February 2020

Reliability Analysis of Power Grids and its Interdependent Infrastructures: An Interaction Graph-based Approach

Upama Nakarmi
University of South Florida

Reliability Analysis of Power Grids and its Interdependent Infrastructures: An Interaction Graph-based Approach

by

Upama Nakarmi

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Electrical Engineering
College of Engineering
University of South Florida

Major Professor: Mahshid Rahnamay Naeini, Ph.D.
Ismail Uysal, Ph.D.
Yasin Yilmaz, Ph.D.
Kaiqi Xiong, Ph.D.
Xinming Ou, Ph.D.

Date of Approval:
March 26, 2020

Keywords: Cascading Failures, Community Structures, Critical Components, Cascade Size, Electric Vehicle Charging Infrastructure

Dedication

This dissertation is dedicated to my extremely supportive father, Rajendra Dhoj, my loving mother, Bidhya, my adorable sisters, Garima and Mona, and my encouraging partner Sulav. I love you all.

Acknowledgments

I wish to express my gratitude and sincere appreciation to several individuals, who have helped and supported me throughout my graduate studies. I would first like to express my deepest gratitude to my mentor and major advisor, Dr. Mahshid Rahnamay Naeini, who has for the past five years provided me with invaluable guidance, encouragement, patience, support, and help in my research and academic work. In retrospect, her constant guidance and feedback throughout my PhD as well as her advice on developing professionally has irrevocably shaped my scientific career.

I would also like to extend my deepest gratitude to Dr. Ismail Uysal, Dr. Kaiqi Xiong, Dr. Xinming Ou, and Dr. Yasin Yilmaz for the considerable time and effort dedicated towards serving on my committee. My sincere gratitude also goes to Dr. Hana Khamfroush for her consultation in one of the publications. My gratitude also extends to my friends and fellow lab-mates, Jakir, Hasnat, Sonali Sen, Richard, and Sraddhanjali for the stimulating discussions, suggestions, and the experiences shared via various assignments and projects. I am also extremely grateful to the staff and faculty at USF, for all the help in the academic activities.

Finally, I am grateful to have been funded over the course of my PhD by Dr. Naeini's project supported by the National Science Foundation under Grant No. 1761471. This gratitude also extends to Dr. Naeini's project supported by the Defense Threat Reduction Agency's Basic Research Program under Grant No. HDTRA1-13-1-0020.

Table of Contents

List of Tables

List of Figures

Abstract

Large blackouts with significant societal and economic impacts result from cascade of failures in the transmission network of power grids. Understanding and mitigating cascading failures in power grids is challenging due to the large number of components and their complex interactions, wherein, in addition to the physical topology of the system, the physics of power flow and functional dependencies among components largely affect the spatial distribution and propagation of failures. In this dissertation, data-driven interaction graphs, which help in capturing the underlying interactions and influences among the components during cascading failures, are used for capturing the non-local nature of propagation of failures as well as for simplifying the modeling and analysis of cascades. Particularly, influence and correlation graphs are constructed for revealing and comparing various types of interactions/influences during cascades.

In addition, as a step towards analyzing cascades, community structures in the interaction graphs, which bear critical information about cascade processes and the role of system components during cascades are identified. The key idea behind using community structures for analyzing cascades is that a cascade entering a community is likely to reach to most of the other members of the same community while less likely to reach to other communities. Thus, community structures significantly impact cascade behavior by trapping failures within communities. Further, a centrality measure based on the community structures is proposed to identify critical components of the system, which their protection can help in containing failures within a community and prevent the propagation of failures to large sections of the power grid. Various criticality evaluation techniques,

including data-driven, epidemic simulation-based, power system simulation-based and graph-based, have been used to verify the importance of the identified critical components in the cascade process and compare them with those identified by traditional centrality measures. Moreover, it has been shown that the loading level of the power grid impacts the interaction graph and consequently, the community structure and criticality of the components in the cascade process.

Furthermore, a Markov chain model is designed based on the community structures embedded in the data-driven interaction graphs of power grids. This model exploits the properties of community structures in interactions to enable the probabilistic analysis of cascade sizes in power grids. The trapping property of communities is extensively used to show that the probability distribution of cascade sizes exhibit power-law behavior as observed in previous studies and historical data.

Finally, an integrated framework based on the influence model, a networked Markov chain framework, is proposed for modeling the integrated power grid and transportation infrastructures, through one source of their interdependency i.e., electric-vehicle (EV) charging stations. The interactions based on the rules and policies governing their internal and interaction dynamics is captured. Particularly, the proposed integrated framework is used to design an algorithm for assigning dynamic charging prices for the EV charging infrastructure with the goal of increasing the likelihood of having balanced charging and electric infrastructures. The proposed scheme for charging prices is traffic and power aware as the states and interactions of transportation and power infrastructures are captured in the integrated framework. Finally, the critical role of cyber infrastructure in enabling such collaborative solutions is also discussed.

Chapter 1: Introduction

By the year 2050, 68% of the world's population is expected to reside in urban areas, with North America currently leading the race in terms of the most urbanized geographic location till date [2]. In these increasingly urban societies, goods and services are expected to be conveniently available at all times with high reliability. Major systems required for the smooth operation of an urban society are the *critical infrastructures*, which include power, energy (gas and oil), communication, transportation, emergency services, water, and food supply. Critical infrastructures are highly interdependent and collaborative systems due to the services and influences that they receive from one another. The complex mesh of interdependencies within standalone infrastructures as well as interdependencies shared among these infrastructures could be both problematic, by introducing vulnerabilities and compromising reliability, and beneficial, by introducing opportunities for increasing the effective operation of these systems.

The work in this dissertation aims to study and understand the reliability challenges of one of the most critical infrastructure, i.e., electric power grids. This dissertation also analyzes the reliability of the power grid infrastructure in conjunction with another critical infrastructure, i.e., the transportation infrastructure, through one form of their interdependency: the emerging new technology of electric vehicle (EV) charging infrastructure.

Portions of this chapter were published in IEEE PES [3], IEEE TNSE [4], Smartgreens [5] and Springer [6]. Copyright permissions from the publishers are included in Appendix A.

Power grids are large-scaled and complex interconnected networks that supply electricity from generators to consumers via a channel of high voltage transmission lines. In power grid infrastructures, the complex nature of influences and interactions between various components contribute in the reliability of the overall infrastructure. For instance, failure of few components may influence failure of other components and lead to large-scale blackouts. In this dissertation, we investigate the leading cause of large-scale blackouts, i.e. cascading failures [7]. We also analyze the risk posed by the EV charging infrastructure to the power grid infrastructure in terms of the increased levels of stress in power grids.

While this dissertation focuses on two interdependent infrastructures i.e., power grid and transportation infrastructures, the modeling and reliability analysis methods discussed in this dissertation is applicable to other interdependent infrastructures as well. The upcoming sections in this chapter will introduce these infrastructures as well as the specific reliability challenges investigated throughout this dissertation in detail.

1.1 Power Grid Infrastructure

An electric/power grid is an interconnected network that supplies electricity from producers to consumers. It consists of producers such as generating stations that generate electricity, high voltage transmission lines that carry power over long distances, distribution lines that supply electricity to individual consumers, and substations for stepping the voltage up or down during transmission, as shown in Figure 1.1. In this dissertation, the network of high voltage transmission lines is termed as the transmission network. In the transmission network, nodes represent generating stations $\boldsymbol{\mathcal{G}}$, load buses $\boldsymbol{\mathcal{L}}$, combinations of load and generating buses, and transmission buses

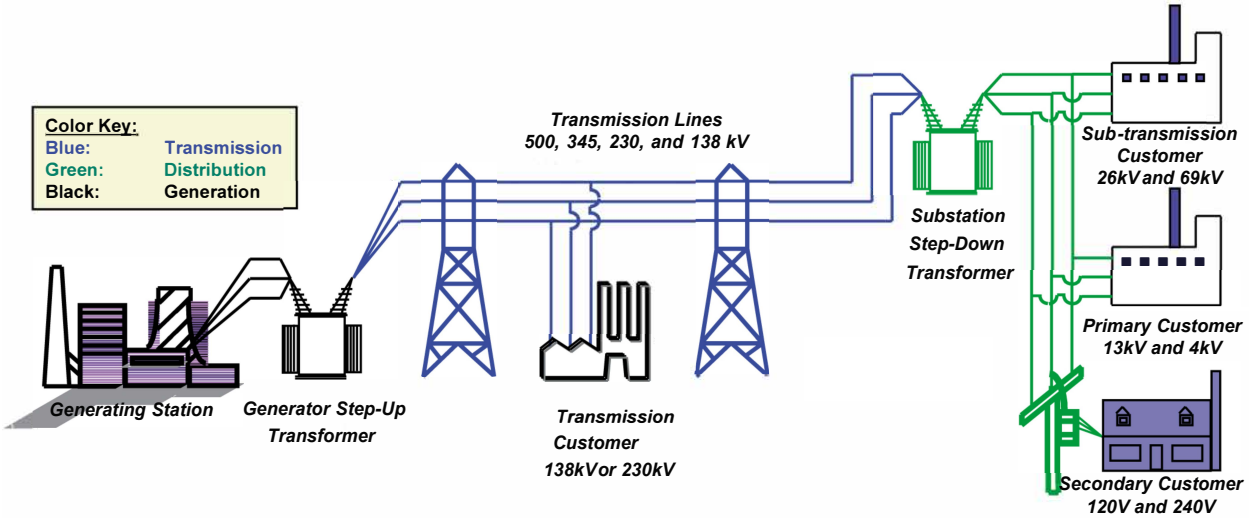


Figure 1.1: Structure of power grids. (Figure is borrowed from [1])

whose sole purpose is to transfer power. As mentioned, our reliability study is focused on cascading failures and in general, cascading failures can be attributed to transmission networks of power grids. The general mechanism of cascades is due to the redistribution of power flows caused by few initial failures, which overloads and consequently, causes dependent failures of other components in the system, such as transmission lines. In this dissertation, we focus on the transmission network of power grids. An example of transmission network, that is used in this dissertation for studies, is the IEEE 118 bus test system (Figure 1.2) with 118 buses, including substations and generators, and 186 transmission lines.

1.1.1 Power Grid's Interdependent Infrastructures

Power grids are critical infrastructures with interdependency with other critical infrastructures such as energy, communication and transportation systems. In this section, we discuss the co-dependency between the power grid infrastructure and the transportation infrastructure. Particularly, the increase in the number of hybrid electric transportation systems, including plugin

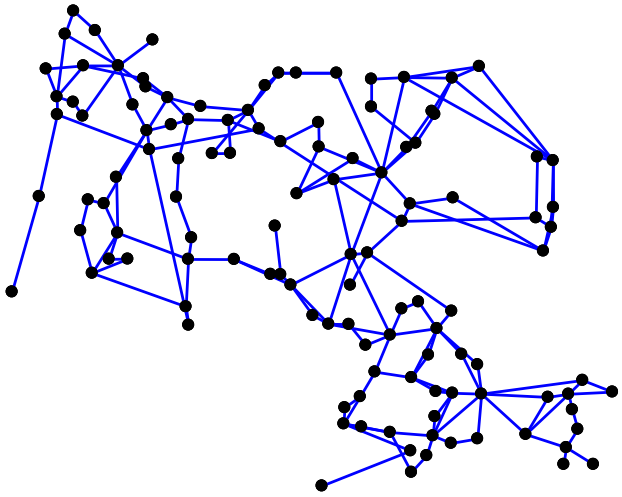


Figure 1.2: Topology of the IEEE 118-bus system. Nodes (black dots) represent generating stations, load buses, transmission buses, and combinations of generating stations and load buses; and edges (blue lines) represent transmission lines.

hybrid EV's and hybrid electric trains have introduced new interdependencies between the power and transportation infrastructures [8], [9], [10], [11], [12], [13], [14]. For instance, vehicle-to-grid (V2G) technology allow EV's to discharge their energy to the power grid using bi-directional power electronic DC/AC interfaces, which can help in stabilizing the power grid during disturbance and power shortage [15], [16].

Another source of interdependency between power and transportation infrastructures comes from the EV charging infrastructure. The EV charging infrastructures are slowly emerging in cities [17], similar to the traditional gas stations. On one hand, in the charging infrastructure, traffic patterns and population distribution can affect the power demand in the electric grid at various times and locations. On the other hand, the demand on the power grid can affect the charging price and consequently, affect the traffic pattern in the transportation system.

Such interdependencies are important as, for instance, during the peak-energy-consumption hours, inappropriate energy pricing signals at charging stations that motivate EV users to use

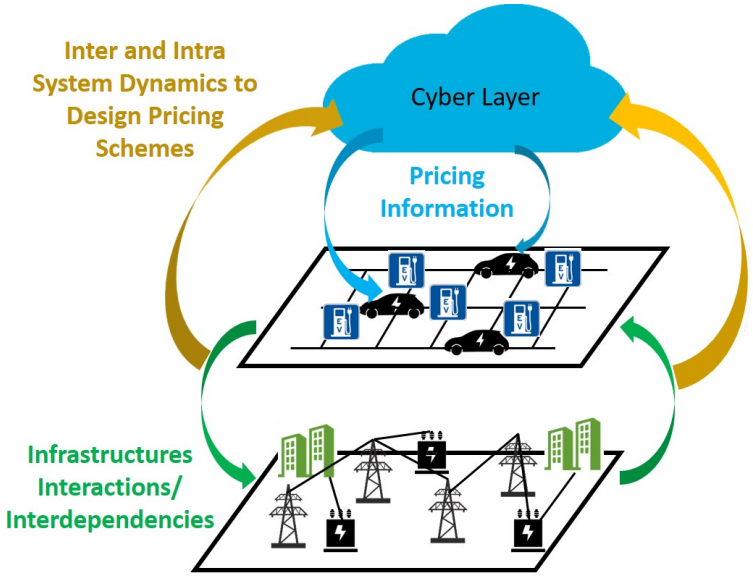


Figure 1.3: Interdependent networks of power (bottom) and transportation (top) infrastructures via EV charging stations and the role of cyber infrastructure.

specific charging stations, along with other factors, can lead to power demand profiles that result in instability of the electric grid and in worse cases power outages [18]. As such, it is essential to design and operate these charging infrastructures while considering the interdependency between power and transportation systems and the state of these systems. For instance, designing pricing incentives can provide a controlling mechanism for interdependency and reliable operation of these systems. The incentives will be communicated to the users through the cyber infrastructure. In general, the cyber infrastructure plays a key role in enabling such collaboration and cooperation among infrastructures, while also explicitly benefiting from the reliable power system as the source of electricity. Figure 1.3 shows the structure of intra-system and inter-system interdependencies within and among these infrastructures respectively and the pricing mechanism relayed by the cyber infrastructure.

1.1.2 Reliability Challenges in Power Grids

Current power grids face multiple challenges including growing electricity demand, aging of the incumbent power grid, introduction of intermittent power sources such as renewable energy to new sources of load such as electric vehicles, and distributed storage technologies. One such challenge faced by power grids are large scale blackouts, which is a rare probability event; however, occurrence of such events can lead to consequences of catastrophic scales. For example, a wide-area blackout occurred in the Pacific Southwest on September 8, 2011, which lasted for around twelve hours and affected around 5 million people [19]. The blackout that piqued the interest of power system researchers to investigate blackouts and their causes was the US Northeast blackout of August 14, 2003, which lasted for more than two days and affected around 50 million people [7].

Large blackouts with significant societal and economic impacts result from cascade of failures in the transmission network of power grids [7]. Understanding and mitigating cascading failures in power grids is challenging due to the large number of components and their complex interactions, which contribute in the cascade process. Intensive research efforts have been focused on understanding the underlying interactions in cascades, which for example, enable predicting the propagation path of failures and identifying critical/vulnerable components in the power grid. While large blackouts are infrequent, the power law behavior exhibited by the blackout size distribution (e.g. measured in terms of unserved energy, numbers of customers with no service, number of transmission lines tripped) warrants the need to study such events [20].

The North American Electric Reliability Corporation (NERC), which oversees the reliability of the bulk transmission system in North America stated "Cascading outages are defined as the uncontrolled loss of any system facilities or load, whether because of thermal overload, voltage

collapse, or loss of synchronism, except those occurring as a result of fault isolation" [21]. However, the IEEE Task Force on Understanding, Prediction, Mitigation, and Restoration of Cascading Failures [22] inform that all cascading outages do not eventually lead to blackouts and thus, state "A cascading outage is a sequence of events in which an initial disturbance, or set of disturbances, triggers a sequence of one or more dependent component outages". In the latter definition, large blackouts may or may not be consequences of cascading outages. However, in most cases, cascading outages lead to wide-area blackouts [7, 19].

In general, *cascading failures* can be defined as a sequence of interdependent outage events, initiated by few outages or disturbances [22, 23]. The initiating events can be attributed to various external factors or random events such as natural disasters, vegetation disturbances (e.g. tree contact), human errors, software/hardware errors, and so on. In recent years, cyber/physical attacks on power grids, such as the case of the Ukrainian cyber attack of 2015 [24] are also precursors to cascading failures. After the occurrence of the initiating events, the dependent sequence of outages result from the internal mechanisms such as voltage and angular instability, line overloads, hidden failures as well as errors related to maintenance, operation and human judgement [22].

1.2 Cascade Models of Power Grids

Modeling and studying cascading failures include a diverse field of techniques and approaches (see [22, 23] for a review). They include topological models, high level statistical models, deterministic and probabilistic models, simulation-based models for analyzing quasi-steady and dynamic behavior of the system, or hybrid and interdependent models with other systems (e.g., communication systems), and so on (see [25]). Many of these methods have been bench-marked and validated

as well as cross validated [26, 27]. In this section, we restate the cascade model adopted in this dissertation from the studies in [28, 29].

1.2.1 Cascade Phenomena in Networks

It is noteworthy to mention that cascading phenomena are not unique to power systems and occur in different forms (e.g., information diffusion, disease epidemics, viral marketing, etc.) on various real world networks (e.g., social networks and communication networks). Such processes have been the focus of many research studies. In this dissertation, we specifically focus on the cascading failure phenomena in power grid systems.

1.2.2 Cascade Model for the Power Grid

Cascade model with the quasi-static approach for simulating cascading failures for the power grid and the optimal DC power flow model has been adopted from the studies in [28, 29]. Here, we briefly restate the details of the model from [28, 29].

1.2.2.1 Transmission Line Overloading and Failure Mechanism

A system with $\mathcal{V}$ nodes (substations) interconnected by n transmission lines is considered. The main mechanism of cascading failures is transmission line overloading due to power flow redistribution after failures. Specifically, if failures occur in the power grid, the assumption is that power will be redistributed based on the optimal DC power flow formulation as described in the immediate next section. If the new power-flow distribution overloads lines based on their power-flow capacity, more failures will occur in the power grid. This process iterates until no more failures occur in the system. Note that as discussed in [28–30], a transmission line k has a capacity, say C_k, that can be

governed by the thermal limit, the voltage drop limit, or the steady-state stability limit of the line. Constant capacities for the lines are assumed, where for the IEEE 118-bus system are estimated using the power flow through the lines based on the original setting of the model. The estimated capacities are then rounded to the closest capacity from the set $\mathcal{C} = \{20, 80, 200, 500, 800\}$ in MW [30].

Following the models presented in [28, 29, 31], a threshold α_k for the power flow through the kth line above which the protection relay (e.g., circuit breaker or impedance protective relay) trips the line is considered. Various factors and mechanisms, which specify this threshold for transmission lines are discussed in [28, 29, 32, 33]. Thus, a line is considered overloaded when the power flow through the line exceeds $(1 - \alpha_k)C_k$. Moreover, the initial triggering events for cascading failures are considered to have more than one failure (two or three initial failures are considered in this case) as the N-1 security is ensured in all *loading levels* of the power grid. The power-grid *loading level* denoted by r is an operating characteristic of the power grid, defined as the ratio of the total demand to the generation-capacity of the power grid. The parameter r represents the level of stress over the grid in terms of the loading level of its components. Other examples of operating characteristics and their role on the power system cascades can be found in [28, 29, 34, 35].

1.2.2.2 Power-Flow Formulation

Next, the DC optimal power flow formulation, which is used for calculating power flow re-distributions after failures is restated from [28, 29]. In the transmission network, sets $\boldsymbol{\mathcal{L}}$ and $\boldsymbol{\mathcal{G}}$ are the set of load buses and the set of generator buses, respectively. The notation L_i represents the

demand at the load bus i. The DC power-flow equations [36] can be summarized as

$$F = AP, \tag{1.1}$$

where P is a power vector whose components are the input power of nodes in the grid (except the reference generator), F is a vector whose n components are the power flow through the transmission lines, and A is a matrix whose elements can be calculated in terms of the connectivity of transmission lines in the power grid and the impedance of lines. Note that this system of equations does not have a unique solution as the elements of vector P including the generator injections denoted by g_is for $i \in \mathcal{G}$ and served loads denoted by ℓ_j for $j \in \mathcal{L}$ are assumed to be unknown and need to be identified. A standard optimization approach with the objective of minimizing the cost function below for variables g_i and ℓ_j is considered(i.e., load shedding mechanism in the power flow optimization is considered):

$$Cost = \sum_{i \in \mathcal{G}} w^g{}_i g_i + \sum_{j \in \mathcal{L}} w^\ell{}_j \ell_j. \tag{1.2}$$

with respect to the following constraints:

1. DC power flow equations: $F = AP$.

2. Capacity of generators: ${G_i}^{\min} \leq g_i \leq {G_i}^{\max}$, $i \in \mathcal{G}$.

3. Load to be served at buses: $L_j \leq \ell_j \leq 0$, $j \in \mathcal{L}$.

4. Power flow capacity of transmission lines: $|F_k| \leq C_k$ for $k \in \{1, ..., n\}$.

5. Power balance constraints (power generated and consumed must be balanced):

$\sum_{i \in \mathcal{G}} g_i + \sum_{j \in \mathcal{L}} \ell_j = 0$.

Note that the quantities ℓ_js are negative and g_is are positive by definition. In this cost function, ${w^g}_i$ and ${w^\ell}_j$ are positive values representing the generation cost for every node $i \in \mathcal{G}$ and the load-shedding price for every node $j \in \mathcal{L}$, respectively. High prices for load shedding is assumed so that a load is to be curtailed only when the constraints of the optimization cannot be satisfied otherwise (such as in case of generation inadequacy or transmission capacity limitations). Also, a minimum ${G_i}^{\min}$ and maximum range ${G_i}^{\max}$ for the generators' injections, where the minimum and maximum can be set based on the generators' last setting and their possible range of variations is assumed. It is assumed that technologies, such as pumped-storage hydro-power, can make such assumptions possible and provide the flexibility in relative adjustments of generation during the short process of cascades. In the simulations, ${G_i}^{\max}$ is set based on the maximum capacity of the generators available in the IEEE 118-bus model and ${G_i}^{\min}$ is set to be 80 percent of the generator's last injection setting. Note that if the power grid separates into islands due to failures, the presented DC power flow optimization is run on each island separately. In cases where the balance of power generated and consumed cannot be achieved for an island, the components in the island are assumed to have failed.

1.2.2.3 Cascading Failure Simulations in Power Grids

Following the approach presented in [28, 29], in this dissertation, MATPOWER [37], a package of MATLAB m-files are used for solving the optimal power flow and simulating cascading failures based on a quasi-static approach that focuses on transmission line overloads as the mechanism for propagation of failures. For studies that require cascade data, a large dataset of cascade scenarios was generated by triggering two or three random initial failures in the system. Note that in this dissertation, we refer to the cascade as any number of successive failures after the initial disturbance

(even few more failures). The order of failures in each cascade scenario in the cascade dataset is stored. The power grid loading level, r (the ratio of the total demand over total generation capacity of the system) is used to simulate cascades under different settings and evaluate the role of operating characteristics in cascade analyses. For each analysis, at least 16,000 unique cascading failure scenarios have been simulated.

1.3 Key Contributions of this Dissertation

The main contributions of this dissertation can be broadly categorized based on the two critical infrastructures analyzed and studied throughout. The first contribution lies in the improvement of the reliability of the power grid infrastructure susceptible to cascading failures. The second contribution is towards increasing the reliability of the power grid, considering its interdependency with the transportation infrastructure via EV charging stations. In both reliability studies, the goal is to leverage the maximal amount of information from datasets of the respective infrastructures that implicitly capture various physical details of the infrastructure being studied. In this section, we discuss the key contributions as follows.

1. Studying the role of community structures in cascade processes of power grids: The first contribution is the comprehensive study and investigation of the role of community structures in the control and mitigation of cascading failures in power grids. While structures and communities were shown to impact cascade processes in various other networks [38–42], the detailed analysis and investigation of cascading failures in a power grid network, undertaken in this dissertation, is the first study of its kind. A critical property of community structures utilized in this dissertation for controlling cascades is the trapping property of communities, which suggests that a cascade entering a commu-

nity is likely to reach to most of the other members of the community, while less likely to reach to other communities. We studied the community structures of data-driven interaction graphs using community detection approaches, while considering both directed influence graphs and undirected correlation graphs. An important observation from this study was that applying different community detection techniques led to variation in the identified structures and each revealed various aspects of cascade processes.

2. Analyzing reliability of power grids by identifying critical components of cascade processes using community structures of interaction graphs: The second contribution is the utilization of the community structures in the interaction graphs of power grids for reliability analysis. Specifically, this dissertation focused on identifying critical components in the power grid's cascade processes. For this purpose, a novel *community-based centrality measure* was developed to capture various aspects of community structures of interaction graphs of power grids. The importance of the community-based centrality measure's identified critical components was also compared with critical components identified using standard centrality measures including betweenness, closeness, eigen, and degree, using multiple verification techniques such as data-driven, SI epidemic simulation-based, power system simulation-based, and graph-based approaches. While each of the standard centrality measures shed light on different aspects of the criticality of components, they are not designed to identify critical components in cascade and epidemic processes. Our key observation was that in most cases, our community-based centrality measure performed better than these measures in identifying critical components in the cascade. Further, we also compared the critical components identified by the community-based centrality measure with those identified by the influence-based measure of the study in [43]. This comparison also showed that the community-based centrality measure performed

better in most cases. The results of these experiments imply that protecting the critical components identified using our community-based centrality can help in reducing the risk of large cascades and blackouts.

3. Investigating impact of power system's loading level on criticality of components using community structures of interaction graphs: The third contribution is the investigation of the impact of power system's loading level on the interaction graph and the community structures. For two different loading levels i.e. normal and stressed levels, we observed different interaction graphs and consequently, different community structures in the graphs, which led to different ranking of components using the community-based centrality measure. Overall, our key observation was that depending on the condition and the operating settings of the power grid, the critical components of the system varied. Therefore, it is important to perform criticality analysis with considerations about the system's state and operating settings.

4. Characterizing and predicting size of cascades using Markov chain framework derived from community structures of interaction graphs: The fourth contribution is the development of a Markov chain framework to study the failure propagation process between communities and to characterize the likelihood of various cascade sizes. For this purpose, we formulated a Markov chain model based on the community structures in the interaction graphs of the power grid. The states of the community-based Markov chain model tracked the size of cascades. The main idea used was that the groups of components that formed communities, provided an estimate measure of cascade size. Thus, using the community-based Markov chain, the distribution of cascade sizes was characterized using the size of communities. Additionally, depending on the initial conditions such as the community from which the cascade started, cascade size distribution was characterized. A key finding from

this framework was the power-law behavior in the distribution of cascade sizes, which suggested the importance of community structures of interaction graphs in cascade behavior. Power law behavior has been observed in historical data as well as previous studies of cascading failures [20].

5. Analyzing reliability of power grids and its interdependent transportation infrastructure using an integrated framework: The fifth and the final contribution is the study and reliability analysis of the coupled power grid infrastructure and transportation infrastructure via interdependency links between EV charging stations and power grids. An integrated framework of power grids and EV charging stations modeled the interactions between the infrastructures based on the networked Markov chain framework i.e. the influence model. This integrated framework captured the state and stochastic dynamics of inter and intra-system interactions. Particularly, intra-system dynamics or self-influence transition probabilities were leveraged from real taxi dataset that contained movement of vehicles in a particular location. Constraint-based influences were used to activate and deactivate influences among the charging stations and substations to probabilistically force the components to transition to desirable states. For this purpose, a novel pricing algorithm was designed for charging stations (which was relayed to EV drivers by the cyber infrastructure) to motivate EV drivers to travel to appropriate locations such that the influences could be activated and deactivated. In general, this integrated framework can be applied to other networked systems as well as infrastructures to probabilistically capture interdependency aspects and improve the overall reliability of the coupled systems.

1.4 Structure of this Dissertation

For straightforward readability of this dissertation, each chapter is commenced with a brief introduction about the details of the overall chapter including discussion on the appropriate related

work and also concluded with a brief summary and key observations of the chapter. While Chapter 1 provides an overview of the overall dissertation, the remaining chapters from Chapter 2 to Chapter 5 are each focused on different aspects of reliability analysis of standalone power grids and Chapter 6 focuses on the reliability analysis of the power grid infrastructure in conjunction with the interdependent EV charging infrastructure.

Chapter 2 is focused on constructing the interaction graphs of power grids using the cascade dataset obtained from the cascade model discussed in Chapter 1. The inadequacy of the physical topology-based interaction graphs of power grids is discussed and a literature review on the two categories: data-driven and electric distance-based methods for constructing interaction graphs is provided. Then, results and statistics on the influence-based and correlation-based interaction graphs of the IEEE 118 test bus system is discussed along with key observations and conclusions. Additionally, appropriate thresholds applied to interactions graphs is discussed as well.

Then, Chapter 3 proceeds with the utilization of the influence-based and correlation-based interaction graphs, constructed in Chapter 2, to comprehensively study the role of community structures in the interaction graphs using disjoint and overlapped community structures over the directed and undirected interaction graphs. Key observations, discussion, and statistics for the IEEE 118 test bus system is provided.

Finally, Chapter 4 uses the community structures derived in Chapter 3 to perform the reliability study of the power system by identifying critical components in the cascade process. A novel community-based centrality measure is used to identify and rank the critical components of the IEEE 118 test bus system (which are mostly the bridge or overlap nodes). Further, the criticality of the identified components is verified by comparing with critical components identified

by standard centrality measures as well as an existing influence-based method [44] using data-driven, SI epidemic simulation, power simulation, and graph-based efficiency analyses. Additionally, the role of operating settings, i.e. loading levels in power grids' centrality is discussed.

In Chapter 5, the community structures derived in Chapter 3 is utilized to construct a Markov chain framework for characterizing and predicting size of cascades. Numerical analysis-based results are provided for the IEEE 118 test bus system, which includes the probability distribution of average cascade sizes as well as probability distribution of cascade sizes given the cascade started from a unique community.

In Chapter 6, the reliability analysis of the power grid infrastructure along with the EV charging infrastructure is provided. First, individual interaction graphs for the EV charging infrastructure and power grid infrastructure are constructed. Then, the two interaction graphs are combined together to form an integrated framework, in which all reliability studies are performed. This includes the constraint-based influence model analysis as well as the design of optimal charging prices for EV charging stations. Note that substations considered in this study is a generalization of the details of a power system.

1.5 Publications Resulting from this Dissertation

The following publications are associated with this dissertation.

1. Upama Nakarmi, Mahshid Rahnamay-Naeini, Md Jakir Hossain, and Md Abul Hasnat, "Interaction Graphs for Cascading Failure Analysis in Power Grids: A Survey", In Review, Energies, 2020.

2. Upama Nakarmi and Mahshid Rahnamay-Naeini, "A Markov Chain Approach for Cascade Size Analysis in Power Grids based on Community Structures in Interaction Graphs", In Review, Probabilistic Methods Applied to Power Systems (PMAPS), 2020.

3. Upama Nakarmi, Mahshid Rahnamay-Naeini, and Hana Khamfroush, "Critical Component Analysis in Cascading Failures for Power Grids Using Community Structures in Interaction Graphs", IEEE Transactions on Network Science and Engineering, 2019.

4. Upama Nakarmi and Mahshid Rahnamay-Naeini, "Analyzing Power Grids' Cascading Failures and Critical Components using Interaction Graphs", IEEE Power & Energy Society General Meeting (PESGM), 2018.

5. Upama Nakarmi and Mahshid Rahnamay-Naeini, "An Influence-based Model for Smart City's Interdependent Infrastructures: Application in Pricing Design for EV Charging Infrastructures", Springer, Cham, 2017.

6. Upama Nakarmi and Mahshid Rahnamay-Naeini, "Towards Integrated Infrastructures for Smart City Services: A Story of Traffic and Energy Aware Pricing Policy for Charging Infrastructures", Smartgreens, 2017.

Chapter 2: Interaction Graphs of Power Grids

As discussed in Chapter 1, we analyze *cascading failures* in power grids, which are complex phenomena caused due to complex interactions among components occurring within a short period of time [23]. Various studies and models have been developed to understand and control cascading failures including methods based on power system simulation [20, 46], deterministic analytical models [47], and probabilistic models [29, 48, 49].

Understanding and analyzing properties of each individual component in a power grid during cascading failures is challenging due to the large size and complex and sometimes hidden interactions among the components. Therefore, an alternative to reduce the complexity of cascading failure studies is by analyzing the global properties of the power grid by modeling the infrastructure as a graph, where nodes represent the individual components of the power grid and edges represent the interactions that the components have among themselves. In the past two decades, graph-based methods [3, 4, 44, 50–97] have attracted a lot of attention due to the simplicity of the models and ability to describe the propagation behavior of the failures on the graph of the system [98, 99]. Thus, in this dissertation, we will focus on graph-based models of power grids and provide a literature review on the various methods and research studies undertaken as well as ongoing in this specific area. In addition to graph representations of power grids, combination of power grids' graph with

Portions of this chapter were published in IEEE PES [3] and IEEE TNSE [4]. Copyright permissions from the publishers are included in Appendix A. Portions of this chapter are also available as preprint in arXiv [45].

graphs of other dependent infrastructures can represent the complex interdependency/interactions among the power grid and its dependent infrastructures (which will be discussed in Chapter 6).

We will first discuss the physical topology-based graphs of power grids as many initial graph-based models were developed based on the physical topology of the power system, where the connections among the nodes represented the actual physical connections among the components of the system [100, 101]. However, studies in [96, 102] showed the lack of strong connection between the physical topology of the system and failure propagation in cascading failures in power grids. In general, influences and interactions among the components of the system during cascade processes may occur both locally and at distance due to the physics of electricity governing the power flow dynamics as well as other functional and cyber dependencies among the components of the system. For instance, historical as well as simulation data verify that failure of a critical transmission line in the power grid may cause overload/failure of another transmission line that may or may not be topologically close. Therefore, graph models based on the physical topology of the system are not adequate in describing the propagation behavior of failures in power grids. Hence, new methods are emerging to reveal the complex and hidden interactions that may not be readily available from physical topology of the power system. These new approaches are focused on extracting and modeling the underlying graph of interactions among the components of the system.

In this dissertation, we use the term *interaction graphs* to refer to these models. We broadly categorize methods for constructing interaction graphs into two main classes: *data-driven approaches* and *electric distance-based approaches*. As the name implies, the *data-driven approaches* for building interaction graphs rely on data collected from the system (historical and real data or simulation data) for inferring and characterizing interactions among the components of the system. We further

define three categories for data-driven interaction graphs based on the method used for analyzing the data. These include: (1) methods based on outage sequence analysis [3, 4, 44, 50–71], (2) risk-graph methods [72–75]; and (3) correlation-based methods [3, 4, 76, 103]. The category of outage sequence analysis can be further divided into four sub-categories including (i) consecutive failure-based methods [50–57], (ii) generation-based methods [58–63], (iii) influence-based methods [3, 4, 44, 64], and (iv) multiple and simultaneous failure methods [65–71].

On the other hand, *electric distance-based approaches* exploit properties based on physics of power and electricity governed by Kirchoff's laws to define interactions among the components. Thus, the interactions are represented by *electrical distances*, which illustrate the properties of the electrical interactions based on power flows among the components. We define two sub-categories for *electric distance-based* interaction graphs based on the electrical properties utilized for creating the graphs. These include: (1) methods that define the interactions based on changes in the power flow due to changes in physical attributes of components caused by outage conditions [77–81] and (2) methods that define the interactions among components during normal or non-outage operating conditions [82–97, 104]. The category of defining interactions during non-outage operating conditions can be classified into: (i) impedance-based methods, which define interactions by considering a single impedance measure among the components connected over multiple paths [82–93] and (ii) sensitivities in components' states due to changes in voltage magnitudes and voltage phase angles [94–97, 104]. Figure 2.1 shows the taxonomy of the reviewed methods for constructing various types of interaction graphs. We will discuss these methods in detail.

Next, we use the cascade dataset generated in Section 1.2.2 of Chapter 1 to extract the interactions/relations among the components of the power grid during the cascade process. As the

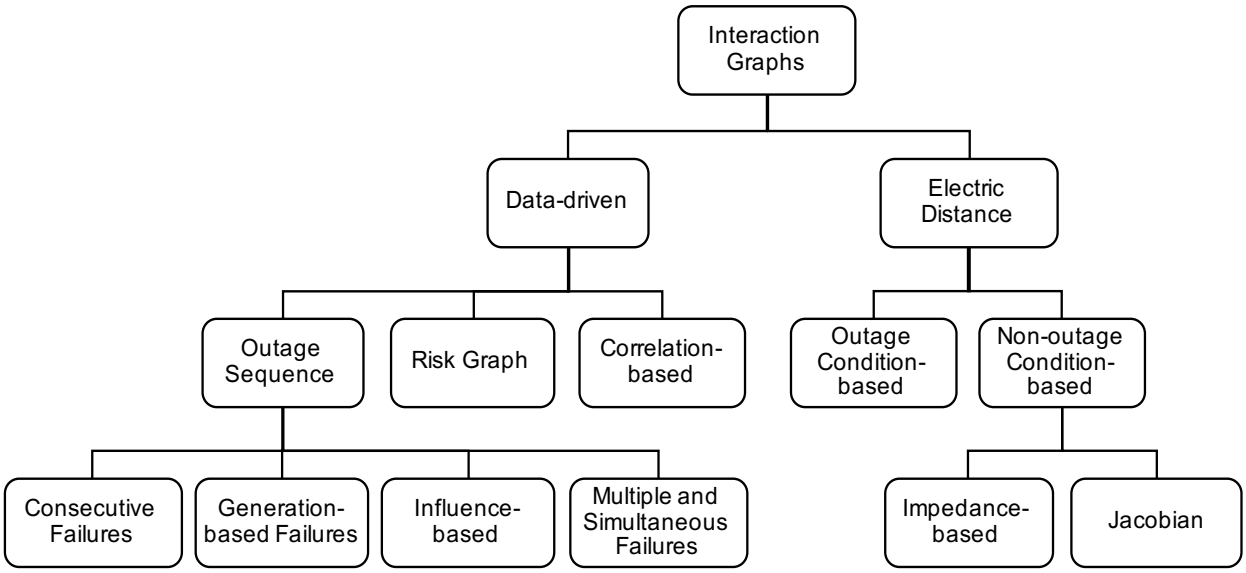


Figure 2.1: Taxonomy of methods for constructing interaction graphs.

cascade model is based on transmission line overloading, which is the main attribute for causing cascading failures in power grids, we focus on extracting interactions among the transmission lines of the power grid. Hence, in the interaction graphs, the nodes represent the transmission lines in the physical topology and the edges between the transmission lines will be the interactions that we derive from the cascade dataset. We will specifically use two techniques to extract the relations and construct the graph of interactions, including the influence-based [43] and correlation-based [105] approaches. Finally, we will present the interaction graphs for IEEE 118 bus system.

2.1 Physical Topology-based Graphs of Power Grids

Initial graph-based studies of power grids, such as [98–101], were based on the physical topology of the power grid. In general, a power grid can simply be represented by a graph, $G = (V, E)$, where V represents the set of generator, transmission, substation, or load buses and E represents the set of power lines [106].

Physical topology-based graphs of power grids shows the physical connectivity among the components of the system. Various studies have been performed on such graphs by analyzing their global structural properties [100, 101], such as average path length, clustering coefficient, and degree distribution, for analyzing power grids with respect to standard complex networks such as small world, random, and scale-free graphs. Particularly, the study in [101] compares the average path length and clustering coefficient of real-world power grids to their equivalent random and scale-free network models. However, the study concluded that real-world power grids differed significantly from standard network models as the clustering coefficient and the average path length of real-world grids were significantly greater than that of their complex network model counterparts.

Some studies performed on the physical topology also focus on properties of the electrical connections [107] identified using centrality measures such as degree, eigenvector, closeness, and betweenness (for a review and definition of centrality measures refer to [108]). However, it has been discussed that physical graphs may be inadequate in representing and capturing the interactions among the components of the power grid [96, 102] specifically for analyzing cascading failures. This limitation is due to the inability of the physical graphs to capture the dynamics of interactions at-distance in cascading failures, for instance, due to Kirchoff's and Ohm's laws.

Recently, some studies consider the physical and electrical properties of power grids to generate synthetic power grid networks that consider the heterogeneity of the components in terms of their operating voltages [106, 109]. In such graphs, each vertex is associated with a voltage rating such that transmission lines are represented as edges between vertices of the same voltage level and voltage transformers are represented as edges between vertices with different voltage levels. However, analysis of such graphs for cascading failures scenarios is an open research problem.

Table 2.1: Classification of existing studies using our data-driven taxonomy

Category	**Subcategory**	**Further Subcategory**	**Works**
Data-driven Interaction Graphs	Outage Sequence	Consecutive Failures	[50], [51], [52], [53], [54], [55], [56], [57]
		Generation-based Failures	[58], [59], [60], [61], [62], [63]
		Influence-based	[44], [64], [3], [4]
		Multiple and Simultaneous Failures	[65], [66], [67], [68], [69], [70], [71]
	Risk-graph		[72], [73], [74], [75]
	Correlation-based		[3], [4], [76]

2.2 Data-driven Interaction Graphs

Various *data-driven* approaches have been proposed for inferring and modeling interactions among the components of the power grid. These approaches rely on data from simulation or historical outage datasets. As the historical datasets are limited, majority of the studies use simulation data. Data-driven methods build a graph of interactions for the system, denoted by $G_i = (V_i, E_i)$, in which the set of vertices V_i are the components of the system that their interactions are of interest, such as the set of buses or transmission lines. Further, the set E_i represents the set of interactions/influences among the components, which may be directed, undirected, weighted (representing the strength of interactions or influences) or unweighted depending on the analysis of interest. We categorize and discuss data-driven methods for modeling interaction graphs for studying cascading failures in power grids into five categories as shown in Table 2.1.

2.2.1 Interaction Graphs based on Outage Sequences in Cascading Failures

This class of methods rely on cascade data in the form of sequence of failures in each cascade.

For instance, the sequence $l_5 \rightarrow l_7 \rightarrow l_3 \rightarrow l_6$ represents an example of sequence of transmission line failures in a cascade scenario. These methods are based on analysis of sequence of failures for extracting interactions and focus on the cause and effect interactions among failure of components. Methods in this category use various techniques and statistics to analyze such data as discussed next.

2.2.1.1 Interaction Graph based on Consecutive Failures

In this category of outage sequence analysis, only direct consecutive failures in a sequence are used for deriving the interaction links among the components of the system. In other words, two components in the system have a directed interaction link, $e_{i,j} \in E_i$, if they appear as successive outages in the order $l_i \rightarrow l_j$ in a cascade scenario in the dataset. For instance, if the sequence $l_5 \rightarrow l_7 \rightarrow l_3 \rightarrow l_6$ represents an example sequence of transmission line failures in a cascade scenario, the following directed interaction edges will belong to graph G_i, i.e., $\{e_{5,7}, e_{7,3},$ and $e_{3,6}\} \in E_i$. The strength of interactions among the components in this case can be characterized using the statistics of occurrences of pairs of successive outages in cascade scenarios in the dataset. For instance, the work in [50] assigns weights to the interaction edges by statistical analysis of number of times that a pair of successive line outages occur in the cascade dataset (i.e., $|l_a \rightarrow l_b|$/(total number of successive pairs), where $|l_a \rightarrow l_b|$ is the number of times failures l_a and l_b occur successively in the cascade dataset). These weights can be interpreted as the probability of occurrence of each pairs of successive line outages. Examples of studies using this method to develop the power grid's graph of interactions include [50–57], where they consider transmission lines in the system as the vertices V_i of the interaction graph G_i.

In the study presented in [110], the sequences of consecutive failures are called *fault chains*. For creating the dataset of fault chains, in the first step, a single transmission line is tripped as the initiator of cascading failure in the simulation and in the subsequent steps, the most overloaded component due to power flow re-distributions is considered as the next failure in the overall sequence of consecutive failures. In the studies in [51–53, 55–57], for a power grid with n transmission lines, n fault chains are created and the edges among consecutive failures in each chain is weighted based on power flow changes in a line after the failure. In the work in [54], fault chains are created by considering multiple initial failures such that, a system with n transmission lines may have more than n fault chains. Finally, a fault chain graph is developed by combining all fault chains together into a single graph where the vertices are all the components that have failed in the fault chains and the edges between the vertices exist if the outages have successively occurred in the fault chains. For pairs of outages $(i \rightarrow j)$ that have reoccurred in multiple fault chains, their combined edge weight in the fault chain graph is averaged.

2.2.1.2 Interaction Graph using Generation-based Failures

The method based on the consecutive failures discussed in Section 2.2.1.1 focuses on one to one impact that the outage of a line has on the outage of another line. However, in cascading failures, instead of pair-wise interactions among successive failures, a group of failures may contribute to failures of other components. Therefore, it is important to consider the effects of groups of failures and characterize interactions among the components based on the effects among groups of components. The works presented in [58–62] define such groups as *generation* of failures within a cascade process, which are failures that occur within short temporal distance of each other. In these works, the sequence of failures in the cascade are divided into sequence of generations and the failure

induced cause and effect relationships are considered between consecutive generations. Specifically, outages occurring in generation $m+1$ are assumed to be caused by outages in generation m.

The interactions based on successive generations are defined in different ways in the literature. For instance, the authors in [44] assume that all components in generation m have interactions with all components in generation $m+1$, i.e., if generation m has n_1 number of components and generation $m+1$ has n_2 number of components, then the number of interactions between generation m and $m+1$ will be $n_1 \times n_2$. But some studies argue that considering all possible pairs of interactions among components of two consecutive generations overestimate the interactions among components [58–63]. Specifically, all line outages in one generation may not be the cause of a line outage in the next generation. Therefore, in the works presented in [58, 60], the cause of failure of a line k in generation $m+1$ is considered to be due to the failure of a line in generation m with the maximum influence value on the line k. The influence value for component j in generation m is defined as the number of times that the component j has failed in a generation m before the failure of line k in the successive generation $m+1$ in the cascade dataset. For cases where two or more lines in generation m have the same maximum influence values on line k in generation $m+1$, all such components are assumed to interact with line k.

In the works discussed so far in this section, the interaction among component j in generation m and component k in generation $m+1$ will be represented by a directed link $e_{j,k}$. While the works in [58, 60] limit the interactions by only considering the maximum influence values in current generation as the probable cause of component failures in the next generation, the work in [59] gives an improved estimate of the interactions between successive generations using the expectation maximization (EM) algorithm.

The weight of the interaction links can also be defined in various ways. For instance, the weight of the link can be defined as the ratio of number of times that the pair of components appeared in two successive generations over the total number of times that component k has appeared in the dataset. This weight can be interpreted as the likelihood of failure of component k in the next generation given the failure of component j in the current generation. The study in [63] considered both statistical properties as well as the amount of load shed that has occurred between successive generations to assign interaction link weights. However, the study in [63] identified islands formed in the power grids during outages and then, selectively assigned links between components of successive generations only if the generations were located in islands that were direct consequences of one another.

2.2.1.3 Influence-based Interaction Graph

In this method, the interactions among the components are derived based on successive generations in cascades; however, the weights of the interactions are characterized based on the influence model and the branching process probabilistic framework. The influence model is a networked Markov chain framework, originally introduced in [111] and was first applied to cascade dataset in the work presented in [112]. Studies in this category consider transmission lines in the system as the vertices V_i of the interaction graph G_i and the influences/interactions between the lines as the edges E_i.

In [44], authors consider interaction links among all pairs of lines in two successive generations in a cascade using the influence model. The weights of the directed links are derived in two steps. In the first step, a branching process approach is used in which each component can produce a random number of outages in the next generation. The number of induced outages by

each component is assumed to have a Poisson distribution based on the branching process model. Parameter λ_i specifies the propagation rate (mean number of outages) in generation $m+1$ for the outage of component i in generation m. In other words, this step defines the impact of components on the process of cascade by describing how many failures their failure can generate [49]. In the second step, it is assumed that given that component i causes k outages in the next generation, some components are more likely to outage than others. Therefore, they calculate the conditional probability $g[j|i]$, which is the probability of component j failing in generation $m+1$, given the failure of component i in generation m. If only $g[j|i]$ values based on the statistical analysis of data are considered, then the probability of failure of component j given component i failure will be known; however, the expected number of failures from failure of component i is not known. Hence, both steps are important in characterizing the influences among components.

The final step consists of combining the information from the first and second steps into a single influence matrix $\mathbf{H}$ (representing the links of graph of interactions and their weights). The elements of the matrix are defined based on the conditional probability that a particular component j fails in the next generation $m+1$, given that component i has failed in generation m and that generation $m+1$ includes exactly k failures. This probability can be defined as $P(j|i,k) = 1-(1-g[j|i])^k$. Then, the conditional probability $h_{i,j,m}$ that component j fails in generation $m+1$, given that component i failed in generation m, over all possible values of k represents the actual elements of $\mathbf{H}$, and is found by multiplying $P(j|i,k)$ with the probability of k failures occurring as follows:

$$h_{i,j,m} = \sum_{k=0}^{\infty}(1-(1-g[j|i]^k))\frac{\lambda_{i,m}^k}{k!}e^{-\lambda_{i,m}}. \tag{2.1}$$

Based on the influence graph, cascading failures can start with a line outage at a node of the graph and propagate probabilistically along the directed links in the graph. Examples of other works, which have used the influence-based approach to derive the graph of interactions for power grids include [3, 4, 64]. In this dissertation, we will use this technique in deriving the graph of interactions for cascade analysis, as will be discussed in Section 2.4.

2.2.1.4 Interaction Graph of Multiple and Simultaneous Failures

This class of methods also use the sequence of failures to model the interactions among the components of the system; however, they consider the interactions among multiple simultaneous failures.

Specifically, in the study presented in [65], a Markovian graph was developed with the goal of addressing the problem of capturing the effect of multiple simultaneous outages within generations on the characterization of the interactions among the components of the successive generations in a cascade. In this case, the nodes of the graph represent the states of the Markov chain defined as the set of line outages in a generation of the cascade and the links represent the transition among the states (i.e., interactions between successive generations of outages). Hence, each node in the graph may represent the outage of a single line or multiple lines. Markovian interaction graphs also consider a node with a null state, which represents the state where the cascade stops. This state occurs at the end of all cascade scenarios. The transition probabilities among the states (i.e., the weight of the links) from state i to state j can be estimated by counting the number of consecutive states in which state i and state j occur in all the cascades and dividing by the number of occurrences of state i. Other studies that consider the interaction among multiple failures at the same time are presented in [66–71].

2.2.2 Risk Graphs for Interaction Graph

The work presented in [72] introduces the *risk-based* interaction graphs, which describes the interactions or relationships among the nodes (i.e., buses/substations) of the power grid based on effects of their simultaneous failures in causing damage in the system. This graph is not solely focused on analysis of interactions among components during cascading failures. Instead, it is focused on the vulnerability analysis of the power grid and the effect of failures is assessed using metrics such as *net-ability*, which measures the effectiveness of a power grid subjected to failures, based on power system attributes including power injection limitation and impedance among the components.

Construction of risk graphs are done in two steps. The first step includes generating and tracking the sets of strongest node combinations whose simultaneous failures have significant effects on the power grid [72, 73, 73]. In the second step, these sets of strong node combinations are used to form the *risk graphs*. If a node appears at least once in the sets of strong node combinations, then the node becomes a vertex of the risk graph. Links among nodes in the risk graph exist if they appear in the same set of strong node combination. Both nodes and links in the risk graph are weighted based on the frequency of their appearance in the sets of strong node combinations. This approach results in a weighted but undirected node risk graph, where higher weight values on the links suggest stronger node combinations. Node risk graphs are dependent on the system parameters such as ratio of capacity to the initial load of the nodes in the system. To remove dependencies on system parameters, node risk graphs can be constructed for multiple parameter values and combined together to form the node integrated risk graph using the risk graph additivity property [72, 73]. The aforementioned risk graph can also be extended to a directed risk graph,

where the removal of components in a specific order in strong node combinations are considered [74].

Another similar concept to risk graph is the *double contingency graph* introduced in [75]. While m contingency combinations of attack scenarios for the power grid was studied in the risk graphs, many methods focus on N-2 contingency analysis as the power grid is considered to be N-1 protected [113]. In the double contingency graph, the vertices of the graph are the transmission lines and the links between vertices show pairs of transmission lines whose simultaneous failure as initial triggers can affect the reliability of the system by, for instance, violating the thermal constraint rules in the power grid.

2.2.3 Correlation-based Interaction Graph

The work in [76] presents a graph of interactions for power grids based on correlation among the failures of the components. In the correlation-based interaction graph in [76], vertices represent the transmission lines and the edges represent the pairwise correlation between line failures in the cascade dataset. The correlation dependence between failures are captured in the correlation matrix, whose ijth elements are positive Pearson correlation coefficient between the failure statuses of components i and j in the cascade dataset. The resulting correlation matrix is symmetric and can be interpreted as an undirected and weighted interaction graph, where the nodes are the failed lines, the edges are the interactions between the lines and the weights are the correlation values among the components. Similarly, the studies in [3] and [4] also construct correlation-based interactions graphs from simulated cascade dataset consisting of sequences of transmission line failures. In this dissertation, in addition to the influence-based technique discussed in Section 2.2.1.3, we will also

use this technique in deriving the graph of interactions for cascade analysis, as will be discussed in Section 2.4.

2.3 Electric Distance-based Interaction Graphs

Various *electric distance-based* methods have been proposed for modeling interactions among the components of the power grid using the dynamics of power flow as well as the physical/electrical properties of the system and components. In a power grid, electricity does not flow through the shortest path between two nodes i and j. Instead, it can flow through parallel paths between nodes i and j based on the physical properties of the system and its components as well as the physics of electricity (i.e., Ohm's law). Thus, the electrical interactions/distances between the components may extend beyond the physical topology and the direct connections in the power grid. The concept of *electric distance* was first introduced by Lagonotte et al. [104] in 1989 as a measure of coupling between buses in the power system and was based on sensitivities in the power system due to changes in voltage magnitudes.

Similar to data-driven interaction graphs, electric distance-based methods build a graph of interactions for the system, denoted by $G_i = (V_i, E_i)$, in which the set of vertices V_i are the components of the system that their interactions are of interest, such as the set of buses or transmission lines. Further, the set E_i represents the set of interactions/influences among the components, which may be directed, undirected, weighted (representing the strength of interactions or influences) or unweighted depending on the analysis of interest. We categorize and discuss electric distance-based methods for modeling interaction graphs in power grids into two categories as shown in Table 2.2.

Table 2.2: Classification of existing studies using our electric distance-based taxonomy

Category	**Subcategory**	**Further Subcategory**	**Works**
Electric Distance-based Interaction Graphs	Outage Condition-based		[77], [78], [79], [80], [81]
	Non-outage Condition-based	Impedance-based	[61], [82], [83], [84], [85], [86], [87], [88], [89], [90], [91], [92], [93]
		Jacobian	[86], [94], [95], [96],[97]

2.3.1 Outage Condition-based Interaction Graph

This class of methods for characterizing electric distance-based interaction graphs are focused on interactions among the components of the power grid during outage conditions. For instance, in the study presented in [78], interactions among the components as well as their weights are derived using the changes in the power flows in transmission lines during outage conditions. Thus, the outage induced interaction graph G_i consists of the set of vertices V_i that represents the transmission lines and the set of edges E_i that represents the impact of outage of one line on another. This impact is characterized using LODF [114], where LODF for line $e_{i,j} \in E_i$ is calculated based on the ratio of the impact of outage of line i on line j based on the reactance of all possible spanning tree paths between the lines, over the impact of outage of line i on line j using the reactance of all alternative spanning tree paths that the power can flow (i.e., excluding the spanning tree path of line i) [77].

However, during cascading failures, the impact of a failed line on the remaining lines is not limited to changes in power flows. In a power grid, if two or more lines share a bus, outage of one line may expose the remaining lines (connected through the same bus) to incorrect tripping due to malfunctioning of the protection relays. The exposed lines are prone to failure and increase in power flow in the exposed lines exacerbates their tripping probability causing further outages. Such failures

are known as hidden failures. In the studies in [79–81], vertices V_i represent transmission lines as well as a hidden failure state and edges E_i represent inter-line interactions as well as interactions between lines and the hidden failure state. Thus, the interaction graph G_i will have $n + 1$ nodes where n is the number of transmission lines and the extra one node represents the hidden failure state. The hidden failure node has bidirectional links from itself to every other node in the power grid. However, the hidden failure node does not have influence on itself. The inter-line interaction $e_{i,j} \in E_i$ shows the increase of power flow in line j due to outage of line i. The interaction from the hidden failure node to a line i reflects the tripping probability of line i caused due to the increase in power flow in line i exceeding the line flow limits. And, interaction from line i to the hidden failure node reflects the average tripping probability of all the other remaining lines.

2.3.2 Non-outage Condition-based Interaction Graphs

As the name suggests, this class of methods for constructing the electric distance-based interaction graphs are focused on interactions among the components of the power grid during normal operating conditions. We broadly categorize non-outage condition-based interaction graphs into two categories: impedance-based and Jacobian as discussed next.

2.3.2.1 Impedance-based Interaction Graph

Impedance-based electric distance interaction graphs G_i consist of vertices V_i that represent buses and edges E_i that represent electrical interactions between pairs of buses weighted by their corresponding impedance-based electrical distances.

Inverse admittance matrix, more commonly known as the impedance matrix $\mathbf{Z}$, is one of the simplest forms of representing electrical interactions between pairs of buses in the system and is

found by inverting the system admittance matrix $\mathbf{Y}$, i.e. $\mathbf{Z} = \mathbf{Y}^{-1}$. Matrix $\mathbf{Z}$ shows the relationship between the nodal bus voltage vector and the nodal current injection vector. However, unlike matrix $\mathbf{Y}$, which is sparse, impedance matrix $\mathbf{Z}$ is non-sparse as it represents the changes in nodal voltage throughout the system due to a single nodal current injection between a pair of nodes in the system. Therefore, edges in the impedance-based interaction graph are the connections between the elements in the $\mathbf{Z}$ matrix with weights between buses i and j corresponding to their absolute value of the impedance, i.e. $|Z_{ij}|$ [82–90]. Smaller magnitudes of impedance represent shorter electric distance between buses. Note that the individual elements Z_{ij} in matrix Z are complex valued.

Note that studies in [82–90] are not focused on cascading failures but their concept of formulating impedance-based electrical distances can be extended for studying cascade processes. For example, in the study in [61], transmission lines are considered as the nodes of the interaction graph and thus, impedance-based electric distance between pairs of transmission lines are assigned as weights of the interaction links. To find electric distance between transmission lines i and j, where line i connects bus i_s to i_d and line j connects bus j_s to j_d, the minimum of the four possible Thevenin equivalent impedance's between the pairs of buses, i.e., $Z_{i_s j_s}$, $Z_{i_s j_d}$, $Z_{i_d j_s}$, and $Z_{i_d j_d}$ is taken.

Cascading failures can also be studied by representing interactions between pairs of nodes by effective resistances between the nodes. Effective resistance, R_{ij}, between nodes i and j, also known as *Klein resistance distance* [115], is the equivalent resistance of all parallel paths between the nodes. It shows the potential difference between nodes i and j due to unit current injection at node i and withdrawal at node j. For cascading failure analysis, as the impedance of a transmission line in a high voltage transmission network is dominated by the imaginary part of impedance, i.e. reactance,

the effective resistance between nodes can be formulated in terms of their reactance. Thus, in the studies in [91–93], effective resistance between nodes i and j is found as $R_{ij} = Q^{+}_{ii} - 2Q^{+}_{ij} + Q^{+}_{jj}$ where, Q_{ij} is the row i and column j element of $\mathbf{Q}^{+}$, which is the penrose pseudo-inverse of the Laplacian matrix $\mathbf{Q}$. Matrix $\mathbf{Q}$ is defined as the difference between the weighted diagonal degree matrix and the weighted adjacency matrix derived from the physical topology, and shows the relationship between the buses and transmission lines in the grid. In the studies in [91–93], the weights of the edges in the physical topology required for finding the weighted diagonal degree and adjacency matrix are the susceptance (i.e., the imaginary part of impedance) values between the nodes. Thus, edges E_i in the effective resistance interaction graph reflect the electrical connections between the buses with weights between nodes i and j being the corresponding R_{ij} values.

2.3.2.2 Jacobian Interaction Graph

Electric distance-based interaction graphs can also be constructed using the sensitivity matrix of the power grid during normal operating conditions. In such interaction graphs G_i, vertices V_i represent the buses and the edges E_i represent the electrical interactions in terms of sensitivities between the buses. These sensitivities can be found using the Jacobian matrix, which is obtained during Newton Raphson-based load flow computation. Jacobian sensitivity matrix $\mathbf{J}$, shows the effect of complex power injection at a bus on the voltage magnitude and voltage phase angles of other buses. Matrix $\mathbf{J}$ consists of four sub-matrices.

The seminal work of electrical distances by Lagonotte et al. in [104] focused on using the Jacobian sub-matrix ($\mathbf{J_{QV}}$), also known as the voltage sensitivity matrix $\boldsymbol{\partial \mathbf{V}/\partial \mathbf{Q}}$, to find the electric distance between buses. Similarly, the study in [94, 95] also used the voltage sensitivity matrix. In the studies in [94, 104], the matrix of maximum attenuations was found, which consisted of

columns of voltage sensitivity matrix divided by the diagonal values. Finally, electrical interactions $e_{i,j} \in E_i$ between buses i and j weighted by their electric distance was derived as the logarithm of the individual elements of the attenuation matrix. Note that studies in [94, 95] analyzed the risk of cascading failures by studying the voltage collapse phenomenon, which is a sequential process during which large parts of the power grid may suffer due to low voltages [116].

2.4 Interaction Graphs for IEEE 118-bus System

The simulated cascade dataset generated using the cascade model discussed in Section 1.2.2 of Chapter 1, is used to find the logical graphs of interactions, H and CR based on the influence-based and correlation-based techniques discussed in Section 2.2.1.3 and Section 2.2.3 respectively. The physical topology of the IEEE 118 test bus system shown in Figure 1.2 of Chapter 1, is converted to a 186 node *line graph* shown in Figure 2.2, where the 186 nodes represent the edges (transmission lines) in the original topology of Figure 1.2 and edges represent common buses between the transmission lines.

We use the overall cascade dataset (with size of 16,000 cascade scenarios) without considering operating characteristics (i.e., r parameter). We constructed the directed graph of interactions based on influence-based approach (i.e., graph H) and the undirected graph of interactions using correlation-based approach (i.e., graph CR). Excluding the self loops, total number of possible links in the interaction graph is 34,410 for the undirected correlation CR matrix and 68,820 for the directed influence H matrix. However, based on the cascade dataset, the graph of interactions for CR and H consist of 34,396 and 32,504 links, respectively. Both graphs H and CR are dense graphs. However, there are 0.05% missing links out of the total possible undirected links in graph CR (due

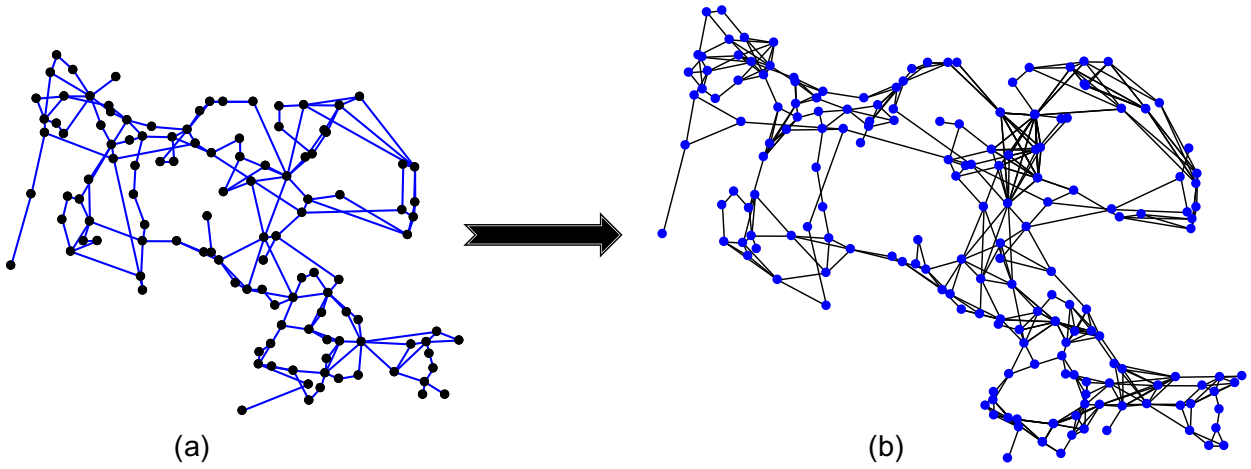


Figure 2.2: IEEE 118-bus system converted to a 186 node *line graph*. (a) Topology of the IEEE 118-bus system shown in Figure 1.2 of Chapter 1, where black nodes represent the buses of the power grid and the blue links represent the transmission lines between the buses. (b) Line graph of the topology shown in (a), where blue nodes in (b) represent the transmission lines in (a) and black links in (b) show the common buses between the transmission lines in (a).

to zero correlation) and 52.7% missing links out of the total possible directed links in graph H (due to zero influence).

2.4.1 Applying Threshold to Interaction Graphs

Influence-based and correlation-based approaches pick even the smallest interactions based on the cascade data, due to which the small interactions may act as noise [117]. Hence, to focus on the major interactions in the system, we apply thresholds to only consider interactions with strength larger than a threshold value [118]. Since, thresholds can result in islands in the interaction graph, we focus on the Largest Connected Components (LCC). Identification of LCC in interactions graphs with major strength of interactions can be used in prediction of the largest cascade sizes. For comparisons between the interaction graphs based on influence-based and correlation-based approaches, we choose thresholds such that the size of LCC is comparable in both networks. For

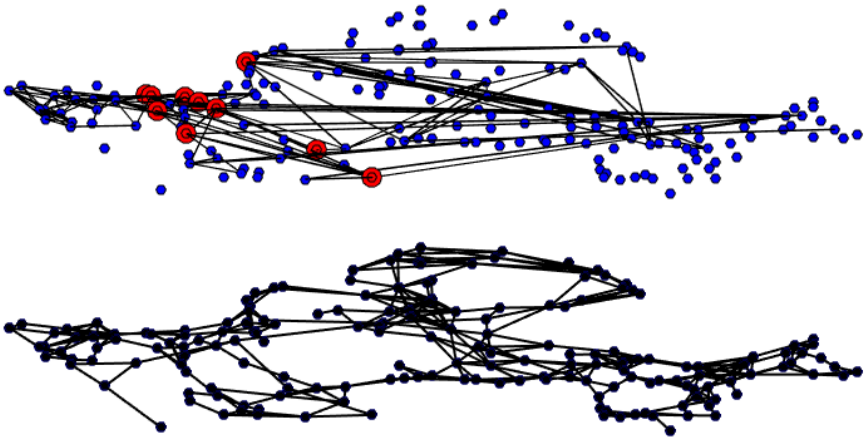


Figure 2.3: Interaction graphs for H (Threshold ≥ 0.7) over the *line graph* (shown in black).

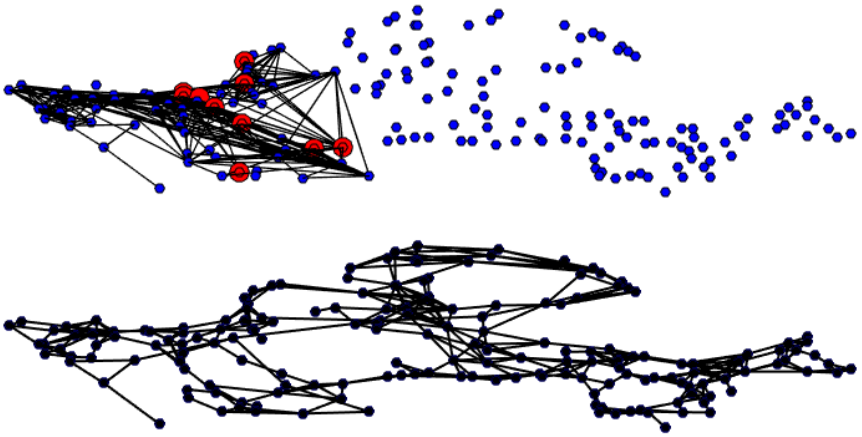


Figure 2.4: Interaction graphs for CR (Threshold ≥ 0.7) over the *line graph* (shown in black).

instance, Figure 2.3 and Figure 2.4 show the graph of interactions atop the line graph (shown in Figure 2.2 (b)) for the influence-based and correlation-based approaches for threshold 0.7, for which the size of LCC for H and CR are 57 and 59, respectively. The strong interactions on the same dataset of cascades is shown among different set of components in Figure 2.3 and Figure 2.4, which emphasize the role of the technique in extracting the graph of interactions. We also observed that the size of LCC for $H \geq 0.6$ and $CR \geq 0.4$ are similar. However, in these thresholds too, the strong interactions among components are shown in different sets of components. Similar

to the different techniques that lead to different graph of interactions on the cascade process, we observed that operating settings, e.g. different loading levels for the power grid r, also result in different graphs of interactions. In the upcoming chapters, we will show that different techniques and operating settings lead to different structures in interaction graphs and consequently, different criticality of the components. These properties will have a central role in determining propagation path of failures and help in predicting cascade sizes. Overall, these studies contribute in mitigating large scale cascading outages.

2.5 Summary

In this chapter, we reasoned the importance of graph representations of power grids conditioned to cascading failures. We discussed the physical topology-based graphs of power grids and the inadequacy of such graphs in providing accurate representations of cascading failure scenarios. We learned that accurate modeling of cascading failures by graphs should be undertaken using interaction graphs rather than physical topology-based graphs. Thus, we reviewed existing work on various techniques deployed for constructing interaction graphs in two distinct categories: data-driven and electric distance-based. We specifically used two data-driven methods: influence-based and correlation-based for deriving interaction graphs from our cascade dataset produced in Chapter 1. We showed that depending on the technique of deriving interaction graphs, different graphs are produced from the same dataset. Such graphs have the same number of nodes but differ in the interaction links among components, in terms of the number of links and their direction as well as their inherent meaning. For example, the influence-based interaction graphs are directed and weighted, where the weights represent the probability of failure of the child node given the failure of the parent node. In contrast, the correlation-based interaction graph is undirected and weighted,

where the weights represent the correlation with which two components fail in the same cascade. We also showed that applying thresholds to interaction graphs is necessary to focus on the influential interactions and remove unwanted noise. However, we observed that applying the same threshold to the influence-based and correlation-based interaction graphs produce different graphs with interaction links among different sets of components. Further, we also showed that operating settings, such as loading levels of the power grid impact the produced interaction graph. We postulate that different techniques of deriving interaction graphs will consequently lead to different structures and patterns in the interaction graphs and thus, different criticality of components, as will be discussed in detail in Chapter 3 and Chapter 4 respectively. We will also produce interaction graphs for the interdependent power and transportation infrastructure for the purpose of reliability analysis, as will be discussed in Chapter 6.

Chapter 3: Structures and Patterns in Interaction Graphs of Power Grids

Structures and patterns in various networks are important in describing the spread of various processes such as infectious diseases, behaviors, rumors etc. [119–121]. In the case of large scale power grids, structures present in the interaction graphs of power grids can be used to study the impact of cascading failures in the transmission network and utilize the graph structure to mitigate large cascades. For example, graph structures considered in the studies in [77, 78] for cascade mitigation are tree partitions. In the study in [77], tree structures present in the outage condition-based interaction graph (details discussed in Section 2.3.1) showed that transmission line failures could not propagate across common areas of tree partitions. Further, the extended work of [77] in [78] found the critical components of the tree partitions, known as bridges. The failure of bridge lines played a crucial role in the propagation of cascading failures. However, failure of non-bridge components did not propagate failures and the impact was more likely to be contained inside smaller regions/cells. This important property of bridge lines was used in mitigation of cascading failures by switching off transmission lines that caused negligible network congestion as well as improved the robustness of the system.

Purpose of analyzing structures present in the interaction graphs are not limited to mitigation of cascading failures. For instance, in the study in [83], network structures were used for

Portions of this chapter were published in IEEE PES [3] and IEEE TNSE [4]. Copyright permissions from the publishers are included in Appendix A.

contingency analysis. The impedance interaction graph (details discussed in Section 2.3.1) in [83] was pruned by removing edges above an operator defined threshold. Then, the common structure between the pruned impedance interaction graph and the topological graph of the power grid were analyzed and verified to be the contingencies that violate transmission line limits and cause overloads. Similarly, the study in [85] identified zonal patterns in the impedance interaction graph for reliability assessment of zones for load deliverability analysis.

In general, there are various ways and purposes for defining the structures and patterns of connections in graphs. In this dissertation, we analyze the *community structures* present in the influence-based and correlation-based interaction graphs, derived in Chapter 2 using the cascade dataset generated in Chapter 1. Many research works have studied the impact of structures and communities in networks on cascade processes [38–42]. Communities are defined as groups of nodes that are densely connected among themselves while having scarce connections to other groups. Specifically, in these studies, it has been shown that in addition to the microscopic and macroscopic properties of networks, community structures within the network have significant impact on cascade behavior. According to the findings in [38, 40, 42], in general a cascade entering a community will reach to most of the other members of the community, while it is more difficult for it to reach to other communities. In other words, the communities act as traps for the cascade processes over the networks [41, 122, 123]. This important property can have important practical implications in controlling the cascade processes by limiting propagation of cascading failures inside community structures.

However, even with community detection, different techniques can reveal different structures. To show this point, we study the community structure of the graph of interactions using disjoint and

overlapped community detection algorithms while considering both directed and undirected graphs. For community detection, we specifically apply Infomap [124, 125], which considers link directions, and conductance-based community detection [126], which does not consider link directions. These community detection techniques are different in the nature of their structure detection approach ,as discussed in Section 3.2.3. We will show that applying different community detection techniques can lead to variation in the identified structures and each can reveal various aspects of cascade processes.

3.1 Related Prior Work

In this section, we briefly review the existing studies for analyzing various roles of community structures in networks. It has been shown that the community structure in a network is an important player in percolation process, spreading behavior, epidemic model or information diffusion in the network [38–42]. The extent to which communities impact the epidemics are particularly studied in [38]. The authors in [38] discussed that while the set of communities are of crucial importance, the exact internal structure of the communities barely influence the behavior of the percolation processes across networks unless it is a targeted attack that starts the spread. Other works have discussed that communities act as traps for the epidemic processes over networks [38], [41]. The work in [38] also discussed that the inter-community edges are important for the spread of epidemics. As suggested by [39], on average, epidemics in networks with strong community structures exhibit greater variance in the final size.

The community structure of networks has also been studied for targeted immunization to prevent spread of diseases [39], [40]. Specifically, [39] exploits the concept of community bridges

(i.e., nodes that connect to multiple communities) and presents a stochastic algorithm for finding bridging nodes using random walk over the network. The authors in [39] showed that immunization of identified bridging nodes are more effective than those simply targeting highly connected nodes. In this chapter, we use a similar concept to protect the critical components identified based on the community structure to reduce the risk of large cascades.

3.2 Community Structures in Interaction Graphs of Power Grids

In this section, we will start by discussing some preliminaries of community structures, including the types of community structures and their importance in cascade behavior. Then, we will discuss two specific community detection algorithms that we use in this dissertation. Finally, we will show the results obtained from the community detection algorithms applied to the interaction graphs derived in Chapter 2 using the cascade dataset generated in Chapter 1.

3.2.1 Preliminaries

Communities are densely connected groups of components with scarce connections to components of other groups [127]. Community structures identified in graphs can be of two types (1) *overlapped*, such that a component may be a member of more than one community and (2) *disjoint*, such that a component is a member of a single community. Community structures are identified in graphs using community detection techniques (see [127–129] for a survey of such approaches). These approaches utilize the inherent patterns and properties of the graph, such as the weights of the links and their directions. In general, the goal of the community detection algorithm is to identify the set of communities and the membership labels for the components of the graph. Figure 3.1 shows examples of disjoint and overlapped community structures. As seen in the figure, bridge nodes of

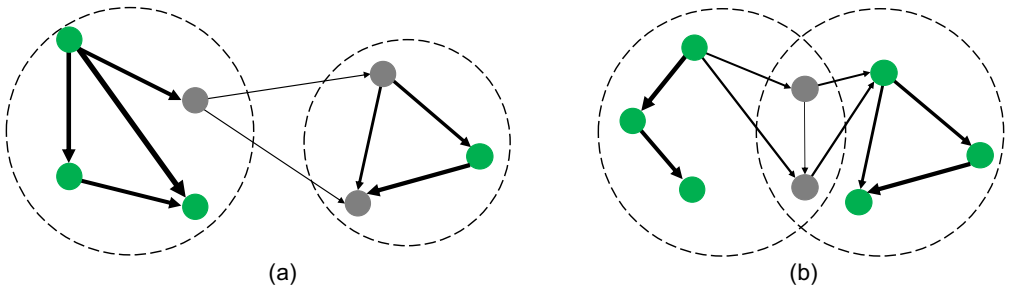


Figure 3.1: Examples of community structures shown by dashed lines. (a) Two disjoint communities with weak interactions between the bridge nodes (bridge nodes shown in grey). (b) Two overlapped communities with weak interactions between the overlap nodes (overlapped nodes shown in grey).

disjoint communities tend to have weak interactions with components of other communities while having strong interactions with components within the same community. Similarly, overlap nodes of overlapped communities tend to have strong interactions among other components within the same community but weak interactions among overlap nodes of different communities.

Studies in network science have shown that communities play important roles in defining the propagation behaviors in networks [123]. Particularly, the propagations tend to stay within communities due to tight internal interactions and weak external interactions. In this dissertation, we focus on overlapped as well as disjoint community structures in the interaction graphs of the power grid as we believe they can provide a new perspective in defining the key players in the cascade process.

3.2.2 Identifying Community Structures in Interaction Graphs

In this section, we evaluate the community structure of the graph of interactions to subsequently characterize their role in the cascade processes in power grids. Consider the interaction graph, constructed using the influence-based and correlation-based techniques discussed in Section

2.2.1.3 and Section 2.2.3 of Chapter 1, respectively. Let the interaction graphs be denoted by $\mathcal{IG} = (N_{IG}, E_{IG})$, where N_{IG} is the set of the nodes in the graph (i.e., the set of transmission lines in the power system) and E_{IG} is the set of interaction links in the graph. Community detection algorithms identify the set of communities, $\boldsymbol{\mathcal{C}} = \{C_1, C_2, \cdots, C_K\}$, where $\boldsymbol{K}$ is the number of communities in the graph, based on the patterns and weights of interactions. Then they assign membership labels to the nodes in N_{IG}. We denote the set of community labels for node $n_i \in N_{IG}$ by $CL(n_i) = \{C_v, \cdots, C_u\}$, where $C_v, C_u \in \boldsymbol{\mathcal{C}}$. For the overlapped community structures $|CL(n_i)| \geq 1$, while for disjoint community structures, $|CL(n_i)| = 1$, where $|.|$ is the cardinality of the set.

In this dissertation, we study the overlapped as well as disjoint community structures in the power grid's interaction graphs as they allow the modeling of the spread of failures from one community to the next using the overlap or bridge nodes. We specifically adopt the overlapped community detection algorithm presented in [126] (with minor modifications discussed in [130]), which considers weights but not the direction of interactions. It uses the belonging degree and conductance to locally optimize the conductance-based utility function for communities by evaluating a new community formed with an addition of a new neighboring node. We refer to this community detection algorithm as the *conductance-based community detection* in this dissertation. A detailed review of this approach is presented in [126, 130]. In order to utilize the weights as well as the direction of interactions in our analyses, we adopt the *Infomap community detection* approach [131]. In Infomap, we exploit both disjoint and overlapped community detection algorithms, using variations of its random walk algorithm, to evaluate and compare the role of disjoint and overlapped community structures in the cascade processes of power grids. For more details on this algorithm, refer to [131].

In overlapped communities of directed as well as undirected interaction graphs, *common/ overlap* nodes (nodes that belong to multiple communities) contribute in the spread of failures during cascades. In disjoint communities of undirected interaction graphs, *bridge* nodes (nodes that have links to components in other communities) contribute in the spread of failures during cascades. In the case of directed interaction graphs with disjoint communities, there are two types of bridge nodes: *o-bridge* nodes that have outgoing links towards nodes of other communities and *i-bridge* nodes that have incoming links from nodes of other communities. While, incoming and outgoing properties are both important factors in determining spread of failures, we consider the former as the initiator of failure propagation during cascades. Note, a node may have both incoming and outgoing links, however, the spread of failures to other communities occur through the outgoing links only.

Next, we briefly review the conductance-based community detection and the Infomap community detection algorithms. Note that the two approaches used in this dissertation for community detection are two examples of community detection algorithms and evaluating the performance of Infomap or conductance-based methods in identifying communities is not the focus of this dissertation. Instead, we use these methods to show that various community structures can lead to each component having a distinct role in the propagation characteristics of cascades.

3.2.2.1 Conductance-based Community Detection

We adopt the weighted, overlapped community detection algorithm presented in [126] (with minor modifications) to characterize the overlapped community structure in the graph of interactions. Since, communities are overlapped, for some of the i and j pairs of communities, $C_i \cap C_j \neq \emptyset$.

The conductance-based community detection uses the *belonging degree* of a node to a community and the overall *conductance* of a community to find community structures in graphs.

1. Belonging Degree: The set of neighbors and the degree for node i is denoted by N_i and D_i, respectively. The degree of node i is defined as $D_i = \sum_{j \in N_i} w_{ij}$, where w_{ij} is the weight of the link from node i to node j. Thus, the belonging degree of node i to a community C_k is defined as

$$B(i, C_k) = \frac{\sum_{j \in C_k} w_{ij}}{D_i} \tag{3.1}$$

2. Conductance: The conductance ϕ_{C_k} of a community $C_k \in \boldsymbol{\mathcal{C}}$ is defined as

$$\phi_{C_k} = \frac{cut(C_k, G \backslash C_k)}{w_{C_k}}, \tag{3.2}$$

where $cut(C_k, G \backslash C_k)$ denotes the sum of weights of edges adjacent to the nodes in the community except edges inside the community itself and w_{C_k} denotes the sum of the weights of all edges connected to nodes in the community. Smaller values of conductance for communities are preferred.

3. Conductance-based Community Detection Algorithm: The community detection algorithm presented in the study in [126] is shown in Algorithm 1 with two minor modifications: (1) to update the conductance as in line 8 and (2) to remove the analyzed neighbors from the neighboring set as in line 9 and 11. Thus, all neighbors of an initially identified community are either appended or skipped. In this algorithm, the edges within community C are denoted by E_c. While, the appending process of a node to a community is different from the algorithm presented in [126], the identified communities are in agreement with the concept of tight interactions within communities and weak interactions outside communities.

Algorithm 1 Community detection algorithm based on [126]

Input: Graph $G = (V, E)$
Output: Overlapped Communities $\mathcal{C}$
Initialize: $\boldsymbol{\mathcal{C}} = \emptyset$
while $E \neq \emptyset$ **do**
  $C = \{i, j\}$ where $e(i, j) = \underset{(u,v)\in E}{\operatorname{argmax}} w_{ij}$
  while $N_c \neq \emptyset$ **do**
    $C' = C \cup \underset{w\in N_c}{\operatorname{argmax}} B(w, c)$
    if $\Phi(C') < \Phi(C)$ **then**
      $C = C'$
      $\Phi(C) = \Phi(C')$
      $N_c = N_c \backslash C$
    else
      $N_c = N_c \backslash C$
    end if
  end while
  $E = E \backslash E_c$
  $\boldsymbol{\mathcal{C}} = \mathcal{C} \cup C$
end while

3.2.2.2 Infomap Community Detection

The Infomap community detection algorithm [124] is an information-theoretic algorithm, where the problem of finding the community structure of a network is a problem of optimally compressing the information accumulated by a random walk on the network. Each node or community in the network is represented by a code-word such that the random walk, represented by a stream of codewords is of shortest possible length, known as the minimum description length. A simple method of assigning codewords to nodes is by using Huffman codes, where the frequency of visit by the random walk is used to assign the codewords.

The frequency of visits depends on the structural regularities of the network. Specifically, for a network having a community structure, a random walk will most likely visit the node inside the

same community in the next step, i.e., the propagation rate of the random walk inside the community is higher compared to outside the community and a random walk entering a community persists in the community for a longer time. This regularity in the network is used to solve the compression problem by reusing codewords, where each community is represented by a unique codeword but codewords representing individual nodes inside the communities may be reused. However, assigning codewords to nodes and communities is not the main goal and the actual codewords are not devised. In fact, only the theoretical limit on the shortest possible description length to represent the nodes and communities in the network is required and, it can be found by Shannon's source coding theorem. Further, using Shannon's theorem in conjunction with the structural regularity of the network, the method finds the $\boldsymbol{K}$ partitions (communities) in the network with efficient membership of nodes to each partition.

Mathematically, the average description length of a single step taken by the random walk is given by the map equation as follows

$$\mathcal{L}(\boldsymbol{\mathcal{C}}) = q_{\curvearrowright}\mathcal{H}(\mathcal{Q}) + \sum_{i=1}^{\boldsymbol{K}} p_{\circlearrowright}^{i}\mathcal{H}(\mathcal{P}^{i}), \tag{3.3}$$

where $\boldsymbol{\mathcal{C}}$ is the set of disjoint communities in a network (we will discuss overlapped communities later in this section) and $\mathcal{L}(\boldsymbol{\mathcal{C}})$ is the lower bound on the code length of a single step of the random walk through the network. Equation (3.3) consists of two parts: the cost of movement between communities, and the cost of movement inside communities. These movements are represented by $\boldsymbol{K}+1$ codebooks: one index codebook and $\boldsymbol{K}$ module codebooks. The index codebook consists of the $\boldsymbol{K}$ unique codewords given to $\boldsymbol{K}$ communities while the module codebook for community C_i consists of the codewords given to nodes inside that community including one code to exit the

community. Specifically, in Equation (3.3), $\mathcal{H}(\mathcal{Q})$ is the entropy of the index codes given to the communities weighted by $q_{\curvearrowright}$, which is the rate of use of the index codebook and $\mathcal{Q}$ is the normalized probability distribution of $q_{\curvearrowright}$. If $q_{i\curvearrowright}$ is the probability of the random walk exiting the community i then, $q_{\curvearrowright} = \sum_{i=1}^{K} q_{i\curvearrowright}$. Next, $\mathcal{H}(\mathcal{P}^i)$ is the entropy of the module codenames given to nodes within a community C_i, including the entropy to exit the community and $\mathcal{P}^i$ is the normalized probability distribution. It is weighted by $p^i_{\circlearrowright}$, which is the fraction of within community movements that occur in community C_i including the probability of exiting the community, i.e. $\sum_{i=1}^{K} p^i_{\circlearrowright} = 1 + q_{\curvearrowright}$. The map equation is discussed in detail in [131].

As the Infomap community detection algorithm is based on random walk, it can be applied to directed, undirected, weighted and unweighted graphs. It also has variations that allow identification of overlapping communities. In this case, instead of assigning only one codeword to a node, multiple codewords can be assigned by considering the origin of the random walk. Solving the map equation for overlapping modules gives the number of communities, the membership of nodes to communities and also the degree by which a node belongs to a module by looking at how often a random walk switches modules [132].

3.2.3 Community Structure Analysis and Results

In this section, we will present the analysis and results of the community structures identified in the influence-based and correlation-based interaction graphs of IEEE 118 test bus system constructed in Chapter 2 using the cascade dataset generated in Chapter 1. Particularly, we will use the conductance-based community detection and the Infomap community detection to identify the overlapped as well as disjoint community structures in the interaction graphs. We will show that depending on the type of interaction graph, the algorithm used in identifying communities, and

the operating settings of the power grid used for generating the cascade dataset, the structures and patterns identified in the graphs will differ and lead to different conclusions about the criticality of components.

3.2.3.1 Community Structures in IEEE 118-bus System

We apply the Infomap disjoint and Infomap overlap community detection as well as the conductance-based community detection to the influence-based (H) and correlation-based (CR) graphs for the IEEE 118-bus system. We briefly discussed the thresholded interaction graphs in Section 2.4.1 of Chapter 1. Removal of the small weighted interactions from the interaction graphs have dual purpose. In addition to showing only the strong interactions between the components, removal of small weighted links ensures that community detection algorithms can detect properties that bear the macroscopic and microscopic features relative to the scale of the interaction graphs.

If community detection algorithms, discussed in Section 3.2.2, are applied to interaction graphs directly, they will not be able to reveal the dominant structures in the graph due to links with small weights acting as noise [117]. Based on our observations, these community detection algorithms result in very few communities (one or two communities for interaction graphs H and CR) with majority of the components belonging to one community. As such, we use a threshold for the weight of the links, such that the resulted communities are not too small or too large in number or size; such that the resulting communities can bear properties between microscopic and macroscopic properties relative to the scale of the graph [118].

We have tested various thresholds and found that the best thresholds based on the quality of the identified communities are 0.6 for graph H and 0.7 for graph CR. To make a comparable case

Table 3.1: Properties of the thresholded graphs of H and CR

Graph	No. of vertices in LCC	No. of edges	No. of communities: Conductance-based	No. of communities: Infomap Disjoint	No. of communities: Infomap Overlap
$H \geq 0.6$	143	1160	45	12	12
$H \geq 0.7$	57	612	12	4	7
$CR \geq 0.7$	59	636	12	5	5

for the graph CR with threshold 0.7, we also evaluate graph H with threshold 0.7, which provides similar number of components in the largest connected component (LCC) and similar number of communities as presented in Table 3.1. As the thresholds can result in disconnected graphs; we only use the LCC of the interaction graphs in our analyses similar to the work in [118].

For the influence graph H with threshold 0.6, we obtain 12, 12 and 45 communities from Infomap disjoint, Infomap overlap and conductance-based community detection, respectively. We have depicted the structures of these communities in Figure 3.2 over the *line graph* of IEEE 118-bus system shown in Figure 2.2 of Chapter 2. The colors applied to the nodes are classifying the transmission lines into different communities each represented by a color. If a component belongs to multiple communities, as in the case of overlapped communities, multiple colors are assigned to the node. In the case of disjoint communities, the bridge nodes are assigned multiple colors depending on the communities that they bridge. The small blue nodes represent the transmission lines that have weak interactions (below the defined threshold) and consequently, did not participate in the community detection process. We observed that depending on the interaction graph derived from the cascade dataset and the community detection algorithm applied to the thresholded interaction graphs, the identified community structures also differ.

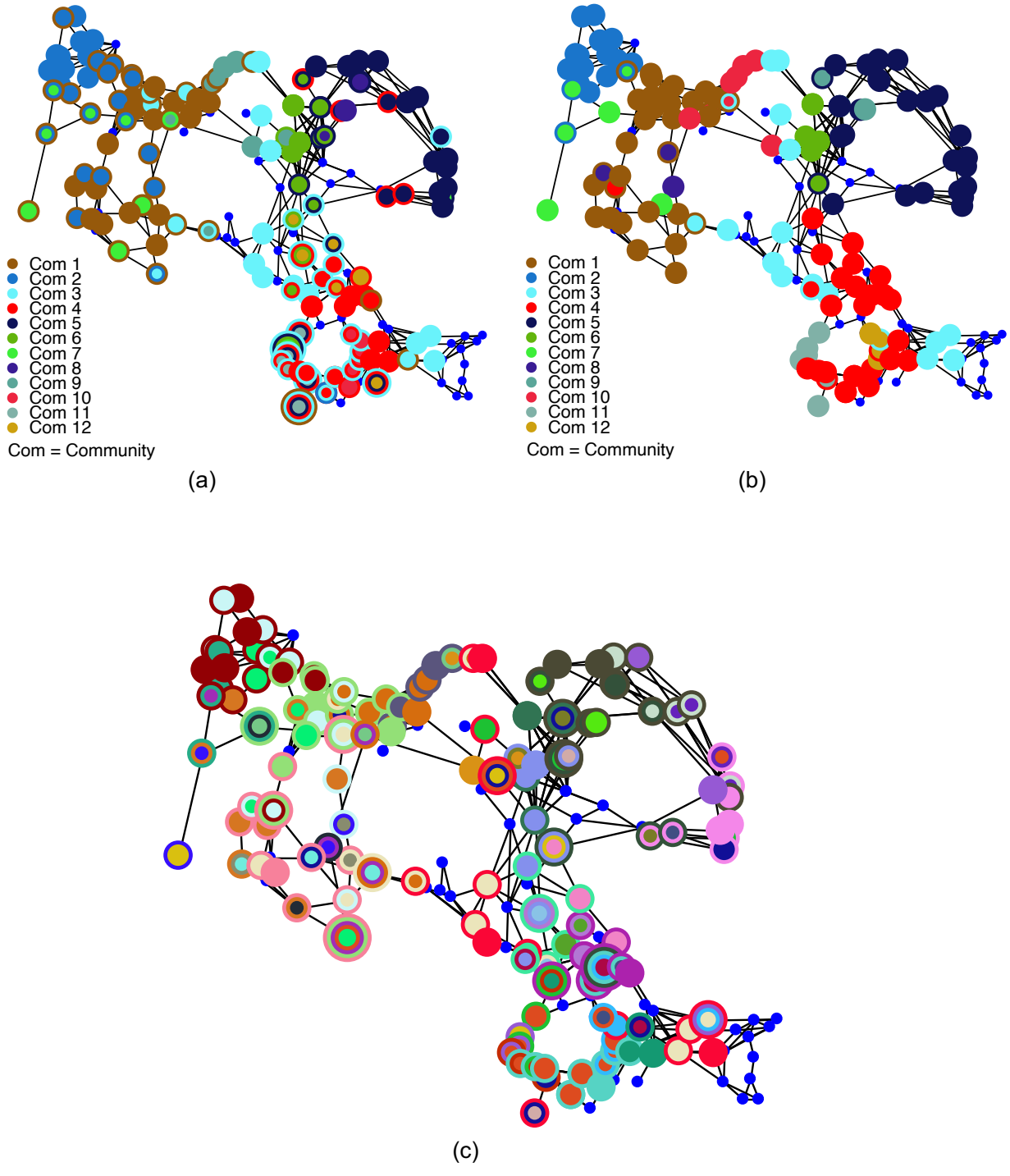


Figure 3.2: Community structure of graph H with threshold 0.6 over the line graph of IEEE 118-bus system based on (a) Infomap disjoint, (b) Infomap overlap, and (c) Conductance-based community detection algorithms.

3.2.3.2 Community Structures' Statistics in IEEE 118-bus System

In our first study, to analyze the community structures of interaction graphs, we use the conductance-based community detection algorithm discussed in Section 3.2.2.1 for identifying overlapped community structures. We applied this algorithm over the LCC of the influence-based and correlation-based interaction graphs H and CR after applying threshold 0.7 and found 12 communities in each, respectively. However, the structure of communities and their OR are quite different. Our analysis show that some communities do not overlap suggesting that failures inside such communities rarely propagate to other communities. Table 3.2 shows the OR for the communities identified in H with threshold 0.7. Table 3.3 shows the number of nodes in each community for the three community detection algorithms. Since, the conductance-based algorithm yields 45 communities, for simplicity we show only the 12 highly populated communities in Table 3.3. To provide a more detailed view on the community structures, in Table 3.4, we show the overlap/bridge size between pairs of communities C_i and C_j, denoted by $|O(i,j)|$. The size of bridges for the disjoint communities is counted as the total number of nodes that act as bridges between two communities C_i and C_j. For overlapped communities, the size of overlaps is the number of nodes that fall in the overlap region (i.e., belong to both communities). The number of communities with n overlap/bridge components is denoted by $U_n = \sum_{C_i,C_j \in \mathcal{C}} \delta(|O(i,j)| - n)$, where δ is the delta function, such that $\delta(i)$ equals to one if $i = 0$ and zero otherwise. In Table 3.4, for Infomap overlap, there are no pairs of communities that have overlap size greater than 3, hence U_4 and above is zero. The last column of Table 3.4 shows the pairs of communities that have overlap/bridge size greater than 4. Tables 3.3 and 3.4 show that the shape and size of communities as well as the size of overlap/bridge between communities differ depending on the applied algorithm.

Table 3.2: Overlap Ratio *(OR)* of communities in H with threshold 0.7

Community	3	4	6	7	8	9	10	11	12
1	-	0.067	-	0.053	0.389	-	-	-	-
2	0.1	-	0.08	0.105	0.13	0.187	-	0.058	0.133
3	-	-	0.105	0.154	0.111	-	-	0.091	-
5	-	-	0.227	-	-	-	0.2	-	-
6	-	-	-	0.053	-	0.058	0.062	0.062	-
7	-	-	-	-	0.187	-	-	-	-
8	-	-	-	-	-	0.062	-	-	-
9	-	-	-	-	-	-	-	0.125	-

Table 3.3: Community sizes for graph H with threshold 0.6

Community	No. of nodes (Conductance-based)	No. of nodes (Infomap Disjoint)	No. of nodes (Infomap Overlap)
1	19	31	32
2	19	17	17
3	18	16	18
4	15	21	29
5	14	26	28
6	14	7	7
7	14	3	6
8	14	2	3
9	13	6	3
10	13	5	8
11	10	6	6
12	10	3	5

Table 3.4: Community overlap sizes for graph H with threshold 0.6

Algorithm	U_0	U_1	U_2	U_3	U_4	$\|O(i,j)\| > 4$
Infomap Disjoint	97	11	5	2	2	O(3,6)=5, O(3,10)=5, O(4,10)=5, O(4,12)=5, O(5,11)=5, O(3,12)=6, O(4,11)=6, O(3,11)=7, O(1,7)=8, O(5,6)=8, O(1,3)=9, O(3,5)=9, O(4,5)=10, O(1,2)=18, O(3,4)=20
Infomap Overlap	118	9	4	1	0	0
Conductance-based	1789	139	25	15	7	O(3,28)=5, O(5,8)=5, O(10,30)=6, O(1,27)=7, O(7,15)=7

3.3 Summary

In this chapter, we studied the role of community structures in the power grid's interaction graphs during cascading failures. To characterize the community structures in power grid's graph of interactions, we used both disjoint and overlapped community detection techniques.

We observed that applying different community detection techniques led to variation in the identified structures and each structure could reveal various aspects of cascade processes. We used the motivation that propagations are more likely to stay within communities to identify the overlapped and bridge components, that can contribute in propagation of failures from one community to another and consequently, contributing in larger cascades. We postulate that protecting the critical components such as bridge and overlap nodes can help in reducing the risk of large cascades and blackouts. In Chapter 4, we will specifically show that protection of critical components identified using community structures will reveal key players of the cascade process that are not identified using traditional metrics.

Chapter 4: Reliability Analysis of Power Grids Using Community Structures

Occurrence of cascading failures in power grids has motivated power system researchers to use graph-based cascading failure models for various reliability studies. In this dissertation, our reliability analysis studies on power grids focuses on identifying critical components of the cascade processes. Identification of critical components have also been undertaken in other areas like epidemics, viral marketing, and disease spread [40, 133]. Particularly, in power grids, identifying critical components in the cascade process has been conducted using different approaches including power system models and simulations [51, 134–136], graph-based analysis of physical structures and topology of power systems [89, 137], and interaction graphs [43, 77, 78, 138, 139]. For instance, in [135] and [136], the authors focused on synchronization, *phase-locked* states and flow exchanges that ensured stable operation of the power grid. Power flow re-routing caused by outages can de-synchronize the grid. Thus, critical lines in the grid were predicted by identifying lines whose removal caused de-synchronization in the power system. In the latter work, the authors also considered the secondary outages caused by the transient dynamics.

The closest to the study and analysis presented in this chapter are those, which use power grid's network properties, standard centrality measures, or define new centrality measures to rank the components based on their importance in the cascade process [107, 137, 140, 141]. For instance,

Portions of this chapter were published in IEEE PES [3] and IEEE TNSE [4]. Copyright permissions from the publishers are included in Appendix A. Portions of this chapter are also available as preprint in arXiv [45].

the work in [140] uses the coreness of the components in the power grid's network (where coreness is defined by the largest integer c for which a node belongs to a c-core–the largest sub-network with all nodes having at least degree c) to identify the critical set of components. As another example, in [141], a centrality measure is introduced based on the electrical betweenness and eigenvector centrality. Combination of these two measures yield the electrical centrality, which accounted for the electrical as well as the topological properties of a power grid. Moreover, in [77, 78], the authors identify tree partitions in the power grid's graph of interactions and show that transmission line failures cannot propagate across common areas of tree partitions. Although the latter use different type of structures and evaluation mechanism for the identified critical components, it is the closest work in concept to the work presented in this chapter. Since our work is focused on using interaction graphs for criticality analysis, we will also provide a literature review on the reliability analysis performed using interaction graphs of power grids.

Next, we will exploit the community structure of the graph of interactions to introduce a new centrality measure, named *community-based centrality* to specify the criticality of the components (specifically, transmission lines in the power grid) in the cascade process. The main idea behind this new measure is that communities formed in the interaction graph reveal group of components that are likely to contain failures within themselves during the cascade (due to tight influences/interactions); however, if a component belongs to multiple communities (in the case of overlapped communities) or bridges two communities (in the case of disjoint communities) then it will be critical in the cascade process in the sense that it serves as a gateway to spread failures from one community to another. As the scale of the communities are smaller than that of the whole network, identifying and protecting such components can help in containing failures within a com-

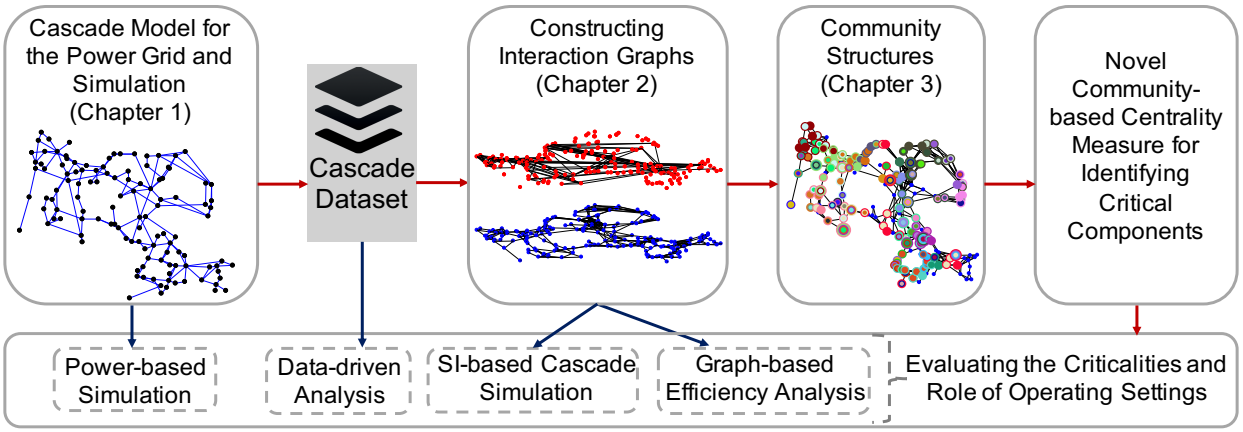


Figure 4.1: Key components for identifying critical components of power grids during cascade processes using community structures in interaction graphs (shown with red arrows) and evaluating the role of critical components (shown with blue arrows).

munity and reduce the risk of large blackouts. Figure 4.1 shows the key components of identifying critical components with relation to the previous chapters.

To verify the criticality of the identified components based on community-based centrality, we will use data-driven, SI (Susceptible and Infectious) epidemic [142] simulation-based, power system simulation-based, and graph-based approaches. We will also compare the role of identified critical components with those identified using standard centrality measures including betweenness, closeness, eigenvector, and degree centralities as well as the influence-based method presented in [43]. We postulate that our results will confirm that critical components identified based on the community structures will prove to be key players in cascading failures in power grids compared to those identified with other centrality measures. We will specifically show that protecting the critical components identified using community-based centrality can help in reducing the risk of large cascades while, the performance in improving reliability also depends on the community detection algorithm and the type of interaction graph.

4.1 Related Prior Work

Interaction graphs discussed in Chapter 2 can be used for various analysis; specifically related to the reliability of the power grid; including analyzing role of components and finding critical ones that contribute heavily in a cascade process, predicting distribution of cascades sizes, and studying patterns and structures that reveal connections and properties of the components in the power grid that extend beyond physical topology-based graphs. Thus, we divide the reliability studies performed using the interaction graphs into various categories and discuss them below.

4.1.1 Critical Component Analysis

We classify studies that identify and analyze role of critical components in power grid's reliability into three broad categories that include 1) using pre-existing as well as novel measures to find critical buses/transmission lines, 2) evaluating attack strategies that cause significant damage in the power grid, and 3) employing mitigation measures such as upgrading transmission lines or adding new components to protect the identified critical components.

4.1.1.1 Critical Component Identification

This class of reliability analyses focuses on finding critical buses/transmission lines by analyzing structural properties of interaction graphs using standard centrality measures such as degree, betweenness etc. (for a review of standard centrality measures refer to [108]) or by defining novel interaction graph based metrics.

Critical Component Identification using Standard Centrality Measures: In the studies presented in [51–53, 55–57], *fault chain*-based interaction graphs are found to be scale free graphs,

indicating that most nodes possess low degrees but a limited number of nodes possess high in and out degrees. Thus, in the *fault chain*-based interaction graphs, vertices with higher degrees are assumed to be the critical components of the system. Similar conclusions are obtained by authors in the studies in [84, 88], where the inverse admittance interaction graph is observed to be scale-free and consisting of limited number of nodes with high degrees, which are considered as the critical components of the system. These are examples of works that consider the degree centrality measure to identify critical components of the system. Other centrality measures such as betweenness, eigenvector, and PageRank have also been considered on interaction graph-based representations of power grids including [79, 81, 95] to find critical components of the system.

Critical Component Identification using New Centrality Measures: In addition to studies that rely on standard centrality measures; some works develop new centrality measures in the context of power grids and the developed interaction graphs to analyze criticality of the components. For instance, in the generation-based interaction graphs [58–61, 63], out-strength measure, which is the sum of the weights of the interaction links originating from a node, is used to find critical transmission lines. Such lines are the ones whose failure at any stage of the cascade including the initial stage or propagation stage induces failure in significant number of other transmission lines. Outages in the initial stages are caused by external factors such as bad weather conditions, improper vegetation management, and exogenous events, whereas outages in the propagation stage is caused due to power flow re-distributions, hidden failures, and other interactions between components. Influence-based [44, 64] and multiple and simultaneous failure [65] interaction graphs are also used to find critical transmission lines but they explicitly focus on lines whose failure during the propagation stage of cascading failures cause large cascades. Particularly, the studies in [44, 64] use a cascade

probability vector derived using the influence-based interaction graph to quantify the probability of failure of lines during propagation stage of cascades and defines critical lines as the ones whose corresponding entries in the probability vector have higher values. Similarly, the study in [65] finds the probability distribution of states of the multiple and simultaneous failures interaction graph and defines critical lines as the ones that belong to states with higher probability of occurrence. Influence-based and correlation-based interaction graphs constructed in the studies in [3, 4] are also used to find the critical transmission lines during cascade processes by using a *community-centrality* measure. As the name suggests, the measure quantifies the criticality of transmission lines based on their community membership, where critical lines are the ones that belong to multiple communities or act as bridges between communities. Note that communities are defined as groups of vertices with strong connections among themselves and few connections outside(for definition of communities and a review of community detection methods on graphs refer to [127]).

Identification of critical lines is not limited to data-driven interaction graphs. Multiple studies use electric distance-based interaction graphs for such analysis as well. Effective resistance between components in the effective resistance-based interaction graph can be summed for all node pairs in the graph to find the *effective graph resistance* metric of the power grid. Effective graph resistance metric was initially defined in the study in [115] as *Kirchhoff index* and used in the study in [143] as a robustness metric. Lower values of this metric suggests that the power grid is robust to cascading failures. Effective graph resistance can also be found using the eigenvalues and eigenvectors of the Laplacian matrix of the grid [144]. In the study in [91], critical transmission lines are found by measuring the changes in *effective graph resistance* before and after the removal of the line. In a similar manner, impedance-based interaction graphs constructed in the study in

[90] and [87] are also used to find critical transmission lines by measuring the changes in *net-ability* metric before and after the removal of a line. Net-ability reflects the performance of a network by quantifying the ability of a generator to transfer power to a load within the power flow limits.

4.1.1.2 Studying the Effect of Line Upgrades and Line Additions

While identification of critical components in the power grid is necessary, assessing the impact of modifications and protection of such critical components in the overall power grid is the next step in the study. In the studies presented in [44] and [64], an influence interaction graph-based metric is used to quantify the impact of upgrading the critical lines (for example, by improving vegetation management around the lines or by improving protection systems) on cascade propagation. The work in [65] uses the multiple and simultaneous failure interaction graph to do a similar study. In both interaction graphs, the authors conclude that upgrading lines that take part in propagation of cascades reduces the risk of large cascades compared to upgrade of lines that initiate cascades. While the studies in [44, 64] investigate the performance of the power grid networks after line upgrades, the studies in [92] and [93] uses *effective graph resistance* metric to study the impact of adding transmission lines in optimal locations of the power grid. However, in [93] the authors warn that placing an additional line between a pair of nodes does not necessarily imply increased robustness of the grid. Infact, grid robustness may decrease after adding additional lines (due to Braess's paradox [145]), if the additions are done haphazardly.

4.1.1.3 Analyzing Response to Attack/Failure Scenarios

In addition to identifying critical components and characterizing the impact of their modifications in the reliability of power grids, the study of the response of power grids to attacks and

failures is also necessary. Such studies can be used to find critical components and attack strategies that threaten the reliability of the overall power grid. In the studies in [72–74], node integrated risk graphs are used to find groups of transmission lines whose removal from the graph causes the largest drop in *net-ability* of the power grid, as discussed in Section 2.2.2. These groups can be found in real-time independent of system parameters. The study in [93] also analyzes the robustness of power grids to deliberate attacks using the *effective graph resistance* metric, as discussed in Section 4.1.1.1.

4.2 Formulation of Community-based Centrality Measure

While standard centrality measures, such as closeness, betweenness, eigenvector, and degree centralities [146], can reveal various aspects of criticality of components based on topological properties, new approaches are needed to identify key players in processes on networks such as the cascade processes. As discussed in Chapter 3, it has been shown that community structures in networks play a key role in the cascade processes on the network. In this section, we introduce a novel *community-based centrality* measure to identify critical components in the cascade process of power grids based on the idea of trapping failures in communities.

The main idea behind the community-based centrality is that overlapped nodes (in the case of overlapped communities) or bridge nodes (in the case of disjoint communities) are critical in the cascade process as they heavily contribute in the spread of failures from one community to another. However, this condition is not enough in defining the centrality for various cases; for example, in the case of the following scenarios:

1. (S1) a bridge or overlapped node of multiple communities, where the communities themselves do not have a central role in the propagation of failures;

2. (S2) multiple nodes have the same properties as far as community membership (e.g., they belong to the same communities); however, their centrality is different due to the microscopic connectivity properties of the nodes.

To capture these factors in the novel community-based centrality measure, we proceed as follows. In the first step, we address scenario (S1) by defining a community-based centrality measure using the disjoint and overlapped community structures identified by the conductance-based and infomap community detection algorithms over the influence-based and correlation-based interaction graphs. Then, in the second step, we further develop and improve the community-based centrality measure defined in the first step, such that both scenarios (S1) and (S2) can be addressed. Using this weighted community-based centrality measure and the community structures identified by the Infomap overlap, Infomap disjoint, and conductance-based community detection algorithms over the influence-based and correlation-based interaction graphs, we identify key components in the cascade process of power grids.

To define the community-based centrality measure, we first create an augmented graph of communities to define the centrality of communities in the network and then sum up the centrality of communities for the nodes they belong to (i.e., nodes get their centrality from the centrality of all the communities they belong to). In the augmented graph, the communities are represented by nodes and the strength/weights between the community-nodes are defined using the overlap/bridge nodes shared between the communities as well as the strength of interactions between the overlap/bridge nodes. An example of an augmented graph of the community structures identified using the Infomap

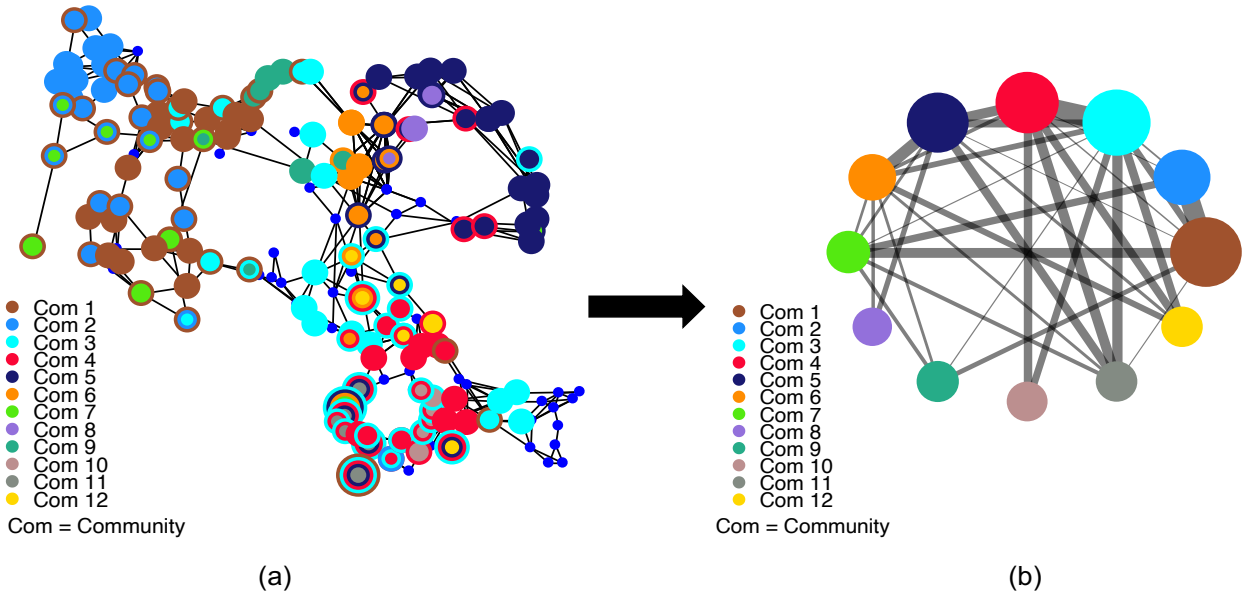


Figure 4.2: Example of converting community structures of interaction graphs to augmented graphs of community structures. (a) Infomap disjoint-based community structure of graph H with threshold 0.6 over the line graph of IEEE 118-bus system. (b) Augmented graph of Infomap disjoint-based community structures shown in (a).

disjoint algorithm for the influence-based interaction graph with threshold 0.6 is shown in Figure 4.2. Next, to define the weighted community-based centrality measure, we adjust the community-based centrality measure using the local interaction weights of the nodes. We will discuss both these measures in detail.

4.2.1 Community-based Centrality Measure

We form the augmented graph of the communities denoted by $\mathcal{AG} = (N_{AG}, E_{AG})$, where N_{AG} is the set of communities (i.e., $N_{AG} = \mathcal{C}$) and E_{AG} is the set of links among communities defined as following. For networks with overlapped communities, E_{AG} represents connections due to common (overlapped) nodes among the communities (i.e., $e_{rs} \in E_{AG}$ if there exists a node $n_i \in N_{IG}$ such that C_r and C_s both belong to $CL(n_i)$). For network structures with disjoint communities, E_{AG}

represents connections due to bridging links among the communities (i.e., $e_{rs} \in E_{AG}$ if there exists a link $e_{ij} \in E_{IG}$ such that $C_r \in CL(n_i)$ and $C_s \in CL(n_j)$, where n_i and n_j are the two end nodes of line e_{ij} in the interaction graph $\mathcal{IG}$). Next, to capture the strength of interactions among communities, we consider their closeness centrality. Closeness centrality over the relatively small graph of N_{AG} can show how fast failures from a community can get to another community. To evaluate closeness centrality of communities, we assign weights (representing distance between communities) to the links E_{AG} in the $\mathcal{AG}$. Specifically, distance among the communities is defined to depend on the overlap/bridge size between pairs of communities. We define $R_{\{C_i,C_j\}} = |O(i,j)|/|C_i \cup C_j|$, where $|C_i \cup C_j|$ denotes the total number of unique components in communities C_i and C_j. An effective distance inversely proportional to $R_{\{C_i,C_j\}}$ is assigned to the links among the nodes of the $\mathcal{AG}$, i.e. communities with larger overlap/bridge size will be assigned smaller link weights suggesting smaller distance between the communities. The rest of the communities with no overlaps/bridges are assigned equal larger weights for their links than any overlapped/disjoint communities (i.e., suggesting a larger distance between the communities). The weighted $\mathcal{AG}$ is then used to find closeness centrality for each community. Using the identified centrality values for the communities, we define initial measure of importance for node $n_i \in \mathcal{IG}$ as

$$I_{n_i} = \sum_{k \text{ such that } n_i \in C_k} M_k, \tag{4.1}$$

where M_k is the closeness centrality of community k. This initial measure of importance for nodes in the interaction graphs addresses scenario (S1). Next, we discuss the critical components identified using the community-based centrality measure and the verification of the critical components using a data-based analysis.

4.2.1.1 Critical Components Identification

We used the conductance-based community detection algorithm over the influence-based and correlation-based interaction graphs with various thresholds but similar Largest Connected Component (LCC) size (Table 4.1). Note that the size of LCC for H and CR with threshold 0.7 is 57 and 59 respectively. Similarly, the size of LCC is similar for H and CR with threshold 0.6 and 0.4 respectively. Then, using the community-based centrality measure, we identified ten most influential nodes in the cascade process of power grids. We observe that overlap components that belong to multiple communities are in higher ranks. For instance, nodes 32 and 151 in H with threshold 0.6 belong to five different communities, which contributed to their importance.

We also derived the influence-based H and correlation-based CR interaction graphs for various operating settings, e.g. power grid loading levels, r (the ratio of the total demand over total generation capacity of the system) to simulate cascades under different settings and evaluate the role of operating characteristics in cascade analyses. Note, for each analysis, we have simulated at least 20,000 unique cascading failure scenarios. Then, we applied the conductance-based community detection algorithm to the H and CR interaction graphs with various thresholds but comparable LCC sizes. We observed that similar to the different techniques that lead to different graph of interactions on the cascade process, different loading levels r for the power grid also result in different graphs of interactions and ranking of critical components. Table 4.2, shows the top 5 critical components identified for the system for different r values. These results also show that depending on the condition and the operating settings of the power grid, the critical components of the system may vary. Therefore, it is important to perform criticality study with considerations about the power grid's conditions.

Table 4.1: Critical components based on community-based centrality (I)

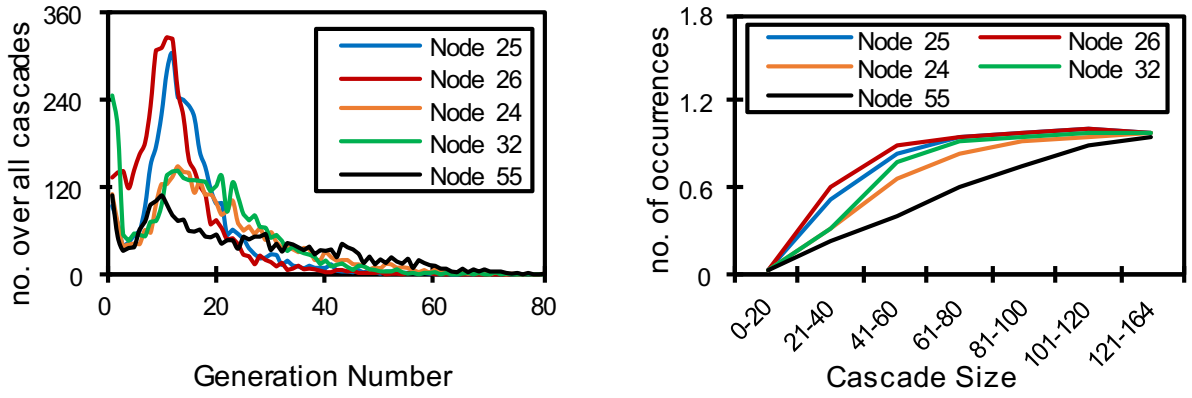
Rank	$H \geq 0.6$	$H \geq 0.7$	$CR \geq 0.4$	$CR \geq 0.7$
1	32	25	138	44
2	151	26	119	26
3	96	24	82	25
4	167	32	73	43
5	76	55	64	27

Table 4.2: Critical components based on community-based centrality (I) for different power grid loading levels

	H (Threshold = 0.3)				*CR (Threshold = 0.3)*			
Rank	r=0.6	r=0.7	r=0.8	r=0.9	r=0.6	r=0.7	r=0.8	r=0.9
1	9	62	45	8	44	68	26	20
2	58	44	51	9	36	71	36	36
3	35	16	52	45	21	69	38	54
4	27	181	9	58	51	75	25	31
5	61	7	102	48	50	50	39	33

4.2.1.2 Verification of Criticality

In order to verify that the identified critical components based on the community-based centrality measure reveal the actual influential components in the cascade process, we have done a set of analyses based on our cascade dataset. Particularly, we have focused on the 5 most critical components identified for H with a threshold of 0.7 (i.e., nodes 24, 25, 26, 32 and 55 according to Table 4.1). Figure 4.3 (a) shows that these five nodes appear in the early generations of the cascades, which shows their contribution in the progress of cascade. Moreover, the results in Figure 4.3 (b) show that among all cascade sizes observed in the dataset, these nodes tend to be a part of larger cascades. In other words, the occurrence of these nodes in cascades increase with the size of cascade. These results confirm the criticality of the identified nodes using the new community-based centrality measure.



(a) (b)

Figure 4.3: Data-driven verification of community-based critical components I. (a) Number of times critical components failed in various generations and (b) number of occurrences of the critical components in different cascade sizes for $H \geq 0.7$.

4.2.2 Weighted Community-based Centrality Measure

While, the community-based centrality measure ranks the nodes such that the important nodes are the ones that help in spreading the failures to more central communities (addressing Scenario (S1) mentioned above), it cannot address Scenario (S2) as it does not differentiate among the nodes with the same community membership. To capture the microscopic properties of the nodes to differentiate local centralities, we define a weighted and scaled version of I_{n_i} for node $n_i \in C_r$ as following

$$CM_{n_i} = I_{n_i} * \sum_{e_{ij} \in E_{IG} \text{ s.t. } n_j \in C_s, r \neq s} w_{n_i n_j}, \tag{4.2}$$

where e_{ij} is the link connecting node n_i to n_j and $w_{n_i n_j}$: *(1) for the case of disjoint communities of directed graphs:* is the weight of the outgoing link from o-bridge node n_i to i-bridge node n_j, *(2) for the case of disjoint communities of undirected graphs:* is the weight of the link connecting bridge node n_i to bridge node n_j and *(3) for the case of overlapped communities:* weight of the link connecting overlap node n_i to overlap node n_j.

Table 4.3: Critical Components based on weighted community-based centrality (CM) for Conductance-based, Infomap Disjoint and Infomap Overlap community detection algorithms.

Rank	Conductance-based			Infomap Disjoint			Infomap Overlap		
	$H \geq 0.6$	$H \geq 0.7$	$CR \geq 0.7$	$H \geq 0.6$	$H \geq 0.7$	$CR \geq 0.7$	$H \geq 0.6$	$H \geq 0.7$	$CR \geq 0.7$
1	30	43	43	30	30	39	30	30	20
2	96	30	31	32	25	25	40	17	30
3	32	168	25	76	32	32	186	35	44
4	148	39	27	70	18	21	75	13	8
5	167	25	30	179	29	20	73	20	-

Table 4.3 shows the top 5 critical components identified using the community-based centrality measure, i.e., CM. These identified components belong to either the bridge or overlap region in community structures. In the case of Infomap overlap, for graph CR with threshold 0.7, there are only four components in the overlap region thus, only four components are shown in Table 4.3.

4.3 Evaluation of Weighted Community-based Centrality Measure

In order to show that community structures in the graph of interactions play central roles in cascade processes and to verify the importance of the critical transmission lines identified using the community-based centrality measure CM, we use four methods: (1) data-driven analysis, (2) SI-based simulation, (3) power system simulation and (4) graph-based efficiency analysis. We compare the centrality of the identified components based on the community-based centrality with those found using the standard centrality measures including betweenness, closeness, eigenvector and degree centralities, as well as the influence-based criticality measure discussed in [43]. Specifically, the influence-based criticality measure ranks the components based on their probability of failure at any stage of the cascade, which can be calculated for the influence-based interaction graph [43]. While each of these measures reveal various aspects of criticality, we will show that the community-based centrality can reveal aspects about cascade process that were not captured by other measures.

4.3.1 Data Driven Analysis of Centrality

In this section, we use two statistical properties of the identified central components using the cascade datasets to verify their importance.

4.3.1.1 Critical Components' Occurrence in Critical Cascade Stages

In general, cascading failures in power grids have different stages including: (1) *precursor stage*, where the trigger events and initial failures occur in the system; the progress of failures are relatively slow and it is the best time to respond and prevent a large cascade and blackout, (2) *escalation stage*, where due to critical component failures, the next generations of failures are accumulating quickly and the control of the system becomes very difficult, and (3) *phase-out stage*, where the progress of cascade slows down again as the cascade phases out and comes to an end. These stages can be observed from historical and simulation cascade data from power grids [29, 64, 147]. As such, the failures that occur at the end of the precursor phase and at the beginning of the escalation phase can be considered to have the most contributions in aggravating the failures and fueling the cascade. Hence, we use the cascade data to evaluate when in the cascade process the identified central components appeared. To do this, following [148], we define generations within each cascade as groups of failed components. As mentioned in [148], useful insights can be obtained by grouping components into generations, which suggests that initial failures are mostly the failures occurring in the precursor stage and the subsequent failures are the ones belonging to the escalation and phase-out stages. Depending on the rate of propagation of failures between various generations and stages in the same cascade, the impact of the failures on the subsequently failing components can be determined. In our cascade dataset, the range of number of observed generations is 1 to 96.

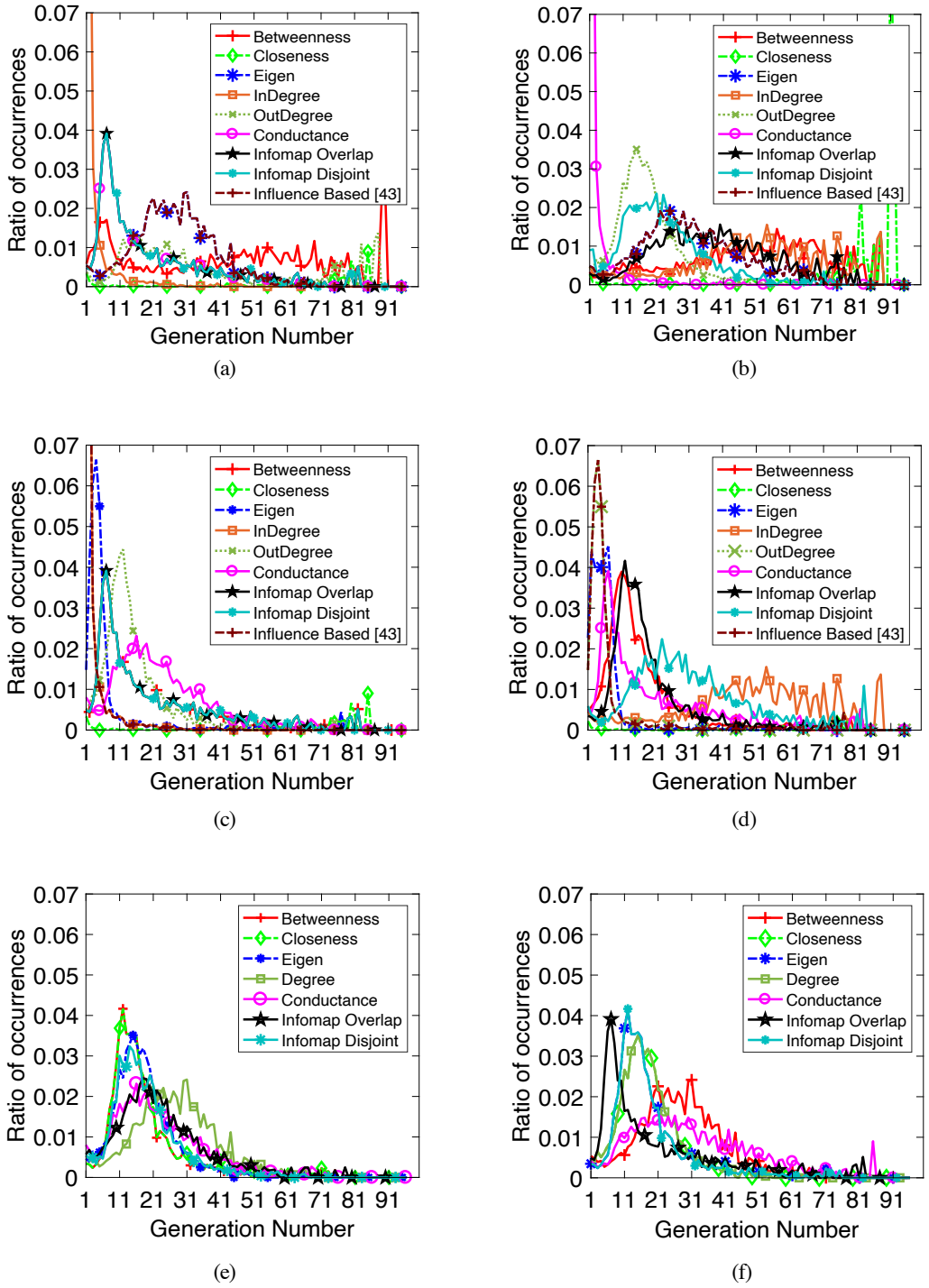


Figure 4.4: Ratio of occurrences of central component failures in various cascade generations over the total number of generations with the same index. Results for rank 1 central component are presented for graph (a) $H \geq 0.6$, (c) $H \geq 0.7$, and (e) $CR \geq 0.7$; and for rank 2 central components for graph (b) $H \geq 0.6$, (d) $H \geq 0.7$, and (f) $CR \geq 0.7$.

We have assumed that the generations 1 to 3 represent the precursor phase, the generations 4 to 50 represent the escalation phase and the generations 51 to 96 represent the phase-out stage.

In Figure 4.4, we present the results for this evaluation measure for three graphs of interactions; all constructed based on the same cascade dataset. Specifically, Figure 4.4 shows the ratio of occurrence of the failures of central components with rank 1 and 2 for various cascade generations over the total number of generations with the same index in the cascade dataset. In Figure 4.4a, the rank 1 central component identified by the conductance-based, Infomap overlap and Infomap disjoint community detection approaches is the same (component 30 with rank 1 as seen in Table 4.3). It can be seen that the peak of the occurrence lies in the early cascade generations. Similarly, for the influence-based graph of interactions with threshold 0.7 and the correlation-based graph of interactions with threshold 0.7 (i.e., Figures 4.4c and 4.4e respectively), all three community-based approaches identify components that their occurrence peak is in the early generations. While the central components identified by other standard centrality measures, and also by the influence based measure [43] show the peak at slightly different generations, they all mainly agree with the same trend of higher peaks at early to middle generations. Also, we can observe that depending on the graph of interactions, the agreement among these measures vary. A similar behavior can be observed for the rank 2 central component as shown in Figures 4.4b, 4.4d and 4.4f.

4.3.1.2 Critical Components' Occurrence in Various Cascade Sizes

As another data-driven measure, we also evaluate the severity of the cascades that the identified central components were a part of. Specifically, we evaluate the ratio of occurrence of central component failures in various cascade sizes. To this end, higher occurrence ratios in larger cascade sizes can imply critical contribution of the component in the cascade. In Figure

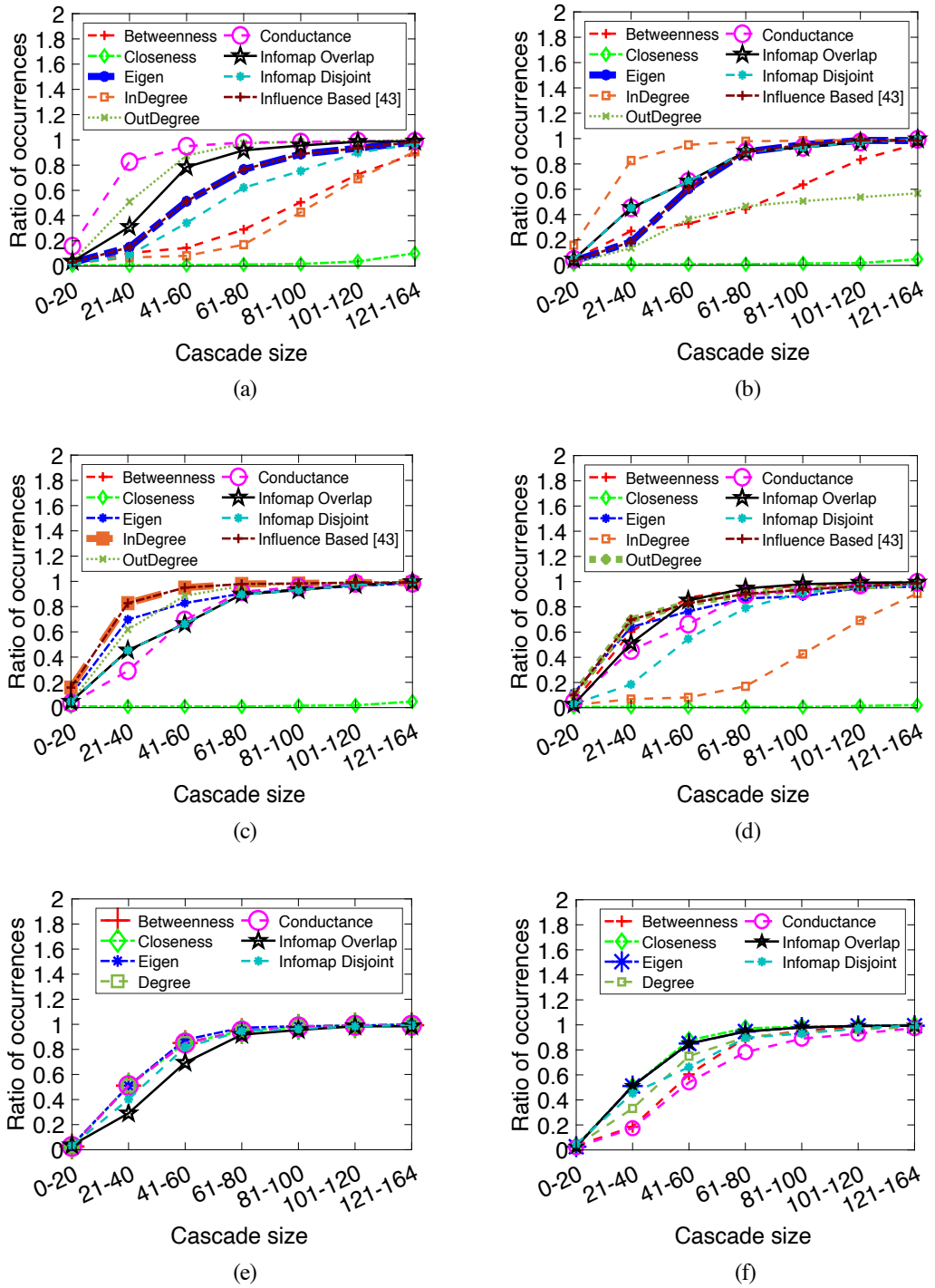


Figure 4.5: Ratio of occurrences of central component failures in various cascade sizes over the total number of cascade scenarios. Results for rank 1 central component are presented for graph (a) $H \geq 0.6$, (c) $H \geq 0.7$, and (e) $CR \geq 0.7$; and results for rank 2 central component are presented for graph (b) $H \geq 0.6$, (d) $H \geq 0.7$, and (f) $CR \geq 0.7$.

4.5, we present the results for cascade size of critical components for three graphs of interactions; all constructed based on the same cascade dataset. Figure 4.5 presents the ratio of occurrences of failure of central components with rank 1 and 2 centralities over the total number of cascade scenarios for various cascade sizes. The results presented in Figure 4.5 suggest that among all cascade sizes observed in the dataset, the identified central components tend to be a part of larger cascades. Some points to note from the results are as following. The community-based centrality measures show a consistent performance in identifying components that appear in larger cascade sizes while other centrality measure such as degree-based, betweenness, eigenvector centralities, and the influence-based measure [43] show more fluctuations among different graphs and ranks.

4.3.2 SI Simulation-based Verification of Centrality

In this section, we evaluate the impact of the identified central components in the cascade process by protecting them from failures and characterizing the effect of their protection on reducing the risk of failure propagation. To do so, we use the SI epidemic model on the graph of interactions to simulate the cascade process. We particularly consider two possible states for each node (component) of our network including: (1) functional, and (2) failed. We use the weights of the links connecting the nodes (identified through the construction of the interaction graphs in Chapter 2) to probabilistically propagate failure from one node to another. To protect a component from failures, we simply remove it and its adjacent links from the graph of interactions; meaning that they cannot be failed and they do not participate in propagation of failures. Protecting components happen before the start of the cascade process and is enforced as an initial condition. To evaluate the role of the identified central components, we set them as protected and run the SI simulation to characterize the average size of failures. We trigger the failures using two random initial failures

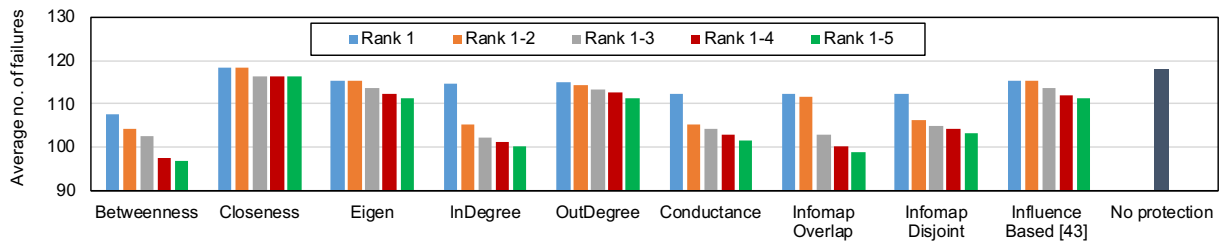
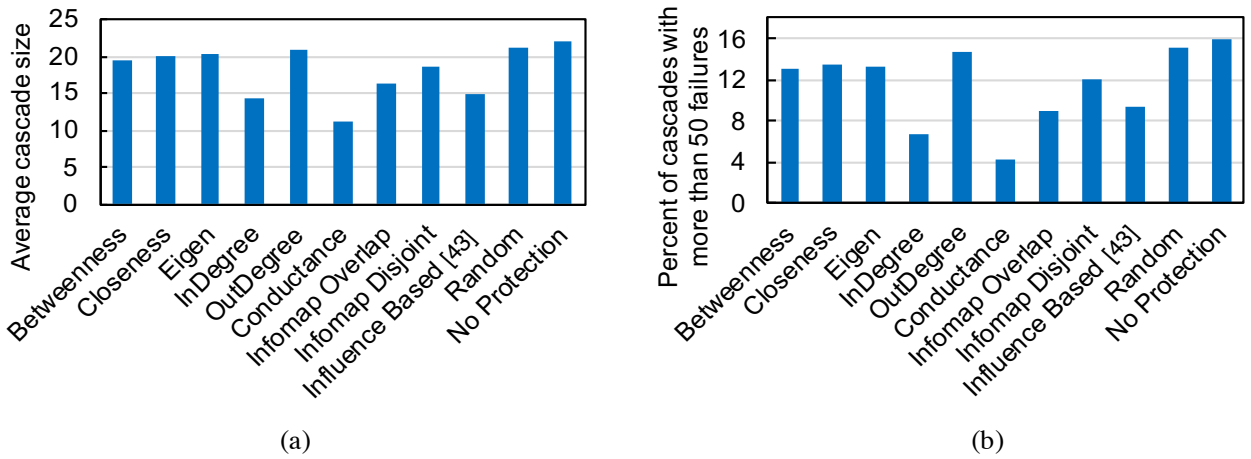


Figure 4.6: SI simulation of cascade over influence interaction graph H with threshold 0.6. For each centrality measure, (from left to right) the first bar represents protection of component with rank 1, the second bar represents protection of components with rank 1 and rank 2, and so on.

and repeat the simulations 100 times for each initial failure scenario. We use 1000 different two initial failures scenarios and apply the same initial failure scenarios for all centrality measures as the trigger event.

We specifically protect nodes identified as central using different centrality approaches in the order of rank 1; rank 1 and rank 2; rank 1, rank 2 and rank 3 and so on. The results are presented in Figure 4.6 for graph H with threshold 0.6. Similar behavior has been observed for other graphs of interactions. From Figure 4.6, it can be observed that protection of central components based on betweenness centrality provides the best performance in reducing the risk of propagations with Infomap overlap-based centrality closely following. The in-degree, conductance, and Infomap disjoint also follow next while the total improvement in the average number of failures is not significant in all cases due to small number of protected components.

In the next experiment, we protected a larger number of central components identified from the community-based centrality measures. Specifically, we protected 16 central components, which is the number of overlapped components identified from Infomap overlap method (we used the same number for the other two methods as the number of overlapped and bridge components for them are large, see Table 3.4). We observed an improvement of 33.5%, 29.8%, and 26.4% in the average



(a) (b)

Figure 4.7: Power simulation of cascade with protection of top five central components showing (a) average cascade size and (b) percentage of cascade sizes larger than 50.

size of failures compared to the unprotected scenario as shown in Figure 4.6 for conductance-based, Infomap overlap and Infomap disjoint centralities, respectively. We also protected 16 randomly selected components and obtained a 2.3% improvement in the average size of failures compared to the unprotected scenario, which is smaller than that of community-based centralities; suggesting the importance of protecting central nodes in reducing the risk of failures.

4.3.3 Power Simulation-based Verification of Centrality

In this evaluation method, we use the power system simulation of cascading failures, as discussed in Section 1.2.2, to evaluate the role of protecting central components in reducing the risk of cascading failures in the power grid. As discussed in Section 1.2.2, transmission line overloading is the main mechanism for propagation of failures in our power simulations; as such, by adequately increasing the transmission capacities of lines for handling larger power flows, the central lines will be protected from failures. A similar approach was used in [138] based on the branch capacity expansion model [149]. As such, similar to the SI simulation-based model, we protect the first five

central components (transmission lines) by doubling their capacities and run the power system simulation to characterize the effects of such protection on the risk of cascade. We trigger the failures using two random initial failures and use 1000 different two initial failure scenarios. Note that we apply the same initial failure scenarios for all centrality measures so that the results are comparable. We specifically look at the average size of cascade (Figure 4.7a) and percentage of cascades with more than 50 failures in the simulation (Figure 4.7b) to evaluate how protecting the central components affect the risk of failures. Inspection of the capacity upgrade for critical components based on Figure 4.7 suggests that protecting central components identified based on community-based centrality measures can reduce the risk of cascading failures better than protection of components identified using other methods. Specifically, protecting the components identified based on the conductance-based method shows the best performance in reducing the risk of cascade. The in-degree and influence-based measure [43] also perform well followed closely by the Infomap overlap and Infomap disjoint measures. These results suggest that certain criticality analyses based on centrality measures, such as conductance-based centrality measure, can have real-world implications for upgrading power grids and improving their resilience.

4.3.4 Graph-based Efficiency Analysis of Centrality

Here, we use graph-based efficiency properties to evaluate the effect of failure of central components on the overall connectivity of the graph. In particular, we use the global efficiency measure introduced in [150] as following:

$$Ef = 1/(n(n-1)) * \sum_{i \neq j} 1/d_{ij}, \tag{4.3}$$

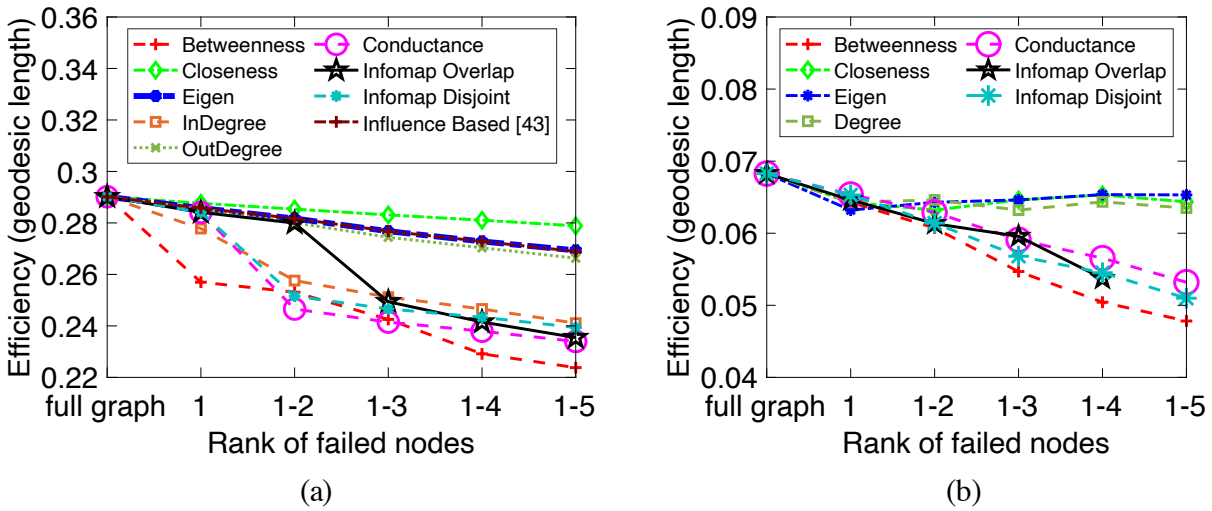


Figure 4.8: Graph efficiency-based analysis with failure of the top five central components for graph (a) $H \geq 0.6$, and (b) $CR \geq 0.7$.

where n is the number of components in the graph and d_{ij} is the shortest path (geodesic distance) between any two nodes i and j. To show the role of central nodes (components) on the efficiency of the network, we remove (i.e., fail) central nodes starting from rank 1, rank 1 and rank 2 and so on. To this end, the graph efficiency should drop as the number of failed nodes increases. We can see from Figure 4.8 that central components identified using the community-based approaches have higher drops in efficiency with failure of identified central nodes compared to other centrality measures except for the betweenness centrality for both the influence and correlation interaction graphs. We observe that betweenness centrality dominates the drop in efficiency, which is due to the definition of the global efficiency measure based on the geodesic distances.

4.3.5 Conclusions from Results and Complexity Analysis

Based on the presented results, betweenness, indegree, conductance-based, Infomap community-based, and influence-based [43] centrality measures each perform better for different evaluation methods, while the community-based centrality measures are more consistent and show less varia-

tions in performance among the presented evaluation methods. The computational complexity of calculating the betweenness centrality is $\Theta(|V|^3)$ in general, while for sparse graphs can be reduced to $O(|V|^2 log(|V|) + |V||E|)$ (where $|V|$ is the number of nodes and $|E|$ is the number of edges in the graph); however, the interaction graphs are not usually sparse. The computational complexity of indegree centrality is $\Theta(|V||E|)$ and $\Theta(|V|)$ for list and matrix implementation of the graph, respectively. Finally, the computational complexity for conductance-based centrality is $O(|V|^2)$ [126] and Infomap-based centrality is $O(|E|)$ [151]. Note that the majority of the computational complexity for the proposed community-based centrality measure is in finding communities while the rest of the calculations are in the order of $O(1)$. Therefore, there is a tradeoff between computational complexity and performance of the centrality measures in some cases; however, the community-based centrality measures have better computational complexity than the betweenness centrality. While comparison between different centralities offers important insights about performance of these measures in identifying criticality of components, the main goal of the study and presented results is to show that community structures in interaction graphs play a key role in the cascade process in power grids similar to many other real-world networks. The community structures are not only useful for identifying critical components and preventing the spread of failures but also for other applications such as providing information on the possible sizes of cascades based on the sizes of the communities.

4.4 Role of Operating Characteristics in Power Grids in Centrality

Similar to the different techniques that lead to different graph of interactions for the cascade process, operating settings, as introduced in Section 1.2.2 (such as the loading level for the power grid, r), also results in different graphs of interactions and subsequently different ranking of central

Table 4.4: Critical Components based on Conductance-based, Infomap Disjoint, and Infomap Overlapped Community-based Centrality (CM) for different power grid loading level

Rank	Conductance-based		Infomap Disjoint		Infomap Overlap	
	r=0.6	r=0.9	r=0.6	r=0.9	r=0.6	r=0.9
1	58	27	36	27	36	27
2	36	8	19	8	9	8
3	19	36	38	36	35	36
4	38	55	8	55	96	101
5	8	58	9	58	60	96

components. In this section, we use two very different loading levels for our power grid; specifically, $r = 0.6$ as a normally loaded power system and $r = 0.9$ as a stressed loading, where the power grid works close to its capacity. As expected in the latter setting, our power system simulation leads to larger sizes of cascading failures. Using the cascade model discussed in Section 1.2.2, we generate two cascade datasets for each of these settings (with size of 10,000 scenarios each). Next, we generate the graph of interactions using influence-based approach. We observe two very different graphs from these two datasets with different number of links and communities suggesting that the interactions among the components change as the operating settings of the power system vary. Due to different graphs of interactions, the central components also vary between these cases. The list of critical components identified based on the community-based centrality measure for these cases is presented in Table 4.4. We have also evaluated the role of the identified central components in the cascade process for these two cases and have observed similar trends as those of Section 4.3. Overall, these results show that depending on the condition and the operating settings of the power grid, the critical components of the system may vary. Therefore, it is important to perform criticality analysis with considerations about the system's state and operating settings.

4.5 Summary and Conclusions

In this chapter, we proposed a community-based centrality measure based on the trapping property of community structures and identified critical components and their role using data-driven, epidemic simulation on the graph of interactions, power system simulation, and graph-based approaches. We also compared the role of the identified central components with the central components identified using other standard centrality measures. While each of the measures shed light on different aspects of the criticality of components, they are not designed to identify critical components in cascade and epidemic processes. We showed that in most cases, the community-based centrality measure performed better than these measures in identifying critical components in the cascade. We also compared the critical components identified using the community-based centrality measure with the ones identified using the influence probability vector, designed in the study in [43]. We observed that our community-based centrality measure performed better than the influence probability vector in most cases. Moreover, we showed that the loading level of the power grid impacted the graph of interactions and consequently, the community structures and criticality of the components in the cascade process. This suggests the importance of considering the state and operating settings of power systems in reliability analyses. The presented study suggests that protecting the identified critical components using the community-based centrality can help in reducing the risk of large cascades and blackouts.

Chapter 5: Cascade Size Analysis in Power Grids Using Community Structures

Understanding cascade size distribution in power grids and predicting cascade size when it gets triggered, have been extensively studied in literature. For instance, in the study in [20], blackout data from historical sources as well as from simulations revealed the power-law behavior in the blackout size distribution (e.g. measured in terms of unserved energy or number of tripped transmission lines). This suggests that the likelihood of occurrence of large blackouts is more than what is traditionally expected. Additionally, prediction of cascade sizes given an initial triggering event, can help in estimating the risk of large blackouts and control and mitigate the spread of failures during cascade processes. It also allows for characterization of contributions of components of the system towards large cascades.

In Chapter 3, we studied the structures and patterns of interactions using the community structures in interaction graphs constructed in Chapter 2 based on the influence-based and correlation-based methods. In this chapter, we use the community structures that are present in the interaction graphs to study the failure propagation between communities and to characterize the likelihood of various cascade sizes. For this purpose, we formulate a Markov chain (MC) model based on the community structures in the interaction graphs of the power grid. This model exploits the properties of overlap and bridge nodes of communities (i.e., nodes that belong to multiple communities or have connections to other communities) as well as the strength of influences/interactions of the components to characterize transition probabilities in the MC. The states of the community-

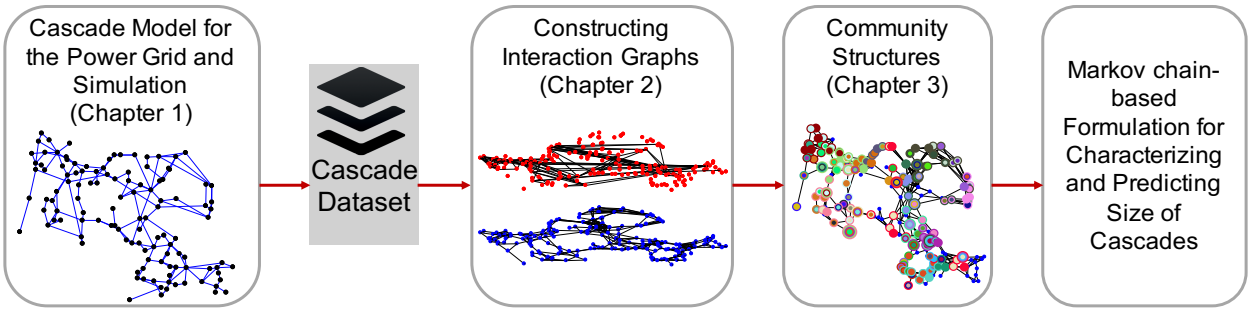


Figure 5.1: Key components for characterizing and predicting cascade sizes using community structures in interaction graphs of power grids.

based MC model allow the tracking of the size of cascades. The main idea behind this model is that the groups of components that form communities, provide an estimate measure of cascade size, as a cascade entering a community is likely to spread failures to other components within the community and less likely to spread to outside the community. Thus, using our community-based MC, the distribution of cascade sizes can be characterized using the size of communities. Additionally, depending on the initial conditions such as the community from which the cascade starts, cascade size distribution can be characterized. As suggested by historical data and previous studies of cascading failures [20], we hope to observe a power-law behavior in the distribution of cascade sizes, which suggests the importance of community structures of interaction graphs in cascade behavior. Figure 5.1 shows the key components required the reliability studies in this chapter.

5.1 Related Prior Work

Forecasting cascade sizes is a challenge in the reliability analysis of power grids. In the study in [76], a correlation graph-based statistical model, known as the *co-susceptibility* model, is used to predict cascade size distributions in transmission network of power grids using individual failure probabilities of transmission lines as well as failure correlations between transmission lines found

from the correlation matrix. The study exploits the idea that groups of components that have higher correlations are likely to fail together and uses the correlation matrix to find such co-susceptible groups which is an approximate estimate of the cascade size given an initial trigger failure. Similar idea is used in the studies in [3, 4], where components within the same community are assumed to be likely to fail together and the size of communities gives an approximation of cascade sizes. The study in [65] also characterizes size of cascades by using the states of the Markov chain to find the probability distribution of the number of generations in a cascade.

In this section, we briefly review the probabilistic [29, 34, 48, 49, 152, 153] and graph-based models [65, 105] related to this work that focus on characterization and prediction of cascade sizes.

In a group of works in literature, it has been discussed that branching process [48, 49] can provide an abstract model for outage distribution in cascading failures. In branching process models, outages are grouped into *generations*, where each generation is a sequence of components that failed within a short time-frame. Each component in a generation can independently produce a random number of child outages based on a Poisson offspring distribution, with a specific propagation rate. Thus, historical transmission line outage data was used to estimate static propagation rates for all generations [48] and varying propagation rates for each generation [49] to predict the probability distribution of subsequent line outages given distribution of initial failures.

Probabilistic models in the studies in [29, 34, 152, 153], are used for stochastic modeling of cascading failures. In the study in [152], a regeneration-based probabilistic approach was used for characterizing the probability of reaching an arbitrary blackout size at any time given initial power grid conditions, which included loading level, maximum capacity of the set of failed lines, and number of failed lines. In the extended study in [29], an analytically tractable Markov chain model

was developed, in which the states represented the critical grid conditions identified in the study in [152]. Additionally, operational characteristics i.e., loading level, load shedding constraints, and line tripping threshold were also considered for determining transition rates. This model was used to analytically predict the probability of blackout in time and also asymptotically determine the probability mass function of blackout size. The study in [153] extended the Markov model in [29] to consider the effect of interdependencies between power and communication systems on cascade size distributions. In the study in [34], the operational characteristics discussed in [29], was used to study specific conditions which led to power-law behavior on the probability mass function of blackout size.

In addition to probabilistic models, graph-based models [65, 105] have also been used for analyzing cascades and their size distribution. In the study in [105], a correlation matrix constructed using cascade data of failure/functional statuses of transmission lines, is used for predicting distribution of cascade sizes. And in the study in [65], a Markov chain constructed using transmission line outage data, where states represented *generations* of line outages, was used to predict distribution of cascades of varying sizes. The work presented in this chapter, belongs to both of the aforementioned categories as it provides a MC model based on the structures embedded in the interaction graphs of power grids for analysis of cascading failures.

5.2 Markov Chain Formulation

Using the concepts discussed in the previous sections, we propose a Markov chain (MC) framework to model the cascade size evolution using the community structures in the interaction graph of the power grid. As mentioned earlier, communities tend to trap failure propagation inside

and thus, provide an estimate on the likelihood of various cascade sizes depending on the size of the community, where the failures started, and the communities that failures can propagate to.

We define the state space of the MC based on all possible combinations of the set of communities $\mathcal{C}$ in the interaction graph $\mathcal{IG}$. We consider two variables as the state variables of the MC: (1) variable S representing the set of communities, which have been involved in the cascade process (communities with failed components) and (2) a binary variable I representing the condition that the cascade is contained within existing communities and thus, cascade stops with an absorbing state for MC (i.e., $I = 1$) or not (i.e., $I = 0$). Let $X(t)$ denote the state of the MC at time $t \geq 0$ using pair $(S(t), I(t))$.

To illustrate the MC model and its state space, consider the example interaction graph for a power grid shown in Fig. 5.2-a. Assume that applying a community detection algorithm on this graph identifies three disjoint communities, named C_1, C_2, and C_3, as shown in the figure. Given these communities, there are fourteen possible states for the MC in which, half of the states are transient and half are absorbing (due to binary variable I) as below. Thus, the state space of the MC for the system in Fig. 5.2-a is: $\mathcal{S} = \{(\{C_1\}, 0), (\{C_1\}, 1), ((\{C_2\}, 0)), (\{C_2\}, 1), (\{C_3\}, 0), (\{C_3\}, 1), (\{C_1, C_2\}, 0), (\{C_1, C_2\}, 1), (\{C_1, C_3\}, 0), (\{C_1, C_3\}, 1), (\{C_2, C_3\}, 0), (\{C_2, C_3\}, 1), (\{C_1, C_2, C_3\}, 0), (\{C_1, C_2, C_3\}, 1)$. For an interaction graph with n number of communities, the number of transient states in the MC is $\sum_{r=1}^{n} \frac{n!}{(n-r)!r!}$, where r represents the number of communities involved in the cascade process. While the number of states can be large for large values of n, the number of communities in interactions graphs are generally much smaller than the number of nodes in the graph and thus, state space explosion will not occur for interaction graphs of power grids.

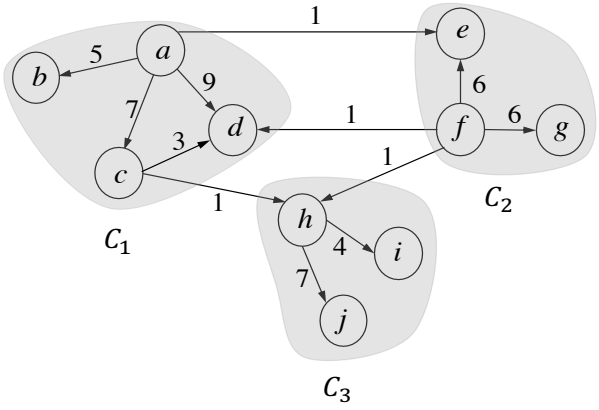
We define the transitions among the MC states by exploiting the connections among the communities. As the connections (overlap/bridge components) among the communities result in propagation of failures among the communities, they cause the state of the MC to change by involving more communities in the cascade process. We assume that in each transition of the MC, at most one new community will get involved in the cascade process (cascade will propagate to a new community through the connections with communities in the current state of the MC). This simplifying assumption can be justified by considering that time has been divided into small instances and only one new community can get involved at each time instance. This assumption help in finding the transition probabilities and make the transition matrix of the MC sparse. We also assume that if the cascade gets contained in the communities that have already been involved in the cascade process based on current state of the MC, then the MC will transit to an absorbing state with same set of communities. Thus, in our MC, three types of state transitions are possible:

1. transition from a transient state to another transient state with one additional community when failures propagate to a new community,

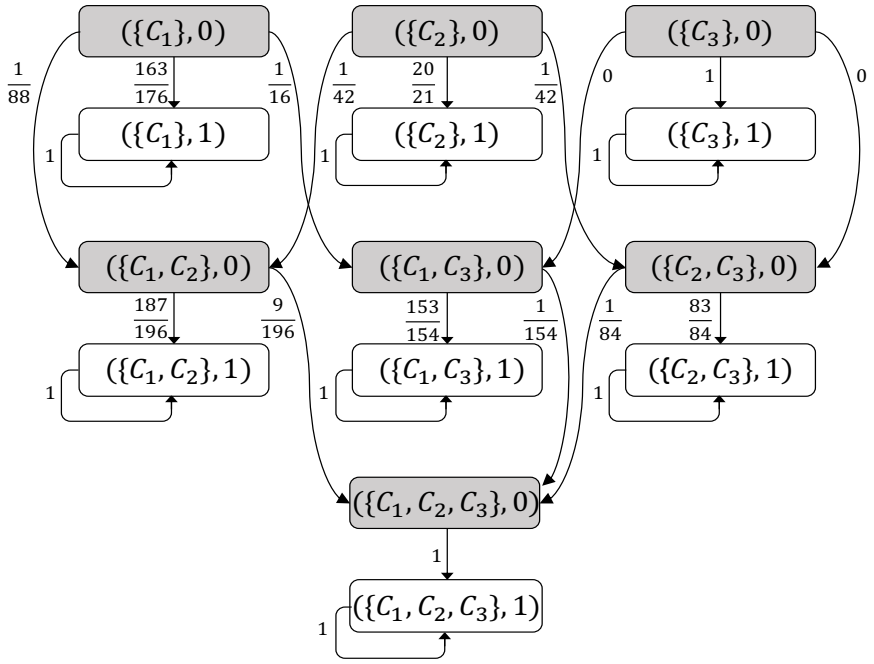
2. transition from a transient state to an absorbing state with the same set of communities when cascade stops, and

3. finally, transition from an absorbing state to itself representing that the state of the system will not change with regards to cascading failures when it stops.

Note, the set of communities in each MC state provides an estimate of the cascade size. Specifically, we assume that the number of failed components at each state of the MC can be found by adding up the sizes of the set of communities in that state.



(a)



(b)

Figure 5.2: Markov chain framework of an example of community structures. (a) Example of three disjoint communities. (b) Markov chain and the transition probabilities derived using the community structure with three disjoint communities and weights of interactions shown in (a).

To formulate the transition probabilities among the states in the MC, we utilize the sizes of the communities, the number of overlap/bridge nodes among communities, and the weight of influences/interaction links between communities. As the first step towards formulating the transition probabilities, we define the contribution of a single node $u \in C_i$ that has interaction links to nodes in community C_j (i.e., node u is an overlap or bridge node), in the propagation of failure from community C_i to community C_j as

$$CFP(u) = \frac{\sum_{v \in C_j, j \neq i} w_{u,v}}{\sum_{q \in L_{\mathcal{IG}}} w_{u,q}}. \tag{5.1}$$

In the context of various types of communities, in equation (1) we have: (i) *for the case of disjoint communities of directed interaction graphs:* $w_{u,v}$ is the weight of the interaction link from o-bridge node u to i-bridge node v and $w_{u,q}$ is the weight of the interaction link from o-bridge node u to node q, (ii) *for the case of disjoint communities of undirected graphs:* $w_{u,v}$ is the weight of the interaction link connecting bridge node u to bridge node v and $w_{u,q}$ is the weight of the interaction link connecting bridge node u to node q, and (iii) *for the case of overlapped communities of directed as well as undirected graphs:* $w_{u,v}$ is the weight of the interaction link connecting overlap node u to overlap node v and $w_{u,q}$ is the weight of the interaction link connecting overlap node u to node q. Note, for all the above discussed cases, node q can be a member of any community including communities C_i and C_j.

In the next step toward formulating the transition probabilities, we need to consider the cumulative effect of all the overlap/bridge nodes among the communities. Specifically, the contribution of: (i) all o-bridge nodes in community C_i in the case of disjoint communities of directed graphs, (ii) all bridge nodes in community C_i in the case of disjoint communities of undirected

graphs, and (iii) all overlap nodes in community C_i in the case of overlapped communities of both directed and undirected graphs; will be added up. For the case of transitioning from a state with a single involved community to a state with two involved communities we have:

$$p((\{C_i\},0) \rightarrow (\{C_i,C_j\},0)) = \frac{\sum_{z\in C_i} CFP(z)}{|C_i|}, \tag{5.2}$$

where $|C_i|$ is the size of community C_i and $p((\{C_i\},0) \rightarrow (\{C_i,C_j\},0))$ is the probability of transition from transient state $X(t) = (\{C_i\},0)$ to transient state $X(t+1) = (\{C_i,C_j\},0)$ with one new additional community. E.g., in the MC of Fig. 5.2-b, contribution of o-bridge node f to failure propagation in community C_1 i.e. $CFP(f)$ is $1/(1+1+6+6) = 1/14$ and probability of transition from transient state $(\{C_2\},0)$ to transient state $(\{C_1,C_2\},0)$ is $1/(14\times 3) = 1/42$.

To generalize, the probability of transition from a transient state $(\{C_i,C_j,\cdots,C_r\},0)$ to a transient state $(\{C_i,C_j,\cdots C_q,C_{r+1}\},0)$ with one new additional community can be defined as:

$$p((\{C_i,C_j,\cdots,C_r\},0) \rightarrow (\{C_i,C_j,\cdots,C_q,C_{r+1}\},0)) = \frac{\sum_{z\in O_{\{C_i,C_j,\cdots,C_r\}}} CFP(z)}{|\{C_i,C_j,\cdots,C_r\}|}, \tag{5.3}$$

where $O_{\{C_i,C_j,\cdots,C_r\}}$ is the set of all o-bridge, bridge, or overlap nodes of the set of communities $\{C_i,C_j,\cdots,C_q\}$, which have interaction links to nodes in community C_{q+1}; and $|\{C_i,C_j,\cdots,C_q\}|$ is the sum of sizes of all communities in the set $\{C_i,C_j,\cdots,C_q\}$. Note, for calculating the sum of sizes of overlapped communities, repeated entries of overlap components are counted only once.

The probability of transition from a transient state $X(t) = (S(t),0)$ to its associated absorbing state $X(t+1) = (S(t+1),1)$ (i.e., absorbing state with the same set of communities as that of the transient state, i.e., $S(t) = S(t+1)$), describes the probability of failure of *non-*

overlap/non-bridge nodes (i.e., nodes that belong to a single community). This type of transition implies that failure of non-overlap/non-bridge nodes of a community spreads failures to other components within the community itself only and failures are contained within involved communities. The probability of transition from transient state $X(t) = (S(t), 0)$ to its associated absorbing state $X(t+1) = (S(t+1), 1)$ will be the complement of probabilities of transition to all other transient states from state $X(t)$. Finally, the only possible transition from an absorbing state will be to itself, such that $p((S(t), 1) \rightarrow (S(t+1), 1))$ for $S(t) = S(t+1)$ is one.

5.3 Evaluation and Results

In this section, we use the community structures derived using the Infomap disjoint, Infomap overlap, and conductance-based community detection algorithms in the influence-based and correlation-based interaction graphs of power grids. The results and analysis for these studies have been discussed in detail in Chapter 2 and Chapter 3. In this chapter, our main focus and discussion is on the characterization of cascade sizes using our community-based MC.

5.3.1 Distribution of Cascade Sizes

To characterize the contribution of community structures in the distribution of cascade sizes, we use the MC formulation discussed in Section 5.2. We numerically calculate the average steady state distribution of the MC as well as cascade size distribution for various initial states of the system depending on where the failures started.

First, we look at the steady state distribution of the MC for the average case when the cascade can start from any community with equal probability. We evaluate this probability distribution for three interaction graphs including interaction graphs based on $H \geq 0.6$, $H \geq 0.7$, and

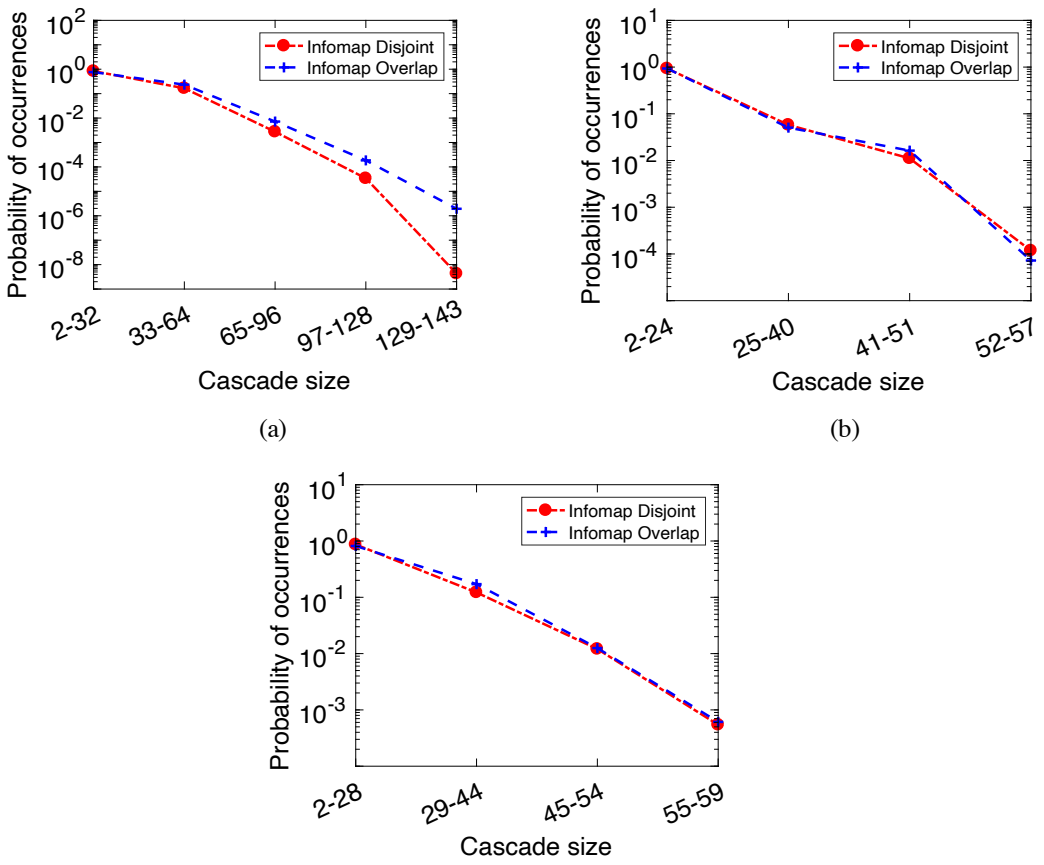


Figure 5.3: Log scaled probabilities of cascade sizes when cascade has equal probability to start from any community, for communities identified using Infomap disjoint (in red) and Infomap overlap (in blue) over the influence-based interaction graph of (a) $H \geq 0.6$, (b) $H \geq 0.7$, and correlation-based interaction graph of (c) $CR \geq 0.7$.

$CR \geq 0.7$. These results are presented in Fig. 5.3. The y-axis in Fig. 5.3 shows the log scaled distribution of probability of occurrences with respect to various cascade sizes in the x-axis. The range of cascade sizes in the x-axis shown in Fig. 5.3 correspond to the LCC of the thresholded graphs shown in Table 3.1 of Chapter 3. Based on these results, we observe that small cascade sizes are more probable compared to large cascade sizes. For example, in Fig. 5.3-a for $H \geq 0.6$, the probability of occurrence of cascade size in the range of 2 to 32 failures is 0.83 (seen as 10^0 due to log scale), while the probability of occurrence of cascade size in the range of 129 to 143 is small.

Table 5.1: Community and overlap/bridge sizes for influence-based interaction graph H with threshold 0.6

Community	Infomap Disjoint		Infomap Overlap	
	No. of nodes	No. of bridges	No. of nodes	No. of overlaps
1	31	14	32	13
2	17	7	17	7
3	16	5	18	6
4	21	11	29	16
5	26	12	28	14
6	7	3	7	3
7	3	3	6	6
8	2	1	3	3
9	6	3	3	1
10	5	5	8	5
11	6	6	6	6
12	3	3	5	4

These results are in agreement with the cascade size distributions found based on other historical and simulation data in [20]. This power-law based distribution of cascade sizes suggests that large cascade sizes are rare but their occurrence cannot be neglected. These large cascade sizes can be attributed to failure propagation caused by the size of overlap/bridge nodes as well as the strength of interaction of these nodes with other communities.

Next, we look at the cascade size distribution conditioned on the community that failures started in. In Fig. 5.4, for communities identified using Infomap disjoint for $H \geq 0.6$, we observe that each community has a distinct role in the probability of occurrence of various cascade sizes. For example, probability of occurrence of large cascade size of range 129 to 143 is highest for community 11 compared to other communities. This can be attributed to the large number of bridge nodes compared to the community size and the strong weight of the interaction links of the bridge nodes. As seen in Table 5.1, community 11 has 6 nodes and all 6 nodes are bridge nodes, as such failures

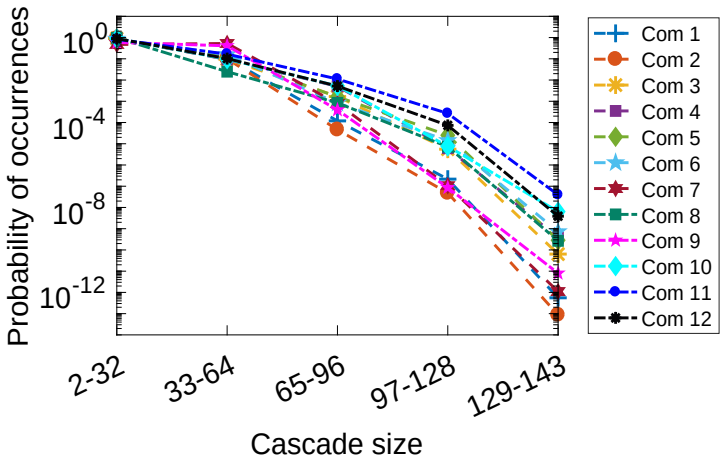


Figure 5.4: Log scaled probabilities of cascade sizes when cascade initiates from any community, for communities identified using Infomap disjoint over the influence-based interaction graph of $H \geq 0.6$.

in community 11 has very high likelihood of spreading to other communities. We observe similar results for other interaction graphs with different thresholds and community structures as well.

To normalize the effect of size of the initial community on cascade size, we also look at the number of ultimate failure per node failure in the initial community. As shown in Fig.5.5, for both Infomap disjoint and Infomap overlap, we observe that communities 1 to 5 induce smaller cascade sizes (from 1 to 10) shown in the expanded figure on the top, whereas communities 6 to 12 induce larger cascade sizes (from 1 to 72) shown in the main figure. This behavior is due to the larger size and smaller number of overlap/bridge nodes for communities 1 to 5 and vice versa situation for communities 6 to 12, as shown in Table 5.1. Thus, the contribution of nodes in smaller sized communities in causing large cascades are more prominent compared to nodes in larger sized communities.

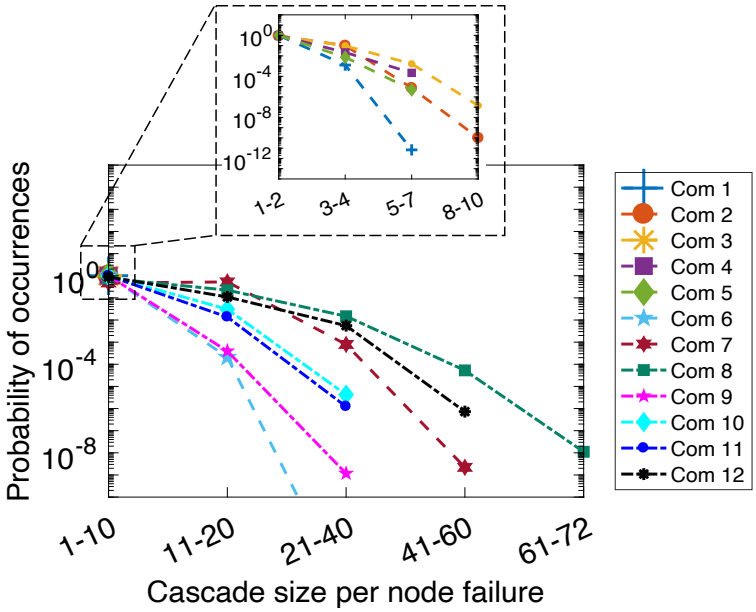


Figure 5.5: Log scaled probabilities of cascade sizes induced by failure of one component belonging to different communities, for communities identified using Infomap disjoint over the influence-based interaction graph of $H \geq 0.6$.

5.4 Summary and Conclusions

In this chapter, we developed a Markov Chain model to track the evolution of the cascading failures through communities embedded in interaction graphs. We discussed that the communities in interaction graphs can reveal important properties related to cascading failures as they tend to trap failures. We used this property and showed that the probability distribution of cascade sizes exhibited power-law behavior as observed in previous studies and historical data. This suggested that community structures affect the behavior of cascading failures and bear important information about cascade processes.

Chapter 6: Reliability Analysis of Interdependent Infrastructures Using Interaction Graphs

In this chapter, an integrated framework for the power grid and electric vehicle (EV) charging infrastructure is proposed and built using an abstract probabilistic framework, which can capture the role of structure of the systems on the dynamics to model the dynamics and specifically, the demand and traffic distribution in EV charging infrastructures. The goal of the model applied to EV charging infrastructures is to identify incentives, when and where they are needed, to design dynamic energy pricing signals based on the state of both the power and transportation systems, such that the incentives help in appropriate distribution of load in both systems and orchestrating their operation. Based on the proposed integrated model, we identify incentives in terms of charging prices using an algorithm based on *topological sort* on the active influence graph of the charging infrastructure. We will show that the identified incentives based on this model lead to higher probabilities of stable and balanced systems both for the power and transportation systems. This study is an effort towards promoting the need and importance of integrated frameworks for modeling and analysis of smart cities.

6.1 Related Prior Work

We review the related work in two main categories. First, we briefly review recent efforts on

Portions of this chapter were published in Smartgreens [5] and Springer, Cham [6]. Copyright permissions from the publishers are included in Appendix A.

modeling, simulation, operation and design of integrated and interdependent infrastructure frameworks for smart cities. Second, as the focus of this study is on charging infrastructures, we review the related work on different aspects of design, operation and optimization of charging infrastructures.

6.1.1 Integrated and Interdependent Infrastructures for Smart Cities

Critical infrastructures and particularly their reliability has been the focus of many recent research efforts. The role of integrated, interdependent infrastructures in designing efficient smart cities has been emphasized by smart-city research community [154]. Moreover, the vision of smart cities has been described in different ways among practitioners and academia [154]. Hall [155] visioned the smart city as a city that monitored and integrated conditions of all of its critical infrastructures to optimize its resources and services to its citizens. In the last decade a large body of work has emerged in modeling and understanding interdependent infrastructures. The general concepts of interdependencies among critical infrastructures, challenges in modeling interdependent systems and their control and recovery mechanisms have been intensively discussed in [156], [157], [158], [159]. These works mainly discuss the intrinsic difficulties in modeling interdependent systems and suggest new methodologies for their modeling and simulation as a single coupled system.

The majority of such integrated, theoretical frameworks has been focused on analyzing the reliability of coupled systems and the negative aspects of the interdependencies among critical infrastructures [160], [161], [162]. For instance, one of the problems of concern in interdependent infrastructures is their reliability to cascading failures and propagation of faults. Recently, many researchers have studied cascading failures in interdependent systems (see for example, [161, 163] and references therein).

The work presented in this chapter is an effort to present an abstract and unified framework to model interactions among infrastructures, which can be used to design various smart-city solutions based on the state of interacting systems, for instance, the pricing mechanism based on the state of the EV charging infrastructure and the electric grid.

6.1.2 Charging Infrastructures

In recent years, a large body of work is focused on optimal placement of EV charging stations [10, 18, 164–170]. In particular, optimization formulations with various criteria have been used for addressing this problem [18, 165–167]. Examples of such criteria include, maximizing sustainability from the environment, economics and society perspective [165], minimizing the trip time of EV's to access charging stations [166], minimizing trip and queuing time [167], and maximizing the coverage of charging stations [18]. In the work presented in [164], [169], the set cover algorithm is used to optimize the location of charging stations from a set of possible locations. In addition, agent-based [168] and game-theoretic approaches [10], [170] have also been adopted in characterizing optimal deployment of charging infrastructures. Reference [171] presents a more detailed review of various approaches used for the optimal deployment of EV charging stations.

Another research aspect of charging infrastructures is their pricing mechanisms. Studies of traditional fueling infrastructures [172], [173] show that the price of fuel impact the behavior of drivers, which suggests that the charging price for EVs can also impact the users' choice and behavior. Specifically, authors in [174], discuss that the optimal placement of charging stations will be insufficient to handle rapid changes in traffic patterns and urbanization, hence an efficient pricing model that also minimize the social cost of traffic congestion and congestion at EV charging stations is needed. As another example, the impact of energy price and the interplay between the

price and other factors, such as cost and emissions, on the charging decisions have been studied in [17]. Besides the studies on the impact of price on charging decisions and traffic patterns, some efforts are focused on designing and optimizing pricing and analyzing their impact on the users' behavior and the system operation. Examples of such efforts include the work presented in [11], which uses a game theoretical approach to study the price competition among EV charging stations with renewable power generators and also discusses the benefits of having renewable resources at charging stations. Similarly, game-theoretic approaches that model a game between the electric grid and their users, specifically for EV charging, in order to design pricing schemes, have been studied, for example in [175]. The model in [175] provides strategies to EV chargers to choose the amount of energy to buy based on a pricing scheme to operate the charging infrastructures at their optimal levels.

The work presented in this chapter is closest to the studies on pricing mechanism design and also the interplay between the electric and EV charging infrastructures. At the same time, it is different in the approach as it considers the stochastic dynamics of the interdependent EV charging infrastructures and the power grid and their interactions in designing the charging prices at stations.

6.2 System Model for Interdependent EV Charging and Power Infrastructures

In this section, we describe our system model for the interdependent EV charging and power infrastructures; however, the model is adequately general to be applied to any interdependent infrastructure with interacting components. We first discuss the system model of the EV charging infrastructure, which receives energy from the power grid infrastructure. Then, we discuss the system model of the power grid infrastructure, where the EV charging infrastructure is assumed to

be the only source of load for the power infrastructure. In our extended system model, in addition to the EV charging infrastructure as one source of load, the power infrastructure model also includes other types of loads (residential or commercial). Finally, we discuss the integrated system model of the EV charging with the system model of the power grid infrastructure. The schematics of the integrated system model of the EV charging infrastructure with the power system model is depicted in Figure 6.1. Similarly, the integrated system model of the EV charging infrastructure with the extended power system model is depicted in Figure 6.2.

As the figures show, our study considers three layers in the integrated system: (1) the power/electric grid layer, (2) the EV charging infrastructure layer, and (3) the cyber layer, which enables the collaborative solution for the pricing between EV charging layer and power grid layer. Our model is mainly focused on the power and the EV charging infrastructures. While the cyber layer is not a part of the theoretical model, we will discuss its key role in Section 6.3. The key interactions among the layers of this system can be summarized as following. The EV charging infrastructure receives energy from the power grid and thus, the load on charging stations may affect the load on power substations. The pricing scheme, which depends on the state of both power and EV charging infrastructures, will be communicated through the cyber layer to the users. Finally, the communicated price will affect the load distribution over the charging infrastructure and subsequently, the load on power substations. In case of the extended model, the residential/commercial loads also effect the load on power substations and consequently, the pricing policy.

6.2.1 System Model of EV Charging Infrastructure

We denote the set of charging stations in a region by $\mathcal{CS} = \{CS_1, CS_2, ..., CS_k\}$. For simplicity, we assume that the charging stations are distributed over a grid region such that each

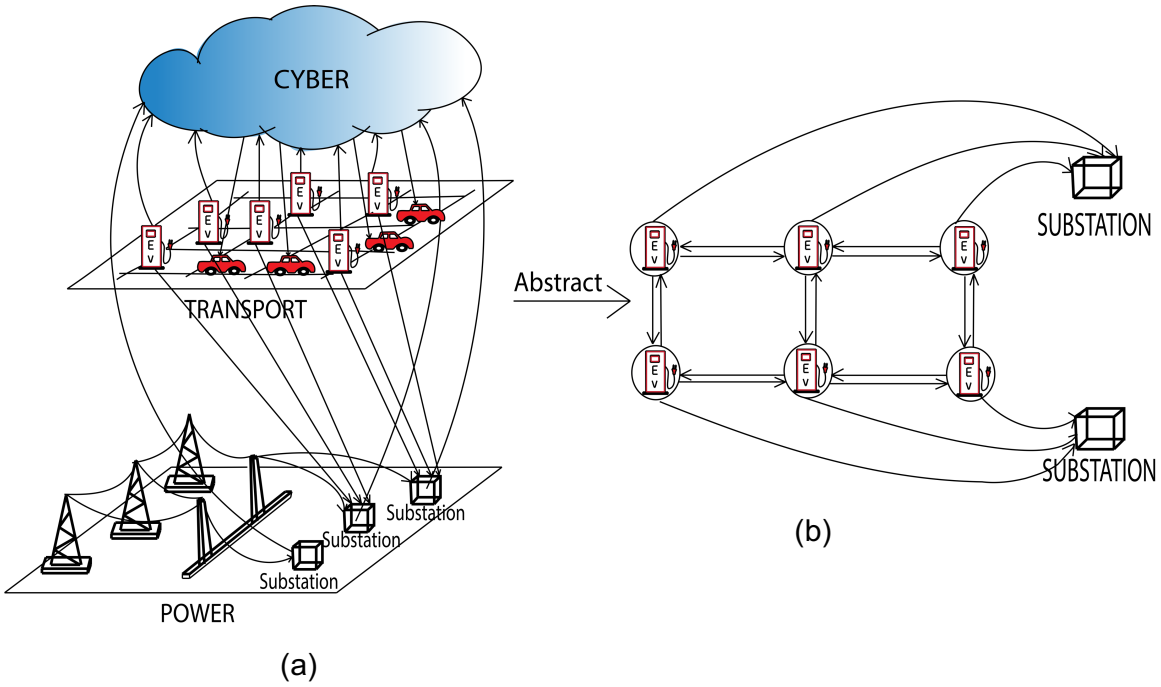


Figure 6.1: Integrated system model of interdependent networks of power and EV charging stations and the role of cyber infrastructure.

cell in the grid holds one charging station as shown in Figure 6.1 (a) and Figure 6.2 (a). The charging stations are connected over a directed graph $G = (\mathcal{CS}, E_c)$, where E_c represents the set of directed links specifying the possibility of travel between charging stations for the users. For instance, $e_{ij} \in E_c$ implies that users in the cell containing the station CS_i can travel to station CS_j for charging. These links help in specifying the constraints on the travel for charging, e.g., based on the distance that the users are willing to travel and the distance that a EV with the need for charging can travel before it runs out of energy. We will explain later that when the right incentives are applied, there is a likelihood for each user to travel to other stations with direct links (a probabilistic behavior). In this dissertation, we focus on a graph, in which charging stations in adjacent cells are connected. Other graphs with different topologies can also be considered and will not change the model.

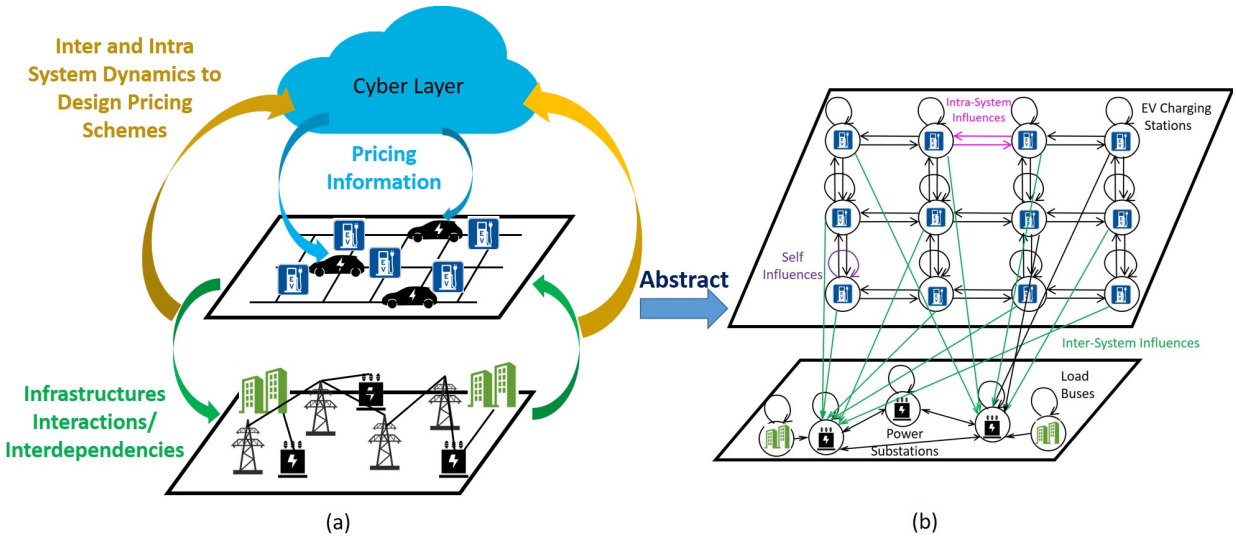


Figure 6.2: Extended integrated system model of interdependent networks of power and EV charging stations and the role of cyber infrastructure.

6.2.2 System Model of Power Grid Infrastructure

In this section, we describe the power infrastructure layer for the system model in Figure 6.1. In this dissertation, the intra-system model for the power infrastructure only considers the power grid substations denoted by $\mathcal{S} = \{S_1, S_2, ..., S_m\}$ and their internal dynamics (as will be explained in Section 6.2.5). Although the presented model for power system still does not capture the complete power grid model and dynamics with generators and power lines; it enables an abstraction for points of contact with the EV charging infrastructure and components which will directly impact it. To model the inter-system interactions between the power and charging infrastructures, we assume that multiple charging stations belong to the distribution network of one substation. And as such, we consider a set of inter-system links denoted by $\mathcal{L}$, where $L_{ij} \in \mathcal{L}$ specifies that charging station CS_i affect the load of substation S_j. In this model, $CS_i \in \mathcal{CS}$ should have a link to one specific $S_j \in \mathcal{S}$ while each S_j can have multiple incoming links from different geographically co-located charging station. Also, note that there will be no links from S_j to any node in $\mathcal{CS}$. Such interactions

and the effects of power substations on charging stations will be indirectly through the incentives communicated by the cyber layer.

6.2.3 Extended Model of Power Grid Infrastructure

Next, we describe the power infrastructure layer for the extended model in Figure 6.2. In this dissertation, we consider the power grid loads (e.g., residential or commercial loads) denoted by $\mathcal{B} = \{B_1, B_2, ..., B_k\}$ and substations denoted by $\mathcal{S} = \{S_1, S_2, ..., S_m\}$ (similar to the set of substations $\mathcal{S}$ defined in the power system model in Section 6.2.2) and their internal dynamics (as will be explained in Section 6.2.5) as well as the intra-system interactions (based on the physics of the electricity that can lead to propagation of voltage or current stresses and instabilities among substations). This represents a simplified model for a power system that is an extension to the system model discussed in Section 6.2.2, which only considered individual substations with no interactions among the substations.

The links among substations and residential/commercial loads are defined as following. The residential/commercial loads and substations are connected over a directed graph $G = (\mathcal{B} \cup \mathcal{S}, E_p)$, where link $e_{ij} \in E_p$ from load node to substation implies that the load is receiving energy from that substation and link $e_{ij} \in E_p$ between two substations imply that they can affect each others stress level and stability. In this model, we assume that there is no link from substation to loads.

The inter-system interactions between the power and charging infrastructures, are the same as the interactions discussed in the system model of Section 6.2.2. We assume that multiple charging stations belong to the distribution network of one substation. And as such, the set of inter-system links denoted by $\mathcal{L}$, where $L_{ij} \in \mathcal{L}$ specify that charging station C_i affects the load of substation S_j.

Similar to the system model, $CS_i \in \mathcal{CS}$ should have a link to one specific $S_j \in \mathcal{S}$ while each S_j can have multiple incoming links from different geographically co-located charging station. Also, there are no links from S_j to any node in $\mathcal{CS}$. Such interactions and the effects of power substations on charging stations is indirectly communicated through the incentives by the cyber layer.

6.2.4 Integrated System Model of EV Charging and Power Infrastructure

Based on the above discussion, the total integrated system of the EV charging and power infrastructure, is named as the *integrated system model* and is denoted by a graph as $G_p = (\mathcal{CS} \cup \mathcal{S}, E_c \cup \mathcal{L})$. Similarly, the total integrated system of the EV charging and extended power infrastructure, is named as the *extended integrated system model* and is denoted by a graph as $G_e = (\mathcal{CS} \cup \mathcal{S} \cup \mathcal{B}, E_c \cup E_p \cup \mathcal{L})$. However, the model for the system is not simply a graph. Next, we will explain how each component in these graphs stochastically and dynamically evolve and interact with other components. We will specifically present a model to capture such dynamics. We have chosen a probabilistic approach for the modeling as various aspects of this system is stochastic. For instance, the state of a charging station (e.g., being busy or not) varies probabilistically at different times of the day and week and due to EV users mobility pattern and behavior. The state of the load in a substation also varies due to stochastic nature of the demand. The interactions among components are also stochastic and as components influence each other depending on their state. For instance, if charging stations, which have a link to substation S_j, become busy and overloaded with lots of demand then, this increased demand will increase the likelihood of S_j to become overloaded and hinder the stability of the power grid. In such cases, it would be ideal to distribute the load in the system using pricing incentives to increase the willingness of EV users to travel to

other charging stations. These stochastic interactions and dynamics will be modeled in an influence framework as explained next.

6.2.5 Influence Model for Integrated Infrastructures

Here, we briefly review the Influence Model (IM) as first introduced in [176], [111] and present an IM-based framework for modeling the integrated charging and power infrastructures.

The IM is a framework consisting of a weighted and directed graph of interconnected nodes, in which, the internal stochastic dynamics of each node is represented by a Markov chain (MC) and the states of the nodes varies in time due to the internal transitions of MCs as well as the external transitional influences from other nodes. The weights on the directed links represent the strength of influences that nodes receive from one another. In the following, we put the IM model in perspective with respect to the integrated charging and power infrastructures. In our model, graph G_p and graph G_e with different types of links and nodes (as introduced in Section 6.2.4) will serve as the underlying graph for the IM. To represent the internal dynamics of nodes, we consider that the state of the charging stations can be abstracted to three levels: (1) underloaded, (2) normal, and (3) overloaded. As such, we define a MC with state space of size three for each $CS_i \in \mathcal{CS}$. These states help in describing the load (in terms of power demand) on a charging station at each time.

In general, the state of a CS_i may change due to departure or arrival of EV users. On the other hand, we model a substation S_j with an internal MC, which has two possible states: normal and stressed. These states specify if a power substation is overloaded and stressed or it is working under normal conditions. We also model the residential/commercial load B_j (in the case of extended

power system model) with an internal MC, which has three possible states including underloaded, normal and overloaded, similar to the CS_i states.

The transition probability matrix of the internal MC for a node, say node $i \in \mathcal{CS} \cup \mathcal{S}$ in the case of integrated system model (node $i \in \mathcal{CS} \cup \mathcal{S} \cup \mathcal{B}$ in the case of extended integrated system model), is denoted by $\mathbf{A}_{ii}$, which is an $m \times m$ row stochastic matrix, where m is the size of the state space. We use a data driven approach to characterize the transition probabilities of these internal MCs based on datasets of system dynamics and simulations as will be explained later. The links in the graphs G_p and G_e specify the influence relation among the nodes.

In particular, for the integrated system model, there are two types of influences in our model: (1) when a charging station influences another charging station, then it means there is a likelihood that it will send users (using proper incentives) to the influenced station; (2) when a charging station influences a power substation, then it means that there is a likelihood that the charging station increases the load on the power substation to a level that could change the state of the power substations (e.g., from normal to stressed). In the case of the extended integrated system model, there are two additional influences in addition to the ones considered in the integrated system model: (1) when a substation influences another substation, then it means that there is a likelihood that the substation stresses the other substation if it is stressed or helps a stressed substation to stabilize if it is normal; and (2) when a residential/commercial load node influences a substation then it means that it increases the load on the power substation to a level that could change the state of the power substations.

The weights on the links also specify the strength of the influence. The influences among the nodes of the network is captured by the *influence matrix* denoted by $\mathbf{D}$, where each element

d_{ij} is a number between 0 and 1 representing the amount of influence that node i receives from node j. The larger the d_{ij} is, the greater the influence that node i receives from node j; with the two extreme cases being $d_{ij} = 0$ meaning that node i does not receive any influence from node j and $d_{ij} = 1$ meaning that the next state of node i deterministically depends on the state of node j. Note that receiving influence from a node itself (self influence), i.e., d_{ii}, specifies how much the state evolution of a node depends on its internal MC. The total influence that a node receives should add up to unity i.e., $\sum_{j=1}^{n} d_{ij} = 1$, and therefore, matrix $\mathbf{D}$ is a row stochastic matrix too.

In IM, the status of a node, say node i, at time t is denoted by $s_i[t]$, a vector of length m, where m is the number of possible states for the node. At each time, all the elements of $s_i[t]$ are 0 except for the one which corresponds to the current state of the node (with value 1). In our model, $s_{i1}[t]$, $s_{i2}[t]$, and $s_{i3}[t]$ correspond to overloaded, normal and underloaded states, respectively, for charging stations and residential/commercial loads. Similarly, $s_{i1}[t]$ and $s_{i2}[t]$ correspond to normal and stressed states for power substations, respectively. The statuses of all the nodes concatenated together as $\mathbf{S}[t] = (s_1[t]s_2[t]...s_n[t])$ described the state of the whole system in time t, where $n = |\mathcal{CS} \cup \mathcal{S}|$ in the case of the system model ($n = |\mathcal{CS} \cup \mathcal{S} \cup \mathcal{B}|$ in the case of the extended system model) and $|.|$ denotes the cardinality of the set.

The influence matrix $\mathbf{D}$ specifies how much two nodes influence each other. In order to specify how the states of the nodes will change due to the influences, we also need state-transition matrices $\mathbf{A}_{ij}$, which capture the probabilities of transiting to various states due to the state of the influencing node. Matrix $\mathbf{A}_{ii}$ represents the special case of self-influence, which is described by the internal MC of the node. Note that the $\mathbf{A}_{ij}$ matrices are row stochastic. In the general IM [176], the collective influences among the nodes in the network is summarized in the total influence matrix

H defined as:

$$\mathbf{H} = \mathbf{D}' \otimes \{\mathbf{A}_{ij}\} = \begin{pmatrix} d'_{11}A_{11} & \cdots & d'_{1n}A_{1n} \\ \vdots & \ddots & \vdots \\ d'_{n1}A_{n1} & \cdots & d'_{nn}A_{nn} \end{pmatrix}, \tag{6.1}$$

where $\mathbf{D}'$ is the transpose of the matrix $\mathbf{D}$ and $\otimes$ is the generalized Kronecker multiplication of matrices [176]. Finally, based on the the total influence matrix $\mathbf{H}$ the evolution equation of the model is defined as

$$\mathbf{p}[t+1] = \mathbf{S}[t]\mathbf{H}, \tag{6.2}$$

where vector $\mathbf{p}[t+1]$ describes the probability of various states for all the nodes in the network in the next time step. Steady state analysis of IM has some similarities with that of MCs and has been discussed for various scenarios in [176], [111]. For a more detailed discussion on the IM please refer to [176], [111].

6.2.5.1 Constraint-based Influence Model

The work in [177] extends the original IM to a constraint or rule-based influence framework such that the influences among the nodes can dynamically get activated and deactivated depending on the state of the system. Also, as explained in [177], influences can change the state of the influencer as well (e.g., EV charging stations can transition from overload to normal due to sending load to another charging station). The study in [177] specifically defined a constraint matrix $\mathbf{C}$, where the entry c_{ij} for $i, j \in \mathcal{CS} \cup \mathcal{S}$ (or $i, j \in \mathcal{CS} \cup \mathcal{S} \cup \mathcal{B}$) is a binary variable specifying whether node i gets influenced by node j or not. In particular, $c_{ij} = 1$ indicates that node i gets influenced by node j and $c_{ij} = 0$ indicates otherwise. Moreover, each node always influences itself based on its

internal MC (i.e., $c_{ii} = 1$ for all $i \in \mathcal{CS} \cup \mathcal{S}$ in the case of the integrated system model ($i \in \mathcal{CS} \cup \mathcal{S} \cup \mathcal{B}$ in the case of the extended integrated system model)). As explained in [177], one can define the value of c_{ij} according to boolean logic to capture the rules of interactions in the network. In other words, c_{ij}s are functions of the state of the nodes. For instance, when a charging station in the EV charging infrastructure is in overloaded state and based on G_p (or G_e), it has a link to another station, which is underloaded, the influence over that link should get activated to motivate the EV users to travel from the overloaded state to underload state. These types of rules can be specified using boolean functions such as the following examples. Function $c_{ij} = s_{i3}s_{j1} + s_{i2}s_{j1}$, where $i, j \in \mathcal{CS}$ specifies the rules that can be applied to the transport layer of the model to show influences from charging station j to charging station i. Specifically, a transport node i will receive influence from transport node j if node i is underloaded and node j is overloaded or if node i is normal and node j is overloaded. Also, the power substations receive influences from the charging stations because overloaded charging stations can cause a power substation to go to overloaded state. Example of boolean function describing this rule is $c_{k\ell} = s_{k1} \prod_{j \in CS_{S_k}} s_{ji} + s_{k2} \prod_{j \in CS_{S_k}} s_{ji}$, where $k \in \mathcal{S}$ and $\ell \in \mathcal{CS}$ and $CS_{S_k} \subseteq \mathcal{CS}$ is the set of charging stations connected to the power substation k. Specifically, a power substation, say k will receive influence from charging station ℓ when all the charging stations connected to the power substation are overloaded. As a power station is generally built with a capacity to accommodate large demand, the power substation will go to a stressed state provided that all the influence links connected to it are activated. This is just one example of influence rule and other conditions to specify the rules are also possible.

Note that as the goal of the integrated study of these two systems is to increase the probability of having power substations in normal conditions and charging stations not overloaded, the

interaction rules defined in $\mathbf{C}$ should support this goal. In order to achieve this goal, the influences among the charging stations should be engineered such that it forces the whole system toward desirable states. The second type of influence, which is from the charging station to power substations cannot be engineered and we assume that when the charging stations, which are receiving power from substations, are overloaded they influence (increase the likelihood) the substation to transit to a stressed state (similarly for the load buses influencing substations in the case of the extended integrated system model).

In [177], the constraint matrix $\mathbf{C}$ and the influence matrix $\mathbf{D}$ are used to define the constraint-based influence matrix denoted by $\mathbf{E}$, as

$$\mathbf{E} = \mathbf{D} \circ \mathbf{C} + \mathbf{I} \circ (\mathbf{D} \times (\mathbf{1} - \mathbf{C}')), \tag{6.3}$$

where $\circ$ is the Hadamard product (aka entrywise product), $\mathbf{1}$ is the matrix with all elements equal to 1 and $\mathbf{C}'$ is the transpose of matrix $\mathbf{C}$. Using $\mathbf{E}$, the IM-based state evolutions can be summarized as

$$\mathbf{H} = \mathbf{E}' \otimes \{\mathbf{A}_{ij}\}, \tag{6.4}$$

and $\mathbf{p}[t+1] = \mathbf{S}[t]\mathbf{H}$.

As discussed in [177], this formulation may or may not allow the asymptotic analysis of the behavior of the system. However, regardless of the existence of the analytical solution, this model can be used for Monte-Carlo simulation of the behavior of the system in order to study how influences and interactions affect the state of the whole system. Based on this formulation, as the state of the system varies in time, various sets of influences get activated. Note that in IM, when

a node influences another node, it may result in state change for the influenced node based on the adjusted transition probabilities that are captured through $\mathbf{H}$ and the formulation of $\mathbf{p}[t+1]$. As such, an activated influence in our model increases the probability of transitioning to a normal state for an underloaded charging station due to receiving load from the influencer (based on our definition of influence). In the real-world, proper incentives for the users are needed to make sure that the influences occur (transfer of load from one charging station to another). As such, to achieve the goal of the system which is increasing the probability of normal states, we use the status of the influence links (active or inactive) to guide the charging price design as discussed next.

Algorithm 2 Algorithm for Price Assignment to Charging Stations

1: **Input** : Graph of active influences, $G_t(\mathcal{CS}, E_a(t))$. A maximum electricity price limit $\mathbf{A}$ and a reduction factor in price, α.
2: **Output** : Charging price in each charging station in $\mathcal{C}$ such that the price of the influencer station is higher than the influenced station.
3: Calculate the topological sort $\mathbf{T}$ for G_t.
4: **for** i=1 to $|\mathcal{C}|$ **do**
5: **if** $|I(\mathbf{T}(i))| = 0$ **then**
6: Price($\mathbf{T}(i)$) = $\mathbf{A}$
7: **else**
8: Price($\mathbf{T}(i)$)=$\sum_{j \in I(\mathbf{T}(i))}$ Price(j)/$|I(\mathbf{T}(i))|$
9: **end if**
10: **end for**
11: Return Price.

6.3 Designing Charging Prices for Improved Reliability

The models described in the previous section need an external factor in real-world scenarios to provoke an EV user to travel from one charging station to another for charging (i.e., activating the described influence between charging stations in real-world). This external factor can be in terms of incentives or hampers that an EV user may get if they move from one cell to another. A good incentive would be lower charging prices (whenever the influence should be active) in the

station, which should receive some load. The lower prices can motivate the EV users to move from their currently occupied cell to the other station. However, not every EV user will respond to such incentives in the same way and thus, not every user will travel from the initial cell. Particularly, the probabilistic nature of the IM helps in capturing the random behavior of the users. Intuitively, the higher the influence strength the more we expect that the users travel to the other station, which can help in characterizing the price reduction that is needed. A key point to notice is that the cyber layer plays a key role in letting the desired influences to occur and to let the system identify its next states based on IM. Specifically, the cyber layer should communicate the lower charging price only to the users in the initial cell of concern. Otherwise, if the reduced price is communicated in the system globally and all the EV users in the city know about the reduced price in a station, this will activate influences among neighbor stations (neighbors are defined as according to G_p or G_e) that should not be activated according to the IM model. Thus, in order to only activate the influences that the IM model identifies for leading the system to a more balanced system in each step, the cyber layer plays a key role in communicating the prices to the right EV users based on their location.

In our model based on IM, whenever the set of activated influence links varies, we need to identify new set of prices for each station such that if station say i has an active influence link to station j, then the price at station i should be higher than that of station j. To identify the set of prices that satisfy this condition in the whole system, we apply Algorithm 2. This algorithm is similar to a constrained graph coloring problem. However, the problem of price assignment to the stations based on the above constraint is solvable with complexity $O(|\mathcal{CS}| + |E|)$, which is because the graph of active influences denoted by $G_t(\mathcal{CS}, E_a(t))$ and obtained from simulation of IM at step

t is a directed and acyclic graph (DAG) (note that $E_a(t) \subseteq E$, also note that $E_a(t)$ does not include self-influences as they do not affect the pricing). This property is due to the rule set with the goal of balancing the load in the system, which never result in a cycle in the graph of active influences. In other words, the rule set in the model is very important to ensure that the load is not circulating in the system and purposely directed to the proper charging stations. Algorithm 2 for price assignment uses a topological sort of the graph and then assigns the prices based on the identified order such that the prices ensure that the stations appearing later in the topological sort have lower prices (as they should receive influences or loads). In our algorithm, we consider a maximum price limit of A and each price reduction occurs by a constant α. The values of A and α are considered fixed for simplicity, but can be variable and adjusted based on other factors in the system. In this algorithm, function $I(.)$ receives a node and returns the set of nodes, which influences the input node.

6.4 Evaluation and Results

In order to demonstrate the process of assigning prices to the charging stations dynamically as the system evolves in time, while trying to lead the EV charging and power system to more balanced states, we use the following two example networks:

1. Figure 6.3 with 12 charging stations, which receive their energy from the two substations.

2. Figure 6.4 with 12 charging stations, three substations, in which two of them directly deliver electricity to the charging stations, and two load buses.

6.4.1 Transition Probabilities of Charging Station Nodes

For both these example networks, we use a data driven approach to extract some of the

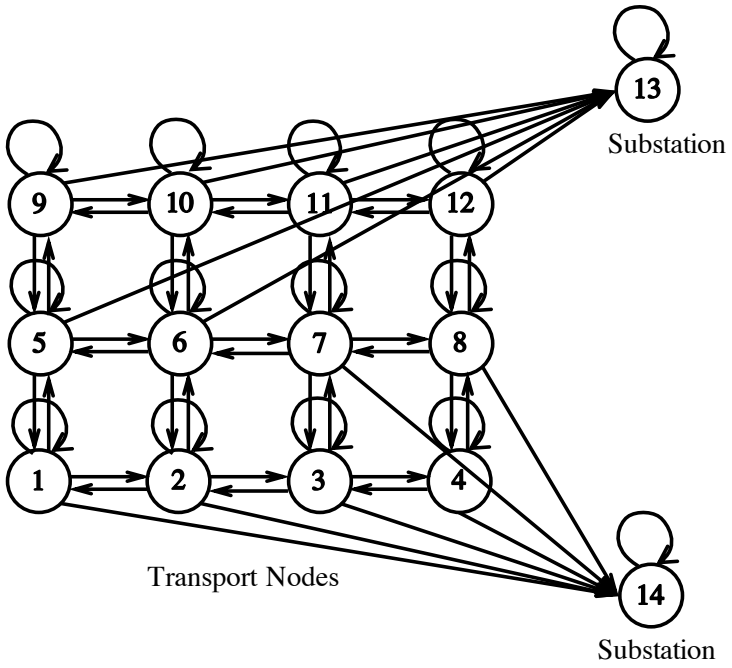


Figure 6.3: Integrated charging and power infrastructures model with twelve charging stations and two substations. (i.e., graph G_p).

parameters of the IM using available data sets of traffic information. Specifically, we used the taxi data in [178], which contains GPS trajectories of 536 taxis in San Francisco, California from May 17, 2009 - July 10, 2009 to estimate the self-influence transition probabilities of the twelve EV charging nodes, denoted by $\mathbf{A}_{ii}$s. An example of $\mathbf{A}_{ii}$ based on the dataset is shown in Equation (6.5), where rows and columns are ordered from overload (O) to normal (N) and then underload (U):

$$\mathbf{A}_{ii} = \begin{array}{c} \\ O \\ N \\ U \end{array} \begin{array}{c} \begin{array}{ccc} O & N & U \end{array} \\ \begin{pmatrix} 0.89473684 & 0.1052632 & 0.00000000 \\ 0.07262570 & 0.8770950 & 0.05027933 \\ 0.07142857 & 0.2142857 & 0.71428571 \end{pmatrix} \end{array} \tag{6.5}$$

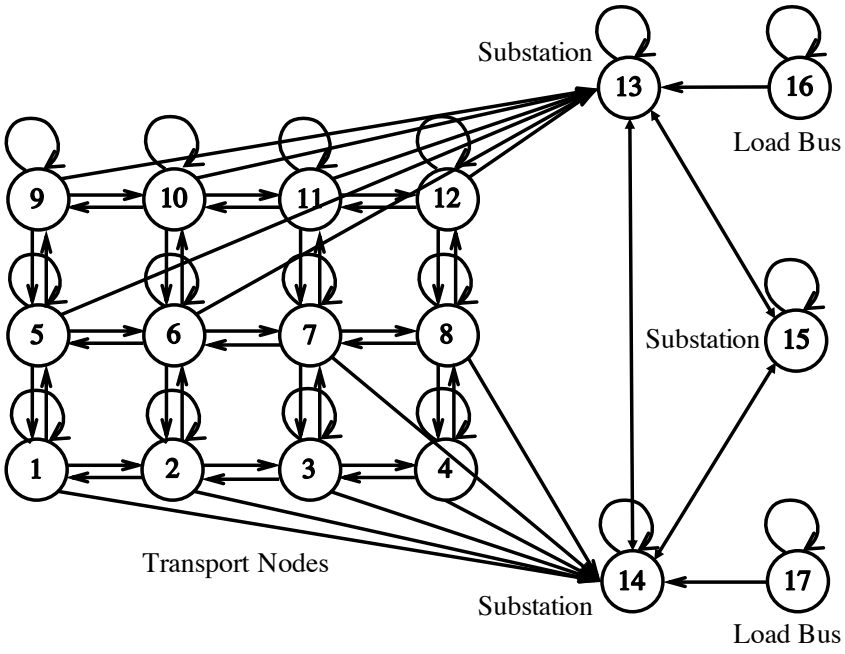


Figure 6.4: Integrated charging and power infrastructures model with twelve charging stations, three substations and two load buses. (i.e., graph G_e).

In addition to $\mathbf{A}_{ii}$s, which characterize the internal dynamics of each charging station, we also need to consider inter-state transition probabilities $\mathbf{A}_{ij}$s to specify how the influences between two stations result in state transitions. An example of $\mathbf{A}_{ij}$ is shown in Equation (6.6), in which each column specifies the probability of transition to overload, normal, and underload, respectively, depending on each row, which specifies the state of the influenced node. For simplicity and due to lack of detailed information in the datasets to characterize this matrix for all cells, we have simplified this matrix to have equal transition probabilities independent of the state of the influenced node (i.e., the same rows). Based on our model and the rules of influences, in order to lead the systems to balanced states, a charging station only tries to send load to another charging station if the other station is not overloaded. As such, the last row of the matrix in Equation (6.6) does not play a role in the analysis.

$$\mathbf{A}_{ij} = \begin{array}{c} \\ O \\ N \\ U \end{array} \begin{array}{c} \begin{array}{ccc} O & N & U \end{array} \\ \begin{pmatrix} 0.2 & 0.5 & 0.3 \\ 0.2 & 0.5 & 0.3 \\ 0.2 & 0.5 & 0.3 \end{pmatrix} \end{array} \tag{6.6}$$

6.4.2 Transition Probabilities of Power Substation Nodes

We discuss the transition probabilities of the states of the power substations in the following two categories for integrated system model and the extended integrated system model.

6.4.2.1 Integrated System Model

For the integrated system model, an example of self-influence transition probabilities $\mathbf{A}_{ii}$ for power substations is shown in Equation (6.7), where rows and columns are ordered from normal (N) to stressed (S). We do not consider the intra-power interactions and matrix $\mathbf{A}_{ii}$ compensates for the missing information of the intra system interactions that can cause state changes.

$$\mathbf{A}_{ii} = \begin{array}{c} \\ N \\ S \end{array} \begin{array}{c} \begin{array}{cc} N & S \end{array} \\ \begin{pmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{pmatrix} \end{array} \tag{6.7}$$

According to Equation (6.7), when the system is stressed (i.e., the second row on the matrix in (6.7)), there is an equal chance to get into normal or stressed state based on internal dynamics. As part of influences in IM-based model, whenever charging stations go back to normal or underloaded states, they can externally help the power substation to transit back to the normal state.

6.4.2.2 Extended Integrated System Model

Similarly, in the case of the extended integrated system model, an example of inter-state transition probabilities $\mathbf{A}_{ij}$ for power substations is shown in Equation (6.8), where rows and columns are ordered from normal (N) to stressed (S) to capture the influence among substations depending on their normal or stressed states:

$$\mathbf{A}_{ij} = \begin{array}{c} \\ N \\ S \end{array} \begin{array}{c} \begin{array}{cc} N & S \end{array} \\ \begin{pmatrix} 0.4 & 0.6 \\ 0.2 & 0.8 \end{pmatrix} \end{array} \tag{6.8}$$

We assume that the internal dynamics of each power substation, i.e., self-influence transition probabilities $\mathbf{A}_{ii}$ can be captured by an identity matrix implying that the internal state of the substations can only change due to influences (i.e., without influences, the probability of transition to the other state is zero). To model the stochastic residential/commercial loads, we consider the $\mathbf{A}_{ii}$ shown in Equation (6.9) to consider the stochastic dynamics of loads (which can also be derived using a data-driven approach using load profile datasets).

$$\mathbf{A}_{ii} = \begin{array}{c} \\ O \\ N \\ U \end{array} \begin{array}{c} \begin{array}{ccc} O & N & U \end{array} \\ \begin{pmatrix} 0.5 & 0.3 & 0.2 \\ 0.25 & 0.5 & 0.25 \\ 0.1 & 0.4 & 0.5 \end{pmatrix} \end{array} \tag{6.9}$$

The set of rules for this study can be described as: (1) for charging stations, node i gets influenced by node j if and only if node i is underloaded and node j is overloaded or node i is in

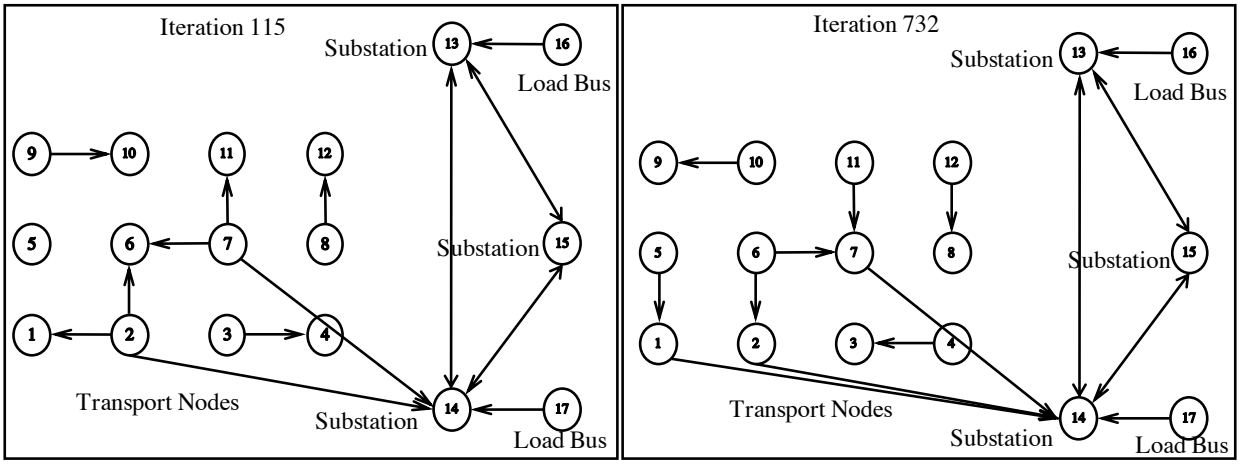


Figure 6.5: Two samples of activated influence links for the extended integrated model shown in Figure 6.4.

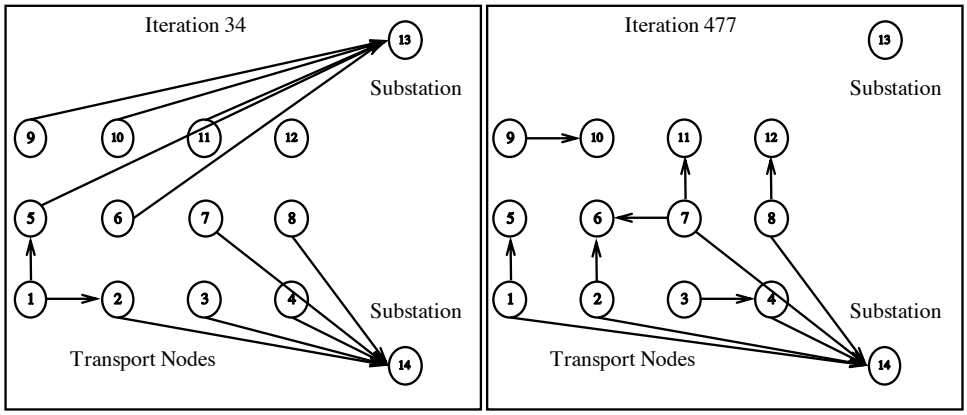


Figure 6.6: Two samples of activated influence links for the extended integrated model shown in Figure 6.3.

normal state and node j is overloaded, and (2) for the influences between the power substations and the charging stations, the power substation gets influenced by a charging station only (in the case of the integrated system model) or charging station as well as load bus (in the case of the extended integrated system model), if the power station is normal and the charging stations or load buses receiving power service from the substation are overloaded or if the power substation is stressed and the charging stations are normal or underloaded.

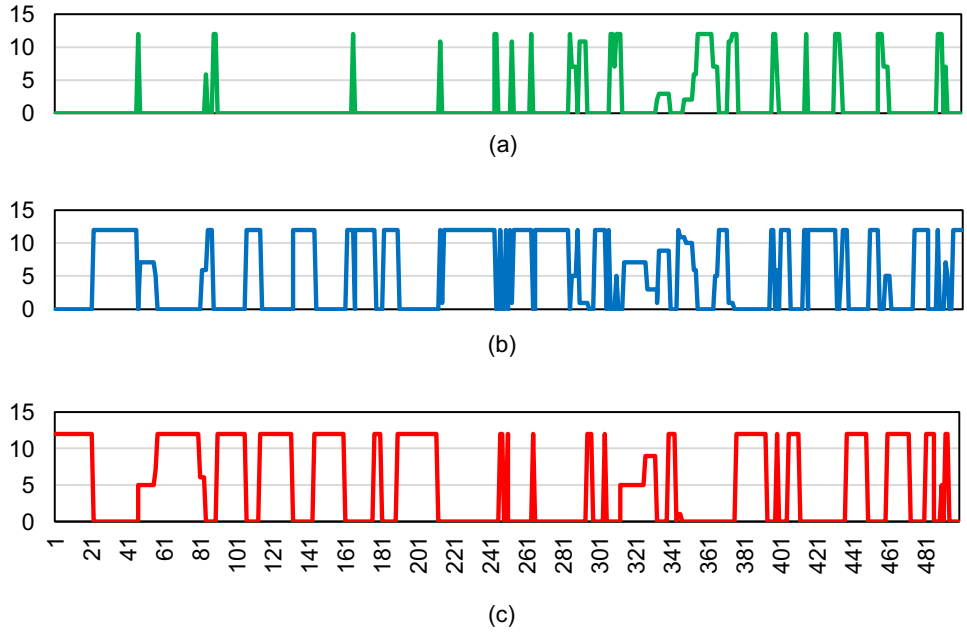


Figure 6.7: Number of charging stations in each iteration in: (a) underloaded, (b) normal, and (c) overloaded states.

6.4.3 Numerical Results

As mentioned earlier, based on the state of the components in the system, the influences among nodes may get activated and deactivated. In Figure 6.5, we show two samples of active influence graphs for the extended integrated system model shown in Figure 6.4. The activated links between charging stations suggest that the load should be transferred from one station to the

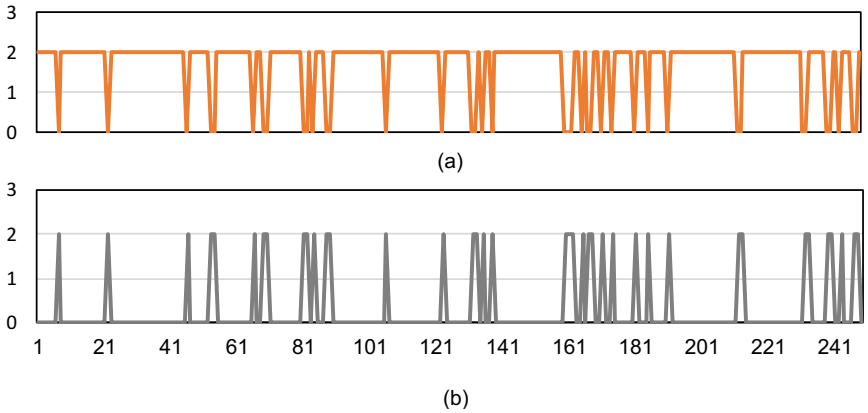


Figure 6.8: Number of power substations in each iteration in: (a) normal, and (b) stressed states.

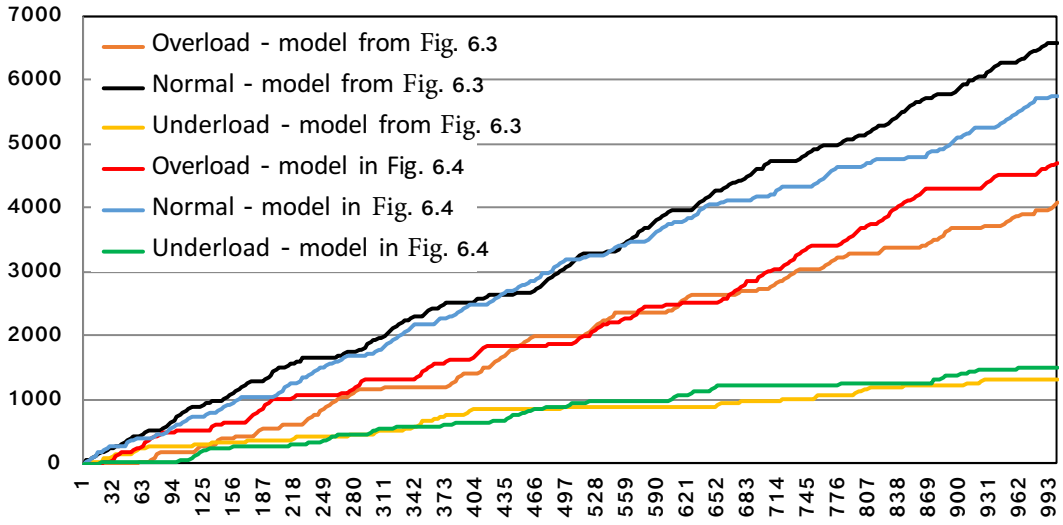


Figure 6.9: Aggregated state distribution for overloaded, normal and underloaded states for charging stations for the system model shown in Figure 6.3 and Figure 6.4.

station on the end of the directed link. Also, Figure 6.6 shows the integrated system model and the activated links for the setting defined in Equation (6.7).

Using the integrated system model shown in Figure 6.3 with the transition probabilities for charging stations discussed in Section 6.4.1 and transition probabilities for power substations discussed in 6.4.2.1, the set of activated influences in each iteration prompts a change of state in the charging stations and power substations, as shown in Figures 6.7 and 6.8. Specifically, Figures 6.7 (a), (b), and (c) show the distribution of the number of charging stations in underload, normal and overload states in each iteration. Similarly, Figures 6.8 (a) and (b) show the distribution of the number of power substations in normal and stressed states in each iteration. Similar observations are obtained for the extended integrated system model shown in Figure 6.4 with the transition probabilities for charging stations discussed in Section 6.4.1 and transition probabilities for power substations as well as residential/commercial loads discussed in 6.4.2.2.

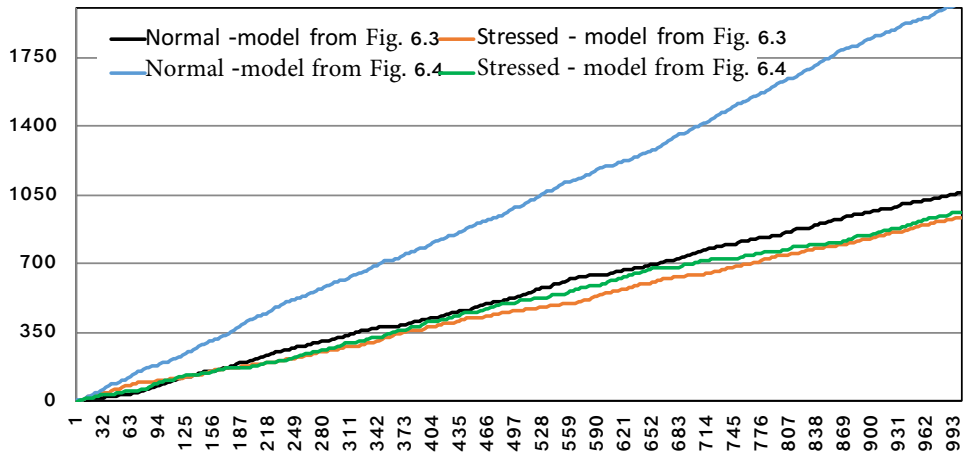


Figure 6.10: Aggregated state distribution for normal and stressed states for the power substations for the system model shown in Figure 6.3 and Figure 6.4.

Although the distributions are fluctuating, it can be observed from Figure 6.9 and Figure 6.10 that the aggregated behavior of the system is helping the system towards being balanced (e.g., the likelihood of normally loaded charging stations and normal power substations is higher than other states). The results in Figures 6.9, 6.10, 6.11, and 6.12 are obtained over 1000 steps of the IM simulation. Figure 6.11 shows the state distribution of the charging stations and power substations with various initial states for the extended integrated system model. Similar observations are obtained for the integrated system model.

6.4.4 Rules of Interactions

An important aspect of the influence model is the set of rules that specify how the nodes should interact and influence each other. To show how the rules of the interactions affect the behavior of the system, here, we have considered other influence rules similar to the rules of interactions defined in [177] as follows:

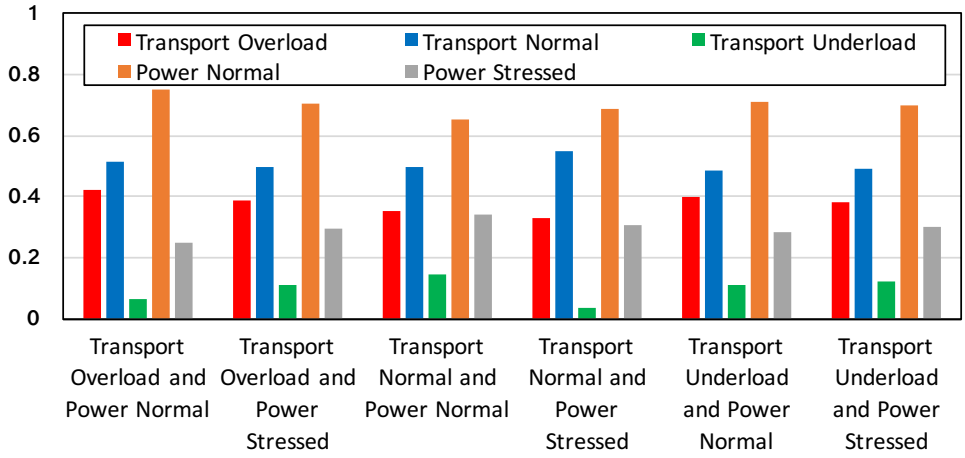


Figure 6.11: State distribution of charging stations and power substations with various initial states for the components using the extended integrated system model in Figure 6.4.

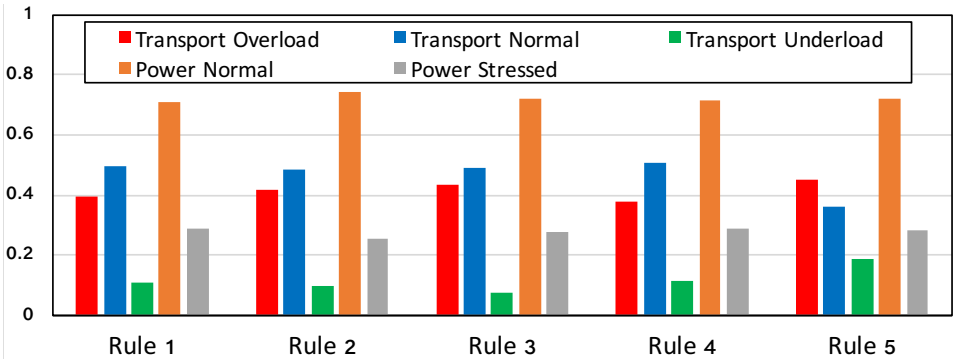


Figure 6.12: State distribution of charging stations and power substations with all charging stations and residential/commercial loads initially overloaded and all power substations initially in normal states for different rules of interactions using the extended integrated system model in Figure 6.4.

1. *Rule 1:* Node i gets influenced by node j if and only if (iff) node i is underloaded and node j is overloaded or node i is in normal state and node j is overloaded.

2. *Rule 2:* Node i gets influenced by (receives workload from) node j iff node i is underloaded and node j is overloaded.

3. *Rule 3:* Node i gets influenced by node j iff node i is underloaded and node j is overloaded or node i is underloaded and node j is in normal state.

4. *Rule 4:* Node i gets influenced by node j iff either node i is underloaded and node j is overloaded, node i is underloaded and node j is in normal state or node i is in normal state and node j is overloaded.

5. *Rule 5:* Node i gets influenced by node j iff either node i is underloaded and node j is overloaded, node i is underloaded and node j is in normal state, node i is in normal state and node j is overloaded or node i is in normal state and node j is in normal state too.

Note that these rules only focus on the interactions/influences among the charging stations. Figure 6.12 shows the state distribution of nodes with all charging stations and residential/commercial loads initially overloaded and all power substations initially normal for different rules applied to the extended integrated system model in Figure 6.4. It can be seen that rule 5 performs the worst among all five rules as the number of overloaded charging stations are higher compared to other rules for the extended integrated system model shown in Figure 6.3. Similar results are observed for the integrated system model shown in Figure 6.4.

6.4.5 Optimal Charging Prices

To design the incentives that enable influences and lead to the above discussed results, we need to design the prices for each charging station. To do so, we have used Algorithm 2 over the active influence graph obtained at each step of the simulation, whenever there is a change in the active influence graphs. Note that Algorithm 2 receives graphs similar to the ones shown in Figure 6.5 where the self-edges are omitted. The price assignment based on this algorithm at each station is shown in Table 6.1 for the extended integrated system model (which considers intra-power system dynamics as well as residential/commercial loads), for the sample steps of our simulation (with Rule

Table 6.1: Charging prices in the EV charging stations over various iterations for the extended integrated system model.

Charging Stations												
Iteration	1	2	3	4	5	6	7	8	9	10	11	12
1	A	A	A	A	A	A	A	A	A	A	A	A
99	A	A	A	A-α	A-α	A-α	A	A	A	A-α	A-α	A-α
368	A-α	A	A	A	A	A	A	A	A	A	A	A
562	A-α	A	A	A	A	A	A	A	A	A	A	A
600	A	A-α	A-α	A	A	A	A-α	A-α	A-α	A	A	A
653	A-α	A	A	A-α	A	A-α	A	A	A	A-α	A-α	A-α
779	A	A-α	A-α	A	A	A	A-α	A-α	A-α	A	A	A
987	A	A	A	A-α	A-α	A-α	A	A	A	A-α	A-α	A-α
993	A	A-α	A	A	A-α	A	A-α	A	A	A	A	A-α

1). It can be observed from the table that all twelve stations have the same price of A initially but the prices vary over the network as the stochastic dynamics of the system change the states of the nodes.

We showed our study of collaborative pricing solution between the EV charging and electric infrastructures based on our IM-based model. Key takeaways from our results include: (1) by designing proper rules of interactions among the integrated systems, the load distribution can be improved in both systems, (2) the pricing assignment based on the obtained active influence graph enables the implementation of appropriate influences and (3) the intra-system and inter-system influences and dynamics both affect the overall behavior and thus, change the obtained pricing schemes.

6.5 Summary and Conclusions

In this chapter, we discussed the interdependence of critical infrastructures of smart cities and thus, the importance of considering collaborative solutions among them for designing services.

To do so, we proposed a synergistic approach towards modeling and analysis of critical interdependent infrastructures that enabled capturing the state and stochastic dynamics of inter and intra-system interactions. To demonstrate the benefit of collaborative solutions, in this dissertation, we focused on interdependent EV charging and power infrastructures and developed an integrated framework for modeling their interactions based on influence model, which is a networked Markov chain framework. We also proposed an algorithm, which assigned prices to charging stations based on the set of active links that lead to more balanced systems. We discussed the role of the cyber infrastructure in enabling this pricing scheme, which considers the state of both charging stations and power systems. The work presented in this chapter is an effort towards stimulating collaboration among various critical infrastructures and analyzing them using integrated models to develop collaborative solutions for smart cities.

Chapter 7: Conclusions and Future Directions

This dissertation is an effort towards improving the reliability of the power grid and its interdependent infrastructures and illustrated several experiments to show the importance of using interaction graphs in such analyses. In power grids, the complex phenomena of cascading failures, which can propel a grid towards large blackouts, was evaluated using community structures within interaction graphs. This dissertation provides key insight about the underlying interactions among the components of the system during cascading failures and that protection of critical components identified using the community structures can help in reducing the risk of blackouts. Further, community structures were used to formulate a Markov chain framework for characterizing and predicting the size of cascades; in which, the results revealed that the community structure-based Markov Chain produced heavy tailed cascade size distributions, typically found in historical data and previous studies of cascading failures [20]. These results suggest that community structures within interactions graphs bear critical information about the cascade process in power grids. Similarly, reliability of the interdependent power grid and EV charging infrastructure was evaluated using the influence model, a networked Markov chain framework, to model the stochastic interactions shared among the infrastructures. We observed that charging prices designed to motivate EV drivers to travel to appropriate locations could lead to balanced states in both systems. Together, these studies illustrate that global properties of various systems studied using interaction graphs can be useful in understanding and improving the reliability of the respective infrastructures and help in

designing dependable and resilient infrastructures. Future work arising from the present studies in this dissertation are discussed next.

In this dissertation, two different data-driven interaction graphs were constructed using influence-based and correlation-based techniques. These graphs heavily rely on cascade data and as such, interactions/influences among components and their strengths may be overestimated or underestimated based on the applied method and the dataset itself. In future, these techniques can be further improved to extract useful statistics and information without overestimation/underestimation problems. Furthermore, these graphs were constructed using simulation data and further analyses can include historical datasets as well. This process can validate the technique used in constructing the graph by showing historical trends such as power law behavior in cascade size distributions.

Also, our studies showed that operating settings, such as the power grid loading level, affected the constructed interaction graphs and consequently, led to different criticality of components for different operating settings. For real-time mitigation of cascading failures, a single and computationally inexpensive interaction graph that models all possible operating setting conditions needs to be created and is an interesting challenge ahead for researchers.

While studying the role of community structures in the interaction graphs, our studies relied on thresholding the dense interactions/influences in the interaction graphs to obtain community structures that could model both microscopic and macroscopic properties relative to the scale of the graph. This process essentially removed less dominant interactions that acted as noise. In future, identifying techniques to reduce the number of interaction links in the graphs can be useful for reducing the computational complexity of constructing the graphs and finding the community structures. We also used two specific community detection techniques, but many other techniques

can be used for identifying different types of patterns and structures in the interaction graphs. This can cross validate the most efficient structure for cascading failure studies.

One of the techniques used for verifying the criticality of the components identified using the community-based centrality measure was the power simulation-based technique, in which, the transmission capacities of the lines were increased for handling larger power flows. This technique can substantially reduce the risk of large blackouts, as shown in our simulations. In future, other cost effective mitigation mechanisms that can be implemented in real time, such as switching off transmission lines during propagation of failures itself should be studied using the community structures by removal of critical components such as bridge or overlap nodes from the interaction graphs.

In our community structure-based Markov chain framework, we numerically analyzed the cascade size distributions based on the trapping property of community structures. However, a complete analytical characterization of the Markov chain framework can be extremely useful in predicting cascade sizes and mitigating the risk of large blackouts in real time for power grids of larger scales and sizes. This is an open research problem and needs to be undertaken in future.

While this dissertation focused on analyzing cascading failures using community structures of standalone power grids, additional analyses of community structures of interdependent infrastructures subjected to cascading failures can model complex inter and intra influences/interactions in the infrastructures. Further, the community-based centrality metric can identify critical components in the interdependent infrastructures. And the general verification techniques such as the data-driven, SI simulation, and graph-based efficiency analysis can be used to verify criticality of components identified in the interdependent infrastructures. Further, the community-based Markov chain framework can also be extended to include additional infrastructures.

References

[1] Nerc blackout-report-presentation-11-19-03. https://www.nerc.com/pa/rrm/ea/August142003BlackoutInvestigationDL/Blackout-Report-Presentation-11-19-03.ppt. Accessed: 03-15-2020.

[2] UN DESA. World urbanization prospects: The 2018 revision (st/esa/ser. a/420). 2019.

[3] Upama Nakarmi and Mahshid Rahnamay-Naeini. Analyzing power grids' cascading failures and critical components using interaction graphs. In *IEEE PESGM*, pages 1–5, 2018.

[4] Upama Nakarmi, Mahshid Rahnamay-Naeini, and Hana Khamfroush. Critical component analysis in cascading failures for power grids using community structures in interaction graphs. *IEEE Trans. on Net. Sci. & Eng.*, 2019.

[5] Upama Nakarmi and Mahshid Rahnamay-Naeini. Towards integrated infrastructures for smart city services: A story of traffic and energy aware pricing policy for charging infrastructures. In *International Conference on Smart Cities and Green ICT Systems (SmartGreens)*. IEEE, 2017.

[6] Upama Nakarmi and Mahshid Rahnamay-Naeini. An influence-based model for smart city's interdependent infrastructures: Application in pricing design for ev charging infrastructures. In *Smart Cities, Green Technologies, and Intelligent Transport Systems*, pages 111–130. Springer, 2017.

[7] U-CPSOT Force, S Abraham, et al. Us-canada power system outage task force. *Final Report on the August 14, 2003 Blackout in the United States and Canada: Causes and Recommendations*, 2004.

[8] Robert Bass and Nicole Zimmerman. Impacts of electric vehicle charging on electric power distribution systems. Technical report, Transportation Research and Education Center (TREC), Portland, OR, USA, 2013.

[9] Chandler E Hatton, Satish K Beella, JC Brezet, and YC Wijnia. Charging stations for urban settings the design of a product platform for electric vehicle infrastructure in Dutch cities. *World Electric Vehicle Journal*, 3(1):134–146, 2009.

[10] Fang He, Di Wu, Yafeng Yin, and Yongpei Guan. Optimal deployment of public charging stations for plug-in hybrid electric vehicles. *Transportation Research Part B: Methodological*, 47:87–101, 2013.

[11] Woongsup Lee, Lin Xiang, Robert Schober, and Vincent WS Wong. Electric vehicle charging stations with renewable power generators: A game theoretical analysis. *IEEE Transactions on Smart Grid*, 6(2):608–617, 2015.

[12] Jian Liu. Electric vehicle charging infrastructure assignment and power grid impacts assessment in Beijing. *Energy Policy*, 51:544–557, 2012.

[13] Imran Rahman, Pandian M Vasant, Balbir Singh Mahinder Singh, and M Abdullah-Al-Wadud. Intelligent energy allocation strategy for PHEV charging station using gravitational search algorithm. In *AIP Conference Proceedings*, volume 1621, pages 52–59, 2014.

[14] Wilfred W Recker and Jee Eun Kang. An activity-based assessment of the potential impacts of plug-in hybrid electric vehicles on energy and emissions using one-day travel data. Technical report, University of California Transportation Center, Berkeley, CA, USA, 2010.

[15] Ryan Liu, Luther Dow, and Edwin Liu. A survey of PEV impacts on electric utilities. In *Innovative Smart Grid Technologies (ISGT)*, pages 1–8. IEEE, 2011.

[16] Jayakrishnan R Pillai and Birgitte Bak-Jensen. Impacts of electric vehicle loads on power distribution systems. In *IEEE Vehicle Power and Propulsion Conference*, pages 1–6. IEEE, 2010.

[17] Ramteen Sioshansi. OR Forum–modeling the impacts of electricity tariffs on plug-in hybrid electric vehicle charging, costs, and emissions. *Operations Research*, 60(3):506–516, 2012.

[18] Sebastian Wagner, Markus Götzinger, and Dirk Neumann. Optimal location of charging stations in smart cities: A points of interest based approach. In *International Conference on Information Systems (ICIS)*, 2013.

[19] J Medina. Human error investigated in california blackout's spread to six million. *The New York Times*, 26, 2011.

[20] Ian Dobson, Benjamin A Carreras, Vickie E Lynch, and David E Newman. Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization. *Chaos: An Interdis. Jrnl. of Nonlinear Sci.*, 17(2):026103, 2007.

[21] Reliability terminology, NERC. http://www.nerc.com/AboutNERC/Documents/Terms\%20AUG13.pdf. Accessed: 03-17-2020.

[22] Marianna Vaiman et al. Risk assessment of cascading outages: Methodologies and challenges. *IEEE Transactions on Power Systems*, 27(2):631, 2012.

[23] Ross Baldick, Badrul Chowdhury, Ian Dobson, Zhaoyang Dong, Bei Gou, David Hawkins, Henry Huang, Manho Joung, Daniel Kirschen, Fangxing Li, et al. Initial review of methods for cascading failure analysis in electric power transmission systems (ieee pes cams task force on understanding, prediction, mitigation and restoration of cascading failures). In *In Proceedings of the IEEE Power & Energy Soc. Gen. Meet. (PESGM) - Conversion and Delivery of Electrical Energy in the 21st Century*, pages 1–8, 2008.

[24] Gaoqi Liang, Steven R Weller, Junhua Zhao, Fengji Luo, and Zhao Yang Dong. The 2015 ukraine blackout: Implications for false data injection attacks. *IEEE Transactions on Power Systems*, 32(4):3317–3318, 2016.

[25] Hengdao Guo, Ciyan Zheng, Herbert Ho-Ching Iu, and Tyrone Fernando. A critical review of cascading failure analysis and modeling of power system. *Renewable and Sustainable Energy Reviews*, 80:9–22, 2017.

[26] Janusz Bialek, Emanuele Ciapessoni, Diego Cirio, Eduardo Cotilla-Sanchez, Chris Dent, Ian Dobson, Pierre Henneaux, Paul Hines, Jorge Jardim, Stephen Miller, et al. Benchmarking and validation of cascading failure analysis tools. *IEEE Transactions on Power Systems*, 31(6):4887–4900, 2016.

[27] Pierre Henneaux, Emanuele Ciapessoni, Diego Cirio, Eduardo Cotilla-Sanchez, Ruisheng Diao, Ian Dobson, Anish Gaikwad, Stephen Miller, Milorad Papic, Andrea Pitto, et al. Benchmarking quasi-steady state cascading outage analysis methodologies. In *2018 IEEE International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, pages 1–6. IEEE, 2018.

[28] Mahshid Rahnamay Naeini. *Stochastic dynamics of cascading failures in electric-cyber infrastructures*. Phd dissertation, Department of Electrical and Computer Engineering, University of New Mexico, NM, USA, 2014.

[29] Mahshid Rahnamay-Naeini, Zhuoyao Wang, Nasir Ghani, Andrea Mammoli, and Majeed M Hayat. Stochastic analysis of cascading-failure dynamics in power grids. *IEEE Transactions on Power Systems*, 29(4):1767–1779, 2014.

[30] J Duncan Glover, Mulukutla S Sarma, and Thomas Overbye. *Power System Analysis & Design, 4th edition*. Cengage Learning, 2008.

[31] Ian Dobson, Benjamin A. Carreras, Vickie E. Lynch, and David E. Newman. Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 17(2):026103, 2007.

[32] Quan Chen and Lamine Mili. Composite power system vulnerability evaluation to cascading failures using importance sampling and antithetic variates. *IEEE Transactions on Power Systems*, 28(3):2321–2330, 2013.

[33] Marija Bockarjova and Goran Andersson. Transmission line conductor temperature impact on state estimation accuracy. In *IEEE Lausanne Power Tech*, pages 701–706. IEEE, 2007.

[34] Mahshid Rahnamay-Naeini and Majeed M Hayat. Impacts of operating characteristics on sensitivity of power grids to cascading failures. In *Power and Energy Society General Meeting (PESGM)*, pages 1–5. IEEE, 2016.

[35] Zhuoyao Wang, Mahshid Rahnamay-Naeini, Joana M Abreu, Rezoan A Shuvro, Pankaz Das, Andrea A Mammoli, Nasir Ghani, and Majeed M Hayat. Impacts of operators' behavior on reliability of power grids during cascading failures. *IEEE Transactions on Power Systems*, 33(6):6013–6024, 2018.

[36] Allen J Wood and B Wollenberg. Power generation operation and control—2nd edition. *Fuel and Energy Abstracts*, 37(3):195, 1996.

[37] Ray Daniel Zimmerman, Carlos Edmundo Murillo-Sánchez, and Robert John Thomas. MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Transactions on Power Systems*, 26(1):12–19, 2011.

[38] Clara Stegehuis, Remco van der Hofstad, and Johan SH van Leeuwaarden. Epidemic spreading on complex networks with community structures. *Scientific Reports*, 6:29748, 2016.

[39] Marcel Salathé and James H Jones. Dynamics and control of diseases in networks with community structure. *PLoS Computational Biology*, 6(4):e1000736, 2010.

[40] Kai Gong et al. An efficient immunization strategy for community networks. *PloS one*, 8(12):e83489, 2013.

[41] J-P Onnela, Jari Saramäki, Jorkki Hyvönen, György Szabó, David Lazer, Kimmo Kaski, János Kertész, and A-L Barabási. Structure and tie strengths in mobile communication networks. *Proceedings of the National Academy of Sciences*, 104(18):7332–7336, 2007.

[42] Chang Zinan and Shao Fei. Epidemic spreading in weighted homogeneous networks with community structure. *Computer Modelling and New Technologies*, 18(11), 2014.

[43] Paul DH Hines, Ian Dobson, and Pooya Rezaei. Cascading power outages propagate locally in an influence graph that is not the actual grid topology. *IEEE Transactions on Power Systems*, 32(2):958–967, 2017.

[44] Paul DH Hines et al. Cascading power outages propagate locally in an influence graph that is not the actual grid topology. *IEEE Trans. on Power Sys.*, 32(2):958–967, 2017.

[45] Upama Nakarmi, Mahshid Rahnamay Naeini, Md Jakir Hossain, and Md Abul Hasnat. Interaction graphs for reliability analysis of power grids: A survey. *arXiv preprint arXiv:1911.00475*, 2019.

[46] Paul Hines, Eduardo Cotilla-Sanchez, and Seth Blumsack. Topological models and critical slowing down: Two approaches to power system blackout risk analysis. In *44th Hawaii Intl. Conf. on Sys. Sci.*, pages 1–10. IEEE, 2011.

[47] Christopher L DeMarco. A phase transition model for cascading network failure. *IEEE Control Sys. Magazine*, 21(6):40–51, 2001.

[48] Hui Ren and Ian Dobson. Using transmission line outage data to estimate cascading failure propagation in an electric power system. *IEEE Trans. on Cir. and Sys. II: Express Briefs*, 55(9):927–931, 2008.

[49] Ian Dobson. Estimating the propagation and extent of cascading line outages from utility data with a branching process. *IEEE Trans. on Power Sys.*, 27(4):2146–2155, 2012.

[50] Paul DH Hines et al. "Dual Graph" and "Random Chemistry" methods for cascading failure analysis. In *46th Hawaii Intl. Conf. on Sys. Sci.*, pages 2141–2150. IEEE, 2013.

[51] Xiaoguang Wei et al. A novel cascading faults graph based transmission network vulnerability assessment method. *IEEE Transactions on Power Systems*, 33(3):2995–3000, 2018.

[52] X. Wei et al. Complex network-based cascading faults graph for the analysis of transmission network vulnerability. *IEEE Trans. on Industrial Informatics*, 15(3):1265–1276, 2018.

[53] Xiaoguang Wei et al. Identification of two vulnerability features: A new framework for electrical networks based on the load redistribution mechanism of complex networks. *Complexity*, 2019, 2019.

[54] Tao Wang et al. Modeling fault propagation paths in power systems: A new framework based on event snp systems with neurotransmitter concentration. *IEEE Access*, 2019.

[55] Tianlei Zang et al. Adjacent graph based vulnerability assessment for electrical networks considering fault adjacent relationships among branches. *IEEE Access*, 7:88927–88936, 2019.

[56] Shenhao Yang et al. A graph-based model for transmission network vulnerability analysis. *IEEE Sys. Jrnl.*, 2019.

[57] Xiaoguang Wei et al. Electrical network operational vulnerability evaluation based on small-world and scale-free properties. *IEEE Access*, 2019.

[58] Junjian Qi, Kai Sun, and Shengwei Mei. An interaction model for simulation and mitigation of cascading failures. *IEEE Trans. on Power Sys.*, 30(2):804–819, 2015.

[59] Junjian Qi, Jianhui Wang, and Kai Sun. Efficient estimation of component interactions for cascading failure analysis by em algorithm. *IEEE Trans. on Power Sys.*, 33(3):3153–3161, 2018.

[60] Wenyun Ju, Junjian Qi, and Kai Sun. Simulation and analysis of cascading failures on an npcc power system test bed. In *Proc. IEEE Power & Energy Soc. Gen. Meeting*, pages 1–5, 2015.

[61] Wenyun Ju, Kai Sun, and Junjian Qi. Multi-layer interaction graph for analysis and mitigation of cascading outages. *IEEE Jrnl. on Emerging and Selected Topics in Circuits and Sys.*, 7(2):239–249, 2017.

[62] Changsheng Chen, Wenyun Ju, Kai Sun, and Shiying Ma. Mitigation of cascading outages using a dynamic interaction graph-based optimal power flow model. *IEEE Access*, 7:168637–168648, 2019.

[63] Chao Luo et al. Identify critical branches with cascading failure chain statistics and hypertext-induced topic search algorithm. In *Proc. IEEE Power & Energy Soc. Gen. Meeting*, pages 1–5, 2017.

[64] Kai Zhou et al. Can an influence graph driven by outage data determine transmission line upgrades that mitigate cascading blackouts? In *International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, 2018.

[65] Kai Zhou, Ian Dobson, Zhaoyu Wang, Alexander Roitershtein, and Arka P Ghosh. A markovian influence graph formed from utility line outage data to mitigate cascading. *arXiv preprint arXiv:1902.00686*, 2019.

[66] Zhi-gang Wu, Qing Zhong, and Yao Zhang. State transition graph of cascading electrical power grids. In *2007 IEEE Power Engineering Society General Meeting*, pages 1–7. IEEE, 2007.

[67] Liang Chang and Zhigang Wu. Performance and reliability of electrical power grids under cascading failures. *International Journal of Electrical Power & Energy Systems*, 33(8):1410–1419, 2011.

[68] Linzhi Li, Hao Wu, and Yonghua Song. Temporal difference learning based critical component identifying method with cascading failure data in power systems. In *2018 IEEE Power & Energy Society General Meeting (PESGM)*, pages 1–5. IEEE, 2018.

[69] Linzhi Li, Hao Wu, Yonghua Song, Dunwen Song, and Yi Liu. Quantify the impact of line capacity temporary expansion on blackout risk by the state-failure–network method. *IEEE Access*, 7:183049–183060, 2019.

[70] Linzhi Li, Hao Wu, Yonghua Song, and Yi Liu. A state-failure-network method to identify critical components in power systems. *Electric Power Systems Research*, 181:106192, 2020.

[71] Lu Liu, Linzhi Li, and Hao Wu. Identifying critical patterns of cascading failure in power systems based on sequential pattern mining with gap constraints. In *Proceedings of PURPLE MOUNTAIN FORUM 2019-International Forum on Smart Grid Protection and Control*, pages 837–855. Springer, 2020.

[72] Yihai Zhu, Jun Yan, Yan Sun, and Haibo He. Risk-aware vulnerability analysis of electric grids from attacker's perspective. In *Innovative Smart Grid Technologies (ISGT), 2013 IEEE PES*, pages 1–6. IEEE, 2013.

[73] Yihai Zhu, Jun Yan, Yan Lindsay Sun, and Haibo He. Revealing cascading failure vulnerability in power grids using risk-graph. *IEEE Trans. on Parallel and Distributed Sys.*, 25(12):3274–3284, 2014.

[74] Yihai Zhu, Jun Yan, Yufei Tang, Yan Lindsay Sun, and Haibo He. Resilience analysis of power grids under the sequential attack. *IEEE Trans. on Information Forensics and Security*, 9(12):2340–2354, 2014.

[75] PA Kaplunovich and Konstantin S Turitsyn. Statistical properties and classification of n-2 contingencies in large scale power grids. In *47th Hawaii Intl. Conf. on Sys. Sci.*, pages 2517–2526. IEEE, 2014.

[76] Yang Yang, Takashi Nishikawa, and Adilson E Motter. Vulnerability and cosusceptibility determine the size of network cascades. *Physical review letters*, 118(4):048301, 2017.

[77] Linqi Guo, Chen Liang, and Steven H Low. Monotonicity properties and spectral characterization of power redistribution in cascading failures. *ACM SIGMETRICS Performance Evaluation Review*, 45(2):103–106, 2017.

[78] Linqi Guo et al. Failure localization in power systems via tree partitions. In *IEEE Conference on Decision and Control (CDC)*, pages 6832–6839. IEEE, 2018.

[79] Zhiyuan Ma et al. Fast searching strategy for critical cascading paths toward blackouts. *IEEE Access*, 6:36874–36886, 2018.

[80] Z. Ma et al. Speeding up simulations of cascading blackout in power sys. by identifying high influential lines. In *Proc. IEEE Power & Energy Soc. Gen. Meeting*, pages 1–5, 2017.

[81] Zhiyuan Ma, Chen Shen, Feng Liu, and Shengwei Mei. Fast screening of vulnerable transmission lines in power grids: A PageRank-based approach. *IEEE Trans. on Smart Grid*, 10(2):1982–1991, 2017.

[82] Enrico Zio and Lucia R Golea. Analyzing the topological, electrical and reliability characteristics of a power transmission system for identifying its critical elements. *Reliability Eng. & System Safety*, 101:67–74, 2012.

[83] Shiva Poudel, Zhen Ni, and Wei Sun. Electrical distance approach for searching vulnerable branches during contingencies. *IEEE Trans. on Smart Grid*, 9(4):3373–3382, 2018.

[84] Paul Hines and Seth Blumsack. A centrality measure for electrical networks. In *Hawaii International Conference on System Sciences*, 2008.

[85] Seth Blumsack et al. Defining power network zones from measures of electrical distance. In *Proc. IEEE Power & Energy Soc. Gen. Meeting*, pages 1–8, 2009.

[86] Paul Cuffe and Andrew Keane. Visualizing the electrical structure of power sys. *IEEE Sys. Jrnl.*, 11(3):1810–1821, 2015.

[87] Ettore Bompard, Roberto Napoli, and Fei Xue. Analysis of structural vulnerabilities in power transmission grids. *Intl. Jrnl. of Critical Infrastructure Protection*, 2(1-2):5–12, 2009.

[88] Kai Wang et al. An electrical betweenness approach for vulnerability assessment of power grids considering the capacity of generators and load. *Physica A: Stats. Mechanics and its Appl.*, 390(23-24):4692–4701, 2011.

[89] Ettore Bompard, Enrico Pons, and Di Wu. Extended topological metrics for the analysis of power grid vulnerability. *IEEE Systems Journal*, 6(3):481–487, 2012.

[90] Sergio Arianos et al. Power grid vulnerability: A complex network approach. *Chaos: An Interdisciplinary Jrnl. of Nonlinear Sci.*, 19(1):013119, 2009.

[91] Yakup Koc et al. Structural vulnerability assessment of electric power grids. In *International Conference on Networking, Sensing and Control (ICNSC)*, pages 386–391. IEEE, 2014.

[92] Yakup Koç et al. The impact of the topology on cascading failures in a power grid model. *Physica A: Stats. Mechanics and its Appl.*, 402:169–179, 2014.

[93] Xiangrong Wang et al. A network approach for power grid robustness against cascading failures. In *7th Intl. Workshop on Reliable Networks Design and Modeling*, pages 208–214. IEEE, 2015.

[94] Claudia Caro-Ruiz and Eduardo Mojica-Nava. Voltage collapse analysis in a graph theoretical framework. In *Proc. IEEE PES Innovative Smart Grid Technologies Latin America*, pages 667–672, 2015.

[95] C. Caro-Ruiz and E. Mojica-Nava. Centrality measures for voltage instability analysis in power networks. In *Proc. IEEE 2nd Colombian Conf. on Automatic Control*, pages 1–6, 2015.

[96] Eduardo Cotilla-Sanchez, Paul DH Hines, Clayton Barrows, and Seth Blumsack. Comparing the topological and electrical structure of the north american electric power infrastructure. *IEEE Sys. Jrnl.*, 6(4):616–626, 2012.

[97] Eduardo Cotilla-Sanchez et al. Multi-attribute partitioning of power networks based on electrical distance. *IEEE Trans. on Power Sys.*, 28(4):4979–4987, 2013.

[98] Sergey V Buldyrev, Roni Parshani, Gerald Paul, H Eugene Stanley, and Shlomo Havlin. Catastrophic cascade of failures in interdependent networks. *Nature*, 464(7291):1025–1028, 2010.

[99] Paolo Crucitti, Vito Latora, and Massimo Marchiori. Model for cascading failures in complex networks. *Physical Review E*, 69(4):045104, 2004.

[100] Réka Albert, István Albert, and Gary L Nakarado. Structural vulnerability of the North American power grid. *Phys. review E*, 69(2):25103, 2004.

[101] Ake J Holmgren. Using graph models to analyze the vulnerability of electric power networks. *Risk Analysis*, 26(4):955–969, 2006.

[102] Paul Hines, Eduardo Cotilla-Sanchez, and Seth Blumsack. Do topological models provide good information about electricity infrastructure vulnerability? *Chaos: An Interdis. Jrnl. of Nonli. Sci.*, 20(3):033122, 2010.

[103] MD Abul Hasnat and Mahshid Rahnamay-Naeini. A data-driven dynamic state estimation for smart grids under dos attack using state correlations. In *North American Power Symposium*. IEEE, 2019.

[104] P Lagonotte et al. Structural analysis of the electrical system: Application to secondary voltage control in france. *IEEE Trans. on Power Sys.*, 4(2):479–486, 1989.

[105] Yang Yang, Takashi Nishikawa, and Adilson E. Motter. Vulnerability and cosusceptibility determine the size of network cascades. *Physical Review Letters*, 118:048301, 2017.

[106] Sinan G Aksoy, Emilie Purvine, Eduardo Cotilla-Sanchez, and Mahantesh Halappanavar. A generative graph model for electrical infrastructure networks. *Journal of Complex Networks*, 7(1):128–162, 2019.

[107] Zhifang Wang, Anna Scaglione, and Robert J Thomas. Electrical centrality measures for electric power grid vulnerability analysis. In *IEEE Conference on Decision and Control (CDC)*, pages 5792–5797. IEEE, 2010.

[108] Stephen P Borgatti and Martin G Everett. A graph-theoretic perspective on centrality. *Social networks*, 28(4):466–484, 2006.

[109] Mahantesh Halappanavar, Eduardo Cotilla-Sanchez, Emilie Hogan, Daniel Duncan, Paul DH Hines, et al. A network-of-networks model for electrical infrastructure networks. *arXiv preprint arXiv:1512.01436*, 2015.

[110] Ansi Wang, Yi Luo, Guangyu Tu, and Pei Liu. Vulnerability assessment scheme for power system transmission networks based on the fault chain theory. *IEEE Trans. on power Sys.*, 26(1):442–450, 2011.

[111] Chalee Asavathiratham, Sandip Roy, Bernard Lesieutre, and George Verghese. The influence model. *IEEE Control Systems*, 21(6):52–64, 2001.

[112] Sandip Roy et al. Network models: growth, dynamics, and failure. In *34th Annual Hawaii Intl. Conf. on Sys. Sci.*, pages 728–737. Citeseer, 2001.

[113] (NERC) North American Electricity Reliability Council. Reliability guideline: Methods for establishing IROLs. Technical report, September 2018.

[114] Jiachun Guo, Yong Fu, Zuyi Li, and Mohammad Shahidehpour. Direct calculation of line outage distribution factors. *IEEE Transactions on Power Systems*, 24(3):1633–1634, 2009.

[115] Douglas J Klein and Milan Randić. Resistance distance. *Jrnl. of mathematical chemistry*, 12(1):81–95, 1993.

[116] Prabha Kundur, John Paserba, Venkat Ajjarapu, Göran Andersson, Anjan Bose, Claudio Canizares, Nikos Hatziargyriou, David Hill, Alex Stankovic, Carson Taylor, et al. Definition and classification of power system stability ieee/cigre joint task force on stability terms and definitions. *IEEE transactions on Power Systems*, 19(3):1387–1401, 2004.

[117] Bernardo A Huberman, Daniel M Romero, and Fang Wu. Social networks that matter: Twitter under the microscope. *arXiv preprint arXiv:0812.1045*, 2008.

[118] Joshua R Tyler, Dennis M Wilkinson, and Bernardo A Huberman. E-mail as spectroscopy: Automated discovery of community structure within organizations. *The Information Society*, 21(2):143–153, 2005.

[119] Roy M Anderson, B Anderson, and Robert M May. *Infectious diseases of humans: dynamics and control*. Oxford university press, 1992.

[120] Damon Centola. The spread of behavior in an online social network experiment. *Sci.*, 329(5996):1194–1197, 2010.

[121] Daryl J Daley and David G Kendall. Epidemics and rumours. *Nature*, 204(4963):1118–1118, 1964.

[122] Mark S Granovetter. The strength of weak ties. In *Social Networks*, pages 347–367. Elsevier, 1977.

[123] Lilian Weng, Filippo Menczer, and Yong-Yeol Ahn. Virality prediction and community structure in social networks. *Scientific Reports*, 3:2522, 2013.

[124] Martin Rosvall and Carl T Bergstrom. Maps of information flow reveal community structure in complex networks. *arXiv preprint physics.soc-ph/0707.0609*, 2007.

[125] Martin Rosvall and Carl T Bergstrom. Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, 105(4):1118–1123, 2008.

[126] Zongqing Lu, Yonggang Wen, and Guohong Cao. Community detection in weighted networks: Algorithms and applications. In *IEEE International Conference on Pervasive Computing and Communications (PerCom)*, 2013.

[127] Santo Fortunato and Darko Hric. Community detection in networks: A user guide. *Physics Reports*, 659:1–44, 2016.

[128] Jierui Xie, Stephen Kelley, and Boleslaw K Szymanski. Overlapping community detection in networks: The state-of-the-art and comparative study. *ACM Computing Surveys*, 45(4), 2013.

[129] Fragkiskos D Malliaros and Michalis Vazirgiannis. Clustering and community detection in directed networks: A survey. *Physics Reports*, 533(4):95–142, 2013.

[130] Upama Nakarmi and Mahshid Rahnamay-Naeini. Analyzing power grids' cascading failures and critical components using interaction graphs. In *2018 IEEE Power & Energy Society General Meeting (PESGM)*, pages 1–5. IEEE, 2018.

[131] Martin Rosvall et al. The map equation. *The European Physical Journal Special Topics*, 178(1):13–23, 2009.

[132] Alcides Viamontes Esquivel and Martin Rosvall. Compression of flow can reveal overlapping-module organization in networks. *Physical Review X*, 1(2):021025, 2011.

[133] Jure Leskovec et al. The dynamics of viral marketing. *ACM Transactions on the Web (TWEB)*, 1(1):5, 2007.

[134] Ziqi Wang, Jinghan He, Alexandru Nechifor, Dahai Zhang, and Peter Crossley. Identification of critical transmission lines in complex power networks. *Energies*, 10(9):1294, 2017.

[135] Dirk Witthaut et al. Critical links and nonlocal rerouting in complex supply networks. *Physical Review Letters*, 116(13):138701, 2016.

[136] Benjamin Schäfer et al. Dynamically induced cascading failures in power grids. *Nature communications*, 9(1):1975, 2018.

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

[138] Chao Luo and Jun Yang. Identify critical branches with cascading failure chain statistics and hypertext-induced topic search algorithm. *arXiv preprint arXiv:1704.06917*, 2017.

[139] Junjian Qi and Shengwei Mei. A cascading failure model by quantifying interactions. *arXiv preprint arXiv:1301.2055*, 2013.

[140] Yang Yang, Takashi Nishikawa, and Adilson E Motter. Small vulnerable sets determine large network cascades in power grids. *Science*, 358(6365):eaan3184, 2017.

[141] Bin Liu, Zhen Li, Xi Chen, Yuehui Huang, and Xiangdong Liu. Recognition and vulnerability analysis of key nodes in power grid based on complex network centrality. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 65(3):346–350, 2018.

[142] Alessandro Vespignani. Modelling dynamical processes in complex socio-technical systems. *Nature physics*, 8(1):32, 2012.

[143] Wendy Ellens, FM Spieksma, P Van Mieghem, A Jamakovic, and RE Kooij. Effective graph resistance. *Linear algebra and its applications*, 435(10):2491–2506, 2011.

[144] Ivan Gutman and Bojan Mohar. The quasi-wiener and the Kirchhoff indices coincide. *Jrnl. of Chemical Information and Computer Sci.*, 36(5):982–985, 1996.

[145] Richard Steinberg and Willard I Zangwill. The prevalence of Braess' paradox. *Transportation Sci.*, 17(3):301–318, 1983.

[146] Duanbing Chen, Linyuan Lü, Ming-Sheng Shang, Yi-Cheng Zhang, and Tao Zhou. Identifying influential nodes in complex networks. *Physica a: Statistical mechanics and its applications*, 391(4):1777–1787, 2012.

[147] U.S.-Canada Power System Outage Task Force. Final report on the august 14, 2003 blackout in the United States and Canada: causes and recommendations, 2004.

[148] Ian Dobson. Estimating the propagation and extent of cascading line outages from utility data with a branching process. *IEEE Transactions on Power Systems*, 27(4):2146–2155, 2012.

[149] Jun Yan et al. Multi-contingency cascading analysis of smart grid based on self-organizing map. *IEEE Transactions on Information Forensics and Security*, 8(4):646–656, 2013.

[150] Sergio Arianos, Ettore Bompard, Anna Carbone, and Fei Xue. Power grid vulnerability: A complex network approach. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 19(1):013119, 2009.

[151] Zhao Yang, René Algesheimer, and Claudio J Tessone. A comparative analysis of community detection algorithms on artificial networks. *Scientific Reports*, 6:30750, 2016.

[152] Mahshid Rahnamay-Naeini, Zhuoyao Wang, Andrea Mammoli, and Majeed M Hayat. A probabilistic model for the dynamics of cascading failures and blackouts in power grids. In *IEEE PESGM*, pages 1–8, 2012.

[153] Mahshid Rahnamay-Naeini and Majeed M Hayat. Cascading failures in interdependent infrastructures: An interdependent markov-chain approach. *IEEE Trans. on Smart Grid*, 7(4):1997–2006, 2016.

[154] Hafedh Chourabi, Taewoo Nam, Shawn Walker, J Ramon Gil-Garcia, Sehl Mellouli, Karine Nahon, Theresa A Pardo, and Hans Jochen Scholl. Understanding smart cities: An integrative framework. In *Hawaii International Conference on System Science (HICSS)*, pages 2289–2297. IEEE, 2012.

[155] B Bowerman, J Braverman, J Taylor, H Todosow, and U Von Wimmersperg. The vision of a smart city. In *International Life Extension Technology Workshop*, volume 28, 2000.

[156] Massoud Amin. Toward secure and resilient interdependent infrastructures. *Journal of Infrastructure Systems*, 8(3):67–75, 2002.

[157] Richard G Little. Controlling cascading failure: Understanding the vulnerabilities of interconnected infrastructures. *Journal of Urban Technology*, 9(1):109–123, 2002.

[158] Hyeung-Sik J Min, Walter Beyeler, Theresa Brown, Young Jun Son, and Albert T Jones. Toward modeling and simulation of critical national infrastructure interdependencies. *Iie Transactions*, 39(1):57–71, 2007.

[159] Steven M Rinaldi. Modeling and simulating critical infrastructures and their interdependencies. In *Annual Hawaii international conference on System Sciences*, pages 1–8. IEEE, 2004.

[160] Arun Das, Joydeep Banerjee, and Arunabha Sen. Root cause analysis of failures in interdependent power-communication networks. In *IEEE Military Communications Conference*, pages 910–915. IEEE, 2014.

[161] Jia Shao, Sergey V Buldyrev, Shlomo Havlin, and H Eugene Stanley. Cascade of failures in coupled network systems with multiple support-dependence relations. *Physical Review E*, 83(3):036116, 2011.

[162] Dong-Hoon Shin, Dajun Qian, and Junshan Zhang. Cascading effects in interdependent networks. *IEEE Network*, 28(4):82–87, 2014.

[163] Mahshid Rahnamay-Naeini and Majeed Hayat. Cascading failures in interdependent infrastructures: An interdependent Markov-chain approach. *IEEE Transactions on Smart Grid*, 7(4):1997–2006, 2016.

[164] Chuyue Chen and Guowei Hua. A new model for optimal deployment of electric vehicle charging and battery swapping stations. *International Journal of Control & Automation*, 8(5), 2014.

[165] Sen Guo and Huiru Zhao. Optimal site selection of electric vehicle charging station by using fuzzy TOPSIS based on sustainability perspective. *Applied Energy*, 158:390–402, 2015.

[166] Andrea Hess, Francesco Malandrino, Moritz Bastian Reinhardt, Claudio Casetti, Karin Anna Hummel, and Jose M Barceló-Ordinas. Optimal deployment of charging stations for electric vehicular networks. In *Workshop on Urban networking*, pages 1–6. ACM, 2012.

[167] Yanhua Li, Jun Luo, Chi-Yin Chow, Kam-Lam Chan, Ye Ding, and Fan Zhang. Growing the charging station network for electric vehicles with trajectory data analytics. In *IEEE International Conference on Data Engineering*, pages 1376–1387. IEEE, 2015.

[168] Timothy M Sweda and Diego Klabjan. Agent-based information system for electric vehicle charging infrastructure deployment. *Journal of Infrastructure Systems*, 21(2):04014043, 2014.

[169] Mohammad M Vazifeh, Hongmou Zhang, Paolo Santi, and Carlo Ratti. Optimizing the deployment of electric vehicle charging stations using pervasive mobility data. *arXiv preprint arXiv:1511.00615*, 2015.

[170] Yanhai Xiong, Jiarui Gan, Bo An, Chunyan Miao, and Ana LC Bazzan. Optimal electric vehicle charging station placement. In *International Joint Conference on Artificial Intelligence (IJCAI)*, pages 2662–2668, 2015.

[171] Md Mainul Islam, Hussain Shareef, and Azah Mohamed. A review of techniques for optimal placement and sizing of electric vehicle charging stations. *Electrotechnical Review*, 91(8):122–126, 2015.

[172] Kate Walsh, Cathy A Enz, and Linda Canina. The impact of gasoline price fluctuations on lodging demand for US brand hotels. *International Journal of Hospitality Management*, 23(5):505–521, 2004.

[173] Claude Weis, Kay Axhausen, Robert Schlich, and René Zbinden. Models of mode choice and mobility tool ownership beyond 2008 fuel prices. *Transportation Research Record: Journal of the Transportation Research Board*, 2157(1):86–94, 2010.

[174] Yanhai Xiong, Jiarui Gan, Bo An, Chunyan Miao, and Yeng Chai Soh. Optimal pricing for efficient electric vehicle charging station management. In *International Conference on Autonomous Agents & Multiagent Systems*, pages 749–757. International Foundation for Autonomous Agents and Multiagent Systems, 2016.

[175] Wayes Tushar, Walid Saad, H Vincent Poor, and David B Smith. Economics of electric vehicle charging: A game theoretic approach. *IEEE Transactions on Smart Grid*, 3(4):1767–1778, 2012.

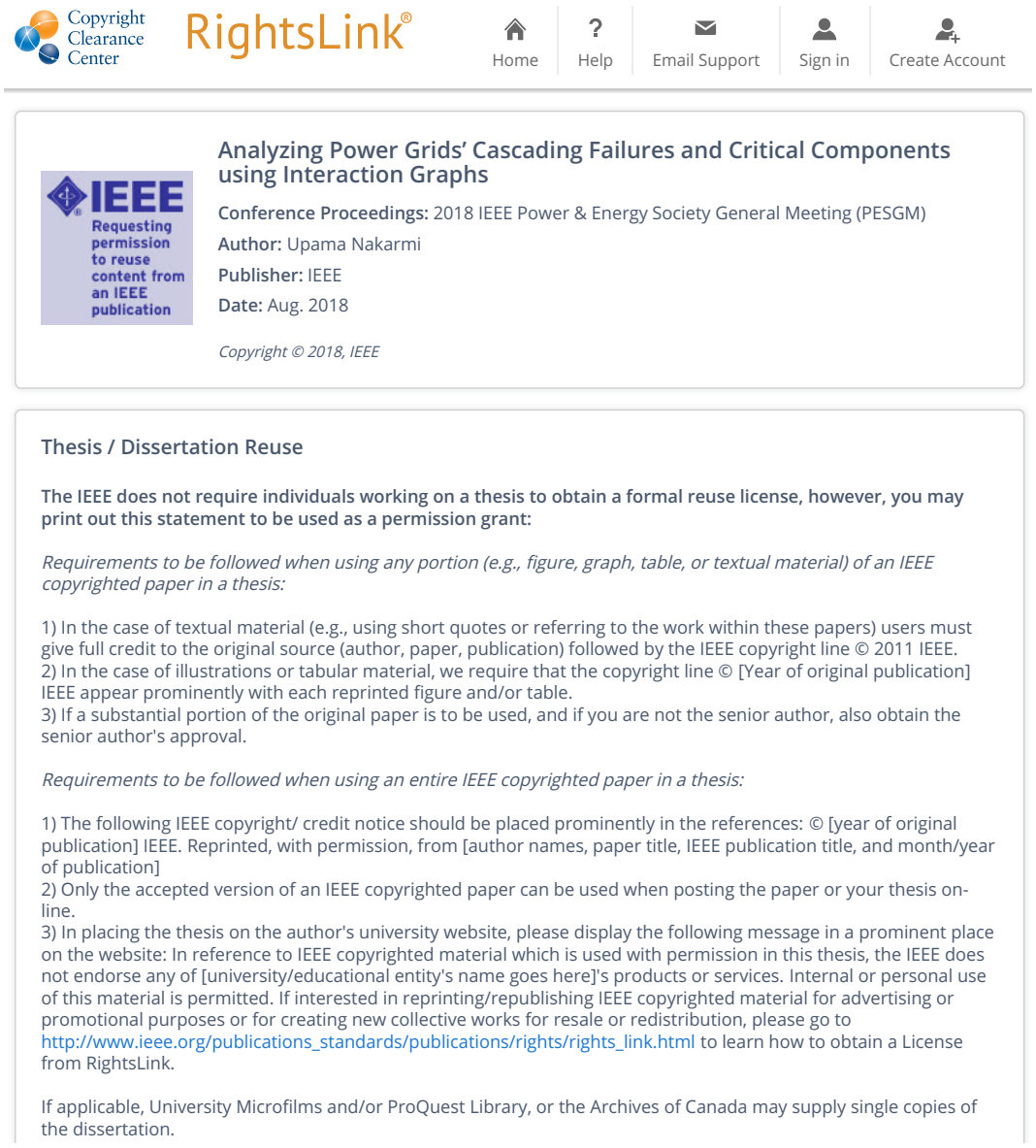
[176] Chalee Asavathiratham. *The influence model: A tractable representation for the dynamics of networked markov chains*. PhD thesis, Citeseer, 2000.

[177] Ehsan Siavashi and Mahshid Rahnamay-Naeini. The dynamic, constraint-based influence model and its application in stochastic modeling of load balancing in computing networks. *International Journal of Network Science*, 2017.

[178] Michal Piorkowski, Natasa Sarafijanovic-Djukic, and Matthias Grossglauser. CRAWDAD data set epfl/mobility (v. 2009-02-24), 2009.

Appendix A: Copyright Permissions

The following copyright permission is for the use of material in Chapter 1, Chapter 3, and Chapter 4 which have appeared in [3].



Copyright Clearance Center RightsLink®

Home | Help | Email Support | Sign in | Create Account

IEEE Requesting permission to reuse content from an IEEE publication

Analyzing Power Grids' Cascading Failures and Critical Components using Interaction Graphs

Conference Proceedings: 2018 IEEE Power & Energy Society General Meeting (PESGM)

Author: Upama Nakarmi

Publisher: IEEE

Date: Aug. 2018

Copyright © 2018, IEEE

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

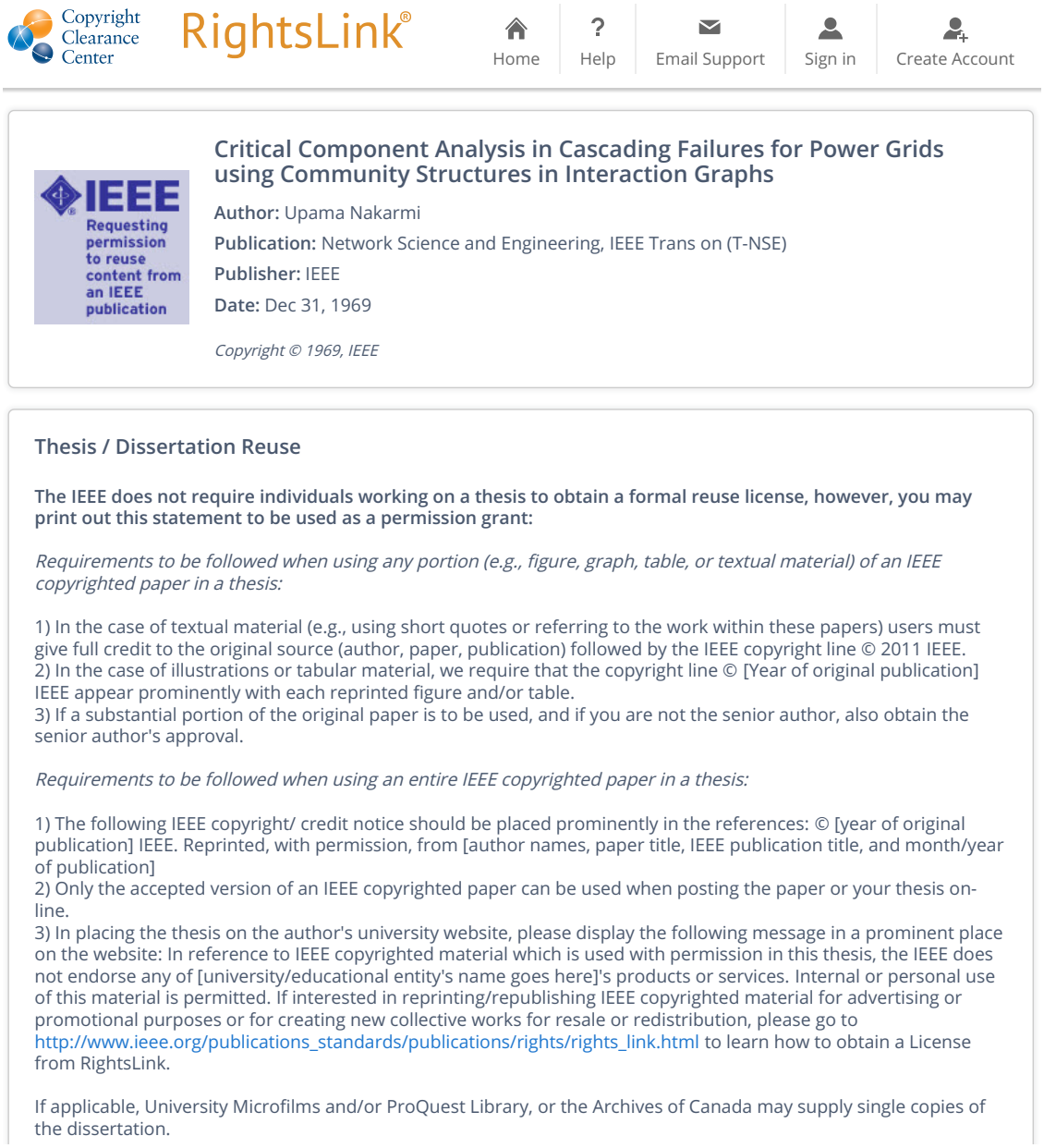
1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

The following copyright permission is for the use of material in Chapter 1, Chapter 3, and Chapter 4 which have appeared in [4].



Copyright Clearance Center RightsLink®

Home | Help | Email Support | Sign in | Create Account

IEEE Requesting permission to reuse content from an IEEE publication

Critical Component Analysis in Cascading Failures for Power Grids using Community Structures in Interaction Graphs

Author: Upama Nakarmi

Publication: Network Science and Engineering, IEEE Trans on (T-NSE)

Publisher: IEEE

Date: Dec 31, 1969

Copyright © 1969, IEEE

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

The following copyright permission is for the use of material in Chapter 6 which have appeared in [5].

USF

Upama Nakarmi <unakarmi@mail.usf.edu>

Permission to reuse Conference Paper in Dissertation

SCITEPRESS Enquiries <info@scitepress.org> Mon, Feb 10, 2020 at 2:08 PM
To: unakarmi@mail.usf.edu
Cc: mahshidr@usf.edu

Dear Upama Nakarmi,

We authorize the paper's reprinting for your dissertation, as long as all the bibliography information from its publication is there too. Any paper version can be there although we prefer to see there the actual published version.

--
Kind regards,
Cláudia Ferreira
SCITEPRESS Team

SCITEPRESS Office
Avenida de S. Francisco Xavier Lote 7 Cv. C, 2900-616 Setubal - Portugal
Tel.: +351 265 520 184/5
Fax: +351 265 520 186
http://www.scitepress.org

On Monday, February 10th 2020, 7:59 pm CET (+0100), Upama Nakarmi wrote:

> Hello,
>
> My name is Upama Nakarmi. I am a PhD candidate at the University of South Florida. My advisor is Dr. Mahshid Rahnamay Naeini. I am requesting an authorization of publication of the following work in a dissertation titled 'Reliability Analysis of Power Grids and its Interdependent Infrastructures using Stochastic Interaction Graphs' with an expected presentation date of March 2020:
>
> Paper Title: **Towards Integrated Infrastructures for Smart City Services: A Story of Traffic and Energy Aware Pricing Policy for Charging Infrastructures**
> Authors: Upama Nakarmi, Mahshid Rahnamay Naeini
> DOI: 10.5220/0006303202080218
> Publisher: SciTePress
> Proceedings of Smartgreens 2017
>
> Thank you.
> --

The following copyright permission is for the use of material in Chapter 6 which have appeared in [6].

SPRINGER NATURE LICENSE
TERMS AND CONDITIONS

Feb 06, 2020

This Agreement between University of South Florida ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	4763250281434
License date	Feb 06, 2020
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Springer eBook
Licensed Content Title	An Influence-Based Model for Smart City's Interdependent Infrastructures: Application in Pricing Design for EV Charging Infrastructures
Licensed Content Author	Upama Nakarmi, Mahshid Rahnamay-Naeini
Licensed Content Date	Jan 1, 2019
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	electronic
Portion	full article/chapter

The following screenshot is for the use of material in Chapter 2 and Chapter 4, which have appeared as preprints in [45].