of cascading failures in power grid. *IEEE Energytech*, pages 1–5, 2011.

[145] . Avaliable at : https://www.smartgrid.gov/the_smart_grid.

[146] . Avaliable at : http://www.intelligentutility.com/article/13/01/smart-grid-distribution-control-center-and-operations.

[147] . Avaliable at : http://www.csc.com/utilities/insights/76161-smart_grid_technology_the_muscle_behind_the_csc_smart_grid.

[148] E. A. Leicht and R. M. D'Souza. Percolation on interacting networks. arxiv.org/abs/0907.0894, 2009.

[149] Yuan Wang, Ahmed Barnawi, Rodrigo Fernandes de Mello, and Ivan Stojmenovic. Localized ant colony of robots for redeployment in wireless sensor networks. *Journal of Multiple-Valued Logic and Soft Computing*, 23(1-2):35–51, 2014.

[150] Xu Li, Rafael Falcon, Amiya Nayak, and Ivan Stojmenovic. Servicing wireless sensor networks by mobile robots. *IEEE Communications Magazine*, 50 (7):147–154, 2012.

[151] URL http://wiki.gephi.org/index.php/Datasets.

[152] R. Parshani, C. Rozenblat, D. Ietri, C. Ducruet, and S. Havlin. Inter-similarity between coupled networks. *Europhysics Letters*, 92(6), 2011.