Spring 1-1-2019

Optimizing Interconnectivity among Networks under Attacks

Pankaz Das
Doctoral Student, Electrical and Computer Engineering

Pankaz Das

Candidate

Electrical and Computer Engineering

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Dr. Majeed M. Hayat ,Chairperson

Dr. Francesco Sorrentino

Dr. Manel Martinez-Ramon

Dr. Balu Santhanam

Dr. Mahshid Rahnamay-Naeini (USF, FL, USA)

Optimizing Interconnectivity among Networks under Attacks

by

Pankaz Das

B.Sc., Electronics & Communication Engineering, KUET, BD, 2010
M.Sc., Electronics Engineering, Kookmin University, S. Korea, 2013
M.S., Electrical Engineering, University of New Mexico, USA, 2018

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Engineering

The University of New Mexico

Albuquerque, New Mexico

December, 2018

Dedication

To my parents, Kazal Das and Parul Das. I am standing on their shoulders.

Acknowledgments

First and foremost, I would like to thank God Almighty for giving me the strength, knowledge, perseverance, ability and opportunity to complete my Ph.D. successfully. Without his blessings, this achievement would not have been possible.

I would like to express my deepest gratitude to my advisor, Professor Majeed M. Hayat, for his excellent guidance and continuous support throughout my graduate studies. He has supported from all aspects including generating ideas, formulating problems, writing and presenting research outcomes, to name a few. I have leaned a lot from him and I am really grateful to him for taking me as his student while I was passing a critical time (financially and mentally) at UNM.

I would like to thank my committee members Dr. Manel Martinez-Ramon, Dr. Balu Santhanam, Dr. Francesco Sorrentino, Dr. Mahshid Rahnamay-Naeini (USF). In particular, I would like to thank Dr. Francesco Sorrentino and Dr. Mahshid Rahnamay-Naeini who helped me a lot with their guidance, invaluable discussions and constructive comments on my works. My special thanks to my friend and colleague Rezoan A. Shurvo for his tremendous encouragement and suggestions for last 3 years. Without his support the journey would have been a very difficult one. I am also delighted to have some friends who made this Ph.D. life very eventful: Muntasir A. Kabir, Sheikh A. Bakir, Md. Mottaleb Hossian, Manish Battrai, Abraham P. Vinod. I would like to acknowledge the helpful faculty and stuff at UNM. This work was supported by the Defense Threat Reduction Agency's Basic Research Program under grant No. HDTRA1-13-1-0020 and NSFs grant No. 2GA11 NSF CRISP.

I am really indebted to my cousin Ashutosh Das for taking care of all my family issues in Bangladesh during my doctoral studies. I would like to thank my sister Konika Das, cousins, brother-in-law, uncles and the people of my village (Dengapara) and my relatives who have been encouraging me since my childhood.

Finally, I thank my parents for everything just I have. I thank my father, my childhood motivator and financial supporter, for instilling in me the dream of becoming a highly educated person. I thank my mother, my childhood instructor, for her strict regulation during my childhood days. Without their encouragement and strict disciplinary actions, it would have been completely impossible to come this far and I would have lost like many others in my village.

Optimizing Interconnectivity among Networks under Attacks

by

Pankaz Das

B.Sc., Electronics & Communication Engineering, KUET, BD, 2010

M.Sc., Electronics Engineering, Kookmin University, S. Korea, 2013
M.S., Electrical Engineering, University of New Mexico, USA, 2018

Ph.D., Engineering, University of New Mexico, 2018

Abstract

Networks may need to be interconnected for various reasons such as inter- organizational communication, redundant connectivity, increasing data-rate and minimizing delay or packet-loss, etc. However, the trustworthiness of an added interconnection link cannot be taken for granted due to the presence of attackers who may compromise the security of an interconnected network by intercepting the interconnections. Namely, an intercepted interconnection link may not be secured due to the data manipulations by attackers. In the first part of this dissertation, the number of interconnections between the two networks is optimized for maximizing the data-rate and minimizing the packet-loss under the threat of security attacks. The optimization of the interconnectivity considering the security attack is formulated using a rate-distortion optimization setting, as originally introduced by Claude E. Shannon in the information theory. In particular, each intercepted interconnection is modeled as a noisy communication channel where the attackers may manipulate the data by flipping and erasing of data bits, and then the total capacity for any given number of interconnections is calculated. By exploiting such formulation, the optimal number of interconnections between two networks is found under network administrators data-rate and packet-loss requirement, and most importantly, without compromising the data security. It is concluded analytically and verified by simulations under certain conditions, increasing interconnections beyond an optimal number would not be beneficial concerning the data-rates and packet-loss. In the second part of this dissertation, the vulnerability of the interconnected network is analyzed by a probabilistic model that maps the intensity of physical attacks to network component failure distributions. In addition, assuming the network is susceptible to the attack propagation, the resiliency of the network is modeled by the influence model and epidemic model. Finally, a stochastic model is proposed to track the node failure dynamics in a network considering dependency with power failures. Besides, the cascading failure in the power grid is analyzed with a data-driven model that reproduces the evolution of power-transmission line failure in power grids. To summarize, the optimal interconnectivity among networks is analyzed under security attacks, and the dynamic interactions in an interconnected network are investigated under various physical and logical attacks.

The proper application of this work would add minimum number of inter-network connections between two networks without compromising the data security. The optimal number interconnections would meet network administrator's requirement and minimize cost (both security and monetary) associated with unnecessary connections. This work can also be used to estimate the reliability of a communication network under different types of physical attacks independently and also by incorporating the dynamics of power failures.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview and motivation

Modern communication networks are interdependent due to the service or support (both infrastructural and logical) they receive from one another. These networks are interconnected through the communication gateways, which we termed asthe multilevel (ML) network. A *ML network* is composed of several interconnected and interdependent sub-networks of varying sizes ranging from small to medium to large, with varying security levels and capacities. Although almost all network infrastructures (military, commercial or private) can be represented by a single topological level concept, it would be precise to model it using the concept of a ML network. Therefore, the representation of such ML network architecture is important, because it can represent the physical infrastructure and attributes of many of todays special communication networks, such as communication networks that serve the military. The infrastructural dependency among networks is primarily needed for efficient communication between different networks. A similar terminology named *multilevel network* is described in [1], where the term *multilevel* implies all five abstract layers

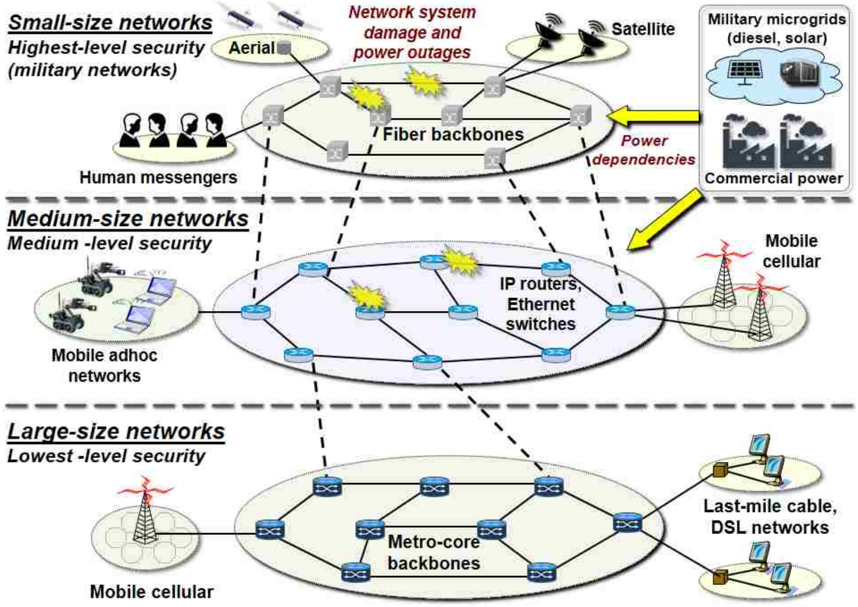


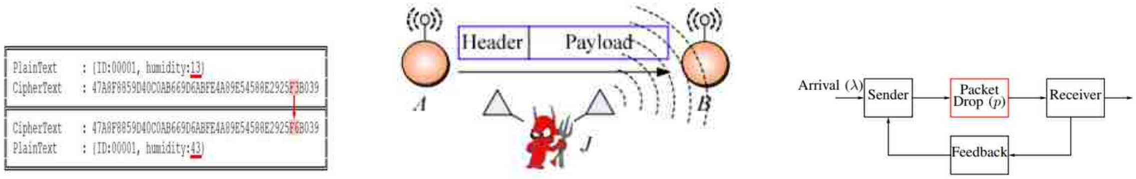
Figure 1.1: A physical topology for a ML communication network.

of OSI (open system interconnection) model. The authors in [2] have introduced a framework termed *multi-provider network*. Akin to this effort, we have defined each level of our ML network as the physical infrastructure of different networks. Figure 1.1 shows a prototype of a ML communication network of the types described here for military communication. We categorize the physical infrastructure of an interdependent communication system into three scales: small, medium and large. The specifications of each scale are given in Table 1.1. For instance, the small, medium and large networks can be military networks, wide area networks and commercial networks, respectively.

Because of these interconnections many network-security issues may arise which we describe in the following. At first, a secure network may be exposed to security attacks through interconnections, such as an interconnection may be intercepted by intruders/attackers [3]. Some example such attacks such as bit-flipping attack [3],

Table 1.1: Specifications of each level in a ML network

Network size / Parameter	Small	Medium	Large
Number of nodes	∼ 100	∼ 1000	∼ 10000
Bandwidth	low	high	high
Security	high	low	low



```
PlainText     : (ID:00001, humidity:13)
CipherText    : 47A8F8859D40C0AB669D6ABFE4A89E54588E2925F3B039

CipherText    : 47A8F8859D40C0AB669D6ABFE4A89E54588E2925F6B039
PlainText     : (ID:00001, humidity:43)
```

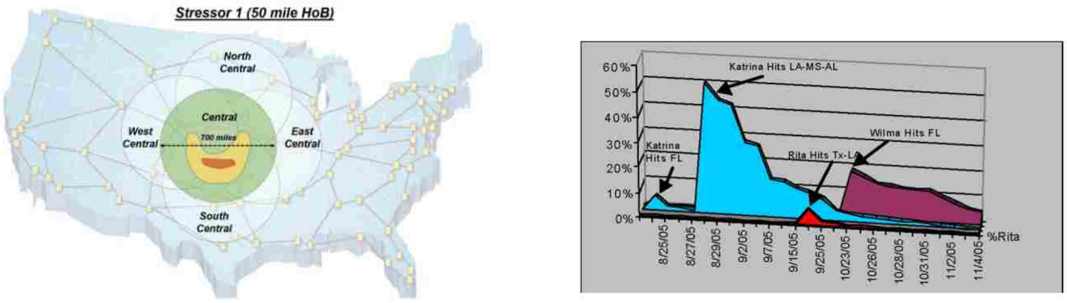
(a) (b) (c)

Figure 1.2: Different types of logical attacks on networks: a) bit-flipping attack b) jamming attack c) packet-dropping channel

.

jamming attack [4] and packet-dropping scenarios [5] shown in Fig. 1.2.

Apart form the above logical security issues, since the ML communication network is a physical infrastructures various physical attacks can directly damage components of a communication networks. For example, man-made attacks such as WMDs (weapons of mass destructions), HEMPs (high-altitude electromagnetic pulses) and natural disasters including earthquake (2006 earthquake in Taiwan), hurricane (Katarina in the USA and Mexico in 2005), tornadoes, flood, lightening, snow/strom, wildfires (2006 wildfires in California) are shown to damage components of communication networks. Figure 1.3a shows the emulation of a man-made attack (i.e.,low-altitude EMP attack) on the central region of the USA [6]. This attack is shown to be directly failed the components of the communication network. Figure 1.3b shows the percent of wirelines and wireless subscribers that were out of service in the affected areas in the USA and Mexico due to the 2005 hurricanes [7]. Here the horizontal-axis represents the date, the vertical-axis is the percentage of subscribers out of service. From the figure we see, for Katrina, the peak reaches nearly 50% of



(a) (b)

Figure 1.3: Effect of different types of physical attacks on networks: a) emulation of a low-altitude EMP attack on the central region of the USA b) percentage of wirelines and wireless subscribers that were out of service in the USA and Mexico due to 2005 hurricanes

.

the total subscribers, which were approximately 3 millions subscribers.

Moreover, attacks may propagate among networks through interconnections. For example, security attacks such as Malwares (Viruses, Worms) can propagate from one node to another node [8] and a secure network node may be compromised due to the propagation of threat/risk from less secure networks. An example of such attacks is WannaCry ransomware attack in 2017, that affects 200,000 computers across 150 countries, total damages ranging from hundreds of millions to billions of dollars (see Wikipedia/News articles). In addition, node removal (failure) can increase the stress/traffic to other nodes, which may result in congestion [9]. Congestion-induced Internet collapse occurred in October 1986, when the speed of the connection between the Lawrence Berkeley Laboratory and the University of California at Berkeley, which are located only 200 meter apart, dropped by a factor of 100 [10].

Further, the functionality of a ML communication network is dependent on the uninterpretable power supply, i.e., proper operation of the power grid. However, the power-grids are also vulnerable to physical attacks and natural disasters. For instance, during northeastern USA and Canada blackout on August, 2003, 3,175 communication networks of 1700 organizations (business entities, government, education institutions) suffered from abnormal connectivity outages. Of those, more than 2,000 networks suffered severe connectivity outages for longer than 4 hours, and over 1,400 networks for longer than 12 hours (some even longer than 48 hours) [11].

As we have seen the properly functionality of a communication network is heavily dependent on various real-world physical and logical attacks. The vulnerability analysis of the communication network to attacks is the main motivation of this dissertation. Hence, this work is mainly about the analyzing the dynamics of communication network under different types of attacks. We characterize how different types of attacks affect the functionality of a network and considering these attacks how we can find optimal interconnection among networks. In addition, the vulnerability and resiliency analysis of the ML network considering the attacks and the influence of power failures can significantly improve the reliability analysis of a realistic ML communication network.

1.2 Literature review

In this section, we review the works related each chapters of this dissertation. In addition, we also point out the limitations of these works and briefly describe contributions of this dissertations where necessary.

1.2.1 Prior work on optimizing inter-network connections among networks

The literature related to the optimal inter-network connection can broadly be divided into two parts: combating security attacks and optimizing interconnectivity among networks. First, we describe the security issues occur due to interconnections. Generally, an interconnected system comprised with individually secure systems is not secure [12], because different networks usually have distinct security policies, control structures, and infrastructural vulnerabilities, as in the case of interconnected military and commercial networks for example [13]. Hence, by compromising a lower secure system an attacker may compromise and access data from a highly secure system through their interconnections [12]. In [13] Shake *et. al.* suggested using highly secure gateway nodes for interconnecting a military network with commercial networks. By adding some overhead to IPSec (IP Security) protocol for the military data (packet) that are routed through commercial networks, a security measure to overcome the attack issues was proposed in [14]. These complex authentication procedures and cryptographic encryption algorithms can minimize the security vulnerabilities of networks. However, these methods come with associated costs of large overhead bits, which, in turn, leads to the reduction of spectral efficiency [15] and also may increase packet (information)-loss due to the limited transmission capacity of the communication links. In addition, an error in the encrypted data may result in serious decryption issues as the authors described in [16]: *a good cryptographic algorithm should be such that a one-bit error in the input to the decryptor should result in about 50% of the decrypted bits to be wrong.* Moreover, the distortion due to packet-loss can be recovered by retransmitting the lost packets (as done in ARQ (Automatic Repeat Request) protocol), which again results in a reduction of the effective throughput of the networks. Security schemes for the physical layer rather than the upper OSI (Open Systems Interconnection) layers were broadly discussed in [17], and Harrison *et. al.* proposed to use error-correction coding for implementing security measures in the physical layer [18], which is closely related to the work presented in the chapter 2 from the security aspect. It is important to note that the security issues may also occur due to attack on nodes of a network and these attacks may propagate between networks through interconnections [19]. However, here capture a scenario where the attackers directly intercept the interconnections; hence, compromise the security of a node indirectly, which is a realistic approach for modeling the security aspect of the interconnected networks.

On the other hand, the literature of interconnectivity among different networks has mainly been focused on military and commercial systems [13, 14]. However, in all those works it was assumed that these interconnections are secure, and hence the authors did not present any analytical procedures on how interconnection should be made considering the security issues that may occur due to the added interconnections. To fill this gap, here we propose a mathematical model to incorporate the security interceptions introduced by the added interconnections and solve for the optimal number of interconnections between two communication networks analytically. Notably, there are analytical works on the optimization of interconnection in cyber-physical systems, where the interdependency may lead to the cascading failures in power grids, and subsequent failures of communication nodes which may directly depend on power nodes [20–23]. These works have used the percolation theory to optimize the interconnection in a cyber-physical system based on different constraints [21–23]. However, the dynamics of the interconnectivity problem between communication networks are different than the cyber-physical system.

1.2.2 Prior work on the vulnerability analysis of ML networks under attacks

Most existing analyzes of network failures start from a given (fixed) network topology and then focus on various stressors [24,25]. Moreover, many of prior studies on physical stressors were limited to single or geographically-isolated small-size attacks or disasters [26,27]. The impact of large-scale stressors on networks was first introduced in [28]. Inherently, large-scale stressors lead to geographical-correlation amongst the failures of network components within the stressor's geographical proximity. Geographical correlations among large number of component failures is modeled in [28] based on the geographical proximity of network components from stressor centers. However, the authors only considered a single stressor event at a time, and failure of any network component within the range of stressor was assumed to be deterministic. However, multiple attack/disaster events can occur simultaneously and their effects on network-component failure are not deterministic. A probabilistic failure model with multiple circular stressor events has been proposed in [25]. However, the authors assigned a fixed failure-probability to all network components, which may not be realistic. Further, assumption of only circular-shaped stressors does not completely capture the impacts of various stressors on a network. A more realistic Gaussian shape function for modeling range and intensity of several stressors was used in [29]. In addition, the authors also used a Strauss point process [30] to characterize the geographical correlation among stressors. There are other works on network failures where SRLG information was used to characterize the correlated link failures in a network [31]. Following a stressor event and by assuming that all components belong to an SRLG will fail with some (fixed) probability, a simple approach to model correlated failures in a network was proposed in [32]. Correlated failures in networks were also studied based on attacks in the logical layers [33, 34]. In contrast to all previous works, we propose a probabilistic model for capturing the intensity of various stressors. Furthermore, we devise a new formulation to compute the failure probabilities of all components in a communication network considering their coupled vulnerabilities.

1.2.3 Prior work on the propagation of attacks in a network

Attack propagation among different networks has particularly been studied between military and commercial networks. In [35] the authors discussed the crucial differences between the commercial and military network. While interconnecting a military network with commercial networks, Shake suggested to use highly secure gateway nodes [13]. In [14] authors proposed to overcome security issues by adding some overhead (IPSec protocol) to the military data (packet) that are routed through commercial networks. Moreover, some of the major challenges in realizing a global network (a network infrastructure that allows military to get information whenever and wherever they need) were summarized in [36]. The authors in [37] proposed a layered network protocol different than the OSI or TCP/IP to establish a global communication network for performing military communications across the world.

Surprisingly, the mathematical modeling of propagation of security risk through interconnections has not been studied much in literature. A relatively close work is [38], where the authors tested different network topologies (e.g., stars, cliques, cycles) to design the intra-connectivity structure of a network in order to maximize the resiliency and connectivity. In contrast, our work deals with the interconnection among different independent subnetworks, specifically, between a highly secure network and networks with relatively low level of security. Further, we have used constrained optimization formulations for maximizing efficient interconnections among subnetworks.

On the other hand, there are analytical works on the propagation of failures in cyber-physical networks, such as one closely related to our work is [39], where the authors optimized the interdependency in cyber-physical network using the evil-rain influence model. The major departure in this work from [39] is that, rather than assuming all the nodes are vulnerable to attack, we have considered nodes in the less secure network are vulnerable and the vulnerability is assumed to propagate to the secure network.

1.2.4 Prior work on the influence power-failures to the dynamics of networks

A substantial amount of work has been done on the topic of network reliability as well as on cascading failures in interdependent systems (a brief review of the latter can be found in [40]). Initially most of the works were concentrated on a single network such as power grids, communication and computer networks, transportation networks, etc. [28, 41, 42]. However, the interdependency among cyber-physical networks can lead to failure dynamics that cannot be captured by analyzing the networks independently [43–45]. Motivated by the massive power and communication blackout in Italy on 28 September 2003, Buldyrev *et al.* modeled the network failure due to interdependency between power and communication networks based on percolation theory [20]. In [46], the authors looked at the robustness of the coupled communication and power networks. Although the authors in [20, 46] have observed the role of interdependency between two networks to design a robust network, they did not consider the intra-dependency within. On the other hand, an influence model has been used to capture the intra-dependency and the cascading failures within a power grid [47]. Later Rahnamy-Naeini *et al.* [39] proposed a cascading resilient interdependent power and communication network model based on the influence-model framework and optimized the inter-dependency between two interdependent networks. It was assumed in [39] that the total influence imposed on a communication node distributed among its neighbors and supporting power node. However, as described in the previous section, a communication node can independently fail due to power disruption as well as the failure of communication components, which cannot be captured by the framework given in [39]. Therefore, in this work we focus on the vulnerability of communication networks with two independent influences, one from the neighboring communication nodes within the network and the other from the supporting power nodes. To the best of our knowledge, these two types of influences have not been considered independently to analyze failure dynamics in communication network.

1.2.5 Prior work on the cascading failures in power grids

Cascading failures in the power grid have been studied extensively. A review of the state-of-the-art models of cascading failure along with their advantages and disadvantages can be found in [48]. Here we briefly describe the probabilistic models that are closely related to the work presented in this chapter 6. In [49] a branching process is used to parametrically model the dynamics of cascading failures, where parameters of the branching process are estimated from real-world blackout data. Later, in [50] the authors have modified the model presented in [49] by replacing the constant parameter of the branching process with a more realistic variable propagation rate of the cascading failure. However, in both papers, the authors assumed all the transmission lines have an identical impact on the evolution of cascading failure. In [41] Rahnamy-Naeini *et al.* have proposed a Markov chain based abstract model, named SASE (Stochastic Abstract-State Evolution), to analyze cascading failures in a power grid considering operating constraints, such as load-shedding, power grid loading level, and capacity estimation error. The SASE model is an analytically tractable model that can predict the probability of blackout sizes in time as well as the conditional distribution of the blackout sizes given the initial number of line failures. We follow the similar simulation approach introduced in [41]. However, in contrast to [41] that assumed single line failure for simulating cascading failures, here we allow multiple line failures in our simulation, which mimics real-world cascading failure dynamics more closely [51]. In [51] the authors have used the resistance distance and pseudoinverse of the grid admittance matrix to analyze the cascading failure in the power grid after a line failure. They have also proposed a heuristic algorithm to calculate the set of initial line failures that result in a minimum ratio

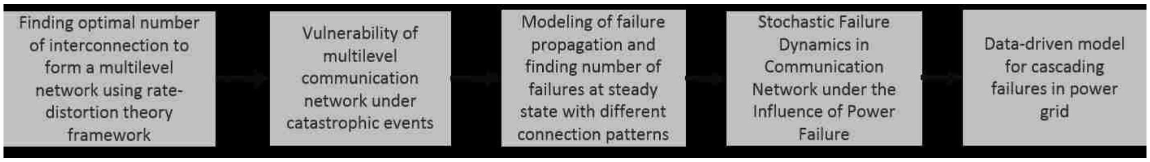

Finding optimal number of interconnection to form a multilevel network using rate-distortion theory framework
Vulnerability of multilevel communication network under catastrophic events
Modeling of failure propagation and finding number of failures at steady state with different connection patterns
Stochastic Failure Dynamics in Communication Network under the Influence of Power Failure
Data-driven model for cascading failures in power grid

Figure 1.4: The contribution of this dissertation.

between the final demand after the cascade stopped and the initial demand when the cascade started. Different from previous works, in chapter 6, we characterize the evolution of cascading failures through the transmission lines of different capacities by a nonlinear parametric model that is based on simulation data.

1.3 Contributions of this dissertation

The main thrust of this dissertation is to analyze the vulnerability of communication network to different types of physical and logical attacks. In addition, we also take into account the dependency of the communication network on the vulnerability of the supporting power grid. We form a ML communication network considering security vulnerabilities of inter-network connections and propagation of security attacks through connections. Then the resiliency of the communication network is also analyzed considering the attack vulnerability of the power grids. Then a separate work on cascading failures in the power grids is characterized through transmission line failure evolution. The chronological dependency of the contributions of this dissertation is shown in Fig. 1.4.

At first we look into problem how to interconnect two networks to form a ML network (especially, a 2-level network) considering security interception on the interconnecting links. The trade-off between security and data-rate enhancement is analyzed by optimizing the number of connections between two networks under the threat of security attacks. The optimization of the inter-network connection for maximizing the data-rate and minimizing the packet loss (represents the degradation of the quality of service) is formulated using a rate-distortion theory setting, as originally introduced by Claude E. Shannon in communication theory. In particular, each added connection that is vulnerable to data manipulation by attackers is modeled as a noisy communication channel. Subsequently, the Shannon entropic capacity for any given number of connections is calculated, and the corresponding packet-loss is modeled based on the required data rate. By exploiting such models, the optimal number of connections between two networks is found under network administrators' data-rate and packet-loss requirements, and most importantly, without compromising the data security. Moreover, the proposed theory identifies condition on attacker's data manipulation probability to differentiate scenarios of vulnerability when increasing the number of connections between two networks. In the vulnerable situation, this work provides a threshold on the optimal number of inter-network connections that results in a point of diminishing returns in increasing the number of connections.

So from the first problem, we will get a interconnected network which we name as the ML network. However, various cyber-physical infrastructures such as communication networks and power grids are known to be vulnerable to large-scale stressors ranging from natural disasters to intentional attacks such as those effected by weapons of mass destruction and high-altitude electromagnetic pulses. The stresses instigated by these events can cause damage to critical components of the network infrastructure. In this second work, a general probabilistic model is developed for assessing the vulnerability of a communication network under various catastrophic events. A scalable ML network framework is proposed to capture the inter-dependencies across various communication networks in the infrastructure. For a given large-scale stressor, the initial-failure probability of each network component is formulated independently and then by taking into account the failure of the components that it depends upon. This enables the modeling of a shared failure among network components. Detailed simulations of a three-level network model are performed and key network-performance metrics are computed including the total network capacity, the maximum flow and the number of node failures. This work paves the way to model and evaluate the reliability of critical communication networks under massive stressor events.

In the second problem, we have seen impact of attacks result in some initial failures in the ML network. Note that, these are not necessarily physical, mostly logical security attacks e.g., cyber attacks. At this point, we ask, what if these initial failures propagate? What is the resiliency of the network in steady state? In this third work, we maximize the interconnectivity among subnetworks under the constraint of security risk. We model the dynamics of security risk propagation by the evil-rain influence model and the SIR (susceptible-infected-recovered) epidemic model. By extensive numerical simulations using different network topologies and interconnection patterns, it is shown that the efficiency of interconnectivity (vulnerability) increases nonlinearly (linearly) with the number of interconnections among subnetworks. Finally, parametric models are proposed to find the number of interconnections for any given efficiency of interconnectivity and resiliency of the secure network. The importance of this work is that, we can know the status of a secure network when that network is susceptible to attack propagation.

Till now, we have a model for initial failures in communication networks under attack and analyzed the propagation of those initial failures and find steady state. In other words, we have modeled the dynamics of communication network independently. However, we have ignored one important part of the physical system, which is power. As we know the functionality of a communication network depends on power network as well and the vulnerability of power networks due to attack can directly affect the supporting power supply components of the communication networks and can interrupt the functionality of the communication components. In other words, both intra-dependency among various components inside a communication network and inter-dependency with supporting power supply nodes can contribute to the propagation of the initial failures. Under the probabilistic initial failures of the network's physical infrastructures, in the fourth work, we develop a stochastic model to track the functional dynamics in a communication network considering dependency among communication nodes as well as with power nodes. A closed-form analytical formula is derived to find the steady state probabilities of all communication nodes in the communication network which are influenced by power node's functionalities. The significance of this work is that it captures effect of power failure on the communication network which is important for any real-world network.

Finally, the previous work assumed power nodes are independent on each other, which is a simplistic assumption. Hence in the final work, we analyze the cascading failure in the power grid by using a data driven model. In this work, a parametric model is proposed that represents the propagation of transmission line failures in the power grid. At first, by optimal DC power-flow simulations of the IEEE 118 bus system, different types of data on the transmission line failures are collected, which can capture the time evolution of the cascade. The probability that a line of a given capacity fails is found to be well approximated by a power-law (the higher is the capacity the lower the probability). Then based on these data, a discrete-time parametric model is proposed to keep track of the failures of lines of different capacities during a cascading failure event. The convergence of the evolution of line failures to the steady state is analyzed. Finally, when compared to real-world data, the model is shown to be capable of reproducing similar trends for the time evolution of the failure dynamics. The significance of this work is that it can reproduce the time-evolution of the failure dynamics of transmission lines in a power grid given the grid topology.

1.4 Organization of the dissertation

This dissertation is organized as follows. In chapter 2, we first form a ML (2-level) network considering the security attacks on the added inter-network connections. We formulate this problem is an interconnection optimization problem between two networks. The constraints of this optimization problem are data-rates and the packet-loss. The outcome of this work is the following: a fundamental theoretical limit on the number of interconnection using rate-distortion theoretic formulation that meets users specified requirement.

The infrastructure of that ML network is vulnerable to large-scale physical attacks ranging from natural disasters to intentional attacks such as WMDs, EMPs. These types of catastrophic events can cause damage to critical components of the network infrastructure. Then in chapter 3, the vulnerability of ML network to different types of physical attacks is analyzed. We develop a probabilistic model for assessing the vulnerability of a communication network under various catastrophic events.

In chapter 4, we dig a bit more into the problem: we model how these initial attacks can propagate between networks and thus how it increases vulnerability of a secure network due to the propagation of a threat/risk from less secure networks. We also model resiliency of the network at steady state, (i.e. number of node survived when the attack propagation is stopped). in summary, we analyze how this initial failures propagate between networks through interconnections and affect the resiliency of the network in chapter 4.

In chapter 5, we extend the dynamical analysis of a communication network under the influence of power failure due to attack. In other words, we develop a stochastic model to track the functionality dynamics of a communication network considering both intra-dependency among communication nodes and inter-dependency with power nodes. We derive a closed form analytical solution to track the functionality dynamics of a communication network considering dependency with power nodes.

In chapter 6, based on the data from optimal DC power-flow simulations of the IEEE 118 bus system, a nonlinear parametric model for simulating the evolution of transmission line failure in power grids is proposed.

Finally, we conclude our works and provide future research directions in chapter 7.

1.5 Publications resulting from the dissertation

A list of our publications, related to this dissertation, is as follows:

1. Pankaz Das, Rezoan A. Shuvro, Mahshid R. Naeini, Nasir Ghani, Majeed M. Hayat. “Optimizing Interconnectivity among Networks for Maximizing Data Rate and Minimizing Packet-loss under Security Attacks ,” to be submitted in IEEE Transactions on Communications, 2018.

2. Pankaz Das, Rezoan A. Shuvro, Mahshid R. Naeini, Nasir Ghani, Majeed M. Hayat, “Stochastic Functionality Dynamics of Communication Network under the Influence of Power Failure, under preparation, 2018.

3. Pankaz Das, Mahshid R. Naeini, Nasir Ghani, Majeed M. Hayat, “On the Vulnerability of Multi-level Communication Network under Catastrophic Events,” IEEE ICNC, San Jose, USA, 2017.

4. Pankaz Das, Rezoan A. Shuvro, Zhuoyao Wang, Mahshid R. Naeini, Nasir Ghani, Majeed M. Hayat. “Stochastic Failure Dynamics in Communication Network under the Influence of Power Failure.” IEEE WiMob, Rome, Italy, 2017.

5. Pankaz Das, Rezoan A. Shuvro, Mahshid R. Naeini, Nasir Ghani, Majeed M. Hayat "Efficient Interconnectivity among Networks under Security Constraint," IEEE MilCom, LA, USA, 2018.

6. Pankaz Das, Rezoan A. Shuvro, Zhuoyao Wang, Majeed M. Hayat, Francesco Sorrentino "A Data-Driven Model for Simulating the Evolution of Transmission Line Failure in Power Grids," IEEE NAPS, ND, USA, 2018.

Chapter 2

Optimizing interconnectivity among networks under security attacks

Interconnections refer to the communication links that connect two or more networks for mutual communication among network users. By the Code of Federal Regulations (47CFR51.5) of the United States, “interconnection is the linking of two or more networks for the mutual exchange of traffic.” The physical linkings between two carrier’s networks or connections between a carrier’s facilities to its customers are also referred as the interconnections in telecommunications networks ([52] Ch. 10). In general, interconnections among independent networks are inevitable for various reasons including expansion of the communication capabilities among geographically-distant networks for seamless communication among users across different networks, establishing redundant/backup connections in case of failures in the primary system, increasing data rate and minimize delay by rerouting through high-bandwidth networks when a network has a limited bandwidth, etc. In particular, interconnectivity between a small-sized network (consisting of a smaller number of nodes with low link-bandwidth) and a large-sized network (comprising a more substantial number of nodes with high link-bandwidth) is required for expanding the communication range and data rate of the small network. For example, communication networks such as military, private, and commercial networks, consist of their secure communication infrastructure, which may, however, interconnect with external networks for the above-mentioned reasons. For instance, a wide range of commercial and non-commercial communication systems and networks are used by military personnel to support the military communications [53].

Clearly, without interconnections, there are no active communication links between the two networks; hence, the data packets of a network that are to be routed through the network will be entirely dropped. In addition, when there are smaller number of interconnections that results in the limited data exchange capability between the networks, the data transmission may face considerable delays. Moreover, in the worst-case scenario, certain data packets may be dropped (idealizing the effect of queuing delay to be considerably large or infinite). To remedy such situations, a larger number of interconnection between networks may be thought of as a viable solution. Therefore, the more interconnections one establish, the more data rate one can achieve, and there will be less likelihood of packet-loss.

However, these interconnections can be intercepted by attackers, and once an interception event occurs, attackers may manipulate the data packets (bits) that are being transmitted through the interconnections. In particular, an intruder may intercept a communication channel between nodes and may add or delete data/control messages of sender and receiver ([8] Ch. 8). For example, the LoRaWAN (Long Range Wide Area Network) technology that is proposed for IoT (Internet of Things) has a security vulnerability which may result in a bit-flipping attack, where the attacker can change encrypted data without decrypting it. In [3], the authors showed an example of bit-flipping in the data that produce wrong information output on the receiver side. Further, bit-flipping attacks are familiar to changes in the destination address, digital signatures, and stream ciphers. Besides bit-flipping, the data bits can be deleted by attackers in certain types of attacks such as in jamming attacks [4], in packet erasure channels [54]. A communication channel considering both insertion and deletion of data bits are described in [55] for the binary channels and in [56] for non-binary channels. Therefore, interconnections come with the price of security vulnerabilities to the system. Further, since each interconnection opens a way of interception for the intruders, the more interconnection a network has, the more prone it is to security attacks. Hence, it becomes important to find the number of interconnection that gives the required data rate and minimize the packet-loss without compromising the security of the network.

In this chapter, we respond to the following question: *how many interconnections two networks should have to maximize inter-network data rates while maintaining the packet-loss below a threshold and limiting the level of compromise in the security of the system?* To answer this question, we first probabilistically model the attacker's interception as a nondecreasing function of the number of interconnections that captures the increase of interceptions due to the added interconnections. After interception attackers may manipulate the data thats goes through the interconnection channel. To model the effect of data manipulation, each intercepted interconnection is modeled as a noisy communication channel, where we think of *noise* as the *data manipulation* by the attackers; in particular, manipulation refers to the flipping and erasing of the data bits. Afterwards, we analytically find the Shannon capacity of the intercepted interconnection channels. From Shannon's channel coding theorem ([57] Ch. 7), if one transmit data below the channel capacity, all the adverse effects on the transmitted data can be corrected at the receiver side using channel coding. Therefore, the gateway nodes always transmit below the Shannon capacity to undo the harmful effect of attackers. Note that here the gateway nodes are those connecting the two networks and knows (or estimates) the capacity of the interconnection channel. Hence, whenever the data rate of the interconnections exceeds the associated Shannon capacity, some data packets (bits) is dropped by the gateway nodes to combat the security attacks. Here we define *packet-loss distortion* as the rate of dropped packets from the source network due to the finite and limited Shannon capacity of the interconnections. Considering a certain tolerance level on the packet-loss distortion, we define the data rate of a network as the maximum transmission rate that can be achieved without compromising security. Finally, we propose two optimization formulations to find the optimal number of interconnections between two networks based on the network administrator's requirement on the data rate and packet-loss distortion. Here the optimality of the number of interconnections is defined with respect to the maximum data rate and minimum packet-loss distortion. Most importantly, the optimization problems take into account the security vulnerabilities (interception occurrence and data manipulations) due to added interconnections and find the optimal number of interconnections without compromising the security at all. As such, we develop an optimal interconnection strategy between two networks for maximum mutual-information exchange and minimize packet-loss with combating the security interceptions of the data. The contributions in this chapter can be summarized as follows:

1) The probabilistic model for the interception of interconnections by attackers is introduced and the effect security attacks after interception is analyzed through the data manipulations. Defining associated system parameters and finding the optimal interconnectivity under the security attacks from the information-theoretic perspective is an exciting application of the state-of-the-art rate-distortion theory and Shannon's channel capacity.

2) Analytic derivation of the closed-form formula for the capacity of an interconnection channel under any given data-manipulation capability of the attackers as well as theoretical limit of the total Shannon capacity of the interconnections when

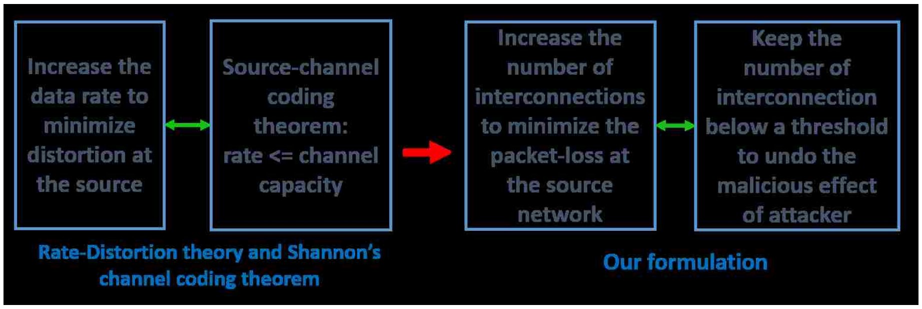

Increase the data rate to minimize distortion at the source
Source-channel coding theorem: rate <= channel capacity
Increase the number of interconnections to minimize the packet-loss at the source network
Keep the number of interconnection below a threshold to undo the malicious effect of attacker
Rate-Distortion theory and Shannon's channel coding theorem
Our formulation

Figure 2.1: The analogy between our formulation and the rate-distortion theory and channel coding theorem in information theory.

maximum number of interconnections tends to infinity.

3) While finding optimal number of interconnections that can withstand under security threats, two optimization sub-problems are formulated and analyzed under two realistic constraints as follows:

a) maximizing the inter-network data rate while keeping the packet-loss below a certain user-specified level,

b) minimizing the packet-loss while satisfying user's data rate constraint.

Figure 2.1 demonstrates the analogy between the formulation proposed in this chapter and the rate-distortion optimization and channel coding concept. In the seminal work of Claude E. Shannon on information theory [57, 58], the rate-distortion optimization provides a bound on the minimum data rate at the source (transmitter) that is needed to satisfy a certain distortion requirement at the destination (receiver), and the more data rate one provides, the less distortion is achievable. However, to transmit data through a noisy channel reliably, the source-channel coding theorem provides a bound on the maximum data rate (termed as *channel capacity*). Combining both theories, to achieve the desired distortion, the source needs to have the data rate given by rate-distortion theory, and that data rate has to be lower than the channel capacity for reliable communication (please see Ch. 7 and Ch. 10 of [57] for details). Analogous to the rate-distortion optimization, here we increase the number of interconnections between the two networks, which, in turn, increase the data rate to minimize the packet-loss distortion between two networks. Therefore, intuitively more interconnections mean higher data rates. However, since the attackers intercept these interconnections for manipulating the data, we provide a maximum bound on the number of interconnections (similar to the channel capacity) that gives the maximum data rate between two networks without compromising the security of the network. To the best of our knowledge, mathematical modeling of the security issues introduced by added interconnections from information theoretical perspective (specially rate-distortion theory and source-channel coding theorem), and then use of that formulation to find the optimal number of interconnections between two networks has not been done heretofore.

This chapter is organized as follows. In Section 2.1, we present a realistic example of the interconnectivity between a small military network and a large commercial network. The adversary model is described in Section 2.2. The key definitions and system parameters introduced and elaborated in Section 2.3. In Section 2.4, we formulate two optimization problems and find the optimal number of interconnections under packet-loss and data-rate constraints. Our conclusions and future work are presented in Section 2.5.

2.1 A realistic example of the interconnectivity between a small and a large network

Consider a large group of military personnel of a country that is stationed abroad for certain operations. Groups of this nature are usually supported by an highly-secure independent communication infrastructure consisting of a small number of

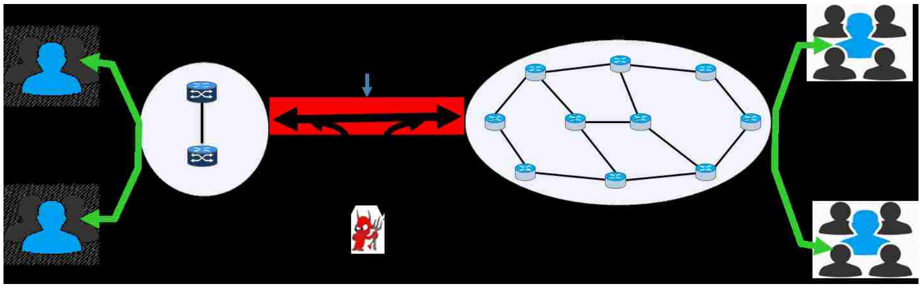


Figure 2.2: The interconnection between a small and large network, and interception of the interconnection by an attacker. An attacker may manipulate (add/delete) data after the interception.

nodes. In addition, such a small-sized network with a small geographical range may have limited connectivity to the outside world for security reasons. However, there may be need for this network to have high-bandwidth connectivity to the outside world for various practical reasons. For example, certain military personnel may need to receive service from Internet service providers, communicate with headquarters and exchange a large amount of data from databases of collaborating agencies when needed, etc. One approach to address such need is to use satellite communication. However, this system have relatively high cost with a substantial end-to-end delay and lower data rate compared to terrestrial/fiber-optic network. Alternatively, the small military network may resort to connecting to readily accessible commercial resources with broad bandwidth to achieve a higher data rate. Furthermore, military networks may also want to keep the commercial network as a redundant communication medium in case of emergency. Nevertheless, as pointed out in the previous subsections, such interconnections with less secure communication networks may increase the vulnerability of military networks to attacks. This realistic scenario is replicated by modeling interconnection between two networks, where a highly secure small network with limited data rate wants to connect with a less secure large network for achieving high data rate. Beyond the military network and commercial network, examples of highly-secure small networks and less-secure large networks may include private networks and intercontinental backbone networks, respectively. Figure 2.2 shows a scenario of the interconnection between a small and large network and the possible security interception on the interconnections by attackers. Note that here we provide a simple example with a small and large network from the application perspective. However, this work may applicable to find optimal interconnection between any networks irrespective of their sizes as long as interconnections are prone to attacks.

2.2 The attacker model

The following assumptions are made to model the attackers who may intercept interconnections and manipulate the data causing the inter-network communication insecure.

1) The attackers may intercept on the bit, packet, signal level of the transmitted data as assumed in [59]. Without loss of generality, we assume attackers work on the bit level since each data packet essentially consists of bits.

2) Interconnections cause a network to be susceptible to attack and the probability of interception by attackers increases as we increase the number of interconnections between a highly secure network and a network with lower security. This assumption is very intuitive and realistic since having more interconnections increases the likelihood of exposure of a network to the attackers. For example, it is shown in [12] that, interconnecting a highly secure system with other systems of lower security opens up paths for the data to be copied from the highly secure system to the lower security system. Since the network is compromised due to the interconnections, our assumption on the increase of security interceptions with more interconnections can be argued to be reasonable.

3) Attackers may have the capability of manipulating data that are transmitted through the interconnection channels ([8] Ch. 8). Our attack model represents the worst-case attack scenario where the attackers may change every bit of a data packet with some probability. However, in intentional attacks, the attackers may manipulate specific burst of data if they know the importance of data, the scenario we are not considering in this chapter.

4) Here we do not consider the effect of noise and interference on the data bits to keep our focus on the security interception issues; namely, the interconnection channel is noiseless and channels do not interfere with each other. However, an additive parameter can be incorporated with the attackers capability parameter to include the effect of noise.

2.3 Formulation of an equivalent channel for the intercept-prone interconnection

Without loss of generality, we consider to interconnect two networks and refer to the source network (Network 1) as the network that needs of connections (e.g., military system), and the auxiliary network (Network 2) as the network that is available to provide connections (e.g., commercial system). Clearly, without interconnections no active communication links exist between two networks; therefore, all the packets of Network 1 that are destined for Network 2 will be dropped. Further, here queuing delay is assumed to be infinite for simplicity. Hence, whenever the data rate of the source network is higher than the aggregated Shannon capacity of the added interconnections, the gateway nodes drop the surpassed packets. The system parameters of this work are modeled and elaborated in the following subsections. Note that all these system parameters described in the following subsections are used to find the optimal number of interconnections between two networks in the Section 2.4.

2.3.1 The number of interconnections ($N_{X_1X_2}$)

It is the number of communication links of certain data rate between the gateway nodes of the two networks. We denote the number of interconnections between Network 1 and Network 2 by $N_{X_1X_2}$, where X_1 and X_2 represent the gateway nodes of networks 1 and 2, respectively. Some of the realistic attributes of these interconnections are as follows. 1) Interconnections are done through the gateway nodes, and a gateway node in a network may connect to any gateway nodes of the other network; namely, a gateway node in Network i may have any number of interconnections from the set $\{0, 1, \cdots, N_i\}$, where N_i is the number of gateway nodes in Network i. Hence the maximum number of allowed interconnections between Network 1 and Network 2 is given by $N_{max} = N_1N_2$. 2) The interconnection (channel) is a discrete memoryless channel (DMC) ([57] Ch. 7), and no feedback from the receiver is allowed, since the feedback does not increase the capacity of a DMC ([57] Ch. 7). 3) All interconnections have a same bandwidth.

Note that all the numerical simulation results shown next are based on the following network parameters: consider two networks that are being interconnected with each other. Network 1 consists of 4 gateway nodes ($N_1 = 4$), and Network 2 consists of 25 gateway nodes ($N_2 = 25$). Then the maximum number of interconnections between two networks is $N_{max} = N_1N_2 = 100$.

2.3.2 Interconnection channel formulation

Each interconnection channel is characterized by the following two probabilities.

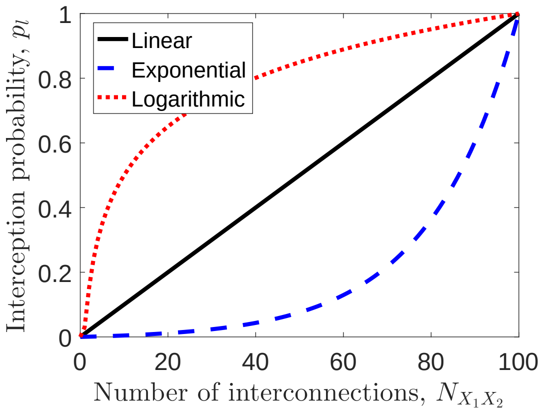


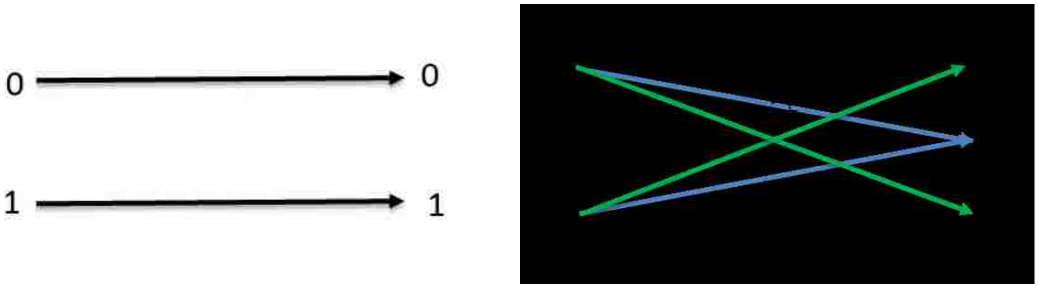
Figure 2.3: The probability of interceptions (p_l) vs. the number of interconnections ($N_{X_1X_2}$) for three types of interception probability models. Here increasing number of interconnection causes an increase in the probability of interceptions.

Interception probability (p_l)

Due to the nature of the attacker model presented in Section 2.2, here each interconnection has a certain likelihood of interception, and more interconnections imply more interceptions by attackers. Therefore, the probability of interception (p_l) of an interconnection can be expressed as a nondecreasing function of the number of interconnections $N_{X_1X_2}$. Since the minimum and maximum number of interconnections are 0, N_{max}, respectively, two extreme cases of the probability of interception are $p_l = 0$ when $N_{X_1X_2} = 0$, and $p_l = 1$ when $N_{X_1X_2} = N_{max}$. Here our assumption on the explicit form of p_l are linear, exponential, and logarithmic as expressed below,

$$\begin{aligned} Linear : p_l &= \frac{N_{X_1X_2}}{N_{max}}; \\ Exponential : p_l &= \frac{e^{\alpha N_{X_1X_2}}}{e^{\alpha N_{max}}}, 0 \leq \alpha \leq 1 = \text{shape of the exponential function;} \\ Logarithmic : p_l &= \frac{\log(N_{X_1X_2})}{\log(N_{max})}, \log 0 := 0. \end{aligned} \tag{2.1}$$

Note that any other form of p_l which is a nondecreasing function of $N_{X_1X_2}$, may also be used depending on the nature of interceptions. Figure 2.3 shows the proba-



(a) Secure interconnection channel

(b) Intercepted interconnection channel

Figure 2.4: The interconnection (channel) model between two networks. (a) Secure interconnection with zero probability of data error ($p_e = 0$); therefore, the capacity is 1 bit. (b) The intercepted interconnection (channel) with the capacity C_l with probability of erasing and flipping are p_ϵ and p_f, respectively. Here, ϵ denotes the erased packet (bits)..

bility of interceptions (p_l) versus the number of interconnections ($N_{X_1X_2}$) for three types of interceptions given by (2.1). From the figure we see that increasing interconnection increases the probability of interception. Note that all the numerical simulation results shown in the subsequent sections are generated considering the linear functional form of p_l.

Probability of data manipulation (p_e)

Clearly, while not intercepted the interconnection can be represented as a noiseless communication channel which we termed as secure interconnection as shown by Fig. 2.4a. In contrast, when the interconnection is intercepted, the attackers may manipulate the data that are flowing through the interconnection channel with some probability. Namely, the data bits in each interconnection can be jammed (erased/blocked), and bits in a packet can also be deliberately flipped in malicious attacks. Therefore, the interception results in a probability of data (bit) flipping (p_f) and data (bit) erasing (p_ϵ). The model for an intercepted interconnection is shown by

Fig. 2.4b. The values of p_f and p_ϵ are determined by the capability of attackers for manipulating the ongoing data transmission through the interconnections. Further, these values may be provided by the network administrators or may be estimated from the historical data of security attacks. Note also that during sniffing, the attackers may copy/eavesdrop the packets for their benefit; however, since it does not directly change the data bits we model this scenario as the zero capability of manipulating. Hence, the total probability of data manipulation (i.e., data error) in an interconnection channel due to the flipping and erasure is $p_e := p_\epsilon + p_f$.

2.3.3 Properties of interconnection channel

The capacity of an interconnection channel is the property of interest here since it reflects how good or bad an interconnection channel is. Here we calculate the following two capacities of interconnection channels to describe their behavior under interception and manipulation.

Average capacity of an interconnection ($\bar{C}$)

Considering two cases of an interconnection whether it is intercepted or not as shown by Fig. 2.4, we can write the capacity of an interconnection as

$$C = \begin{cases} 1, & \text{with probablity } 1 - p_l, \\ C_l, & \text{with probablity } p_l, \end{cases} \tag{2.2}$$

where 1 and C_l (to be defined in Theorem 1) are the Shannon capacity of the secure interconnection channel and intercepted interconnection channel, respectively. Note that the Shannon capacity of an interconnection (channel) is given by the maximum mutual information between the channel input and output, where maximization is taken over the input (symbol) probability distribution [57]. Here we assume a DMC to represent each interconnection channel with binary input, and therefore, according to Shannon's channel capacity formulation, each interconnection can carry 1 bit of data (information) in the secure case. On the other hand, we have following theorem to find the capacity of an intercepted interconnection.

Theorem 2.3.1. *Assume a discrete memoryless interconnection channel where binary input data may be flipped or erased by the attackers with a flipping probability p_f and erasing probability p_ϵ. The Shannon capacity of this channel is $C_l = \big(1-(p_\epsilon+p_f)\big)\log_2\big(1-(p_\epsilon+p_f)\big)+p_f\log_2 p_f-(1-p_\epsilon)\log_2\big(\frac{1-p_\epsilon}{2}\big)$ bits, and this is achieved by a uniform distribution on the input data.*

Proof. Let us denote the input and output discrete random variable of the interconnection channel are $X=\{0,1\}$ and $Y=\{0,\epsilon,1\}$, respectively, where ϵ denotes the erased bit as shown by 2.4b. The mutual information $I(X;Y)$ between X and Y is defined as [57]

$$I(X;Y)=\sum_{x\in X}\sum_{y\in Y}p(y|x)p(x)\log_2\frac{p(y|x)}{p(y)}\text{ bits,}\tag{2.3}$$

where p is the probability mass function of X and Y with $p(\cdot)\geq 0$ and $\sum_{z\in Z}p(z)=1$.

Let $p(X=0)=\alpha$. Then $p(X=1)=1-\alpha$. Using the given flipping and erasing probability, we find $p(Y=0|X=0)=1-(p_f+p_\epsilon)$, $p(Y=0|X=1)=p_f$, $p(Y=\epsilon|X=0)=p_\epsilon$, $p(Y=\epsilon|X=1)=p_\epsilon$, $p(Y=1|X=0)=p_f$, $p(Y=1|X=1)=1-(p_f+p_\epsilon)$. Hence

$$\begin{aligned}p(Y=0)&=\sum_{x\in X}p(Y=0,X=x)=\sum_{x\in X}p(Y=0|X=x)p(X=x)\\&=(1-(p_f+p_\epsilon))\alpha+p_f(1-\alpha)=\alpha+p_f-2\alpha p_f-\alpha p_\epsilon,\\p(Y=1)&=\sum_{x\in X}p(Y=1,X=x)=1-\alpha+2\alpha p_f-p_\epsilon+\alpha p_\epsilon,\\p(Y=\epsilon)&=p_\epsilon.\end{aligned}$$

Substituting these probabilities in (2.3) and after simplifying we get

$$\begin{aligned} I(X;Y) &= (1-p_f-p_\epsilon)\alpha\log_2\frac{(1-(p_f+p_\epsilon))}{\alpha+p_f-2\alpha p_f-\alpha p_\epsilon} \\ &+ p_f(1-\alpha)\log_2\frac{p_f}{\alpha+p_f-2\alpha p_f-\alpha p_\epsilon} \\ &+ p_f\alpha\log_2\frac{p_f}{1-\alpha+2\alpha p_f-p_\epsilon+\alpha p_\epsilon} \\ &+ (1-p_f-p_\epsilon)(1-\alpha)\log_2\frac{(1-p_f-p_\epsilon)}{1-\alpha+2\alpha p_f-p_\epsilon+\alpha p_\epsilon}. \end{aligned} \tag{2.4}$$

The Shannon capacity of a DMC is given by the following formula ([57] Ch. 7)

$$C_l = \max_{p(x)} I(X;Y). \tag{2.5}$$

To maximize $I(X;Y)$, we differentiate (2.4) with respect to $p(x)=\alpha$ and setting the result to zero to get $\alpha=\frac{1}{2}$.

Now substituting $\alpha=\frac{1}{2}$ in (2.4), we find the capacity of the interconnection channel

$$C_l = (1-(p_\epsilon+p_f))\log_2(1-(p_\epsilon+p_f))+p_f\log_2 p_f-(1-p_\epsilon)\log_2(\frac{1-p_\epsilon}{2}) \text{ bits.} \tag{2.6}$$

□

Corolary 2.3.2. *The capacity of an intercepted interconnection $C_l = 1-p_\epsilon$ bits if $p_f=0$ which is the capacity of a binary erasure channel (BEC), and $C_l = 1+p_f\log_2 p_f+(1-p_f)\log_2(1-p_f)$ bits if $p_\epsilon=0$ which is the capacity of a binary symmetric channel (BSC).*

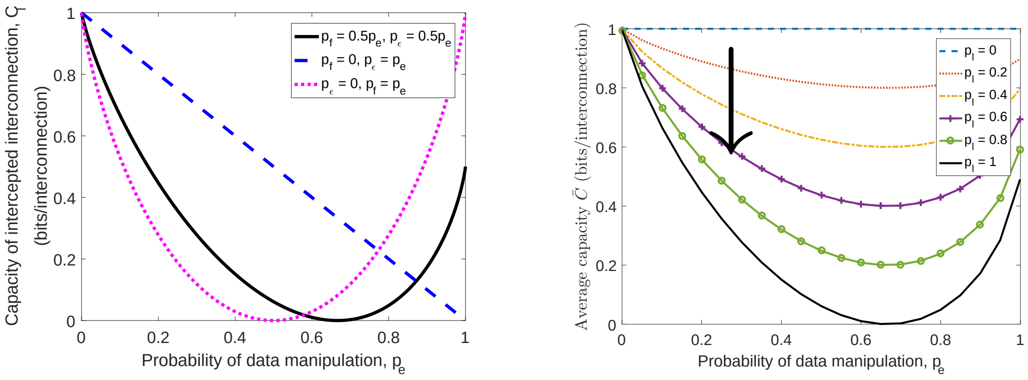
Proof. Substituting $p_f=0$ and $p_\epsilon=0$ in (2.6) it is straightforward to find the capacity of BEC and BSC, respectively. □

Corolary 2.3.3. *The Shannon capacity of an intercepted interconnection (C_l) is no higher than the raw-link (secure) capacity of the interconnection.*

Proof. It can be seen from the Theorem 1 that C_l can be at most equal to 1 when there are no flipping and erasing, i.e., $p_f = 0$ and $p_\epsilon = 0$. For all other any other values of p_f and p_ϵ, $C_l < 1$. However, the capacity of a secure interconnection is always 1. □

Figure 2.5a shows the Shannon capacity of a fully intercepted interconnection C_l concerning the attackers' data manipulation probability p_e for various p_f's and p_ϵ's. Two extreme cases of the interconnection channel, namely BSC ($p_f = 0$), and the BEC ($p_\epsilon = 0$), are shown by the dashed and dotted line, respectively. The solid black line shows the interconnection capacity (calculated by (2.6)) when both erasure and flipping manipulate the data bits equally i.e., $p_f = 0.5p_e$ and $p_\epsilon = 0.5p_e$. In this scenario, the capacity (C_l) is decreasing with the increase of p_e and $C_l = 0$ when $p_e = \frac{2}{3}$. This is because $p_f = \frac{1}{3}$ and $p_\epsilon = \frac{1}{3}$, when $p_e = \frac{2}{3}$. Hence one-third of the data bits are deleted, and one-third of the data bits are flipped, and also one-third of the data bits are received correctly. Since the receiver does not know which data bits are flipped and which are intact, then the uncertainty is maximum at $p_e = \frac{2}{3}$, and hence the capacity is zero. Moreover, C_l starts to increase again after $p_e = \frac{2}{3}$ and reach to 0.5 at $p_e = 1$. This is because the receiver knows all the received data bits are flipped and interprets 0 as 1, and 1 as 0. Hence all the flipped bits can be corrected which is 50% of the total bits since $p_f = 0.5$ when $p_e = 1$. Note that in all the numerical simulations to follow, we assume equal contributions from both flipping and erasure. However, our model can capture any other combination of the flipping and erasing.

Lemma 2.3.4. *Assume a discrete memoryless interconnection channel as shown in Fig. 2.4b, where the probability for data flipping and erasing are given by p_f and p_ϵ, respectively. The Shannon capacity of the interconnection channel is* zero *($C_l = 0$) if the probability of data flipping (p_f) is equal to the probability of data transmission correctly ($1 - p_f - p_\epsilon$), irrespective of the probability of data erasing (p_ϵ).*



(a) C_l vs. p_e

(b) $\bar{C}$ vs. p_e for different p_l's

Figure 2.5: The capacity of an interconnection channel. (a) The capacity of the intercepted interconnection (channel) C_l vs. p_e for various combinations of p_f and p_ϵ. (b) The expected capacity of the intercepted interconnection (channel) $\bar{C}$ vs. p_e for various values of p_l.

Proof. Analytically, it can be proved by substituting $p_f = 1 - p_f - p_\epsilon$ or $p_f = \frac{1-p_\epsilon}{2}$ in (2.6). Intuitively, when the probability of bit-flipping and the probability correctly sent-bits are same, the uncertainty at the receiver side is maximum, i.e., the receiver cannot decide whether a received bit is flipped or not-flipped with equal probability. Since the uncertainty is maximum, therefore, the capacity is zero. It is independent of the portion of erasing since the erased bits are not adding anything to the capacity of the channel. □

Hence, by considering both cases of the capacity of an interconnection given by (2.2), the average Shannon capacity of an interconnection channel with an interception level p_l is

$$\bar{C} = \mathbf{E}[C] = (1 - p_l) + p_l C_l = 1 - p_l(1 - C_l) \text{ bits/interconnection,} \tag{2.7}$$

where $\mathbf{E}$ denotes the expectation operator. Figure 2.5b shows the expected capacity $\bar{C}$ as a function of p_e for different values of p_l. Notice that, $\bar{C}$ decreases as we increase the interception probability (p_l). Hence for any given p_e, the average capacity of interconnection channel ($\bar{C}$) decreases with the probability of interception (p_l). Further, when an interconnection is fully intercepted, i.e., $p_l = 1$, then we get $\bar{C} = C_l$.

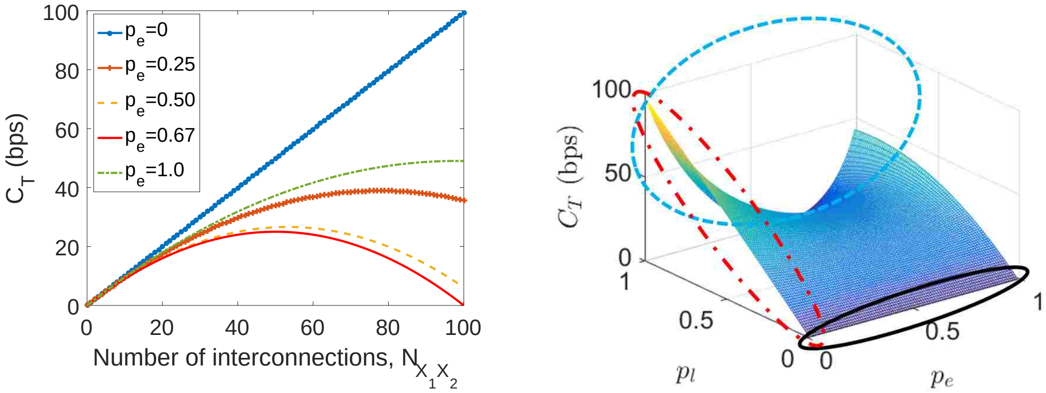
The total expected Shannon capacity (C_T)

In this subsection, we find the total capacity for the given number of interconnections. Since there are $N_{X_1X_2}$ active interconnections, we assume the combined interconnection channel of $N_{X_1X_2}$ interconnections forms a product DMC. In a product DMC, the input and output of the channels are the Cartesian product of input and output of the interconnection channels, respectively, and the channel transition probabilities are independent conditioned on the inputs ([60] Ch. 3). The total expected Shannon capacity of $N_{X_1X_2}$ interconnections (capacity of a product DMC) is given by ([60] Ch. 3)

$$C_T = \sum_{i=1}^{N_{X_1X_2}} \bar{C}_i = N_{X_1X_2}\bar{C} = N_{X_1X_2}(1 - p_l(1 - C_l)) \text{ bits/second (bps)}, \tag{2.8}$$

where $\bar{C}_i = \bar{C}$ is due to the identical bandwidth of all $N_{X_1X_2}$ interconnection which is assumed in Section 2.3 A. We write the the capacity unit as bps since we assume the product DMC is used once in one second, i.e., $N_{X_1X_2}$ interconnections are used in parallel per second.

Figure 2.6a shows the total expected Shannon capacity (C_T) versus the number of interconnections ($N_{X_1X_2}$) for different data manipulation (error) probabilities (p_e). Note the behavior of C_T: initially, as we increase the number of interconnections C_T increases; however, after a certain number of interconnections C_T does not increase with the number of interconnections. In worst-case ($p_e = \frac{2}{3}$), it becomes zero when the interconnections are fully intercepted (i.e., $p_l = 1$). The behavior of C_T is explained as follows. First it can seen from (2.8) C_T is a scaled version of $\bar{C}$. From the definition of $\bar{C}$ given by (2.7), we observe that $\bar{C}$ is a weighted average of 1 and



(a) C_T vs. $N_{X_1X_2}$ for different values of p_e

(b) C_T vs. p_e and p_l

Figure 2.6: (a) The total expected Shannon capacity (C_T) vs. the number of interconnections for various data manipulation (error) probabilities (p_e). (b) The total Shannon capacity (C_T) in 3D vs. p_e and p_l.

C_l; hence, $\bar{C}$ is always between C_l and 1. In the extreme cases, when $p_l = 0$, $\bar{C} = 1$, and when $p_l = 1$, $\bar{C} = C_l \leq 1$. This implies that C_T is always between $N_{X_1X_2}C_l$ and $N_{X_1X_2}$; hence, it is equal to $N_{X_1X_2}$ when $p_l = 0$, and to $N_{X_1X_2}C_l$ when $p_l = 1$. Recall that since p_e does not depend on p_l, C_l has a fixed value for a given p_e. Now when $p_e = \frac{2}{3}$, we get $C_l = 0$ and $C_T = N_{X_1X_2}C_l = 0$ when $p_l = 1$. Note also that there is a certain range of p_e, in which the increase in the number of interconnections results in the decrease of total channel capacity (as depicted in Fig. 2.6b). Finally, the most important but unintuitive observation is that adding more connection beyond the minimum required number is not productive to get more data rates, i.e., C_T. This is due to the security issues introduced by the added interconnections.

Now we describe four extreme cases of the C_T for all ranges of p_e and p_l. First, the total expected Shannon capacity (C_T) vs. p_e and p_l is shown in Fig. 2.6b.

Case 1 ($p_l = 0$): When there are no interconnections ($N_{X_1X_2} = 0$), the interception probability is zero ($p_l = 0$) because the network is not exposed to security risk through interconnections. As a result, attackers cannot manipulate data of a secure network through interconnections. Note also that when $p_l = 0$, C_T does not depend on p_e. This region is shown by the solid black oval in Fig. 2.6b.

Case 2 ($p_e = 0$): When the data manipulation probability is zero ($p_e = 0$), $C_l = 1$, which implies C_T is equal to the $N_{X_1X_2}$. This is because all the interconnection channels are secure without data manipulation error since $p_e = 0$. Intuitively, attackers intercept the interconnections; however, since $p_e = 0$, attackers' capability of changing data (flipping/dropping) through the interconnections is zero. The dash-dotted red oval shows this region in Fig. 2.6b.

Case 3 ($p_l = 1$): When all the interconnections are fully intercepted, C_T depends on the value of p_e (i.e., C_l). For any given p_e, $C_T = N_{X_1X_2}C_l$. The dashed blue oval shows this region in Fig. 2.6b.

Case 4 ($p_e = \frac{2}{3}$): When the data manipulation probability is $\frac{2}{3}$, we observe the worst-case scenario, i.e., $C_l = 0$. Therefore, $\bar{C} = (1 - p_l)$, and $C_T = (1 - p_l)N_{X_1X_2}$. This bound is the minimum capacity we get for any number of interconnections. Because at this value of p_e, attackers manipulate the data such a way the capacity of the intercepted interconnection becomes zero ($C_l = 0$). Hence the total capacity is zero when $p_l = 1$. This is shown by the solid red line in Fig. 2.6a.

The following theorem answers an important question regarding interconnection in the limit: *what if network administrators have capability of making an infinite number of interconnections?*

Theorem 2.3.5. *Assume two networks are interconnected with a finite number of interconnections and hence the intruders can introduce bounded interceptions. With infinite possibilities on the number of interconnections, the probability of interception is zero, and the total expected Shannon capacity is equal to the number of finite interconnections.*

Proof. Here we prove it for the linear interception probability (p_l), the proof for other interception probabilities are similar. Here for a finite number of interconnections, we have $N_{X_1X_2} < \infty$. Moreover, given infinite possibilities on the number of interconnections, we can find p_l by taking the limit on the N_{max} as

$$\lim_{N_{max}\to\infty} p_l = \lim_{N_{max}\to\infty} \frac{N_{X_1X_2}}{N_{max}} = 0,$$

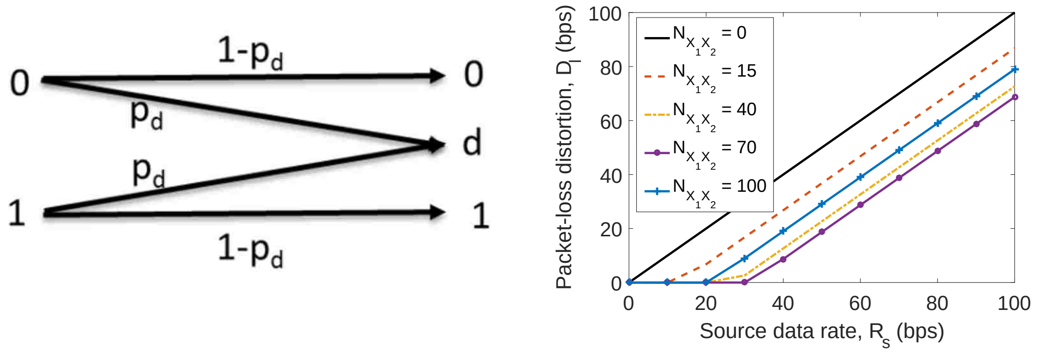
which implies $\bar{C} = 1$ and $C_T = N_{X_1X_2}$ bits/second. □

In other words, the above theorem says if there is an option to have an infinite number of interconnections, then having a finite number of interconnections would result in zero probability of interception. Intuitively, this is because when there are many interconnections, attackers will not have enough knowledge on the interconnections that are used by two networks; hence, the likelihood of an active interconnection being intercepted is zero.

Note that here we assume the values of p_e does not depend on $N_{X_1X_2}$ and p_l. This assumption is not necessary but can be realistic in some scenarios, where attacker's capability of data manipulation is independent of how many interconnections are established, i.e., attackers have a fixed capability of data manipulation on every interconnection. Another way to model p_e can be that it is increasing with $N_{X_1X_2}$, since more interconnections may enable the adversaries to manipulate more data than they can do with fewer connections. In the appendix A, we show C_T by varying p_e as a function p_l i.e., as a function of $N_{X_1X_2}$.

2.3.4 Packet-loss distortion (D_l)

From the previous subsections, we know each interconnection has a Shannon capacity given by the interconnection (channel) model, and we also know the total expected Shannon capacity for $N_{X_1X_2}$ interconnections. The significance of the total expected



(a) Model for the packet-loss

(b) D_l vs. R_s

Figure 2.7: (a) The packet-loss distortion model at the source network, d denotes the lost packet, p_d is the dropping probability of a packet. (b) The total packet-loss distortion (D_l) vs. the source data rate (R_s).

Shannon channel capacity comes from Shannons channel coding theorem: for any $\delta > 0$, if the data rate is less than the channel capacity, then probability of data error can be made less than δ by channel coding of sufficiently large block length; on the other hand, if data rate is grater than the channel capacity, then the probability of error becomes greater than 0.5 when block length goes to infinity [57]. Note that here we interested to undo the malicious effect of attackers to maintain a secure communication. Hence, when the data rate of the source network exceeds the total expected Shannon capacity of the interconnections, some packets will be dropped by the gateway nodes[1]. From this observation, we model the packet-loss distortion as follows: if the source network data (packet) rate is R_s bps and the total expected

[1] Recall that here we are not considering the role of a queue and the associated queuing delay for simplicity, i.e., the delay is assumed to be infinite in this chapter. However, the queuing delay can be incorporated in the model as follows. Each gateway node holds some packets until its queue is full since the node has a queue of finite size. From the basic network theory, we know that the queuing delay increases exponentially with the increase of incoming data rate to the node [8].

Shannon capacity C_T bps, a reasonable formula for the probability of packet-loss is

$$p_d = \begin{cases} 0, & R_s \leq C_T \\ 1 - \frac{C_T}{R_s}, & R_s > C_T. \end{cases} \tag{2.9}$$

Since the loss of a packet (bits) is quantitatively similar to the erasure of a bit in the conventional erasure channel, it is reasonable to model the packet-loss distortion as an erasure channel. Figure 2.7a shows the packet-loss distortion model at the source network, where d denotes the dropped/lost packet. Here for simplicity, we assume each packet consists of one bit, and a packet-loss implies an one-bit drop.

Hence the total packet-loss distortion is given by $D_l = p_d R_s$ bps. Figure 2.7b shows the total packet-loss distortion versus source data rate for a different number of interconnections. Here when there is no interconnection, i.e., $N_{X_1X_2} = 0$, then the gateway nodes drop everything (solid black line). However, as we increase the number of interconnection the distortion (D_l) decreases, but it starts to increase again when the source data rate (R_s) is higher than the total expected Shannon capacity (C_T) of the interconnections. Recall that for the various number of interconnections we get different C_T's.

2.4 Optimal number of interconnections between two networks

In this section, we find the optimal number of interconnections between two networks based on the network administrators specified criteria such as the maximum data rate and minimum packet-loss distortion.

2.4.1 Optimal number of interconnections under the packet-loss constraint

At first, we seek to answer the following question: for any desired packet-loss distortion what is the maximum data rate an interconnected network can achieve, i.e., *what is the optimal number of interconnections that gives the maximum data rate by satisfying the specified packet-loss constraint?*

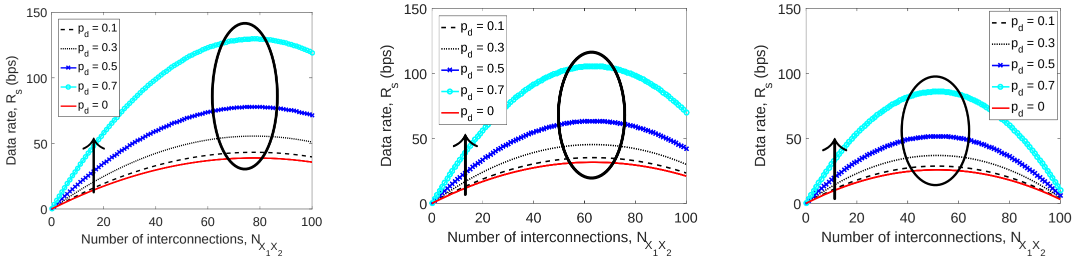
To answer the above question, we define the data rate of the source network considering the normalized packet-loss distortion p_d as

$$R_s(p_e, p_d) = \frac{C_T}{1-p_d} = \frac{\bar{C}N_{X_1X_2}}{1-p_d}, \quad p_d \neq 1, \tag{2.10}$$

where R_s is also a function of p_e, since $\bar{C}$ depends on p_e.

The source network data (packet) rate (R_s) versus the number of interconnections ($N_{X_1X_2}$) for several given normalized packet-loss distortions (p_d) is shown in Fig. 2.8, where we set $p_e = 0.25, 0.35, 0.55$ in three sub-figures, respectively. The solid red line indicates the total Shannon capacity (C_T) for the given interconnections, which is the maximum data rate one can achieve without having any distortions ($p_d = 0$). However, we see if one can tolerate more distortions, then he/she can achieve higher data rates for a given number of interconnections (shown by the uparrow in the figure). In addition, observe that R_s is a concave-like function of $N_{X_1X_2}$ and it achieves a maximum for some values of $N_{X_1X_2}$ (the rate is maximum somewhere inside the black oval shown in the figure). To find the optimal number of interconnection that attains the maximum data-rate for any given distortions we solve the following optimization problem.

In words, we find the optimal number of interconnections that maximizes the data rate of the source network over the number of interconnections and also satisfies any



(a) $p_e = 0.25$ (b) $p_e = 0.35$ (c) $p_e = 0.55$

Figure 2.8: The source network data rate (R_s) with respect to the number of interconnections ($N_{X_1X_2}$) for several normalized packet-loss distortion (p_d) under three p_e's. For any p_e's, the data-rate increases if we increase the packet-loss distortion.

desired distortion constraint (p_d^*). Mathematically,

$$R_s^*(p_e, p_d) = \max_{0 \leq N_{X_1X_2} \leq N_{max}} R_s(p_e, p_d) \text{ s.t. } p_d \leq p_d^*. \tag{2.11}$$

The output argument of this optimization problem is the optimal number of interconnections, call it $N^*_{X_1X_2}$. We solve this optimization problem as follows: note that to maximize $R_s(p_e, p_d)$, we need to minimize its denominator. Since $0 \leq p_d \leq p_d^* < 1$, then $1 \geq 1 - p_d \geq 1 - p_d^* > 0$. Hence as $p_d \to 0, R_s^*(p_e, p_d) \to C_T$, and, as $p_d \to 1, R_s^*(p_e, p_d) \to \infty$. As a result, if we increase p_d the denominator of R_s decreases and we get the higher value of R_s. Therefore, we replace p_d in (2.11) by its maximum value p_d^* to get

$$\begin{aligned} R_s^*(p_e, p_d) &= \max_{0 \leq N_{X_1X_2} \leq N_{max}} R_s(p_e, p_d^*) = \max_{0 \leq N_{X_1X_2} \leq N_{max}} \frac{C_T}{1 - p_d^*} = \max_{0 \leq N_{X_1X_2} \leq N_{max}} \frac{\bar{C} N_{X_1X_2}}{1 - p_d^*} \\ &= \max_{0 \leq N_{X_1X_2} \leq N_{max}} \frac{\Big((1 - p_l) + p_l C_l\Big) N_{X_1X_2}}{1 - p_d^*} \\ &= \max_{0 \leq N_{X_1X_2} \leq N_{max}} \frac{N_{X_1X_2} - \frac{N^2_{X_1X_2}}{N_{max}} + C_l \frac{N^2_{X_1X_2}}{N_{max}}}{1 - p_d^*} \quad (\text{using linear } p_l = \frac{N_{X_1X_2}}{N_{max}}) \\ &= \max_{0 \leq N_{X_1X_2} \leq N_{max}} \frac{N_{X_1X_2} - \frac{N^2_{X_1X_2}}{N_{max}}(1 - C_l)}{1 - p_d^*}. \end{aligned}$$

$$\tag{2.12}$$

Note that the objective function of (2.12) is a quadratic function of $N_{X_1X_2}$ with the optimization variable $N_{X_1X_2}$ is an integer and these types of optimization problems are called mixed-integer quadratic programming (MIQP) problem. Since $N_{X_1X_2}$ is a positive integer, we cannot directly optimize (2.12) due to the NP completeness of this problem. However, MIQPs can be optimized by using branch-and-bound algorithm with the well-known optimization tool Gurobi optimization [61]. The Gurobi optimizer gives a solution in the integer form specifying the gap between the upper bound and lower bound of the optimal solution, which we found to be very small for our problem (less than $10^{-6}\%$ for an $N_{max} = 100$).

On the other hand, we can solve (2.12) analytically if we assume $N_{X_1X_2}$ is a real number. Then, for any given distortion p_d^*, it is easy to show that the objective function of (2.12) is a concave function of the number of interconnections (see appendix B). Since the function is concave, if there exists a locally optimal point it will the global optimal point ([62] Ch. 4 Section 4.2.2). Hence if there exists a number of the interconnection that gives the maximum data rates it will be unique (*uniqueness of the number of interconnections*). In addition, if the objective function is differentiable then the function is continuous on a closed interval; hence, by the *extreme value theorem* the function have a maximum and minimum on the interval (*existence of the number of interconnections*). We first differentiate the objective function of (2.12) with respect to $N_{X_1X_2}$ and set it to zero to get

$$N_{X_1X_2} = \frac{N_{max}}{2(1 - C_l)}.$$

Hence by checking whether the solution $N_{X_1X_2}$ lies in the domain constraint Ω, the optimal number of interconnections for achieving the maximum data rate is

$$N^*_{X_1X_2} = \begin{cases} \frac{N_{max}}{2(1-C_l)}, & \text{if } 0 \leq C_l \leq 0.5, \\ N_{max}, & \text{if } 0.5 < C_l \leq 1 \text{ (due to the extreme value theorem).} \end{cases} \tag{2.13}$$

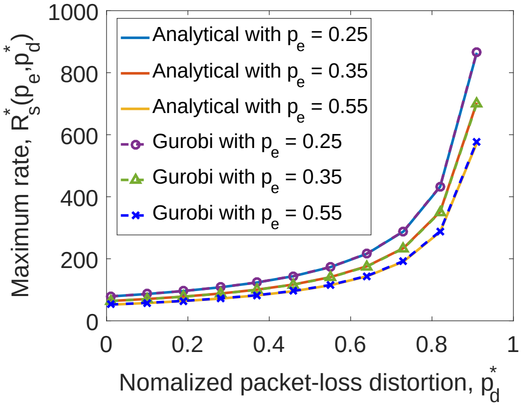


Figure 2.9: The rate-distortion curve that maximum achievable rate for any given distortions. Here the optimal number of interconnections with analytical solution are $N^*_{X_1X_2} = 77.7951, 63.2870, 51.5875$ and with numerical solution by Gurobi optimizer are $N^*_{X_1X_2} = 78, 63, 52$ for $p_e = 0.25, 0.35, 0.55$.

Note that in the analytical approach since we have assumed $N^*_{X_1X_2}$ is a real number, we need to round the $N^*_{X_1X_2}$ to the nearest integer if $N^*_{X_1X_2}$ is not an integer. As mentioned earlier the gap between the optimal solution and integer solution by Gurubi optimizer is quite small, i.e., the rate difference between the optimal analytical solution and the sub-optimal solution found by the Gurobi optimizer is actually negligible when comparing both results in Fig. 2.9.

Interestingly, the optimal number of interconnections that gives the maximum rate is only depends on N_{max} (maximum number of interconnection the networks can have) and p_e (attackers data manipulation capability). However, the maximum rate using these optimal interconnection for any distortion p_d^* does depend on the packet-loss distortion the source network can tolerate and is given by

$$R_s^*(p_e, p_d) = \frac{N^*_{X_1X_2}\bar{C}}{1 - p_d^*} = \begin{cases} \frac{N_{max}}{4(1-C_l)(1-p_d^*)}, & \text{if } 0 \leq C_l \leq 0.5, \\ \frac{N_{max}C_l}{(1-p_d^*)}, & \text{if } 0.5 < C_l \leq 1, \end{cases} \tag{2.14}$$

where C_l is given by (2.6) for a given p_f and p_ϵ.

The analytical and numerical solution result in a rate-distortion curve shown as in Fig. 2.9. Observe that the maximum rate increases with the increase of the normalized distortion. The significance of this rate-distortion curve is that based on the desired packet-loss distortion tolerance it can provide the maximum rate one can achieve without compromising the security at all.

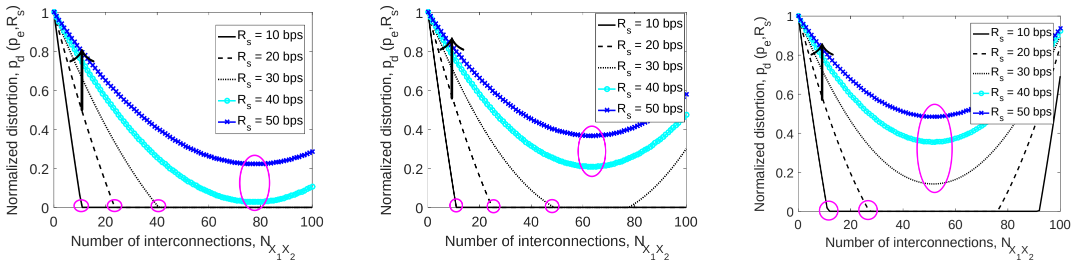
2.4.2 Optimal number of interconnections under data rate constraint

The previous subsection gives us the maximum rate achievable under a given distortion constraint. In this subsection, we are interested in the following question: *for any desired data rate R_s^* given by a network administrator, what is the minimum packet-loss distortion one can achieve over all possible set of interconnections?* Like the previous question, this one is also immensely important from the application perspective, where administrator's primary concern is the data rate. Mathematically, the optimization problem can be expressed as

$$p_d^*(p_e, R_s^*) = \min_{0 \leq N_{X_1X_2} \leq N_{max}} p_d(p_e, R_s) \text{ s.t. } R_s \geq R_s^*, \tag{2.15}$$

where $p_d(p_e, R_s)$ is defined in (2.9).

Figure 2.10 shows the normalized packet-loss distortion (p_d) concerning the number of interconnections ($N_{X_1X_2}$) for various data rates of the source network (R_s) under three p_e's. We see p_d decreases as we increase $N_{X_1X_2}$; however, p_d increases again if the R_s becomes high. As depicted by black uparrow, for any given number of interconnections, if we increase R_s then p_d also increases. Contrary to the rate versus interconnection curve shown in Fig. 2.8, here we observe that the p_d versus $N_{X_1X_2}$ curve is a convex-like function of the number of interconnections. Further, there exists a minimum number of interconnection that specifies the minimum distortion for the source network for a given any given R_s (shown by solid magenta



(a) $p_e = 0.25$ (b) $p_e = 0.35$ (c) $p_e = 0.55$

Figure 2.10: The normalized packet-loss distortion (p_d) vs. the number of interconnections ($N_{X_1X_2}$) for various given rates (R_s) under three p_e's.

ovals in Fig. 2.10). This minimum number of interconnections also varies with the values of p_es. Finally, we see that increasing the number of interconnections beyond the minimum amount is not productive in terms of minimizing the packet-loss distortion. This minimum number of interconnections is termed as the optimal number interconnections, which we found by solving (2.15) analytically as described below.

From the definition of $p_d(p_e, R_s)$ (see (2.9)), we see that p_d is increasing with respect to the rate R_s for a given total Shannon capacity C_T. Therefore, since we want to minimize p_d, we take the minimum of the rate constraint and set $R_s = R_s^*$. Then (2.15) becomes

$$\min_{0 \le N_{X_1X_2} \le N_{max}} p_d(p_e, R_s^*) = \max_{0 \le N_{X_1X_2} \le N_{max}} \begin{cases} 0, & R_s^* \le C_T, \\ 1 - \frac{C_T}{R_s^*}, & R_s^* > C_T. \end{cases} \tag{2.16}$$

In (2.16), C_T is the function of $N_{X_1X_2}$ and it is evident that to minimize $p_d(p_e, R_s^*)$, we need to maximize C_T. From the definition of C_T in (2.8) and it characterization curve shown in Fig. 2.6a, we see that for a given p_e (i.e., C_l) and N_{max}, there exists an $N_{X_1X_2}$ (i.e., $N^*_{X_1X_2}$) that results in the maximum C_T which we denote by C_T^* (to find the C_T^*, similar argument from the previous subsection can be used due to the concave nature of the C_T and assuming $N_{X_1X_2}$ is a real number). Then we have the following two cases:

case I: If $R_s^* \geq C_T^*$, then $N_{X_1X_2}^*$ is the optimal number of interconnections and the corresponding minimum packet-loss distortion is equal to $1 - \frac{C_T^*}{R_s^*}$.

case II: If $R_s^* < C_T^*$, then we will have zero packet-loss distortion (i.e., $p_d^*(p_e, R_s^*) = 0$). Note that C_T is concave function of $N_{X_1X_2}$ given by (2.8) (where C_T^* is the maximum of C_T), and since R_s^* is strictly smaller than C_T^*, then there are at most two points at which C_T and the line R_s^* intersect. To find those two intersection points let us solve the $C_T = R_s^*$. Using linear interception probability $p_l = \frac{N_{X_1X_2}}{N_{max}}$, we can write $C_T = N_{X_1X_2} - \frac{N_{X_1X_2}^2}{N_{max}}(1 - C_l)$. Therefore, from $C_T = R_s^*$, we get the following quadratic equation $N_{X_1X_2}^2(1 - C_l) - N_{max}N_{X_1X_2} + N_{max}R_s^* = 0$. The solution of this quadratic equation is given by $N_{X_1X_2} = \frac{N_{max} \pm \sqrt{N_{max}^2 - 4(1-C_l)N_{max}R_s^*}}{2(1-C_l)}$. Since we want the minimum number of interconnections then the optimal number of solution $N_{X_1X_2}^* = \frac{N_{max} - \sqrt{N_{max}^2 - 4(1-C_l)N_{max}R_s^*}}{2(1-C_l)}, C_l \neq 1$. Note that $N_{X_1X_2}$ is a real and nonnegative number because $N_{max}^2 - 4(1-C_l)N_{max}R_s^* \geq 0$, hence $R_s^* \leq \frac{N_{max}}{4(1-C_l)}, C_l \neq 1$ (this is not assumption and this condition hold automatically when $R_s^* < C_T^*$, see appendix C). Notably, for any N_{max} and p_e (i.e., C_l), we get the upper bound of the data-rate R_s^* that one can support without any packet-loss distortion. Finally, when $C_l = 1$, the quadratic equation becomes $N_{X_1X_2}^2(1 - 1) - N_{max}N_{X_1X_2} + N_{max}R_s^* = 0$, and hence the optimal number interconnection $N_{X_1X_2}^* = R_s^*$.

Since we assume $N_{X_1X_2}$ is a real number in above two cases, we need to round $N_{X_1X_2}^*$ to the nearest integer to find the (sub-optimal) number of interconnections in the integer form. Figure 2.11 shows the numerical results of the minimum distortion versus rate curve for three values of p_e. We know that for each p_e, we have a maximum total capacity C_T^* that results from a unique $N_{X_1X_2}^*$. Thus whenever the rate is lower than C_T^*, we have zero distortion. When the rate crossed C_T^*, the packet-loss distortion monotonically increases with rates.

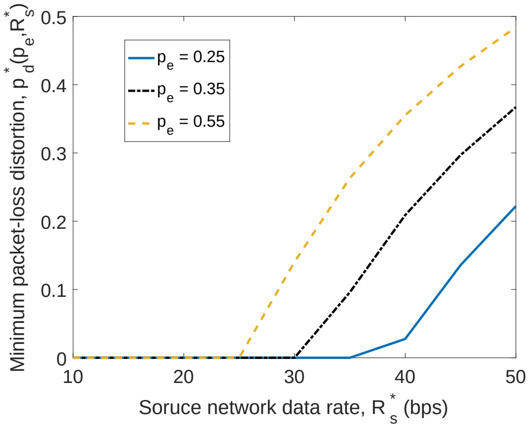


Figure 2.11: The distortion-rate curve

2.5 Summary and conclusions

In this chapter, we have formulated the optimal interconnectivity problem among two networks considering the security interceptions that may be introduced by the added interconnections. We have solved the optimal interconnection problem from two following perspectives: 1) for any given level of allowed packet-loss distortion, we have found the minimum number of interconnections that offers maximum data rate (shown in Fig. 2.8); 2) for any desired data rate, we have also found the minimum number of interconnections that allows communication between two networks with the minimum packet-loss distortion (shown in Fig. 2.10). Since the number of interconnections is optimal, we have shown that adding more connection beyond the minimum required number is not productive to overcome the packet-loss as well as to increase the data rates. In other words, we have proved that, there exists an optimal number of interconnections that specifies the maximum allowable data rate and minimum packet-loss distortion between two networks (shown in Fig. 2.9 and Fig. 2.11). Note that if the attackers ability of data manipulation (flipping/erasing) changes then the number of optimal interconnections will also change. Most importantly, using the Shannon capacity that ensures the existence of a suitable channel coding, our formulation outputs the optimal number of interconnections without compromising data security of the network.

Chapter 3

Vulnerability of communication networks under physical attacks

The functional reliability of many networks largely depends on the geographical topologies of networks as well as on the locations and impact (geographical extent and severity) of stressors [28, 63, 64]. Note that stressors imply those events that impose physical stresses (intense electromagnetic field, heat, pressure, etc.) over a network and can trigger network component failures. As we are dealing with large-scale stresses in this work, stressors and attacks/disasters will be used interchangeably. For example, in the case of a communication network, various network components such as switches, amplifiers/repeaters, multiplexers and links (fibers, copper cables, antennas, etc.) can fail either directly from the stressor-events or indirectly as a result of damage to the components or systems that support the communication network, e.g., outage of power [65, 66]. Based on the geographical extent and severity of stresses, the functionality of various networks can be impaired at different scales. Clearly, the physical topology of the network and the nature of stressors should be taken into consideration together in the analysis of network reliability.

Initial damage from certain stressors, e.g., WMDs, HEMPs and natural disasters, may exhibit a high degree of spatial correlation. Similar to an earlier work [29], we assume a Strauss point process to model the correlated locations of multiple stressors that may occur simultaneously. Furthermore, while in [29] the authors only consider a fixed-form Gaussian degradation function, we generalize it to several degradation functions (e.g., linear, circular and Gaussian) to describe different types of stress influences. This function will also be selected probabilistically to capture uncertainty in the types of stressors. Note that, a degradation function defines the shape, range and intensity of a stressor over a geographical area.

Inherently, the components of ML communication network possess different types of security and tolerance requirements based on their importance in the network. For instance, a military network or a fiber backbone network has extra security and more robustness to stressors than a commercial network due to the significance of these networks. We propose an extended probabilistic SRLG formulation that considers inherent connections among network components to calculate their coupled-vulnerabilities to stressors. The advantage of our new SRLG approach is that it does not require any upper layer (e.g., IP layer) information and allows us to compute the failure probability of each component based on failure of components that it depends on. Finally, we provide analytical and simulation results to demonstrate the overall behavior of a realistic ML network under different types of correlated stressor events.

3.1 Probabilistic model for failures of communication network-component

The goal of this section is to map the spatial distributions of various stressors to a probability distribution for the network components being functional while considering their coupled vulnerabilities.

3.1.1 Modeling correlated stressors

In general, multiple stressors can occur simultaneously either in one geographical location or they can spread over different locations depending upon the nature of stressors. Akin to the work done in [29], we have used the Strauss point process to represent spatially-inhomogeneous and spatially-correlated stressor centers. The Strauss point process enables us to model multiple stressor events simultaneously. The locations of these stressors are spatially correlated with each other on a geographic plane [30].

The spread and intensity of these stressors can be different depending upon the inherent shape and strength of stressors. For instance, a tornado yields different geographical impact than a nuclear attack or an earthquake. Based on literature survey, we have found three degradation functions, namely Gaussian [29], circular [67] and linear [24], that may reasonably characterize various real-world stressors. Figure 3.1 depicts one realization of these three types of degradation functions. A brief description of them is provided below.

Gaussian: Gaussian stressor intensity degrades according to the Gaussian function as the spatial distance from the location of occurrence increases. The variance of the Gaussian function specifies the range and intensity of the stressor on a geographic plane. Many real-world attacks and disasters exhibit a Gaussian nature approximately [29].

Circular: Given a stressor center, a circular degradation function is completely described by two parameters: radius of the circle and intensity of the stressor at the center. Intuitively, the only network component residing within the circle is affected and the intensity at any location is inversely proportional to its distance from the

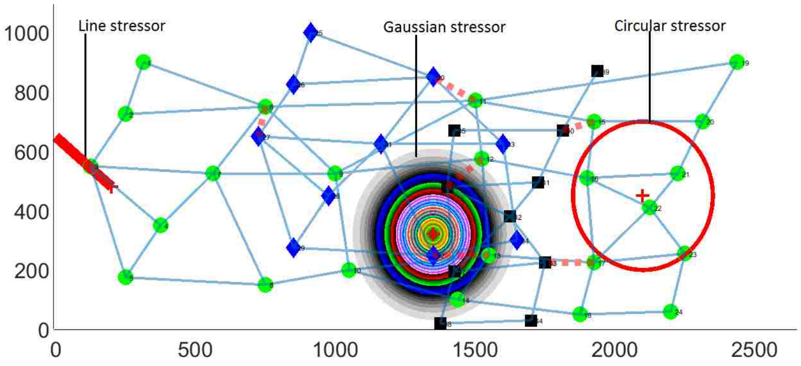


Figure 3.1: Three types of stressors on a simulated ML communication network. Nodes with three colors represent physical nodes of three communication networks which are placed on a (2500 × 1000) plane.

stressor-center.

Linear: We observed from the statistics of tornadoes that it can occur in any geographical location in USA. Typically, a tornado has a radius of 80 meters and length of 3 kilometers [68]. We consider the Poisson point process to model the locations of occurrence of the stressors. Unlike other disasters, a tornado has almost equal strength over the region it spreads. Hence, a uniform intensity all over the line is assumed.

3.1.2 Mapping stressor intensities to the distributions of failures of network components

We denote the stressor(s) event by $W = w$. We adopt the following assumption from [29]: **Assumption 1.** *Upon occurrence of a catastrophic stressor event (e.g., WMD, HEMP, natural disaster, etc.), the initial failure of any network component does not depend on other components.* Due to Assumption 1 and given a stressor event $W = w$, the joint failure probability of all network components can be written as the product of their individual failure probabilities. For nodes we have $p(v_1, v_2, ..., v_N|W =$

$w) = \prod_{i=1}^{N} p(v_i|W = w)$ and for links we have $p((v_1, v_2), ..., (v_{N-1}, v_N)|W = w) = \prod_{(v_i,v_j)\in E} p((v_i, v_j)|W = w)$, where $p(v_i|W = w)$ and $p((v_i, v_j)|W = w)$, denote the failure probability of the ith node and the (v_i, v_j) link, respectively.

Next we find the failure probability of each network component using the following procedure. Clearly, the likelihood of network component failure increases with the intensity of stressor and decreases with component's internal tolerance. Hence, we define the failure probability of ith node as

$$p(v_i|W = w) = \min\left(\frac{I_w(x_i, y_i)}{I_{v_i}(r, c)}, 1\right), \tag{3.1}$$

where $I_w(x_i, y_i) \geq 0$ captures the aggregated intensity of stressor at node v_i's location (x_i, y_i), and $I_{v_i}(r, c) > 0$ is the internal node tolerance. We define $I_{v_i}(r, c)$ by taking into account two realistic physical attributes of a node: $I_{v_i}(r, c) := r + c$, where $r \in (0, r_{max}]$ is a parameter to capture the resistance (e.g., shielding against HEMP [65]) assigned to a node based on its importance (e.g., higher node-degree or a backbone node) in the network. In addition, $c \in (0, c_{max}]$ captures the security requirement of a node being a network component in the ML network. The values of r and c can be estimated from the historical data. Note that, $I_w(x_i, y_i)$ is non-negative due to the fact that stressor intensity can only be positive or zero. Intuitively, all network components possess some resistance to the physical stressors that indicate $I_{v_i}(r, c)$ is a positive quantity. Hence, we have $0 \leq p(v_i|W = w) \leq 1$, thus $p(v_i|W = w)$ is a probability.

In order to find the link failure probability, we first find the stressor intensities over all points on the link. Then we take the maximal stressor intensity to consider maximum impact of the stressor to that link. Since a link can have an infinite number of points, we have taken $L_{(v_i,v_j)}$ number of points on the (v_i, v_j) link to find the maximum stressor intensity. The link failure probability for the (v_i, v_j) link can be written as

$$p((v_i, v_j)|W = w) = \min\left(\frac{\max\limits_{l\in\{1,\ldots,L_{(v_i,v_j)}\}} I_w(x_l, y_l)}{I_{(v_i,v_j)}(r,c)}, 1\right), \tag{3.2}$$

where (x_l, y_l) is the location of the lth point on (v_i, v_j) link. $I_{(v_i,v_j)}(r,c)$ is the tolerance of the (v_i, v_j) link that we define as the average of internal tolerances of the two nodes connected by the link: $I_{(v_i,v_j)}(r,c) := \frac{I_{v_i}(r,c)+I_{v_j}(r,c)}{2}$. The averaging of node tolerances in calculating the link tolerance is realistic. For example, if two nodes are very important then the link connecting them is assumed to have a great importance.

3.1.3 Characterizing coupled vulnerabilities among components of a network

We model the coupled vulnerabilities among network components using a variation of SRLG. First note that the functional vulnerability of a node directly affects the functionality of all links connected to it. Clearly, if a node fails then the links attached to it cannot be used anymore for communication. Therefore, for a stressor event $W = w$, by taking into account the coupled vulnerabilities between nodes and links, we find the failure probability of (v_i, v_j) link as

$$\begin{aligned} &p_{srlg}((v_i, v_j)|W = w) = \mathbf{P}((v_i, v_j) \cup v_i \cup v_j|W = w) \\ &= p((v_i, v_j)) + p(v_i) + p(v_j) - p((v_i, v_j))p(v_i)- \\ &p(v_i)p(v_j) - p((v_i, v_j))p(v_j) + p((v_i, v_j))p(v_i)p(v_j), \end{aligned} \tag{3.3}$$

where the last line follows from Assumption 1. For simplicity of notation, conditioning on the stressor event $W = w$ is removed from the second line. We summarize the link-failure as **Observation 1:** *The increase in failure probability of a communication node increases the failure probability of all links attached to it.* Similarly, link failures can cause node failures as well. For example, a HEMP wave that hampers a link can be carried through the link to the nodes connected to it. We can express the failure probability of the ith node considering all associated link failures as

$$\begin{aligned} p_{srlg}(v_i|W=w) &= \mathbf{P}(v_i \cup (\bigcap_{j\in\mathbb{N}}(v_i,v_j))|W=w) \\ &= p(v_i) + p(\bigcap_{j\in\mathbb{N}}(v_i,v_j)) - p(v_i)p(\bigcap_{j\in\mathbb{N}}(v_i,v_j)), \end{aligned} \tag{3.4}$$

where $\{j \in \mathbb{N} : v_j \in \text{Neighbor}(v_i)\}$ is the index set of the neighbors of node v_i. Again, conditioning on a stressor event $W=w$ is dropped from the notation. We now have **Observation 2:** *The increase in failure probability of all links attached to a node elevates the failure probability of that node.*

3.2 Performance measures

In this section, we define several parameters to evaluate the performance of network under different stressor scenarios.

Definition 1. Total expected capacity (TEC) of network: This metric measures the accumulated average (expected) capacity of all network links. Total capacity of a communication network: $C = \sum_{(v_i,v_j)\in E} C_{ij}$, where C_{ij} is the capacity of the (v_i, v_j) link. C_{ij} is a random variable that we define as

$$C_{ij} = \begin{cases} c_{ij}, & \text{with probablity } p(c_{ij}) = 1 - p((v_i,v_j)|W=w), \\ 0, & \text{with probablity } p(0) = p((v_i,v_j)|W=w), \end{cases}$$

where c_{ij} is the true capacity of the (v_i, v_j) link. We find the TEC of a network by taking conditional expectation ($\mathbf{E}[\cdot|\cdot]$) over C given a stressor $W=w$:

$$TEC = \mathbf{E}[C|W=w] = \sum_{(v_i,v_j)\in E} c_{ij}\left(1 - p((v_i,v_j)|W=w)\right).$$

Definition 2. Total expected number of node failures: Since the number of functional nodes is an important parameter for any network, we calculate the total expected number of node failures among N nodes after the occurrence of a stressor event. We define a random variable that captures the functionality of the ith node as follows

$$X_i = \begin{cases} 1 & \text{if the } i\text{th node fails with probablity } p(v_i|W = w), \\ 0 & \text{if the } i\text{th node is functional.} \end{cases}$$

The total number of node failures can be expressed as $X_T = \sum_{i=1}^{N} X_i$. Then the total expected number of failed nodes is $\mathbf{E}[X_T|W = w] = \sum_{i=1}^{N} p(v_i|W = w)$.

Definition 3. Max-flow between two nodes [28]: This parameter allows us to find the maximum data rate possible between any two fixed nodes in a network.

3.3 Simulation results

For simulation, a three-level communication-network architecture is presented that is scalable both in the number of levels as well as the size of the network in each level. Figure 3.2 depicts a prototype of the physical infrastructure of a ML communication network, which is composed of three real networks: TeliaSonera, Level 3 and Sprint. The physical topology dataset of these three networks are available in [69]. Note that each of TeliaSonera, Level 3 and Sprint networks consists of 21, 99 and 264 nodes, respectively, which are located all over USA. Three connections from both Level 3 and Sprint network are made with the TeliaSonera network based on the geographical distance between nodes and their associated node degrees. We have evaluated the performance of the ML communication network under different types of stressor scenarios. For each scenario we have generated 500 random samples with 2 stressor events. All links have a capacity of 1 Gbps (Gigabits per second) and

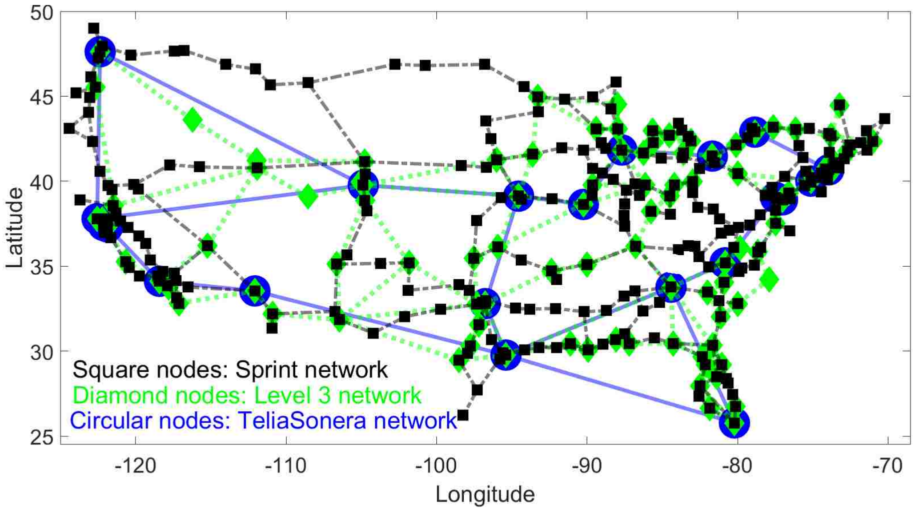

Square nodes: Sprint network
Diamond nodes: Level 3 network
Circular nodes: TeliaSonera network
Latitude
Longitude
50
45
40
35
30
25
-120
-110
-100
-90
-80
-70

Figure 3.2: A physical topology of a ML communication network composed of three real networks: TeliaSonera, Level 3 and Sprint.

intermediate point distance on links is approximately 10 miles. Node tolerances are assigned uniformly in (0, 2]. Moreover, the radius of the circular stressor is assumed to be 200 miles. For a line stressor, the line-direction is considered to be a free parameter within 0-360 degrees, since the line stressor can move to any direction after its occurrence.

Figures 3.3 and 3.4 depict the total expected number of node failures and the TEC of the ML network, respectively, for three different types of stressors. As expected, the TEC decreases and the total expected number of node failures increases with the increase of the parameter value of stressor. Depending on the stressor, the horizontal axis (*parameter of the stressor*) refers to the variance of the Gaussian stressor or the intensity at the center for circular stressor or the length of line for linear stressor. Notice that the network performance becomes worse for all scenarios while we consider the effect of SRLG among the network components. This is because one component failure contributes to the increase of failure probability of other components (Observations 1 and 2). For the particular parameter values assumed in

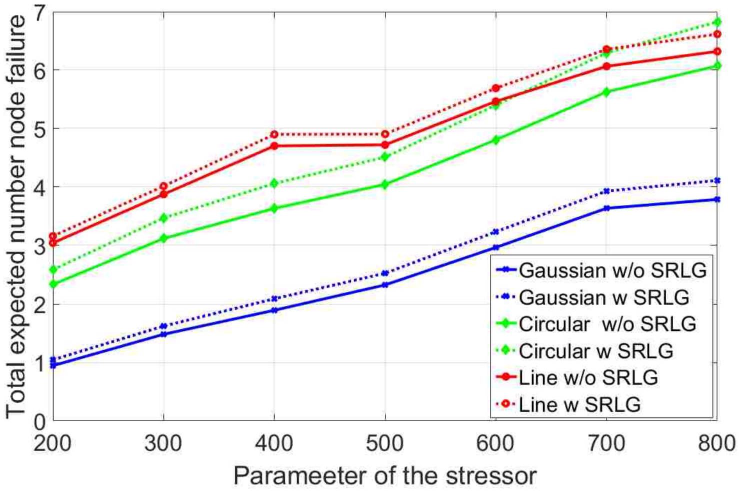


Figure 3.3: Total expected number of failed nodes in the network under different types of stressors with (w) and without (w/o) SRLG effects.

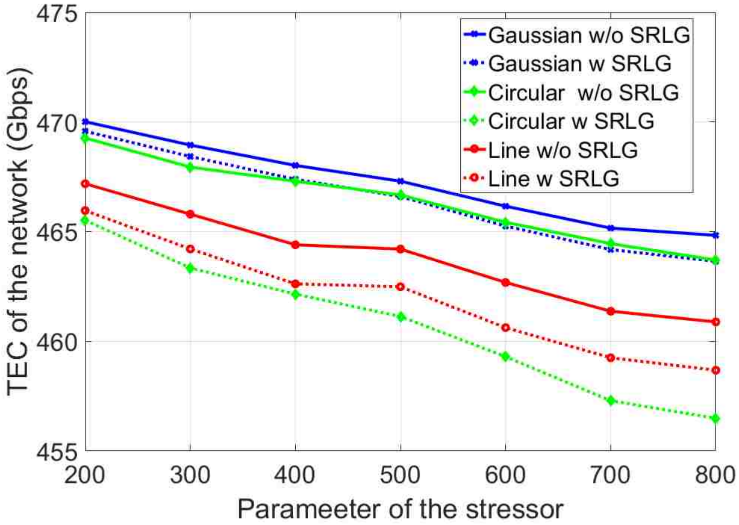


Figure 3.4: TEC of the network under various stressors with and without SRLG.

simulation, the ML communication network is less vulnerable under Gaussian stressor; however, different parameter values may yield different results, which is intuitive.

Figure 3.5 illustrates the Max-flow between two arbitrary fixed nodes (denoted

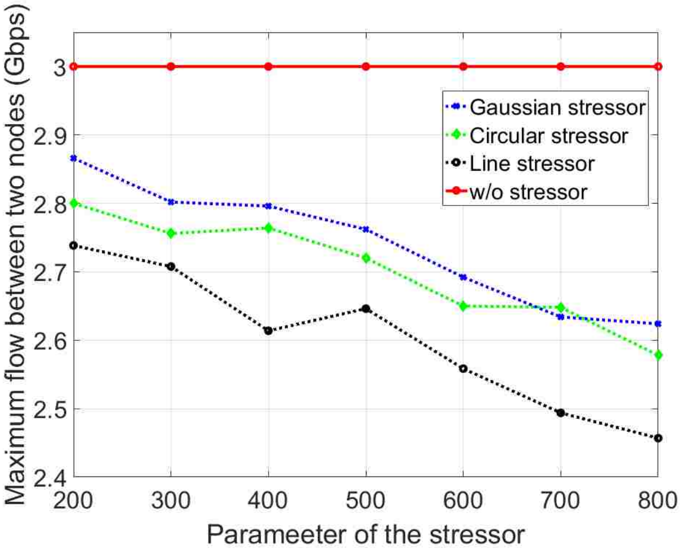


Figure 3.5: Maximum flow between two fixed nodes with SRLG.

by S and D in Fig. 3.2) in the network following a stressor. Here we have directly calculated the Max-flow considering the SRLG. Clearly, Max-flow achievable between these two nodes under normal operation is 3 Gbps, but due to the impact of stressors some nodes/links fail, thus the actual Max-flow between these two nodes is reduced.

3.4 Summary and conclusions

The reliability of a ML network is largely affected by the catastrophic attacks and natural disasters. In this chapter, we have described different types of correlated stressors that can potentially degrade the reliability of a communication network. We have also calculated the coupled-vulnerabilities among network components using a realistic SRLG formulation. Simulation results have shown that the inherent coupling among communication-network components notably increases their vulnerabilities to the large-scale stressors.

Chapter 4

Resiliency of the multilevel network under failure propagation

A communication network (e.g., military, private, commercial network) usually has its own independent and secure network infrastructure. However, an independent secure network may choose to connect with other networks to receive services and use it as a redundant communication medium whenever needed. Therefore, interconnections among independent subnetworks are inevitable for various reasons including the seamless communication among users of different networks, expansion of the communication capabilities among geographically-distant networks, backup communications in case of failures in the primary network, etc. In particular, interconnectivity between a small-sized network (consisting of a smaller number of nodes with lower link-bandwidth) and a large-sized network (consisting of a larger number of nodes with higher link-bandwidth) required for expanding the communication range and data rate of the small network. For example, a wide range of commercial and non-commercial communication systems and networks are used to support the military communications [53].

Owing to interconnections, the subnetworks of an ML network become interdependent through the communication gateways due to the exchange of information among each other. However, different networks usually have distinct security policies, control structures, and infrastructural vulnerabilities, as in the case of military and commercial networks for example [13]. Therefore, interconnections in ML networks can increase the vulnerability of a secure subnetwork due to threat propagation from less secure subnetworks. As a result, the composition of the individually secure system with different security policies is not secure [12]. For example, due to different levels of securities, the inter-operation and data sharing between military and commercial systems through interconnections may increase the probability of breaching security of the military node [13]. Moreover, in a ML network, if a node in a subnetwork is compromised by attackers then there is a possibility that a node in the other subnetwork may be compromised through the interconnected gateways. For example, interconnecting a highly secure network with the public Internet results in an increased vulnerability to the secure network by exposing it to cyber threats such as injection of malwares (viruses, worms), packet sniffing, denial-of-service (DoS) attacks (Section 1.6, [8]). In wireless networks, the internetwork links can be eavesdropped along with a strong possibility of jamming and sniffing [8, 70]. In fact, an adversary may get access to the top-secret data [12], which they can use for their benefit, such as eavesdropping links among nodes to extract critical information, locate the mobile nodes or military troops thus endanger their lives, traffic analysis [13], etc.

Clearly, interconnectivity among subnetworks needs to be addressed in order to compose an efficient and resilient ML network. In this work, we define the *efficiency of interconnectivity* of a ML network and model the *resiliency (vulnerability)* of a secure subnetwork due to the propagation of security risk through interconnections. The dynamics of security risk propagation are captured by two models, namely the evil-rain influence model [71] and the SIR (susceptible-infected-recovered) model [72].

In addition, we formulate two optimization problems that maximize the efficiency of interconnectivity with a constraint on the vulnerability/resiliency. We use different network topologies and interconnection patterns in our simulation and find the resiliency of a secure military network due to interconnections with vulnerable commercial networks. Based on simulation data, we propose two parametric analytic expressions in order to find the optimal number of interconnections that maximizes the efficiency of the ML network under resiliency or vulnerability constraints.

4.1 Vulnerability of a secure network

We have found that the dynamics of security risk propagation among subnetworks can be characterized by the existing evil-rain influence model and SIR epidemic model. These models are also used to model the propagation of a threat/risk in a network [38, 39]. Below we demonstrate how these two models are used to capture the propagation of security risks among subnetworks through interconnections.

4.1.1 The SIR epidemic model

The epidemic model is a dynamical model that captures the spread of a disease in a network of large populations [72]. Among different versions of the epidemic model, we use the SIR model to characterize the dynamics of risk propagation in a network. In the SIR model, all the nodes are susceptible to attack initially, as such one or more nodes can be infected by attackers. The infected node compromises its neighbors with a transmission probability, denoted by τ. Moreover, the infected node will be recovered/removed at the following time step by the recovery mechanism of the SIR model. In the SIR model, the *resiliency* of a network G with N nodes is defined as [38]

$$R(G) = 1 - \frac{\mathbf{E}[N_f] - \mathbf{E}[N_i]}{N - \mathbf{E}[N_i]}, \tag{4.1}$$

where **E**[extend of a cascade] is the average number of nodes eventually infected due to the propagation of security risk from initially compromised nodes. Note that $0 \leq R(G) \leq 1$, where $R(G)$ is 0 if all nodes are compromised.

In our ML network, a node in a less secure subnetwork is attacked initially; then the compromised node infects its neighbors with probability τ. Since the secure subnetwork is also a part of the ML network, the attack also propagates to the secure network.

4.1.2 The evil-rain influence model

The influence model is a networked Markov Chain (MC) framework for modeling interactions among nodes in a network. The internal functional dynamics of each node is captured by an MC and the influence received by a node from its neighbor is between 0 and 1, with the total influence received by a node from all its neighbors summing up to 1 [71]. A special case of the influence model is the "evil-rain model," where two autonomous external nodes, named "source of failures" and "source of repairs," are responsible for injecting failures and reparation in the network, respectively. We model the risk propagation from one node to other node through influences between the nodes; i.e., the probability of propagation of risk between two nodes is equal to the influence among them.

In a network G with N nodes, the expected number of compromised node (vulnerability) is given by [71]

$$V(G) = \mathbf{1}^T(\mathbf{I} - \mathbf{F})^{-1}\mathbf{u}, \tag{4.2}$$

where $\mathbf{u}$ is an N-dimensional vector that represents the external attack probability of each node, $\mathbf{1}$ is a column vector where all elements are 1, and $\mathbf{I}$ is an $N \times N$ identity matrix. Moreover, $\mathbf{F}$ is the interconnection structure that represents influences between nodes as (for 3 subnetworks): $\mathbf{F} = \begin{bmatrix} \mathbf{F}_{11}\mathbf{F}_{12}\mathbf{F}_{13} \\ \mathbf{F}_{21}\mathbf{F}_{22}\mathbf{F}_{23} \\ \mathbf{F}_{31}\mathbf{F}_{32}\mathbf{F}_{33} \end{bmatrix}$, where $\mathbf{F}_{ij}$ denotes the interconnection matrix between subnetwork i and j. $\mathbf{F}_{ij}(l,k) = 0$ indicates that there is no connection between the lth node of network i and kth node network j. Moreover, $\mathbf{F}_{ij}(l,k) = c$, $0 < c \leq 1$, implies that there is a connection with influence strength c between the lth node of network i and kth node network j. The higher the value of strength c the easier is the propagation of security risk from a compromised node to its neighbor, which could be, for instance, due to the lack of security solutions installed in their interface.

Similar to the resiliency in the SIR model, we define the resiliency as the fraction of nodes that are not compromised:

$$R(G) := 1 - \frac{V(G)}{N}. \tag{4.3}$$

4.2 Maximizing interconnectivity in ML networks

In this section, we maximize the efficiency of interconnectivity of a ML network under the resiliency and vulnerability constraints.

4.2.1 Efficiency of interconnectivity

Recall that the interconnectivity among different types of networks is essential for communicating outside their territory, redundant communication medium, etc., thus forming the ML network. The efficiency of connectivity among nodes for a network

G with N nodes is [73]

$$W(G) = \frac{1}{N(N-1)} \sum_{u \in V} \sum_{v \in V - \{u\}} \frac{1}{d(u,v)}, \tag{4.4}$$

where V is the set of N nodes, $d(u,v)$ is the shortest path distance between node u and v and g is the attenuation of the connection which is assumed to be 1. In words, $W(G)$ is the efficiency of information exchange among nodes over the network. The efficiency of connection between node u and v is inversely proportional to the shortest path distance between them. Note that, $d(u,v) = \infty$ implies there is no connection between node u and v, and $d(u,v) = 1$ implies there is a direct connection between u and v. In addition, with no connections among any nodes $W(G) = 0$ i.e., no node can communicate with other nodes in G. Interconnections enable communications among nodes (e.g., in a network where all nodes are directly connected with each other, $W(G) = 1$).

As we are interested in the interconnectivity among subnetworks, we define the efficiency of interconnectivity of a ML network G_m as

$$\hat{W}(G_m) := W(G_m) - W(G_0), \tag{4.5}$$

where G_0 represents the ML network without any interconnections among the subnetworks.

4.2.2 Interconnectivity optimization

Since different mathematical formulations are used in the SIR and evil-rain model for modeling the dynamics of risk propagation, we formulate two optimization problems based on these two models, which are described below.

First, for the SIR model, we maximize the efficiency of interconnectivity with a constraint on the resiliency as

$$\max_{e_{ij}, i \neq j} \hat{W}(G_m) \text{ subject to } R(G_s) \geq R_s, \tag{4.6}$$

where R_s is the minimum resiliency that is required for a secure subnetwork G_s, $e_{ij} \in \{0, 1\}$ represents the connection between node i and j, G_m is the ML network.

Similarly, using the evil-rain model we maximize the efficiency of interconnectivity under the vulnerability constraint,

$$\max_{\mathbf{F}_{ij}, i \neq j} \hat{W}(G_m) \text{ subject to } V(G_s) \leq V_s, \tag{4.7}$$

where V_s is the maximum vulnerability of the secure subnetwork G_s, $\mathbf{F}_{ij}$ is defined in the previous section and here $i \neq j$ since we optimize interconnection between different networks.

Note that, both optimization problems given by (4.6) and (4.7) are nonlinear and non-convex, for which no simple closed-form analytical solution or optimal algorithm exists. However, we recur to data from the numerical simulations in order to solve these optimization problems parametrically, which is demonstrated in the following section.

4.3 Numerical simulation

In this section, we introduce several network topologies and interconnection patterns followed by a description of the generation of ML network. We then find the efficiency of interconnectivity and resiliency for different number and patterns of interconnections using both SIR and evil-rain model and propose our parametric models based on simulation data.

4.3.1 Network topologies and interconnection patterns

We have considered different types of state-of-the-art network graphs to form a ML network. **Erdos-Renyi (ER) graph** [72]: ER graphs are random graphs where each node connects to other nodes independently with a given probability. **Barbasi and Albert (BA) graph** [72]: In a BA graph, each new node connects with some given number (a constant) of nodes with preferential attachment, which results in a scale-free graph. **Telia Carrier (TC) network**: TC network is a real-world physical network topology with 21 nodes and 25 links, which are located over the USA [69].

We have used following link patterns to simulate the interconnectivity among subnetworks. **Assortative Link (AL)**: Here the nodes with highest-degrees in one subnetwork connects to the nodes with highest-degree nodes in the other subnetwork, and so on. **Disassortative Link (DL)**: The highest-degree nodes in one subnetwork connect to the nodes with the lowest-degree in the other subnetwork. **Random Link (RL)**: Here connections among nodes are assigned randomly between two subnetworks. **1-1 Link (1-1)**: Nodes are connected with shortest physical distances, i.e., a node in one subnetwork connects with the closest node in other subnetworks.

4.3.2 Multilevel network generation

We generate a 3-level network similar to the one shown in Fig. 1.1. When all three constituent subnetworks of the 3-level network are the ER graph, we denote it as the ER-ER-ER network. Similarly, we form the BA-BA-BA network and the ER-TC-BA network. The physical topology graph of the TC network can be found in [69]. Since the TC network has 21 nodes, we have used 21 nodes for generating the BA and ER network. Moreover, the TC network is an connected graph and we generate the BA and ER networks so that these networks are also form two connected graphs.

In particular, for ER network, we assign edges between nodes with probability p such that the generated graph is an connected graph (here we use $p = 0.18$ and check whether the graph is connected). An connected BA graph is formed by using the algorithm proposed in [74] with an average node-degree equal to 3.2 and power-law exponent is 2.8. As shown in Fig. 1.1, we assume the level-1 network is the highly secure military network, whereas level-2 and level-3 networks are commercial networks with a lower level of security. While interconnecting these subnetworks to form a 3-level network, we have used the same number of interconnections to connect the military network with two commercial networks. The connection patterns are AL, DL, RL, and 1-1. Here, two commercial networks are used as two backup communication infrastructures for the military network, which can be scaled to any number of networks.

4.3.3 Simulation results

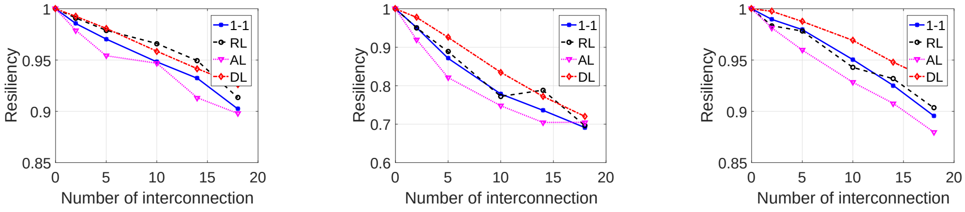
First, we discuss the simulation results of the SIR model. To simulate the SIR model, we assume a node in the less secure subnetwork (commercial network) is attacked (compromised) initially. Then the compromised node propagates the security threats to its neighbors in the 3-level network with a probability τ. Resiliency ($R(G)$) is calculated by using (4.1), where $\mathbf{E}$[extend of a cascade] is computed by averaging over 1000 realizations of the SIR model with one initial failure.

Figure 4.1 shows the resiliency of the military network versus the number of interconnections for the ER-ER-ER, BA-BA-BA and ER-TC-BA network. Here the resiliency decreases as we increase the number of interconnections. This is because with more interconnections the risk easily propagates to the military network from commercial networks. Observe that the DL connection performs better than the AL connection, which is due to the fact that the military nodes with smaller degrees are connected to commercial nodes. Hence, even if an interconnected military node is compromised, due to its smaller degree the probability of compromising many neighbors is low.

Figure 4.2 shows the efficiency of interconnectivity of the 3-level network for different number of interconnections, which can be computed by (4.5). Here, as we increase the number of interconnections the efficiency of interconnectivity increases due to new communication paths between subnetworks. Moreover, with AL connection the efficiency of interconnectivity of the 3-level network is higher than that for the DL connection, which is due to the higher node-degrees of the interconnected nodes.

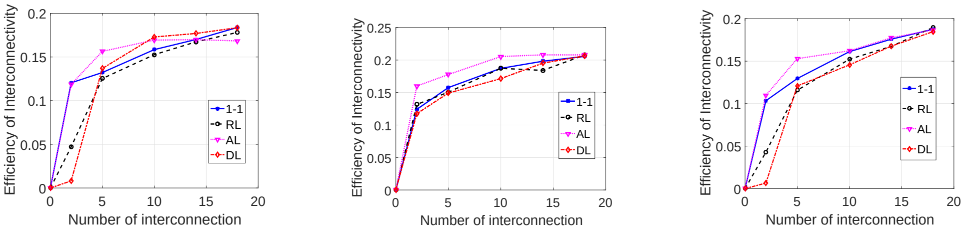
The vulnerability versus the efficiency of interconnectivity for three ML networks is shown in Fig. 4.3. The vulnerability of the military network increases with the efficiency of the network. Moreover, for any given efficiency the higher the value of τ, the vulnerability becomes higher due to the larger transmission probability of risks from the compromised nodes.

Finally, Fig. 4.4 shows the results using the evil-rain model for the ER-ER-ER network due to space limitations. As described in the model, here commercial networks have the "source of failures" with a given probability (0.20 in the simulation), thus failures start from the commercial network and propagate to the military network. However, the military network has the "source of repairs" with a probability (0.20 in the simulation), which prevents the failure of complete network. We can observe the similar trend as in the SIR model. Thus we conclude that the interconnection increases vulnerability of the secure military network. At the same time, the efficiency of interconnectivity among subnetworks also increases.



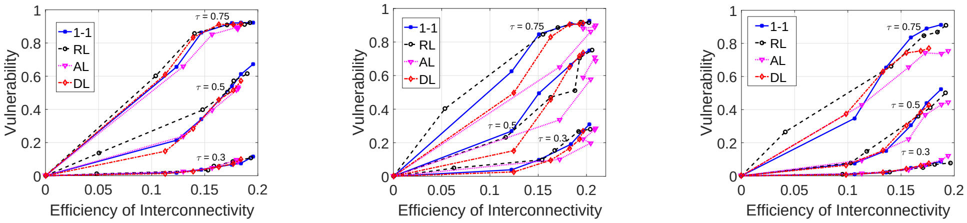
(a) ER-ER-ER network (b) BA-BA-BA network (c) ER-TC-BA network

Figure 4.1: Resiliency versus the number of interconnection for the ER-ER-ER, BA-BA-BA, and ER-TC-BA networks



(a) ER-ER-ER network (b) BA-BA-BA network (c) ER-TC-BA network

Figure 4.2: The efficiency of interconnectivity versus the number of interconnection for the ER-ER-ER, BA-BA-BA, and ER-TC-BA networks

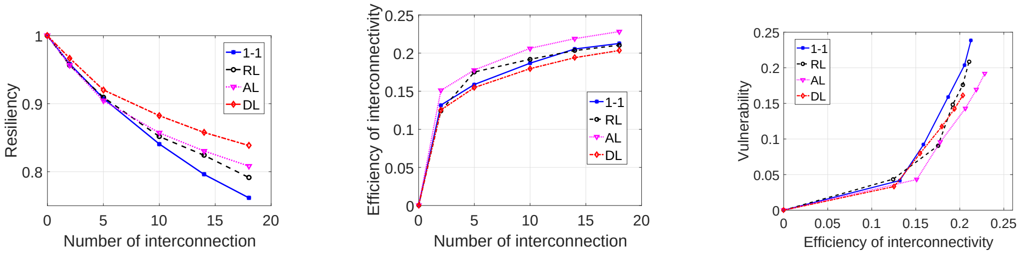


(a) ER-ER-ER network (b) BA-BA-BA network (c) ER-TC-BA network

Figure 4.3: The vulnerability of the military network versus the efficiency of interconnectivity for the ER-ER-ER, BA-BA-BA, and ER-TC-BA networks

4.3.4 Parametric model for the resiliency and efficiency

Motivated by the observed trends in the simulation data, we propose two parametric models for the efficiency of interconnectivity and resiliency for any given number of



(a) Resiliency vs number of interconnection

(b) Efficiency vs number of interconnection

(c) Vulnerability vs efficiency

Figure 4.4: Simulation results using the evil-rain influence model for the ER-ER-ER network

interconnection. From Fig. 4.1, observe that the resiliency is approximately linear with respect to the number of interconnections. We propose a parametric expression for the resiliency ($R(\tau, G)$) with l number of interconnections as,

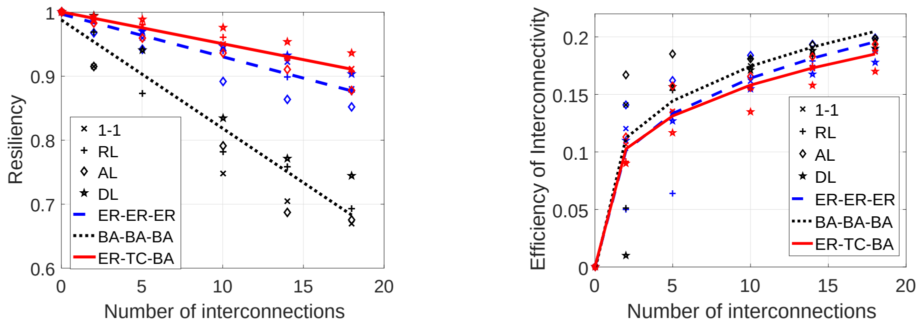
$$R(\tau, G) = a(\tau, G) + lb(\tau, G), \tag{4.8}$$

where $a(\tau, G)$, $b(\tau, G)$ are two parameters estimated from simulation data, G represents the network graph, τ is the transmission probability. We obtained the following values of the optimally fitted parameters: $a = 0.997$, $b = -0.007$ (ER-ER-ER network), $a = 0.998$, $b = -0.017$ (BA-BA-BA network), $a = 1.001$, $b = -0.005$ (ER-TC-BA network), which were then used to generate the fitted lines in the Fig. 4.5(a).

Interestingly, as shown in Fig. 4.2, the efficiency of interconnectivity follows a nonlinear relationship with the number of interconnection (l), which we approximate as the following,

$$\hat{W}(\tau, G) = \alpha(\tau, G) l^{\beta(\tau, G)} + \gamma(\tau, G). \tag{4.9}$$

Here the values of optimally fitted parameters: $\alpha = 0.0828, \beta = 0.2984, \gamma = -4 \times 10^{-4}$ (ER-ER-ER network); $\alpha = 0.0959, \beta = 0.2678, \gamma = -3.2 \times 10^{-3}$ (BA-BA-BA



(a) Resiliency vs the number of interconnection

(b) Efficiency vs the number of interconnection

Figure 4.5: Parametric fitting of the resiliency and efficiency of interconnectivity for different ML networks with four interconnection patterns. Here $\tau = 0.3$.

network); $\alpha = 0.0854, \beta = 0.2678, \gamma = -4.7 \times 10^{-5}$ (ER-TC-BA network); which were then used to find the fitted lines in the Fig. 4.5(b).

The values of parameters in (4.8) and (4.9) are computed by fitting the simulation data so as to minimize the overall mean-square-error (MSE). Higher-order polynomials with added complexities might yield more accurate fitting of the data. Apart from the complexity, the higher-order polynomials might over-fit the data points which is a serious drawback for prediction [75]. Therefore, we trade-off the complexity with slight inaccuracy to keep the model simple and avoid possible over-fitting error. Note that we have derived the parametric model for the SIR model due to space constraints. However, the model parameters can also be tuned for the evil-rain influence model as both models show qualitatively similar trends.

Our parametric models have great importance in deriving key insights. For instance, based on the given constraints (resiliency or vulnerability), one can obtain the number of interconnections by (4.8) and corresponding efficiency of interconnectivity by (4.9), that solves both optimization problems. One can also compute the efficiency of the ML network and resiliency of a secure network provided the number of interconnections. Moreover, the solution is independent of interconnection patterns (AL, DL, RL, 1-1) assuming the MSE tolerance.

4.4 Summary and conclusions

In this chapter, we have analyzed the dynamics of risk propagation in ML networks using the SIR epidemic model and evil-rain influence model. It is shown that increasing the number of interconnections among subnetworks results in a nonlinear increase of efficiency of the ML network; however, interconnectivity also decreases the resiliency of a secure network linearly. In order to maximize the efficiency of interconnection under resiliency/vulnerability constraints, we have proposed two parametric models. These models can be used to find the number of interconnections for exchanging information within the ML network when some subnetworks are vulnerable to security attacks.

Chapter 5

Dynamics of communication network under the influence of power failure

In the previous chapter, we did not consider failures in other infrastructures, such as the power grid, supplying electricity for the communication network, which is not true in reality. Modern network infrastructures are known to be interdependent due to the service they receive from each other [43]. For example, smart grids are dependent on the communication network due to the supervisory control and data acquisition (SCADA) system. On the other hand, communication networks rely on the power grids for their power supply. Similar interdependencies exist in other networks such as traffic systems, water networks, airline networks, water transport networks and sewer networks.

The physical infrastructures of the communication network are known to be vulnerable to various large-scale failures. Different types of natural disasters and attacks (e.g., earthquakes, hurricanes, weapons of mass destruction (WMDs), high-altitude electromagnetic pulses (HEMPs), etc.) can directly damage the critical network components such as switches, routers, amplifiers, electro-optical converters, fiber-optic line driver modules, receiver boxes, etc. [76]. Some real examples of catastrophic disasters that cause a massive infrastructural destruction of communication network include hurricane Katarina in USA and Mexico (2005), earthquake in Taiwan (2006), wildfires in California (2007), rail tunnel fire in Baltimore (2001), 9/11 attack that caused damage of telecommunication hub near world trade center [77, 78]. In addition, a strong electromagnetic pulse (EMP) generated by an HEMP, can produce high electrical currents within a short period of time. Such electrical pulses might get coupled with long-haul metallic conductor lines (e.g., landline communication systems, undersea optical fiber with copper power line, etc.) and fail the components connected to those lines [76, 79, 80].

The above mentioned initial failures can affect the failures of other network components inside a communication network in several ways: a) if all the neighboring nodes of a node fail then the node becomes disconnected from the network (or from the giant component), which implies that it is no longer playing a significant role in the network communication and can be considered as failed [20]; b) an HEMP wave that hampers a node due to high electrical current that can be carried through the links and may fail the neighboring communication nodes [76]; c) node removal (failure) can overload the other nodes, that in turn, might induce congestion in the communication networks [81]. For instance, congestion-induced Internet collapse was occurred in October 1986, when the speed of the connection between the Lawrence Berkeley Laboratory and the University of California at Berkeley, which are located only 200 meter apart, dropped by a factor of 100 [10]. As such, the failure of a communication node inside a network might influence functionality of the neighboring nodes. These types of influences among the communication nodes are termed as the *intra-network influence*.

Moreover, all communication nodes need power (direct or backup) to stay functional, therefore, the vulnerability of power networks can directly affect the supporting power supply components of the communication networks and can interrupt the functionality of the communication components. For example, due to northeastern USA and Canada blackout on 14–16 August 2003, 3175 communication networks of 1700 organizations (business entities, government, education, etc.) suffered from abnormal connectivity (voice/data) outages [11]. In addition, when a direct power supply is not available, assume a battery changes its state from functional to non-functional with a certain probability due to changes its internal state. As a result the communication components that are relying on batteries will be influenced by such a change of state, and therefore will have a higher probability of changing to the non-functional state. For instance, due to mismanagement of backup power supplies, the Delta airlines data center failure occurred on August 8^{th}, 2016, that causes cancellation of more than 300 flights, inconveniencing hundreds of thousands of people all over the world and costing Delta Airlines an estimated 150 million dollar [82]. In this work, we termed these influences as the *inter-network influence*.

Apart from the failure influences described above, in a repairable network the normal functionality of a component can increase the probability of proper functionality of its neighbors. It can also influence the repair-ability of a failed component with some probability. For example, functional nodes can share traffic with each other and can reduce the probability of overloading (congestion) of a communication node due to high traffic.

It is clear from above discussion that the role of both intra- and inter-dependencies should be taken into account to capture actual functional dynamics in a communication network. In this work, we develop a dynamical stochastic model to analyze the failure dynamics in a communication network considering both intra- and inter-dependency that exist in cyber-physical systems. The intra-dependency among communication nodes is captured using the influence model. The inter-dependency of the communication network with the supporting power supply infrastructures is modeled using an independent influence from each power node to the corresponding communication node. Under both influences, we derive an analytical closed-form equation to evaluate the steady-state probability of the communication network nodes. Numerical simulation has been carried out using a Sprint communication network topology to demonstrate the vulnerability of the communication network under intra-influence among communication nodes and the inter-influence from the supporting power supply nodes.

5.1 Influence model for network functionality

5.1.1 Influence model

Influence model is a networked Markov chain framework where the state evolution of each node in the network depends upon its internal Markov chain as well as the state of its neighbors and their influences on the node [71, 83]. Here we describe the influence among communication nodes nodes by a simple example. Figure 5.1 shows the network infrastructure for the communication network with six nodes considering power nodes (batteries) as a special node for the supporting power supply. Notice that the network of interest consists of two types of nodes: communication nodes, denoted by $n_1, n_2, ..., n_6$, and power nodes denoted by $b_1, b_2, ..., b_6$. There are two types of independent influences. First, the intra-network influence, i.e., the influence among communication nodes. Assuming uniform influence among all communication nodes for the network shown in Figure 5.1, we can express the intra-network influence

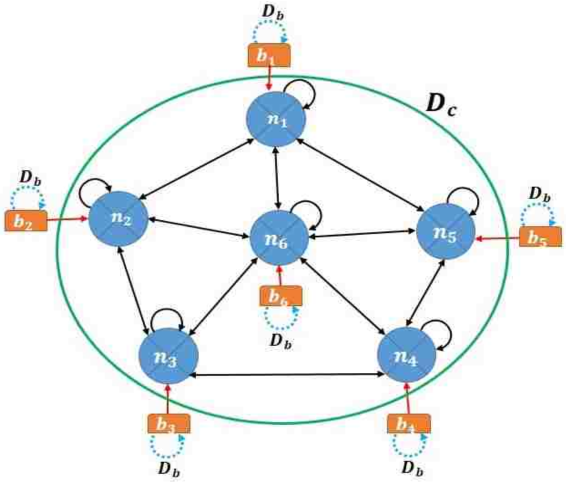


Figure 5.1: An example of the influence model for the communication network with power nodes.

matrix, D_c, as

$$D_c = \begin{bmatrix} 1/4 & 1/4 & 0 & 0 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 0 & 0 & 1/4 \\ 0 & 1/4 & 1/4 & 1/4 & 0 & 1/4 \\ 0 & 0 & 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 0 & 0 & 1/4 & 1/4 & 1/4 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{bmatrix},$$

where $D_{c_{ij}}$ is the influence received by ith communication node from the jth communication node. However, the D_c matrix does not capture the power influence on the communication node (inter-network influence), which we described in the following subsection.

5.1.2 Interdependent influence model

The influence model requires the sum of all the influences imposed on a node to be one. Consider a network where a communication node is influenced by its neighbors as well as the state of the supporting power node. According to the influence model, the total influence needs to be distributed between the power and communication nodes. As a result, even though the supporting power node fails there is still some probability that the communication node will be functional. However, in reality we know that if the supporting power node (direct power or battery) fails then the communication node fails certainly [20]. In particular, the backup supplies (e.g., battery, storage energy) provide power to the corresponding communication nodes for their normal operation when main power source fails by WMDs, HEMPs, or natural disasters. If a backup supply is working then the functionality of the corresponding communication node will depend on the intra-network influence. Whenever the backup supply fails the corresponding communication node will fail too. This scenarios cannot be captured by the traditional influence model. To consider this power influence, we model two separate influences on a communication node: a) intra-network influence (captures the influences among communication networks and adds up to one) and b) inter-network influence (captures the influence from the power node to the communication node and adds up to one). These two influences imposed on a communication node can independently fail a communication node. Interdependent influence model captures both intra- and inter-influences that can affect the functionality of a communication node in a communication network.

5.2 Failure dynamics in interdependent influence model

We assume finitely many nodes and denote the functionality status of ith node, $i \in \{1, ..., N\}$, as 0 if the node is off/failed and 1 if the node is on/functional. We represent the state of the ith communication node at time k without the power influence as $n_i[k]$. Similarly, $b_i[k]$ denotes the state of the ith power node at time k. And

Table 5.1: Functionality state ($s_i[k]$) of the ith communication node considering power influence

$n_i[k]$	$b_i[k]$	$s_i[k]$
0	0	0
0	1	0
1	0	0
1	1	1

$s_i[k]$ represents state of the ith communication node at time k with power influence. To make the analysis simpler, we assign the same self influence D_b to all power nodes; however, D_b can be node dependent (we omit explicit node dependence).

Since the state of a communication node depends on its internal state and the state of the corresponding power node, we define $s_i[k]$ as

$$s_i[k] = n_i[k]b_i[k]. \tag{5.1}$$

Using (5.1), the state of the ith communication node with power influence, $s_i[k]$, can be expressed as shown in table 5.1.

5.2.1 Modeling initial failures in communication networks

Recall that, initial failures in communication networks and power nodes occur due to large-scale stressors such as intensional attacks or natural disasters. At time $k = 0$, following a stressor event $W = w$ (e.g., WMD, HEMP), the conditional expectation of the ith communication node's state is given by

$$\begin{aligned}\mathbf{E}\big[s_i[0]\big|W = w\big] &= \mathsf{P}(s_i[0] = 1|W = w) = \mathbf{E}\big[n_i[0]b_i[0]\big|W = w\big] \\ &= \mathbf{E}\big[n_i[0]\big|W = w\big]\mathbf{E}\big[b_i[0]\big|W = w\big] \\ &= \mathsf{P}(n_i[0] = 1\big|W = w)\mathsf{P}(b_i[0] = 1\big|W = w),\end{aligned} \tag{5.2}$$

where $\mathsf{P}(\cdot)$ defines the probability of an event. The 3rd line follows from the fact that at time $k = 0$ given a stressor $W = w$, the failure probability of all nodes are mutually independent [29]. The initial failure/survival probability, $\mathsf{P}(s_i[0] = 1 \big| W = w)$, under a stressor (attack/disaster) event can be found from a model that captures the impact of the stressors

$$p(v_i|W = w) = \min\left(\frac{I_w(x_i, y_i)}{I_{v_i}(r, c)}, 1\right), \tag{5.3}$$

where $I_w(x_i, y_i) \geq 0$ captures the aggregated intensity of stressor at node v_i's location (x_i, y_i), and $I_{v_i}(r, c) > 0$ is the internal node tolerance (please refer to the chapter 3 for details).

5.2.2 Evolution of the state probabilities

In this subsection, we analyze the stochastic dynamics of the communication network by looking at the evolution of the node state probabilities. Note that, having determined $\mathsf{P}(s_i[0] = 1 \big| W = w)$ using (5.2), the initial state of all communication nodes at time $k = 0$ can be simulated independently using corresponding probability of the nodes. Assembling initial states for all N nodes, we express the state vector for all communication nodes at time $k = 0$ as $\mathbf{s}[0] = \begin{bmatrix} s_1[0] & s_2[0] & ... & s_N[0] \end{bmatrix}^T$, where T is the matrix transpose operator.

Next, in order to find the state of ith communication node with power influence at time $k = 1$, we need to know $\mathsf{P}(s_i[1] = 1)$, which is $\mathbf{E}\big[s_i[1]\big]$. We adopt the recursive relation between the current states and previous states, but with the newly added power influence. First, we proceed to find $\mathbf{E}\big[s_i[1]|\mathbf{s}[0]\big] = \mathsf{P}(s_i[1]\big|\mathbf{s}[0])$ as follows

$$\mathbf{E}\big[s_i[1]\big|\mathbf{s}[0]\big] = \mathbf{E}\big[n_i[1]b_i[1]\big|\mathbf{s}[0]\big] \tag{5.4a}$$

$$= \mathbf{E}\big[n_i[1]\big|\mathbf{s}[0]\big]\mathbf{E}\big[b_i[1]\big|\mathbf{s}[0]\big] \tag{5.4b}$$

$$= \mathbf{d}_i^T\mathbf{s}[0]\mathsf{P}(b_i[1] = 1|\mathbf{s}[0]), \tag{5.4c}$$

where (5.4b) follows from the fact that $n_i[1]$ and $b_i[1]$ are independent given $\mathbf{s}[0]$. Also in (5.4c), $\mathbf{E}\big[n_i[1]\big|\mathbf{s}[0]\big] = \mathbf{d}_i^T\mathbf{s}[0]$, is due to the influence in communication network from previous states, where $\mathbf{d}_i^T$ is the ith row of the D_c matrix. Specifically (5.4c) states that the conditional probability of ith node to be functional at time instant $k = 1$ is equal to the sum of influences from all the functional nodes times functional probability of the corresponding power node at that time.

Note that $\mathsf{P}(b_i[1] = 1|\mathbf{s}[0])$ needs to be calculated in order to compute (5.4). We use the following lemma to find it.

Lemma 5.2.1. $\mathsf{P}(b_i[1] = 1|\mathbf{s}[0]) = D_b\mathsf{P}(b_i[0] = 1)$.

Proof. Recall that the power nodes do not depend on the communication nodes. Therefore, $b_i[1]$ can be found from its previous state $b_i[0]$, which we do not know directly given $\mathbf{s}[0]$. Hence, we apply the law of total probability to find

$$\begin{aligned}&\mathsf{P}(b_i[1] = 1|\mathbf{s}[0]) = \mathsf{P}(b_i[1] = 1|s_i[0])\\&= \mathsf{P}(b_i[1] = 1|s_i[0] = 1)\mathsf{P}(s_i[0] = 1) + \mathsf{P}(b_i[1] = 1|s_i[0] = 0)\mathsf{P}(s_i[0] = 0).\end{aligned} \tag{5.5}$$

For the first term in (5.5), observe from the Table I that $\{s_i[0] = 1\} = \{b_i[0] = 1, n_i[0] = 1\}$. Therefore, $\mathsf{P}(b_i[1] = 1|s_i[0] = 1) = \mathsf{P}(b_i[1] = 1|\{b_i[0] = 1, n_i[0] = 1\}) = p\big(b_i[1] = 1|b_i[0] = 1\big) = D_b$, since a power node only influences itself. For the second term in (5.5), we have

$$\begin{aligned}&\mathsf{P}(b_i[1] = 1|s_i[0] = 0) = \sum_{b\in\{0,1\}} \mathsf{P}(b_i[1] = 1, b_i[0] = b|s_i[0] = 0)\\&= \sum_{b\in\{0,1\}} \mathsf{P}(b_i[1] = 1|s_i[0] = 0, b_i[0] = b)\mathsf{P}(b_i[0] = b|s_i[0] = 0)\\&= \sum_{b\in\{0,1\}} \mathsf{P}(b_i[1] = 1|b_i[0] = b)\mathsf{P}(b_i[0] = b|s_i[0] = 0)\\&= \sum_{b\in\{0,1\}} D_b\mathsf{P}(b_i[0] = b|s_i[0] = 0) = D_b\mathsf{P}(b_i[0] = 1|s_i[0] = 0).\end{aligned} \tag{5.6}$$

With above results we can rewrite (5.5) as

$$\begin{align}
&\mathsf{P}(b_i[1]=1|\mathbf{s}[0])=\mathsf{P}(b_i[1]=1|s_i[0]) \tag{5.7a}\\
&=D_b\mathsf{P}(s_i[0]=1)+D_b\mathsf{P}(b_i[0]=1|s_i[0]=0)\mathsf{P}(s_i[0]=0) \tag{5.7b}\\
&=D_b\big(\mathsf{P}(s_i[0]=1)+\mathsf{P}(b_i[0]=1,s_i[0]=0)\big) \tag{5.7c}\\
&=D_b\big(\mathsf{P}(b_i[0]=1,n_i[0]=1)+\mathsf{P}(b_i[0]=1,n_i[0]=0)\big) \tag{5.7d}\\
&=D_b\mathsf{P}(b_i[0]=1), \tag{5.7e}
\end{align}$$

where (5.7c) follows from Bayes rules and (5.7d) is due to the fact that: $\{s_i[0]=1\}=\{b_i[0]=1,n_i[0]=1\}$, and $\{b_i[0]=1,s_i[0]=0\}=\{b_i[0]=1,n_i[0]=0\}$. □

By substituting (5.7) into (5.4) we obtain

$$\mathbf{E}\big[s_i[1]\big|\mathbf{s}[0]\big]=\mathbf{d}_i^T\mathbf{s}[0]\times D_b\mathsf{P}(b_i[0]=1)=\mathbf{d}_i^T\mathbf{s}[0]\times D_b\mathbf{E}\big[b_i[0]\big]. \tag{5.8}$$

Similarly, we can find the state of all communication nodes at time $k=1$ and express those states in a vector form:

$$\begin{aligned}
\mathbf{E}\big[\mathbf{s}[1]\big|\mathbf{s}[0]\big]&=\begin{bmatrix}\mathbf{d}_1^T\mathbf{s}[0]\times D_b\mathbf{E}[b_1[0]]\\ \mathbf{d}_2^T\mathbf{s}[0]\times D_b\mathbf{E}[b_2[0]]\\ \vdots\\ \mathbf{d}_N^T\mathbf{s}[0]\times D_b\mathbf{E}[b_N[0]]\end{bmatrix}=\begin{bmatrix}D_b\mathbf{E}[b_1[0]]\\ D_b\mathbf{E}[b_2[0]]\\ \vdots\\ D_b\mathbf{E}[b_N[0]]\end{bmatrix}\circ\begin{bmatrix}\mathbf{d}_1^T\mathbf{s}[0]\\ \mathbf{d}_2^T\mathbf{s}[0]\\ \vdots\\ \mathbf{d}_N^T\mathbf{s}[0]\end{bmatrix}\\
&=D_b\left(\begin{bmatrix}\mathbf{E}[b_1[0]]\\ \mathbf{E}[b_2[0]]\\ \vdots\\ \mathbf{E}[b_N[0]]\end{bmatrix}\mathbf{1}^T\circ\begin{bmatrix}\mathbf{d}_1^T\\ \mathbf{d}_2^T\\ \vdots\\ \mathbf{d}_N^T\end{bmatrix}\right)\mathbf{s}[0]\\
&=D_b\big((\mathbf{E}\big[\mathbf{b}[0]\big]\mathbf{1}^T)\circ D_c\big)\mathbf{s}[0]=D_b\mathbf{E}\big[\mathbf{b}[0]\big]\circ(D_c\mathbf{s}[0]),
\end{aligned} \tag{5.9}$$

where ∘ denotes the entry-wise/Hadamard product and **1** denotes all column vector with all elements 1.

Following same procedure as detailed above, we can obtain the state of the ith communication node at time $k = 2$ as follows

$$\begin{aligned}
\mathbf{E}\big[s_i[2]\big|\mathbf{s}[0]\big] &= \mathbf{E}\Big[\mathbf{E}\big[s_i[2]\big|\mathbf{s}[1]\big]\Big|\mathbf{s}[0]\Big] && \text{(5.10a)}\\
&= \mathbf{E}\Big[\mathbf{d}_i^T\mathbf{s}[1] \times D_b\mathbf{E}\big[b_i[1]\big]\Big|\mathbf{s}[0]\Big] && \text{(5.10b)}\\
&= D_b\mathbf{E}[b_i[1]] \times \mathbf{d}_i^T\mathbf{E}\big[\mathbf{s}[1]\big|\mathbf{s}[0]\big] && \text{(5.10c)}\\
&= D_b\mathbf{E}[b_i[1]] \times \mathbf{d}_i^T\big(D_b\mathbf{E}[\mathbf{b}[0]] \circ (D_c\mathbf{s}[0])\big), && \text{(5.10d)}
\end{aligned}$$

where (5.10a) we used smoothing property of conditional expectation, (5.10b) and (5.10d) are follow from (5.8) and (5.9), respectively. Next, as done in (5.9), we can write the state vector of all communication nodes at time $k = 2$ as

$$\begin{aligned}
\mathbf{E}\big[\mathbf{s}[2]\big|\mathbf{s}[0]\big] &= \begin{bmatrix} D_b\mathbf{E}[b_1[1]] \times \mathbf{d}_1^T\big(D_b\mathbf{E}[\mathbf{b}[0]] \circ (D_c\mathbf{s}[0])\big) \\ D_b\mathbf{E}[b_2[1]] \times \mathbf{d}_2^T\big(D_b\mathbf{E}[\mathbf{b}[0]] \circ (D_c\mathbf{s}[0])\big) \\ \vdots \\ D_b\mathbf{E}[b_N[1]] \times \mathbf{d}_N^T\big(D_b\mathbf{E}[\mathbf{b}[0]] \circ (D_c\mathbf{s}[0])\big) \end{bmatrix} \\
&= \begin{bmatrix} D_b\mathbf{E}[b_1[1]]\mathbf{d}_1^T \\ D_b\mathbf{E}[b_2[1]]\mathbf{d}_2^T \\ \vdots \\ D_b\mathbf{E}[b_N[1]]\mathbf{d}_N^T \end{bmatrix} \Big(D_b\big((\mathbf{E}[\mathbf{b}[0]]\mathbf{1}^T) \circ D_c\big)\mathbf{s}[0]\Big) \\
&= \left(D_b \begin{bmatrix} \mathbf{E}[b_1[1]]\mathbf{1}^T \\ \mathbf{E}[b_2[1]]\mathbf{1}^T \\ \vdots \\ \mathbf{E}[b_N[1]]\mathbf{1}^T \end{bmatrix} \circ \begin{bmatrix} \mathbf{d}_1^T \\ \mathbf{d}_2^T \\ \vdots \\ \mathbf{d}_N^T \end{bmatrix}\right)\Big(D_b\big((\mathbf{E}[\mathbf{b}[0]]\mathbf{1}^T) \circ D_c\big)\mathbf{s}[0]\Big) \\
&= D_b^2\big((\mathbf{E}[\mathbf{b}[1]]\mathbf{1}^T) \circ D_c\big)\big((\mathbf{E}[\mathbf{b}[0]]\mathbf{1}^T) \circ D_c\big)\mathbf{s}[0].
\end{aligned} \tag{5.11}$$

Generalizing for all $k \geq 0$ we have

$$\begin{aligned}
\mathbf{E}\big[\mathbf{s}[k+1]\big|\mathbf{s}[0]\big] = D_b^{k+1}\Big[&\big((\mathbf{E}[\mathbf{b}[k]]\mathbf{1}^T) \circ D_c\big)\big((\mathbf{E}[\mathbf{b}[k-1]]\mathbf{1}^T) \circ D_c\big) \\
&...\big((\mathbf{E}[\mathbf{b}[1]]\mathbf{1}^T) \circ D_c\big)\big((\mathbf{E}[\mathbf{b}[0]]\mathbf{1}^T) \circ D_c\big)\Big]\mathbf{s}[0].
\end{aligned} \tag{5.12}$$

To simplify (5.12) and express it using only the initial state, we find $\mathbf{E}\big[\mathbf{b}[k]\big], k > 0$, as

$$\begin{aligned}\mathbf{E}\big[b_i[k]\big] &= \mathsf{P}(b_i[k] = 1) && (5.13a)\\ &= \mathbf{E}\Big[\mathbf{E}\big[b_i[k]\big|b_i[k-1]\big]\Big] && (5.13b)\\ &= \mathbf{E}\big[D_b b_i[k-1]\big] && (5.13c)\\ &= D_b^k \mathbf{E}\big[b_i[0]\big], && (5.13d)\end{aligned}$$

where (5.13b) is due to the smoothing property of conditional expectation, and (5.13d) is the result of applying smoothing property repeatedly. Repeating (5.13) for all i and expressing (5.13) in matrix form we have

$$\mathbf{E}\big[\mathbf{b}[k]\big]\mathbf{1}^T = D_b^k \mathbf{E}\left[\mathbf{b}[0]\mathbf{1}^T\right]. \tag{5.14}$$

Finally, by substituting (5.14) into (5.12) we obtain

$$\begin{aligned}&\mathbf{E}\big[\mathbf{s}[k+1]\big|\mathbf{s}[0]\big]\\ =&D_b^{k+1}\Big[\big((D_b^k\mathbf{E}\left[\mathbf{b}[0]\mathbf{1}^T\right])\circ D_c\big)\big((D_b^{k-1}\mathbf{E}\left[\mathbf{b}[0]\right]\mathbf{1}^T)\circ D_c\big)\dots\\ &\big((D_b^1\mathbf{E}\left[\mathbf{b}[0]\right]\mathbf{1}^T)\circ D_c\big)\big((D_b^0\mathbf{E}[\mathbf{b}[0]]\mathbf{1}^T)\circ D_c\big)\Big]\mathbf{s}[0]\\ &= D_b^{(k+1)(k+2)/2}\big((\mathbf{E}\big[\mathbf{b}[0]\big]\mathbf{1}^T)\circ D_c\big)^{k+1}\mathbf{s}[0].\end{aligned} \tag{5.15}$$

If we assume $D_b = 1$, which means that a power node receives full influence from its previous state only, then in the limit as $k \to \infty$ (5.15) can be expressed as

$$\begin{aligned}\lim_{k\to\infty}\mathbf{E}\big[\mathbf{s}[k+1]\big|\mathbf{s}[0]\big] &= \lim_{k\to\infty}\big((\mathbf{E}[\mathbf{b}[0]]\mathbf{1}^T)\circ D_c\big)^{k+1}\mathbf{s}[0]\\ &= \begin{cases}\mathbf{0}, & \mathbf{0} \le \mathbf{E}[\mathbf{b}[0]] < \mathbf{1}\\ \mathbf{1}\boldsymbol{\pi}_c^T\mathbf{s}[0], & \mathbf{E}\big[\mathbf{b}[0]\big] = \mathbf{1},\end{cases}\end{aligned} \tag{5.16}$$

where $\boldsymbol{\pi}_c$ is the left eigenvector of D_c and is normalized to 1; the justification for this limit is given below.

The result given by (5.16) gives us the steady state probability of each communication node considering impact of the corresponding power node. For a connected network that results in an ergodic network influence matrix, all nodes in the binary influence model always reach a consensus state: all zeros or all ones (*Theorem 3.6* [71]). Here, (5.16) gives us a new but intuitive conclusion: when all the batteries are always functional, the functionality of communication nodes only influenced by their intra-influence; whereas, when all batteries have some positive probability of initial failure, then all the communication nodes will eventually fail. Equations (5.1) - (5.16) results in next theorem which incorporates the influence of batteries in a communication network.

Theorem 5.2.2. *Assume a connected communication network (i.e., ergodic network matrix D_c) that has associated power node with each communication node. Also assume that* $\mathbf{s}[0]$ *is the initial state of all communication nodes with the power influence and* $\boldsymbol{\pi}_c$ *is the normalized left eigenvector of D_c. Then all the communication nodes will eventually fail if all the power nodes have some positive initial failure probabilities. However, if all the power nodes are always functional then all the communication nodes will survive (fail) with equal probability given by* $\boldsymbol{\pi}_c^T\mathbf{s}[0]$ $\left(1-\boldsymbol{\pi}_c^T\mathbf{s}[0]\right)$.

5.3 Simulation results

Figure 5.2 depicts the physical topology of the Sprint network given in [69]. Sprint network consists of 264 nodes and 313 links, which are located all over United States of America (USA) as shown in Fig. 2. We have used this real network topology to validate our analytical model that captures the influences among communication nodes as well as power nodes to track the failure dynamics in communication network. The influence among the network nodes are defined based on the degree of the nodes. For example, if a node has degree d then the influences it receives from its neighbors

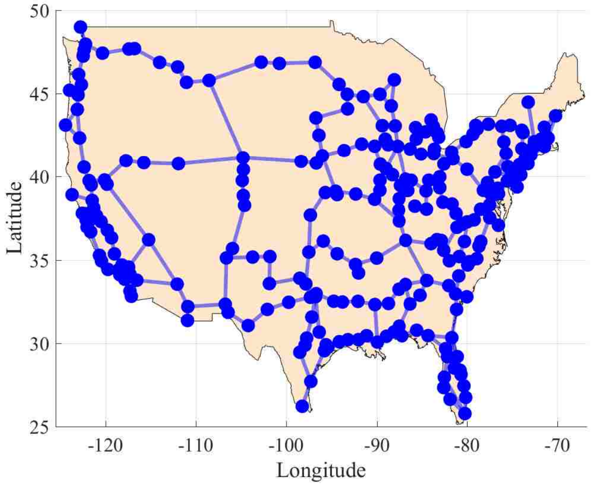


Figure 5.2: Topology of the Sprint network.

are given by $1/(d+1)$. The addition of 1 in the denominator with d is due to the self-loop influence. We find the steady state of the network (all communication nodes functional/non-functional) using Monte-Carlo simulation. All results are averaged over 1000 experiments with 5000 time steps in each experiment. Here we define the failure of all communication nodes as the *network failure*.

At first, we assume that all power nodes are "*on*" throughout the time, i.e., when $\mathbf{E}\big[\mathbf{b}[0]\big] = \mathbf{1}$, to show the impact of intra-influences among communication nodes. (5.16) demonstrates the case when all node power nodes are "*on*", the steady state of the communication nodes depends on eigenvector of D_c (i.e., intra-influences among communication nodes) and initial states of communication nodes. Earlier in this section, we define D_c based on the node-degrees. Here, the initial failure probabilities of all the communication nodes are drawn from an uniform distribution form {[0, 0.1], [0.1, 0.2], ... , [0.9, 1]}. Uniform distribution accounts for various stressor intensities imposed on the communication nodes based on their relative distance from the stressor center as well as tolerance to the stressors. In addition, we use different ranges of the initial failure probabilities to model stressors with various attack strengths. Fig. 5.3 shows the probability of network failure versus average initial failure probability of all communication nodes. As we can see the Monte Carlo simulation result matches closely with our theoretical result given in (5.16). In other words, when all the power nodes are always functional, the network failure probability due to intra-influences and initial failures found in (5.16) validated by the simulation results. In addition, when $\mathbf{E}\big[\mathbf{b}[0]\big] < \mathbf{1}$ (for any non-zero initial failure probabilities of all power nodes), network failure will occur almost surely at the steady state, which is shown in Fig. 5.3. Recall that, this is due to propagation of the initial power failures through their inter-influences to corresponding communication nodes. Knowing that the inter-influences will cause network failure at steady state, we look at the time taken for the network to go to the steady state for different initial failure probabilities of power nodes. This steady state reaching time metric will allow network operator to estimate the time they will get before the complete network failure based intensity of attacks/disasters, take proactive actions to prevent this network collapse. As considered for communication nodes, we assigned uniform initial failure probabilities to the power nodes. Fig. 5.4 illustrates that the time to network failure decreases with the increase of the initial power failure probability. Furthermore, the larger initial failure probabilities of the communication nodes (denoted by p_n) accelerates faster failure of the network. We can see that larger initial damage by catastrophic stressors can cause failure of the network at a faster rate, which becomes worse when both communication and power nodes vulnerable to attack. Therefore, making at least one of the infrastructures robust against massive disasters we can prolong the time of network collapse.

Earlier results in this Section demonstrate the cases of complete network failures considering impact of intra and inter-influences. As described in Section I, in a repairable network intra-influences among communication nodes can reduce overloading neighboring nodes. To incorporate repair-ability, we assign an ergodic transition matrix to each communication nodes. This added repair-ability feature can prevent

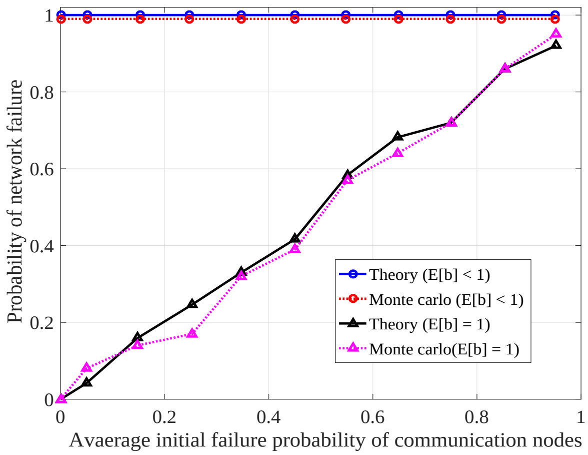


Figure 5.3: Validation of theoretical result by Monte-Carlo simulation: probability of network failure considering impact of initial communication node failures. For $\mathbf{E}\big[\mathbf{b}[0]\big] = \mathbf{1}$, intra-influences among communication nodes are responsible initial failure propagation and results in network failure, which is proportional to initial failure probabilities. When $\mathbf{E}\big[\mathbf{b}[0]\big] < \mathbf{1}$, inter-influences from power nodes cause network failures for any average initial communication failure probabilities.

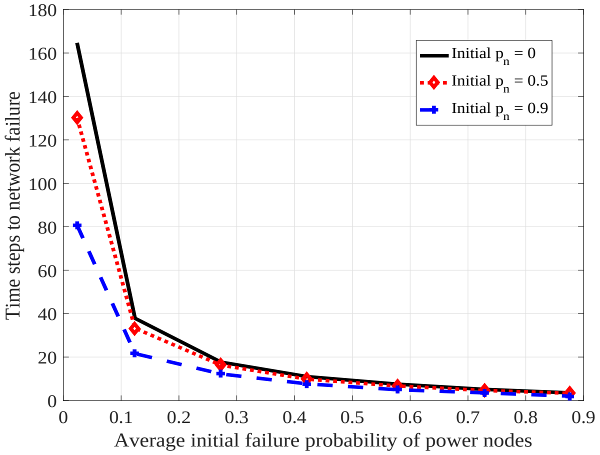


Figure 5.4: Time steps needed for network failure (with intra-influences and inter-influences) based on different average initial failure probabilities of power nodes.

the complete network failure as demonstrated below by simulation. 5.5 depicts the distribution of number of communication node failures considering both intra- and

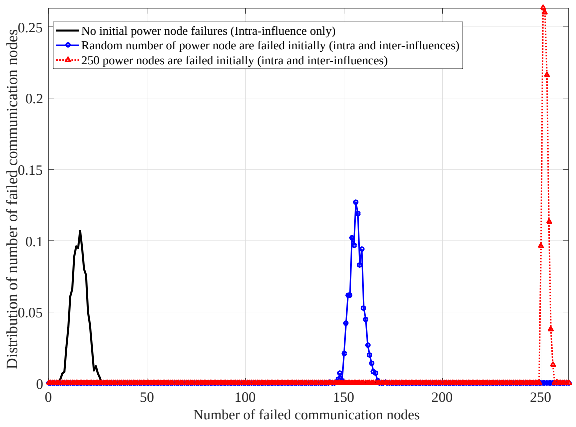


Figure 5.5: Distribution of number of communication node failures with intra influences only and with both with intra- and inter-influences considering repair-ability of the communication nodes.

inter-influences. Without loss generality, in the simulation we use an arbitrary transition matrix ($\begin{bmatrix} 0.95 & 0.05 \\ 0.80 & 0.20 \end{bmatrix}$) for all communication nodes, where first and 2nd states are functional and failed states of communication nodes, respectively. We can see that without inter-influences (solid line: no initial power node failure) distribution has mean failures of 15 communication nodes which are due to intra-influences in communication networks, whereas with inter-influences the mean shifted to larger value (156 and 252 for two cases); however, network does not got to complete failure state due to repair-ability that propagates in the network through intra-influences.

5.4 Summary and conclusions

The initial failure of a communication node can occur in two ways: direct failure of the communication node by a stressor or due to the power outage (direct supply or stored energy it requires for its operation). These initial failures can cause failures of neighboring network components due to their inherit dependency on each other. We have proposed a dynamical model to track the functionality of the communication network following a destructive event. Our model captures intra-dependency among communication nodes as well as the power influence on the communication node to analyze the failure dynamics in communication networks. We provide a tractable mathematical framework to predict the functionality state of the communication network in steady state. We conclude that both intra- and inter-dependencies are responsible for propagating failures in communication networks.

Chapter 6

A data-driven model for simulating the line failure in power grids

The power grid is a critical infrastructure in our society as almost all of our activities rely on electrical power. However, power grids are failure-prone. Man-made attacks such as weapons of mass destruction, high-altitude electromagnetic pulses, and natural disasters such as earthquakes, tornadoes, hurricanes, etc., can cause damage to critical components of any physical infrastructure [84]. For a power-grid, these include failures of transmission lines, generators, and transformers. For instance, in 2003 the initial failure of transmission lines and generating units contributed to extensive power blackouts in the United States and Canada [85]. Besides, power generation and consumption must be balanced for stable operation of a power grid. Initial failures can instigate instability, which, in turn, may cause additional failures (generally transmission lines and generators), and so on. Moreover, failures in communication network and human operator's error may contribute to propagate the initial failures [86, 87]. This phenomenon of continuing and uncontrolled successive failures of grid components is termed as a *cascading failure* [49].

A cascading failure in the power grid is a complex process due to a large number of components, physical attributes, and operating parameters governing the operation of the grid. For example, by considering only two parameters, namely, load growth and power fluctuations, it was shown in [88] that a cascading failure may result in an abrupt breakdown of a power grid. Moreover, in [89] the authors showed that cascading failures cannot be exactly captured by graph-theoretic epidemic and percolation models. The reason is that whenever a transmission line fails the next line failure can occur anywhere in the power grid, not necessarily at the neighboring lines, as typically assumed in an epidemic model. Real-world power outages also exhibit such noncontiguous line failure propagation [90, 91].

In addition to the noncontiguous nature of the transmission line outages, it has been observed that the failure of high-capacity transmission lines usually has a higher impact on the propagation of cascading failures in a power grid than the failure of low-capacity lines [92]. In fact, the system is designed to better protect these high-capacity lines. Hence, it becomes important to relate the propagation of transmission line failures in the power grids to the capacities of the failed lines. In this work, we recur to data from optimal DC power-flow simulations in order to characterize the evolution of cascading failures by also taking into account the capacities of the transmission lines. We perform numerical simulations on the IEEE 118 bus system using MATPOWER [93] under various operating conditions. Then we propose a parametric model that captures the dynamics of the evolution of transmission line failure in a power grid. With the proposed parametric model, we keep track of failures through lines of different capacities.

This chapter is organized as follows. A brief description of the relevant cascading failure models from the literature is given in Section 1.2.5 . In Section 6.1, we present the simulation data generated by MATPOWER on the IEEE 118 bus system. The proposed parametric model is introduced and elaborated in Section 6.2. In Section

6.3, we compare the results obtained from the model with real-world data. Our conclusions and future work are presented in Section 6.4.

6.1 Simulation data

We have used MATPOWER (a MATLAB package) [93] to simulate cascading failures by solving optimal DC power-flow equations on the IEEE 118 bus system. The optimal power-flow solution by MATPOWER is a quasi-static approach for simulating the cascading failures in the power grid, which is extensively used in the literature [41, 94, 95]. We have used the IEEE 118 bus system since it is a simple approximation of real power grids of the United States. The power transmission line capacities are assigned from the set of capacities $C = \{20$ MW, 60 MW, 120 MW, 200 MW, 332 MW$\}$, and the total number of lines with these capacities are $38, 58, 56, 20, 7$, respectively. The cardinality of the set C is denoted by $|C|$, and here $|C| = 5$. In addition to the number of initial line failures (denoted by F_i), some other relevant grid operating characteristics that can affect the cascade are the power-grid loading level (r), the load-shedding constraint level (θ), and the capacity estimation error (e) [41]. Here r is defined as the ratio between the total demand and the total generation capacity of the power grid; $\theta \in [0, 1]$ represents the load controlling capability of the power grid; it is 1 if no load shed can be performed, and it is 0 if all the loads are shed-able whenever needed. Moreover, e controls the capacity estimation error measured by the control center and is defined as $e := \frac{C_j^{opt} - \alpha_j}{C_j^{opt}} \in [0, 0.5]$, where C_j^{opt} is the estimated optimal power-flow through the transmission line of capacity $j \in C$ and $\alpha_j \in C$ is the power-flow threshold of the line above which the protection relay trip the transmission line [41]. Note that here we are following [41] in looking at these parameters.

In our simulations, we start with more than one random initial transmission line

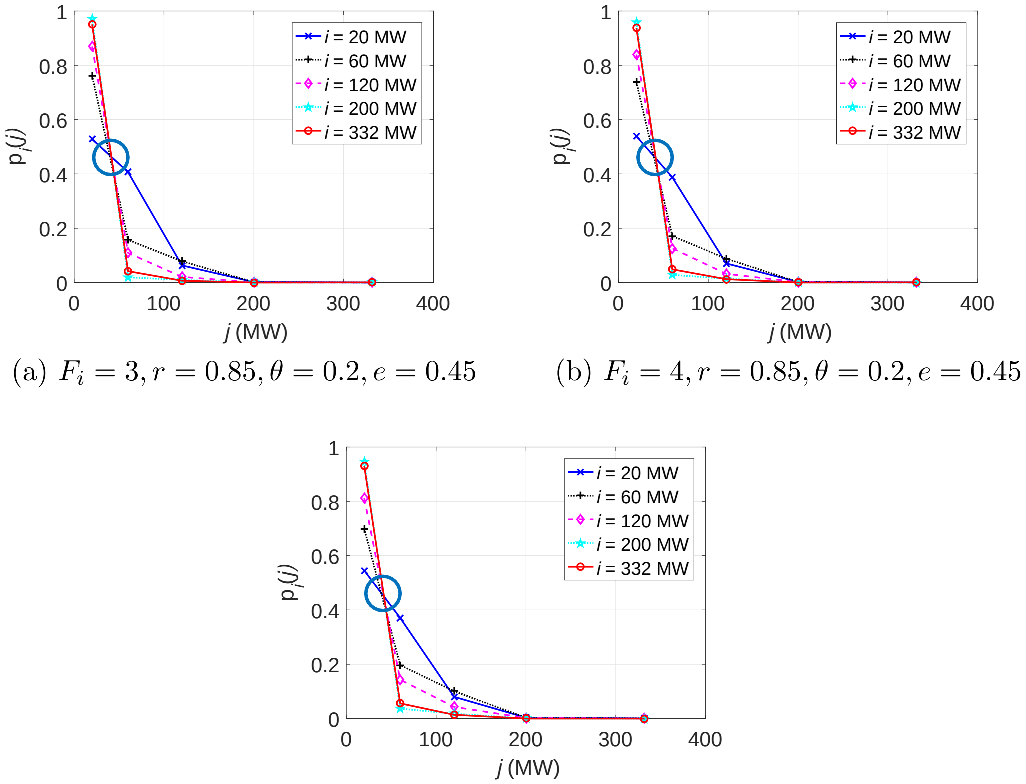
Table 6.1: Number of failures of a line of capacity j at next time step following failure of a line capacity i; NF:= No Failure.

j(MW) i(MW)	20	60	120	200	332	NF
20	24557	17663	3213	115	20	12776
60	23163	5382	2765	59	10	4623
120	13617	2046	523	24	2	1167
200	4234	125	64	1	0.25	8
332	1602	83	21	0.25	1	1
NF	0	0	0	0	0	∞

failure (due to "$N-1$ security," which ensures that if only one transmission line fails the power grid can compensate for that failure without further failures). Then we track the number of failures of transmission lines with different capacities at the following time steps. We run simulations and collect different types of data: 1) the number of failed lines of capacity j ($j \in C$) at time $k+1$ following the failure of a line of capacity i ($i \in C$) at time k ($k = 1,\cdots, K$, where K is the number of time steps); 2) the total number of transmission lines failed when the cascade stops; 3) the time step at which the cascade stops. We describe these data individually in the following three subsections. All these data are then incorporated into a model for the evolution of failures of transmission lines in a power grid, presented in Section IV.

6.1.1 The number of failed lines of capacity j at time step $k+1$ following the failure of a line of capacity i at time step k

Table I shows the number of failures of a line of capacity j at time $k+1$ following the failure of a line of capacity i at time k, for $F_i = 4$, $r = 0.85$, $\theta = 0.2$, $e = 0.45$.



(a) $F_i = 3, r = 0.85, \theta = 0.2, e = 0.45$

(b) $F_i = 4, r = 0.85, \theta = 0.2, e = 0.45$

(c) $F_i = 5, r = 0.85, \theta = 0.2, e = 0.45$

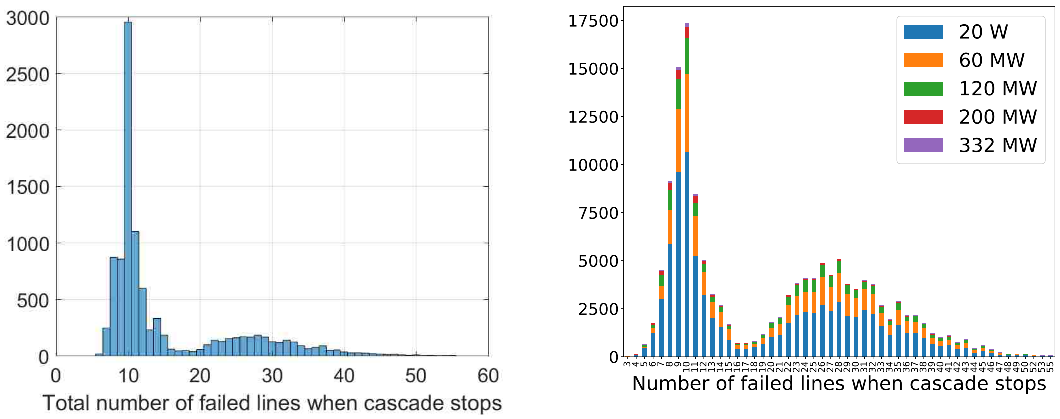
Figure 6.1: Probability of failures of a line of capacity j at time $k+1$ following failure of a line of capacity i at time k.

The simulation data are collected using 10,000 runs with random initial failures of transmission lines. For each run, we track the transmission line failures until the cascade stops. Once the power grid reaches the cascade stop state it remains in that state (i.e., no more failures), therefore in Table I, we use the symbol infinity (∞) to indicate the no failure (NF) state. The numbers in Table I are calculated as follows: if there is only one line failure with capacity $c \in C$ at time k then we use that line as the *source line* (rows in Table I) that causes the next line failures (columns in Table I) at time $k+1$. For example, let a line with capacity 60 MW fail at time k and three lines of capacities 20 MW, 60 MW and 120 MW fail at time $k+1$. Then looking at the second row (60 MW) we add a 1 to the first column (20 MW), a 1 to the second column (60 MW), and a 1 to the third column (120 MW). However, if there is more than one line failure at time k then we assume that the next line failures at time $k+1$ occur due to equal contributions from all of the previous line failures. For instance, if there are m line failures at time step k and l failures at time $k+1$, then each of the m lines (respective capacity rows in the Table I) contributes a fraction $\frac{1}{m}$ to the failing of each of the l lines (respective capacity columns in the Table I).

Figure 6.1 shows the probability of failure of a transmission line with capacity j at time $k+1$ due to the failure of a line with capacity i at time k, denoted by $p_i(j)$, for a given r, θ, e, and different values of F_i. In all the three panels, we see that the probability of failure for the low-capacity lines is larger than for the high-capacity lines. While the decreasing trend of failure probabilities with increasing line capacities follows for all lines, there is a transition occurs when the line capacity increases approximately above 41 MW (circled in Fig. 1). Namely, we see that for line capacities j lower (higher) than 41 MW the lower capacity lines are more vulnerable to failure than the higher (lower) capacity lines. We incorporate these observations in a model that is presented in the next section.

6.1.2 The total number of transmission lines failed when the cascade stops

Figure 6.2(a) is a histogram of the total number of transmission lines failed when the cascade stops and Fig. 6.2(b) shows a histogram of the total number of transmission lines of different capacities failed when the cascade stops. In addition, the ratios between the number of eventually failed transmission lines and the total number of transmission lines for different capacities are shown in Table II. Table II shows that



(a) Histogram of total number of failed lines

(b) with different capacities

Figure 6.2: Histogram of the total number of transmission lines failed when the cascade stops. Here $F_i = 4, r = 0.85, \theta = 0.2, e = 0.45$.

Table 6.2: Ratio between the number of eventually failed lines and total number of lines for different capacities.

Capacity (MW) Parameters	20	60	120	200	332
As in Fig. 6.1(a)	0.71	0.29	0.18	0.1	0.14
As in Fig. 6.1(b)	0.76	0.35	0.18	0.1	0.14
As in Fig. 6.1(c)	0.74	0.31	0.21	0.15	0.28

overall the higher capacity lines are less vulnerable to cascading failures, which can also be seen from Fig. 6.2(b). This observation, that lines of different capacities are more-or-less prone to fail provides the main motivation for this study. As an example, considering the worst-case failure for each transmission line (i.e., for any transmission line we take the largest number of failure from 10,000 runs), the total number of eventually failed lines of 20 MW, 60 MW, 120 MW, 200 MW, 332 MW, capacities are found to be 29, 19, 14, 5, 3, respectively (with the parameters as in Fig. 6.1(c)).

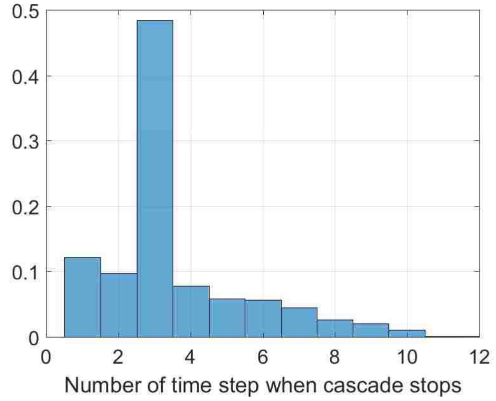


Figure 6.3: Histogram of the number of time step at which the cascade stops. Here $F_i = 4, r = 0.85, \theta = 0.2, e = 0.45$.

6.1.3 The time step at which the cascade stops

In Fig. 6.3 we show the normalized frequency (probability) for the time steps at which the cascade stops. From this data, we see that a large majority of the cascades stops at time step $k = 3$ and the average time at which the cascade stops is equal to 3.55.

6.2 Parametric model

In this section, we propose a parametric model to reproduce the dynamics of transmission line failures that we have seen from the optimal DC power-flow simulations. In this model, we incorporate the three types of data described in Section III A, B, and C.

From Fig. 6.1 we observe two main trends affecting the evolution of line failures. *Observation 1*: The low-capacity lines are more vulnerable to failures than the high-capacity lines, i.e., $p_i(j)$ is a decreasing function of the line capacity j. *Observation 2*: $p_i(j)$ is an increasing function of the capacity i for small j and a decreasing function

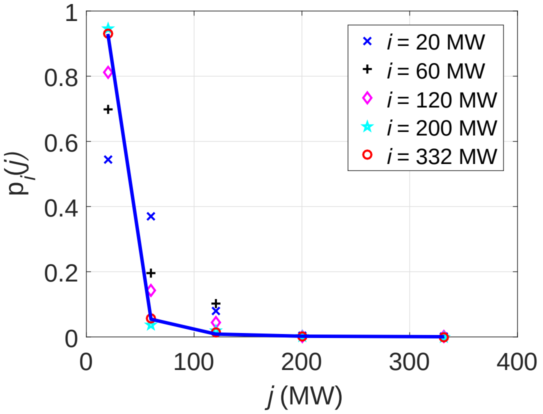


Figure 6.4: Power-law fitting (solid line) for the probability of failures of a line of capacity j at next time step following failure of line of capacity i. Here $F_i = 4, r = 0.85, \theta = 0.2, e = 0.45$.

of the capacity i for large j (due to the *transition* shown in Fig. 6.1).

Based on these observations, we propose a parametric equation to capture the probabilities of line failure evolution,

$$p_i(j) = \alpha j^{\beta}(1 + \gamma i(j - j_{th})), \tag{6.1}$$

where the term $\alpha j^{\beta}, \alpha > 0, \beta < 0$ reflects *Observation 1*. The term $1 + \gamma i(j - j_{th})$ reflects *Observation 2*; this latter term is larger (smaller) than 1 when j is smaller (larger) than j_{th}. The values of α, β, γ, and j_{th} are computed by fitting the data in Fig. 6.4 so as to minimize the overall mean square error. For our data we obtained the following values of the optimally fitted parameters: $\alpha = 2415, \beta = -2.6, \gamma = -9.95 \times 10^{-6}$, and $j_{th} = 41$, which were then used to generate the fitted solid line in Fig. 6.4. Note that the value $j_{th} = 41$ captures the transitions observed in Fig. 6.1.

An important assumption in (6.1) is that the line failure curve follows a power-law. This assumption is consistent with the long tail of the probability curve shown in Fig. 6.4.

6.2.1 Dynamical model for the failures of transmission lines

Our discrete-time model for the evolution of failures of transmission lines of different capacities $j \in C$ is the following,

$$X_j[k+1] = X_j[k] + F_j[k]\sigma\Big(\sum_{i\in C} p_i(j)\big(X_i[k] - X_i[k-1]\big)\Big), \tag{6.2}$$

where $X_j[k]$ is the number of failed lines of capacity j at time k, $F_j[k] = (F_j^* - X_j[k])$ is the number of functioning lines of capacity j at time k. We set the initial conditions in the model as follows: $X_j[k] = 0$ for $k < 0$, and we assume a certain number of initial failures occurring at time $k = 0$. Here F_j^* is the total number of transmission lines of capacity j that eventually fail when the cascade stops (from Section III B). Moreover, $p_i(j)$ is given by (6.1). Finally, $\sigma(y)$ is a sigmoidal function defined as

$$\sigma(y) := \begin{cases} 0, & \text{if } y \leq 0 \\ 1, & \text{if } y \geq \kappa \\ \frac{y}{\kappa}, & \text{otherwise,} \end{cases} \tag{6.3}$$

where κ is a constant to be determined from the simulation data. By definition, $\sigma \in [0, 1]$ and $\sigma(y)$ approaches 0 for small y and approaches 1 for large y. Note that, the product of σ and $F_j[k]$ gives the number of functional line of capacity j that will fail at time k.

Figure 6.5 shows the evolution of failure probabilities of lines with different capacities given by the σ evaluated at $k = 1, 2, ..., 14$. Here we have used $\kappa = 3.55$ in all simulations, since it is the average time found in section 6.1.3. Besides, setting $\kappa = 3.55$ enables the trend of total line failure generated by the model to better replicate the trend of real-world line failure data, as we will show later in Fig. 6.7.

From (6.2), we see that for any line with capacity $j \in C$, the number of failed lines at time step $k+1$ can be obtained by adding the new failed lines at time step k to the

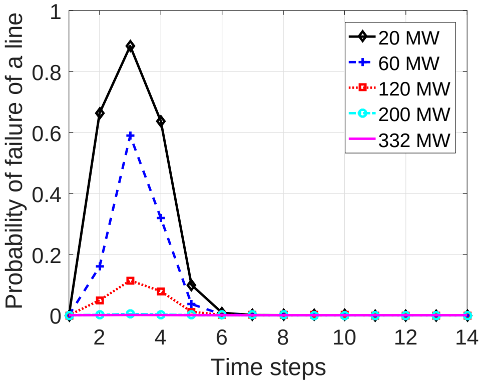


Figure 6.5: Time-evolution of failure probabilities of transmission lines with discrete time steps.

total number of failed lines that failed until time step k. The number of new failed lines of capacity j at time step k is given by $F_j[k]\sigma\Big(\sum_{i\in C} p_i(j)\big(X_i[k]-X_i[k-1]\big)\Big)$. A steady state is reached when there are no new line failures. Hence, the conditions for the steady state for each line of capacity j are the following: either $F_j[k]=0$ for all $j\in C$, indicating there are no more functional lines to fail or $\sum_{i\in C} p_i(j)\big(X_i[k]-X_i[k-1]\big)=0$ indicating the number of new line failures is zero. However, because $p_i(j)>0$, this can only happen for $X_j[k+1]=X_j[k]$ for all js, i.e, no new failures at the previous time step. Iterating this backward in time this is only possible if there are no new failures at the initial time 0.

We rewrite the (6.2) for all $j\in C$,

$$\begin{aligned} X_j[k+1] &= X_j[k] + F_j[k]\sigma\Big(\sum_{i\in C} p_i(j)\big(X_i[k]-Y_i[k]\big)\Big), \\ Y_j[k+1] &= X_j[k]. \end{aligned} \tag{6.4}$$

Theorem 6.2.1. *Assume some $X_j[0]>0, X_j[k]=0$ for $k<0$ and $p_i(j)>0$. The evolution of transmission line failure given by (6.2) asymptotically converges to the equilibrium point $X_j[k]=F_j^*$, $Y_j[k]=F_j^*$ for all $j\in C$.*

Proof. Let us write the state of the system at time k in column vector form as $\mathbf{X}[k] = (X_{20}[k], X_{60}[k], ..., X_{332}[k])$, $\mathbf{Y}[k] = (Y_{20}[k], Y_{60}[k], ..., Y_{332}[k])$, and $\mathbf{Z}[k] = [\mathbf{X}^T[k], \mathbf{Y}^T[k]]^T$, where T denotes the matrix transpose operator. Then (6.4) can be rewritten in vectorial form

$$\mathbf{Z}[k+1] = f(\mathbf{Z}[k]), \tag{6.5}$$

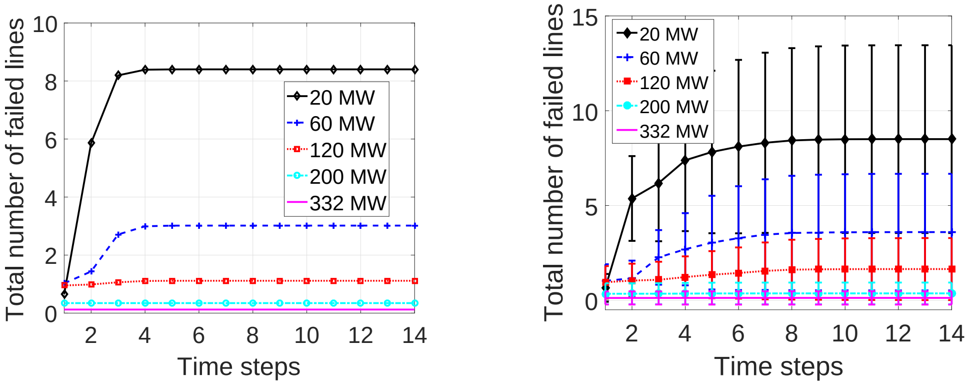
where f is a continuous function. We consider the equilibrium point $X_j = F_j^*, Y_j = F_j^*$ for all j which corresponds to $F_j = 0$ for all j, i.e., there are no more functional lines of any types to fail. We label this equilibrium point as $\bar{\mathbf{Z}} = [\bar{\mathbf{X}}^T, \bar{\mathbf{Y}}^T]^T$, where $\bar{\mathbf{X}} = (F_{20}^*, ..., F_{332}^*)$ and $\bar{\mathbf{Y}} = (F_{20}^*, ..., F_{332}^*)$. In order to show convergence to the equilibrium point $\bar{\mathbf{Z}}$, we need to show there exists a Lyapunov function $\mathbf{V}(\mathbf{Z})$ in a region of attraction $\mathcal{D} := \{\mathbf{Z} \in \mathbb{R}^{|C|\times|C|} : 0 \leq X_j \leq F_j^*, 0 \leq Y_j \leq F_j^* \text{ for all } j \in C\}$ (Theorem 1, Section 9.6 [96]). We define

$$\mathbf{V}(\mathbf{Z}) := \sum_{j\in C} (X_j[k] - F_j^*)^2 + (Y_j[k] - F_j^*)^2. \tag{6.6}$$

It is easy to check the Lyapunov function defined by (6.6) satisfies all three properties of a Lyapunov function for the discrete-time system given by (6.5), namely, it is a continuous function, it has a unique minimum at $\bar{\mathbf{Z}}$, and $\Delta\mathbf{V}(\mathbf{Z}) := \mathbf{V}(\mathbf{f}(\mathbf{Z})) - \mathbf{V}(\mathbf{Z}) < 0$ at any point in $\mathcal{D}$ except $\bar{\mathbf{Z}}$ [96]. To verify the latter condition, note that for $0 \leq X_j[k] < F_j^*$ and $0 \leq Y_j[k] < F_j^*$,

$$\begin{aligned}
\mathbf{V}(\mathbf{f}(\mathbf{Z})) = \mathbf{V}(\mathbf{Z}[k+1]) &= \sum_{j\in C} (X_j[k+1] - F_j^*)^2 + (Y_j[k+1] - F_j^*)^2 \\
&= \sum_{j\in C} \big(F_j^* - X_j[k]\big)^2 \big(1 - \sigma(\cdot)\big)^2 + \big(F_j^* - X_j[k-1]\big)^2 \big(1 - \sigma(\cdot)\big)^2, \text{ (by (6.4))} \\
&< \sum_{j\in C} (F_j^* - X_j[k])^2 + (F_j^* - X_j[k-1])^2 = \mathbf{V}(\mathbf{Z}),
\end{aligned}$$

where we have used $0 < \sigma \leq 1$, since $\sigma = 0$ only when the system is already in the equilibrium state. This proves the asymptotic stability of the equilibrium point $\bar{\mathbf{Z}}$; therefore, the cascade stops at $\bar{\mathbf{Z}}$, i.e., at $F_j[k] = 0$ for all $j \in C$. □



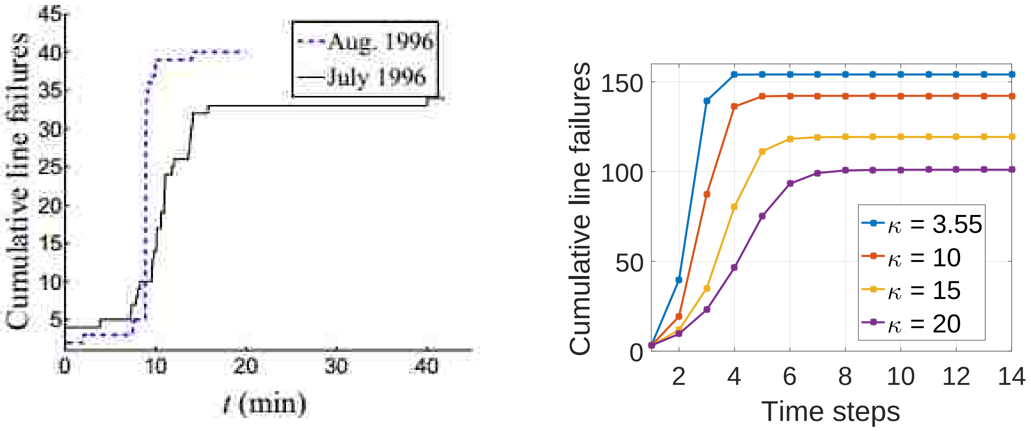
(a) (b)

Figure 6.6: Time-evolution of failures of transmission lines versus the discrete time steps: (a) line failures data using the model in (6.2), (b) optimal power-flow simulation of IEEE 118 bus system using MATPOWER, where the vertical bars are the standard deviation of the number of line failures.

Figure 6.6(a) shows the time-evolution of line failures of different capacities given by our dynamical model. Figure 6.6(b) shows the line failure data from optimal power-flow simulation using MATPOWER with initial conditions $F_i = 4, r = 0.85, \theta = 0.2, e = 0.45$. Here we average the line failures over 10,000 runs where the vertical bars in Fig. 6.6(b) show the standard deviation of the number of line failures. Comparing Fig. 6.6(b) with Fig. 6.6(a), we can see that the trends of average line failure generated by the model and MATPOWER simulation follow a similar pattern.

6.3 Comparison of model results with real-world data

In this section, we compare our model results of the total (cumulative) line failure with real-world data that are collected from two real-world cascading failure events



(a) (b)

Figure 6.7: Total number of failed lines obtained from the proposed model and real-world cascading failure data: (a) total line failures in the July 1996 black (solid line), August 1996 WSCC blackout (dashed lines) (b) total line failures data using the proposed model with different value of κ.

[97, 98]. It can be seen from the Fig. 6.7, the trend of total line failures from the model (Fig. 6.7(b)) is qualitatively similar to the trends observed from the real-world data (Fig. 6.7(a)) (here Fig. 6.7(a) is reproduced from [41].). In Fig. 6.7(b), the parameter κ of (6.3) controls the slope of the total line failure curve in the simulation. Note that the failure behavior in the initial phase of real data is slightly different than the initial phase of our model results. This is because the simplified discrete-time approximation of the model does not correspond exactly with the actual time of the real-world cascading failure event, namely, in our model a time step corresponds to the time when line failures occur whereas for a real-world cascading failure there are some time gaps between the increments of line failures. Note also that the total number of line failures (vertical axis) are different for both figures, which are due to different power-grid topologies used to compare model results with real-world data (as we do not have topologies of the power grids described in [97, 98]).

6.4 Summary and conclusions

In this chapter, we have proposed a data-driven parametric model to characterize the dynamics of the propagation of transmission line failures in the power grid. This model describes the evolution of cascading failure through transmission lines of different capacities. Due to unavailability of all the relevant quantities from the real-world cascading failure data, we have used the data generated by the optimal power-flow simulator (MATPOWER) to feed the parametric model. We have demonstrated that our model can reproduce the trends in the dynamics of line failure in the power grid considering capacities of the transmission lines and compared model results with real-world data. Our model outputs an estimate of the total capacity loss due to the failure of transmission lines during cascading failures at any time step.

Chapter 7

Future Work

This dissertation analyzes the vulnerability of a communication network from many aspects including the interception of inter-network connection, reliability of networks due to physical attacks, propagation of initial failures/attacks in the interconnected network, etc. In addition, the vulnerability of communication network due to interdependency with power grids is investigated and cascading failure in the power is simulated. In this chapter we describe some possible directions of future research that can be evolved from this dissertation.

For the inter-network connection optimization, as of now, we have considered all the interconnections are identical an assumption may not hold in the practical scenario. Thus we are currently working on the optimal interconnection considering the variations in link capacities (or QoS of the links). Moreover, generalizing the binary channel model to arbitrary size packet model and analyzing various natures of data manipulations by intruders after the interception (e.g., manipulations of a burst of the data packet), are also other possible extensions of this work. Analysis with finite queuing delay would also make this work more realistic.

While modeling intensity of physical attacks on the physical infrastructure, the dynamics of network functionality needs to be studied under temporal correlation among different types of stressors in the future. Moreover, robust techniques need to be devised to minimize the impact of catastrophic stressors on the network.

The parametric model for finding the resiliency and efficiency of a secure network is derived from simulation data, not by solving the optimization problem analytically. The analysis of the optimality of the proposed optimization problems and the trade-off between accuracy and complexity of the parametric model are left as future works. Besides, validating the parametric model with data from the real-world network is also a part of our future study. The generalization of this work for modeling the non-homogeneous propagation of security risks (i.e., τ and c may vary based on the types of security risks and links) would better capture the real-world threat propagation in the ML network.

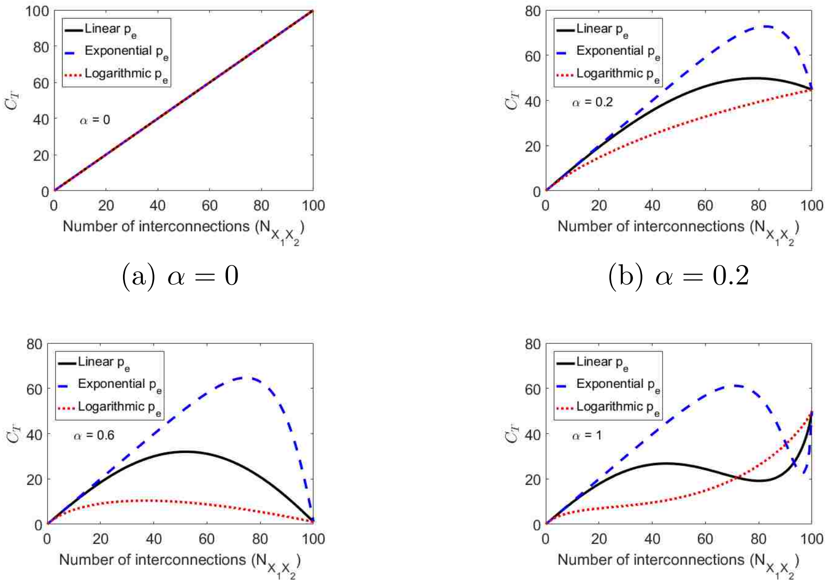
Further, as a future work, the theory developed in this dissertation, specifically the interdependent influence model, can be extended to include more states (e.g., alarm state) of communication nodes and heterogeneous influences among communication network components. Moreover, capturing influences from communication node to power nodes (e.g., smart grids) as well as the interactions among power nodes will also be part of future research to make this model more realistic.

The data-driven model presented in this dissertation provides a simple approach for reproducing the cascading failure behavior in power-grids from the simulation data. The major limitation of our data-driven parametric model is that the model needs to be calibrated when used on different network topologies. However, the model parameters can be estimated by the optimal power-flow simulation given the power grid topologies such as IEEE 300, Polish grid, etc. In addition, incorporation of generator dynamics into the problem would make this problem to capture more realistic cascading failures.

Appendices

Appendix A

C_T vs. $N_{X_1X_2}$ when p_e varies



(a) $\alpha = 0$

(b) $\alpha = 0.2$

(c) $\alpha = 0.6$

(d) $\alpha = 1$

Figure A.1: The total Shannon capacity (C_T) vs. the number of interconnections ($N_{X_1X_2}$) when p_e is a function of p_l.

Here we express p_e as a function of p_l. We define $p_e := \alpha p_l$, where α is a parameter between 0 and 1. Figure A.1 shows the total Shannon capacity (C_T) vs. the number of interconnections ($N_{X_1X_2}$) when p_e is a function of p_l. We see that when p_e varies with $N_{X_1X_2}$ we get a variation in C_T's, which will change the current analysis of finding optimal interconnection.

Appendix B

Proof of the concavity of the objective function of (2.12)

A function $f : D \in \mathbb{R} \to \mathbb{R}$ is said to be a convex function if it satisfies the following two conditions [62]: 1) the domain D of the function f is a convex set, and 2) for any two points $x, y \in D$, the function satisfies the following condition, $f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$,

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y),$$

where θ is in the range $0 \leq \theta \leq 1$. One way to check the concavity of a differentiable function is by differentiating it twice, i.e., f is concave if and only if $f^{''} \leq 0$ and its domain is convex.

Note that the interval $[0, N_{max}]$ a line segment on real line which is a convex set [62]. Now differentiating the objective function of (2.12) twice with respect to $N_{X_1 X_2}$ we get $\frac{2}{1-p_d^*}(-1 + C_l)$. Since $0 \leq C_l \leq 1$, $\frac{2}{1-p_d^*}(-1 + C_l) \leq 0$. Hence, the objective function of (2.12) is concave.

Appendix C

$N_{max}^2 - 4(1 - C_l)N_{max}R_s^* \geq 0$

Since $N_{max}\big(N_{max} - 4(1 - C_l)R_s^*\big) \geq 0$ and $N_{max} \geq 0$, we need to show $N_{max} - 4(1 - C_l)R_s^* \geq 0$, or $N_{max} \geq 4(1 - C_l)R_s^*$. If $R_s^* < C_T^*$, $4(1 - C_l)R_s^* < 4(1 - C_l)C_T^*$, and

$$
\begin{aligned}
4(1 - C_l)C_T^* &= 4(1 - C_l)N_{X_1X_2}^*\bar{C} \text{ (by (2.8) and for optimality } N_{X_1X_2} = N_{X_1X_2}^*\text{)}, \\
&= \begin{cases} 4(1 - C_l)\frac{N_{max}}{2(1-C_l)}\bar{C}, & \text{if } 0 \leq C_l \leq 0.5, \\ 4(1 - C_l)N_{max}\bar{C}, & \text{if } 0.5 < C_l \leq 1 \end{cases} \text{(by (2.13))} \\
&= \begin{cases} 2N_{max}\bar{C}, & \text{if } 0 \leq C_l \leq 0.5, \\ 4(1 - C_l)N_{max}\bar{C}, & \text{if } 0.5 < C_l \leq 1 \end{cases} \\
&= \begin{cases} 2N_{max}(1 - \frac{N_{X_1X_2}^*}{N_{max}}(1 - C_l)), & \text{if } 0 \leq C_l \leq 0.5, \\ 4(1 - C_l)N_{max}(1 - \frac{N_{X_1X_2}^*}{N_{max}}(1 - C_l)), & \text{if } 0.5 < C_l \leq 1 \end{cases} \text{(by (2.7) and linear } p_l\text{)}, \\
&= \begin{cases} 2N_{max}(1 - \frac{1}{2(1-C_l)}(1 - C_l)), & \text{if } 0 \leq C_l \leq 0.5, \\ 4(1 - C_l)N_{max}(1 - (1 - C_l)), & \text{if } 0.5 < C_l \leq 1 \end{cases} \text{(by (2.13))} \\
&= \begin{cases} N_{max}, & \text{if } 0 \leq C_l \leq 0.5, \\ 4(C_l - C_l^2)N_{max}, & \text{if } 0.5 < C_l \leq 1 \end{cases} = \begin{cases} N_{max}, & \text{if } 0 \leq C_l \leq 0.5, \\ \leq N_{max}, & \text{if } 0.5 < C_l \leq 1, (C_l = 0.5) \end{cases}.
\end{aligned}
$$

Hence, from both inequalities we see that $4(1 - C_l)R_s^* \leq N_{max}$.

References

[1] J. P. Sterbenz, D. Hutchison, E. K. Çetinkaya, A. Jabbar, J. P. Rohrer, M. Schöller, and P. Smith, “Resilience and survivability in communication networks: Strategies, principles, and survey of disciplines,” *Computer Networks*, vol. 54, no. 8, pp. 1245–1265, 2010.

[2] E. K. Cetinkaya, M. J. Alenazi, A. M. Peck, J. P. Rohrer, and J. P. Sterbenz, “Multilevel resilience analysis of transportation and communication networks,” *Telecommunication Systems*, vol. 60, no. 4, pp. 515–537, 2015.

[3] J. Lee, D. Hwang, J. Park, and K.-H. Kim, “Risk analysis and countermeasure for bit-flipping attack in LoRaWAN,” in *International Conference on Information Networking (ICOIN).* IEEE, 2017, pp. 549–551.

[4] A. Proano and L. Lazos, “Selective jamming attacks in wireless networks,” in *International Conference on Communications (ICC).* IEEE, 2010, pp. 1–6.

[5] C.-C. Wang, “On the capacity of 1-to-k broadcast packet erasure channels with channel output feedback,” *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 931–956, 2012.

[6] D. Olivera, N. Ghani, T. Lehman, X. Yang, M. M. Hayat, J. Crichigno, and E. Bou-Harb, “Software-defined networking (SDN) testbed for evaluation of large-scale electro-magnetic pulse (EMP) attacks, under review,” *EEE Communications Magazine*, 2017.

[7] G. O’Reilly, A. Jrad, R. Nagarajan, T. Brown, and S. Conrad, “Critical infrastructure analysis of telecom for natural disasters,” in *Telecommunications Network Strategy and Planning Symposium, 2006. NETWORKS 2006. 12th International.* IEEE, 2006, pp. 1–6.

[8] J. F. Kurose and K. W. Ross, *Computer networking: a top-down approach.* Addison-Wesley Reading, 2010.

[9] P. Crucitti, V. Latora, and M. Marchiori, “Model for cascading failures in complex networks,” *Physical Review E*, vol. 69, no. 4, p. 045104, 2004.

[10] V. Jacobson, “Congestion avoidance and control,” in *ACM SIGCOMM computer communication review*, vol. 18, no. 4. ACM, 1988, pp. 314–329.

[11] J. H. Cowie, A. T. Ogielski, B. Premore, E. A. Smith, and T. Underwood, “Impact of the 2003 blackouts on internet communications,” *Preliminary Report, Renesys Corporation (updated March 1, 2004)*, 2003.

[12] S. Bistarelli, S. N. Foley, and B. O’Sullivan, “Detecting and eliminating the cascade vulnerability problem from multilevel security networks using soft constraints,” in *AAAI*, 2004, pp. 808–813.

[13] T. H. Shake, “Security in military/commercial communication gateways,” in *Proceedings IEEE Military Communications Conference (MILCOM)*, vol. 1. IEEE, 1999, pp. 469–474.

[14] R. Di Pietro and G. Me, “Military secure communications over public cellular network infrastructure,” in *Proceedings IEEE Military Communications Conference (MILCOM)*, vol. 1. IEEE, 2002, pp. 400–405.

[15] A. Neri, D. Blasi, L. Gizzi, and P. Campisi, “Joint security and channel coding for OFDM communications,” in *16th European Signal Processing Conference*. IEEE, 2008, pp. 1–5.

[16] N. Živić and O. Rehman, “On using the message digest for error correction in wireless communication networks,” in *2010 IEEE 21st International Symposium on Personal, Indoor and Mobile Radio Communications Workshops*. IEEE, 2010, pp. 491–495.

[17] M. Bloch and J. Barros, *Physical-layer security: from information theory to security engineering*. Cambridge University Press, 2011.

[18] W. K. Harrison, J. Almeida, M. R. Bloch, S. W. McLaughlin, and J. Barros, “Coding for secrecy: An overview of error-control coding techniques for physical-layer security,” *IEEE Signal Processing Magazine*, vol. 30, no. 5, pp. 41–50, 2013.

[19] P. Das, R. A. Shuvro, M. Rahnamay-Naeini, N. Ghani, and M. M. Hayat, “Efficient interconnectivity among networks under security constraint,” in *Proceedings IEEE Military Communications Conference (MILCOM) (Accepted)*. IEEE, 2018.

[20] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, "Catastrophic cascade of failures in interdependent networks," *Nature*, vol. 464, no. 7291, pp. 1025–1028, 2010.

[21] O. Yagan, D. Qian, J. Zhang, and D. Cochran, "Optimal allocation of interconnecting links in cyber-physical systems: Interdependence, cascading failures, and robustness," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 9, pp. 1708–1720, 2012.

[22] S. Chattopadhyay and H. Dai, "Towards optimal link patterns for robustness of interdependent networks against cascading failures," in *2015 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2015, pp. 1–6.

[23] ——, "Designing optimal interlink structures for interdependent networks under budget constraints," in *2017 IEEE International Conference on Communications (ICC)*. IEEE, 2017, pp. 1–6.

[24] S. Neumayer and E. Modiano, "Network reliability with geographically correlated failures," in *Proceedings INFOCOM*. IEEE, 2010, pp. 1–9.

[25] P. K. Agarwal, A. Efrat, S. K. Ganjugunte, D. Hay, S. Sankararaman, and G. Zussman, "Network vulnerability to single, multiple, and probabilistic physical attacks," in *Proceedings MILCOM*. IEEE, 2010, pp. 1824–1829.

[26] A. Narula-Tam, E. Modiano, and A. Brzezinski, "Physical topology design for survivable routing of logical rings in wdm-based networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 8, pp. 1525–1538, 2004.

[27] E. Modiano and A. Narula-Tam, "Survivable lightpath routing: a new approach to the design of wdm-based networks," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 4, pp. 800–809, 2002.

[28] S. Neumayer, G. Zussman, R. Cohen, and E. Modiano, "Assessing the vulnerability of the fiber infrastructure to disasters," *IEEE/ACM Transactions on Networking*, vol. 19, no. 6, pp. 1610–1623, 2011.

[29] M. Rahnamay-Naeini, J. E. Pezoa, G. Azar, N. Ghani, and M. M. Hayat, "Modeling stochastic correlated failures and their effects on network reliability," in *Proceedings of 20th International Conference on Computer Communications and Networks (ICCCN)*. IEEE, 2011, pp. 1–6.

[30] D. J. Strauss, "A model for clustering," *Biometrika*, vol. 62, no. 2, pp. 467–475, 1975.

[31] D. Papadimitriou, F. Poppe, J. Jones, S. Venkatachalam, S. Dharanikota, R. Jain, R. Hartani, D. Griffith, and Y. Xue, "Inference of shared risk link groups," *IETF Draft, OIF Contribution, OIF*, vol. 66, p. 2001, 2001.

[32] H.-W. Lee, E. Modiano, and K. Lee, "Diverse routing in networks with probabilistic failures," *IEEE/ACM Transactions on networking*, vol. 18, no. 6, pp. 1895–1907, 2010.

[33] Z. Kong and E. M. Yeh, "Resilience to degree-dependent and cascading node failures in random geometric networks," *IEEE Transactions on Information Theory*, vol. 56, no. 11, pp. 5533–5546, 2010.

[34] D. Magoni, "Tearing down the internet," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 6, pp. 949–960, 2003.

[35] M. S. Vassiliou *et al.*, "Crucial differences between commercial and military communications technology needs: Why the military still needs its own research," in *IEEE Military Communications Conference*. IEEE, 2013, pp. 342–347.

[36] C. A. Nissen, "Opportunities and challenges in realizing a global (mobile) military information infrastructure," MITRE CORP BEDFORD MA, Tech. Rep., 2006.

[37] B. White, "Layered communications architecture for the global grid," in *Proceedings IEEE Military Communications Conference (MILCOM)*, vol. 1. IEEE, 2001, pp. 506–511.

[38] A. Gutfraind, "Optimizing network topology for cascade resilience," in *Handbook of optimization in complex net.* Springer, 2012, pp. 37–59.

[39] M. Rahnamay-Naeini, "Designing cascade-resilient interdependent networks by optimum allocation of interdependencies," in *International conference on computing, networking and communications (ICNC).* IEEE, 2016, pp. 1–7.

[40] D.-H. Shin, D. Qian, and J. Zhang, "Cascading effects in interdependent networks," *IEEE Network*, vol. 28, no. 4, pp. 82–87, 2014.

[41] M. Rahnamay-Naeini, Z. Wang, N. Ghani, A. Mammoli, and M. M. Hayat, "Stochastic analysis of cascading-failure dynamics in power grids," *IEEE Transactions on Power Systems*, vol. 29, no. 4, pp. 1767–1779, 2014.

[42] L. Zhu, X. Liu, L. Yu, and X. Wu, "Model of cascading failures for communication networks," *International Journal of Computer and Communication Engineering*, vol. 5, no. 5, p. 302, 2016.

[43] S. M. Rinaldi, J. P. Peerenboom, and T. K. Kelly, "Identifying, understanding, and analyzing critical infrastructure interdependencies," *IEEE Control Systems*, vol. 21, no. 6, pp. 11–25, 2001.

[44] M. Amin, "Toward secure and resilient interdependent infrastructures," *Journal of Infrastructure Systems*, vol. 8, no. 3, pp. 67–75, 2002.

[45] V. Rosato, L. Issacharoff, F. Tiriticco, S. Meloni, S. Porcellinis, and R. Setola, "Modelling interdependent infrastructures using interacting dynamical models," *International Journal of Critical Infrastructures*, vol. 4, no. 1-2, pp. 63–79, 2008.

[46] M. Parandehgheibi and E. Modiano, "Robustness of interdependent networks: The case of communication networks and the power grid," in *2013 IEEE Global Communications Conference (GLOBECOM).* IEEE, 2013, pp. 2164–2169.

[47] H. Ma, H. Li, and J. B. Song, "Influence models of cascading failure and frequency oscillation in the power grid," in *Innovative Smart Grid Technologies (ISGT), 2013 IEEE PES.* IEEE, 2013, pp. 1–6.

[48] H. Guo, C. Zheng, H. H.-C. Iu, and T. Fernando, "A critical review of cascading failure analysis and modeling of power system," *Renewable and Sustainable Energy Reviews*, vol. 80, pp. 9–22, 2017.

[49] I. Dobson, B. A. Carreras, and D. E. Newman, "Branching process models for the exponentially increasing portions of cascading failure blackouts," in *38th Annual Hawaii International Conference on System Sciences.* IEEE, 2005, pp. 64a–64a.

[50] I. Dobson, "Estimating the propagation and extent of cascading line outages from utility data with a branching process," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 2146–2155, 2012.

[51] S. Soltan, D. Mazauric, and G. Zussman, "Analysis of failures in power grids," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 2, pp. 288–300, 2017.

[52] B. Buzan and G. Lawson, *The global transformation: history, modernity and the making of international relations.* Cambridge University Press, 2015, vol. 135.

[53] "JIE: How DOD is building a bigger network that's also a smaller target," Author: G. Slabodkin, retrieved from https://defensesystems.com/Articles/2015/02/23/Joint-Information-Environment-JRSS-security.aspx?Page=1, Date of Publication: 2015-02-23, Accessed: 2018-07-17.

[54] D. Vasudevan, V. G. Subramanian, and D. J. Leith, "On ARQ for packet erasure channels with bernoulli arrivals," in *IEEE International Symposium on Information Theory Proceedings (ISIT).* IEEE, 2010, pp. 1793–1797.

[55] M. C. Davey and D. J. MacKay, "Reliable communication over channels with insertions, deletions, and substitutions," *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 687–698, 2001.

[56] R. Yazdani and M. Ardakani, "Reliable communication over non-binary insertion/deletion channels," *IEEE Transactions on Communications*, vol. 60, no. 12, pp. 3597–3608, 2012.

[57] T. M. Cover and J. A. Thomas, *Elements of information theory.* John Wiley & Sons, 2012.

[58] C. E. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE mobile computing and communications review*, vol. 5, no. 1, pp. 3–55, 2001.

[59] S. Gollakota and D. Katabi, "Physical layer wireless security made fast and channel independent," in *2011 Proceedings IEEE INFOCOM.* IEEE, 2011, pp. 1125–1133.

[60] A. El Gamal and Y.-H. Kim, *Network information theory.* Cambridge university press, 2011.

[61] G. Optimization, "Inc., Gurobi optimizer reference manual, 2015," *URL: http://www. gurobi. com*, 2014.

[62] S. Boyd and L. Vandenberghe, *Convex optimization.* Cambridge university press, 2004.

[63] E. K. Çetinkaya, D. Broyles, A. Dandekar, S. Srinivasan, and J. P. Sterbenz, "Modelling communication network challenges for future internet resilience, survivability, and disruption tolerance: A simulation-based approach," *Telecommunication Systems*, vol. 52, no. 2, pp. 751–766, 2013.

[64] H. Haddadi, M. Rio, G. Iannaccone, A. Moore, and R. Mortier, "Network topologies: inference, modeling, and generation," *IEEE Communications Surveys & Tutorials*, vol. 10, no. 2, pp. 48–69, 2008.

[65] C. Wilson, "High altitude electromagnetic pulse (hemp) and high power microwave (hpm) devices: Threat assessments," DTIC Document, Tech. Rep., 2008.

[66] J. Borland, "Analyzing the internet collapse," *ABC News*, 2008.

[67] S. Neumayer and E. Modiano, "Network reliability under random circular cuts," in *IEEE Global Telecommunications Conference*. IEEE, 2011, pp. 1–6.

[68] "Tornado: National oceanic and atmospheric administration: Storm prediction center and wikipedia," http://www.spc.noaa.gov/, accessed: 2016-06-12.

[69] J. P. Sterbenz, J. P. Rohrer, E. K. Cetinkaya, M. JF, and Peck, "Ku-topview network topology tool," http://www.ittc.ku.edu/resilinets/maps/, The University of Kansas, 2010.

[70] V. Paruchuri, A. Durresi, and S. Chellappan, "Secure communications over hybrid military networks," in *IEEE Military Communications Conference*. IEEE, 2008, pp. 1–7.

[71] C. Asavathiratham, "The influence model: A tractable representation for the dynamics of networked markov chains," Ph.D. dissertation, MIT, 2000.

[72] M. Newman, *Networks: an introduction*. Oxford university press, 2010.

[73] V. Latora and M. Marchiori, "Efficient behavior of small-world networks," *Physical review letters*, vol. 87, no. 19, p. 198701, 2001.

[74] K.-I. Goh *et al.*, "Universal behavior of load distribution in scale-free networks," *Physical Review Letters*, vol. 87, no. 27, p. 278701, 2001.

[75] D. A. Pados and P. Papantoni-Kazakos, "A note on the estimation of the generalization error and the prevention of overfitting," in *1994 Int. conf. on neural networks*, vol. 1. IEEE, 1994, pp. 321–326.

[76] G. H. Baker, "Emp knots untied: Some common misconceptions about nuclear emp," *2012 Dupont Summit*, 2012.

[77] A. Townsend, *Telecommunications Infrastructure in Disasters: Preparing Cities for Crisis Communications*. University of New York, 2005. [Online]. Available: https://books.google.com/books?id=vDGHrgEACAAJ

[78] S. Erjongmanee, C. Ji, J. Stokely, and N. Hightower, "Large-scale inference of network-service disruption upon natural disasters," in *Knowledge Discovery from Sensor Data*. Springer, 2010, pp. 134–153.

[79] J. Gilbert, J. Kappenman, W. Radasky, and E. Savage, "The late-time (e3) high-altitude electromagnetic pulse (hemp) and its impact on the us power grid," *Report Meta*, 2010.

[80] E. Savage, J. Gilbert, and W. Radasky, "The early-time (e1) high-altitude electromagnetic pulse (hemp) and its impact on the us power grid," *Report Meta-R-320 for Oak Ridge National Laboratory*, 2010.

[81] A. E. Motter and Y.-C. Lai, "Cascade-based attacks on complex networks," *Physical Review E*, vol. 66, no. 6, p. 065102, 2002.

[82] "Delta airlines system shut down how to avoid a catastrophic failure," https://www.power-solutions.com/pressroom/newsletter/2016/08/delta-airlines-system-shut-avoid-catastrophic-failure, accessed: 2017-06-15.

[83] C. Asavathiratham, S. Roy, B. Lesieutre, and G. Verghese, "The influence model," *IEEE Control Systems*, vol. 21, no. 6, pp. 52–64, 2001.

[84] P. Das, M. Rahnamay-Naeini, N. Ghani, and M. M. Hayat, "On the vulnerability of multi-level communication network under catastrophic events," in *International Conference on Computing, Networking and Communications (ICNC)*. IEEE, 2017, pp. 912–916.

[85] J. Eto, "Blackout 2003: final report on the august 14, 2003 blackout in the united states and canada: causes and recommendations," *Electricity Markets and Policy Group, Energy Analysis and Environmental Impacts Department, US Department of Energy, Washington, DC*, 2004.

[86] R. A. Shuvro, Z. Wangt, P. Das, M. R. Naeini, and M. M. Hayat, "Modeling cascading-failures in power grids including communication and human operator impacts," in *Green Energy and Smart Systems Conference (IGESSC), 2017 IEEE*. IEEE, 2017, pp. 1–6.

[87] Z. Wang, M. R. Naeini, J. Abreu, R. A. Shuvro, P. Das, A. Mammoli, N. Ghani, and M. M. Hayat, "Impacts of operators' behavior on reliability of power grids during cascading failures," *IEEE Trans. on Power Systems*, 2018.

[88] S. Pahwa, C. Scoglio, and A. Scala, "Abruptness of cascade failures in power grids," *Scientific reports*, vol. 4, p. 3694, 2014.

[89] A. Bernstein, D. Bienstock, D. Hay, M. Uzunoglu, and G. Zussman, "Power grid vulnerability to geographically correlated failures-analysis and control implications," in *in Proceedings of the IEEE INFOCOM*. IEEE, 2014, pp. 2634–2642.

[90] A. Bakshi, A. Velayutham, S. Srivastava, K. Agrawal, R. Nayak, S. Soonee, and B. Singh, "Report of the enquiry committee on grid disturbance in northern region on 30th july 2012 and in northern, eastern & north-eastern region on 31st july 2012," *New Delhi, India*, 2012.

[91] N. FERC, "Arizona-southern california outages on september 8, 2011," *Causes and Recommendations*, 2012.

[92] M. Rahnamay-Naeini, Z. Wang, A. Mammoli, and M. M. Hayat, "A probabilistic model for the dynamics of cascading failures and blackouts in power grids," in *IEEE Power and Energy Society General Meeting.* IEEE, 2012, pp. 1–8.

[93] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on power systems*, vol. 26, no. 1, pp. 12–19, 2011.

[94] M. J. Eppstein and P. D. Hines, "A random chemistry algorithm for identifying collections of multiple contingencies that initiate cascading failure," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1698–1705, 2012.

[95] Q. Chen and L. Mili, "Composite power system vulnerability evaluation to cascading failures using importance sampling and antithetic variates," *IEEE Trans. on Power Systems*, vol. 28, no. 3, pp. 2321–2330, 2013.

[96] D. G. Luenberger, *Introduction to dynamic systems: theory, models, and applications.* Wiley New York, 1979, vol. 1.

[97] N. A. E. R. Council(NERC), "1996 system disturbances," *(Available from NERC,116-390 Village Boulevard,Princeton,NJ 08540-5731, USA)*, 2002.

[98] S. Abraham and R. Efford, "Final report on the august 14th blackout in the united states and canada," *US-Canada Power System Outage Task Force*, 2004.