Cascading Failures and Contingency Analysis for Smart Grid Security

Shiva Poudel
South Dakota State University
CASCADING FAILURES AND CONTINGENCY ANALYSIS FOR SMART GRID SECURITY

BY

SHIVA POUDEL

A thesis submitted in partial fulfillment of the requirements for the

Master of Science

Major in Electrical Engineering

South Dakota State University

2016
CASCADING FAILURES AND CONTINGENCY ANALYSIS FOR SMART GRID SECURITY

This thesis is approved as a creditable and independent investigation by a candidate for the Master of Science in Electrical Engineering degree and is acceptable for meeting the thesis requirements for this degree. Acceptance of this thesis does not imply that the conclusions reached by the candidates are necessarily the conclusions of the major department.

Zhen Ni, Ph.D.
The thesis advisor

Steven Hietpas, Ph.D.
Head, Electrical Engineering and Computer Science

Dean, Graduate School
ACKNOWLEDGEMENTS

Foremost, I would like to express the deepest gratitude to my advisor, Dr. Zhen Ni. I have been amazingly fortunate to have an advisor who gave me the freedom to explore on my own, and at the same time the guidance to recover when my steps faltered. Being his student is and always will be one of the greatest honors of my life. Dr. Ni gave many constructive suggestions in my study and his instructions paved the way to this dissertation. In addition, he taught me many skills in writing and revising manuscripts.

Besides my advisor, I would like to thank the rest of my thesis committee: Dr. Reinaldo Tonkoski and Dr. Timothy M. Hansen for their insightful comments and constructive criticism at different stages of my research. Dr. Tonkoski set up high, but achievable, standards for students. I am grateful to Dr. Hansen for practical advice and comments on the countless revision of this manuscript. In addition, many thanks to the graduate faculty representative Dr. Joy Scaria for his valuable time.

I am also indebted to my friends, Manisha, Prajina, Mahesh, Bijen, Lok Raj, Lekhnath, Sudeep, Naresh, Avijit, Venkat, Habib and many others. They have helped me stay sane through these difficult years and their support helped me overcome setbacks and stay focused on my graduate study.

Finally, I must express my very profound gratitude to my parents, sisters, and my brother-in-laws for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. Last but not the least, I love my nephew, Arpit, and niece, Ruby whose love I can’t never thank enough.
CONTENTS

ABBREVIATIONS ............................................................... viii

LIST OF FIGURES .......................................................... xi

LIST OF TABLES ............................................................. xii

ABSTRACT ................................................................. xiii

CHAPTER 1 INTRODUCTION .............................................. 1
  1.1 Background of Smart Grid .......................................... 1
  1.2 Smart Grid Security Challenges ................................... 3
  1.3 Motivations and Contributions ..................................... 6
  1.4 The Structure of Thesis ............................................. 8

CHAPTER 2 UNSUPERVISED LEARNING FOR POWER GRID SECURITY 9
  2.1 Introduction .......................................................... 9
    2.1.1 Attack threats in Smart Grid ............................... 10
    2.1.2 Vulnerability of power grids to cascading failure .......... 11
    2.1.3 Related work .................................................. 15
  2.2 Power Grid Modeling .............................................. 17
    2.2.1 System model ................................................. 17
    2.2.2 Cascading tree ............................................... 22
    2.2.3 Evaluation metric ............................................ 24
2.3 Multi-component Attack Strategies .............................................. 25
  2.3.1 Traditional load based attack strategy ................................. 25
  2.3.2 Unsupervised learning based attack strategy ....................... 26
2.4 Simulation Results ................................................................. 30
  2.4.1 Parameters and environment setup ...................................... 31
  2.4.2 SOM and K-means clustering ............................................ 31
  2.4.3 Performance comparison ................................................ 33
  2.4.4 Transient stability ...................................................... 40
2.5 Summary .............................................................................. 42

CHAPTER 3 ELECTRICAL DISTANCE APPROACH FOR CONTINGENCY ANALYSIS .............................................. 44
3.1 Introduction ........................................................................ 44
  3.1.1 Power system security .................................................. 44
  3.1.2 Contingency analysis ................................................... 45
  3.1.3 Related work ............................................................. 46
3.2 Mathematical Formulation of Contingency Analysis ..................... 48
3.3 Proposed Electrical Distance Method ........................................ 52
  3.3.1 Different structure of power grid .................................... 52
  3.3.2 Measure of electrical distance ....................................... 53
  3.3.3 Formation of new electrical structure ............................... 55
  3.3.4 Discussion on the proposed method ................................ 57
3.4 Simulation Results ................................................................. 59
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGC</td>
<td>Automatic Generation Control</td>
</tr>
<tr>
<td>AMI</td>
<td>Advanced Metering Infrastructure</td>
</tr>
<tr>
<td>ANN</td>
<td>Artificial Neural Network</td>
</tr>
<tr>
<td>BMU</td>
<td>Best Matching Unit</td>
</tr>
<tr>
<td>CA</td>
<td>Contingency Analysis</td>
</tr>
<tr>
<td>CATF</td>
<td>Cyber Attack Task Force</td>
</tr>
<tr>
<td>CI</td>
<td>Computational Intelligence</td>
</tr>
<tr>
<td>CPS</td>
<td>Cyber-Physical System</td>
</tr>
<tr>
<td>DoS</td>
<td>Denial of Service</td>
</tr>
<tr>
<td>EMS</td>
<td>Energy Management System</td>
</tr>
<tr>
<td>ICSEG</td>
<td>Illinois Center for Smarter Electric Grid</td>
</tr>
<tr>
<td>ICT</td>
<td>Information and Communication Technology</td>
</tr>
<tr>
<td>IED</td>
<td>Intelligent Electronic Device</td>
</tr>
<tr>
<td>LODF</td>
<td>Line Outage Distribution Factors</td>
</tr>
<tr>
<td>LP</td>
<td>Linear Programming</td>
</tr>
<tr>
<td>NERC</td>
<td>North American Electric Reliability Council</td>
</tr>
<tr>
<td>OPF</td>
<td>Optimal Power Flow</td>
</tr>
<tr>
<td>PMU</td>
<td>Phasor Measurement Unit</td>
</tr>
<tr>
<td>PoF</td>
<td>Percentage of Failure</td>
</tr>
<tr>
<td>SCADA</td>
<td>Supervisory Control and Data Acquisition</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>SEL</td>
<td>Schweitzer Engineering Laboratories</td>
</tr>
<tr>
<td>SPS</td>
<td>SimPowerSystem</td>
</tr>
<tr>
<td>VPN</td>
<td>Virtual Private Network</td>
</tr>
<tr>
<td>WAMS</td>
<td>Wide Area Measurement System</td>
</tr>
<tr>
<td>WCSS</td>
<td>Within-Cluster Sum of Squares</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>A general schematic diagram of smart grid.</td>
<td>2</td>
</tr>
<tr>
<td>2.1</td>
<td>Illustration of load definition and redistribution according to network</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>connectivity</td>
<td></td>
</tr>
<tr>
<td>2.2</td>
<td>Load burden of a node using power flow information.</td>
<td>21</td>
</tr>
<tr>
<td>2.3</td>
<td>Illustration of cascading failure in successive steps</td>
<td>24</td>
</tr>
<tr>
<td>2.4</td>
<td>Attack strategies on IEEE-39 and IEEE-57 bus system</td>
<td>26</td>
</tr>
<tr>
<td>2.5</td>
<td>Two clusters of IEEE 39 bus system using SOM</td>
<td>32</td>
</tr>
<tr>
<td>2.6</td>
<td>Three clusters of IEEE 57 bus system using K-means</td>
<td>32</td>
</tr>
<tr>
<td>2.7</td>
<td>Performance comparison of traditional load based and clustering based</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>attack for IEEE-39 bus system based on network connectivity.</td>
<td></td>
</tr>
<tr>
<td>2.8</td>
<td>Performance comparison of traditional load based and clustering based</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>attack for IEEE-57 bus system based on network connectivity.</td>
<td></td>
</tr>
<tr>
<td>2.9</td>
<td>Performance comparison of traditional load based and clustering based</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>attack for IEEE-39 bus system based on power flow.</td>
<td></td>
</tr>
<tr>
<td>2.10</td>
<td>Performance comparison of traditional load based and clustering based</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>attack for IEEE-57 bus system based on power flow.</td>
<td></td>
</tr>
<tr>
<td>2.11</td>
<td>Rotor angle of generators for load based attack.</td>
<td>40</td>
</tr>
<tr>
<td>2.12</td>
<td>Rotor angle of generators for SOM based attack.</td>
<td>41</td>
</tr>
<tr>
<td>3.1</td>
<td>Contingency result for outage in power system. The lines identified in circle</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>are overloaded due to outage of line/s identified in rectangle.</td>
<td></td>
</tr>
<tr>
<td>3.2</td>
<td>Single line diagram of W&amp;W 6 bus system [72].</td>
<td>53</td>
</tr>
</tbody>
</table>
Figure 3.3. IEEE-24 bus one-line diagram and node branch representation of the system exactly following the one-line diagram. The number in node are in accordance with the bus number. .......................... 58

Figure 3.4. Node-branch representation of different structure of power grid network; (a) Dense electrical structure showing connection among all node pairs; (b) Important electrical connections with number of connections equal to number of branches in topological structure (Fig. 3.3); (c) Electrical connections that are physically present in benchmark system. .................................................. 60

Figure 3.5. Venn-diagram showing sets of transmission lines belonging to different network structure. The color of marks represents the different structure of power system network and are in analogous to Fig. 3.3 and Fig. 3.4. .......................... 61

Figure 3.6. $N-1$ contingency analysis results for loss of line 3 and line 6. The line identified in circles are overloaded lines due to outage of line identified in rectangle. The number in circle represents the percentage loading of lines. .................................................. 62

Figure 3.7. $N-2$ contingency analysis results for lines 3&6 and lines 6&8. The line identified in circles are overloaded lines due to outage of line identified in rectangle. The number in circle represents the percentage loading of lines. .................................................. 63

Figure 3.8. Voltage profiles for loss of sensitive branches. .................................................. 68

Figure 3.9. Voltage profiles for loss of least sensitive branches. .................................................. 68
Figure 3.10. Voltage distribution for loss of branches. Red mark represents the bus voltage that is out of normal operating range and number refers to the buses whose voltage is mostly affected by loss of particular branch. 69

Figure 4.1. Model of different types of electric power utility [88]. 72

Figure 4.2. Application areas of the CPS testbed. 77

Figure 4.3. Schematic architect of the proposed cyber-physical microgrid testbed. 79

Figure 4.4. Hardware setup in Micro-grid lab at SDSU. 81

Figure 4.5. U.S. Very Inverse Curve [109]. This curve is drawn based on equation (4.1). 82

Figure 4.6. Current in transmission line 7-8. Current increases after 5.29s and due to relay settings, the line 7-8 is opened at 15.30s after which the current is zero. 83

Figure 4.7. Impact of attack on system voltage. 84

Figure 4.8. Impact of attack on generation: (a) Real power generation from three generators; (b) Reactive power generation from three generators. 85

Figure 4.9. OPF mitigation simulation for load curtailment. 87
**LIST OF TABLES**

Table 2.1. History of Major United States Blackouts .......... 12
Table 2.2. Cascading events for IEEE 39 bus system at tolerance of $\alpha = 1.6.$ .... 34
Table 2.4. Cascading events for IEEE 39 bus system at tolerance of $\alpha = 1.25.$ .... 37
Table 3.1. $Y_{bus}$ for W&W 6 bus system. ......................... 53
Table 3.2. Equivalent $Z_{bus}$ for W&W 6 bus system. Each entry represents the amount of coupling between pair of nodes. ......................... 55
Table 3.3. Contingency analysis for IEEE-24 bus system showing line index and number of times it is violated for $N - 1, N - 2,$ and $N - 3$ contingencies 62
Table 3.4. Descriptive statistics for power system test cases ................... 64
Table 3.5. Contingency analysis in different benchmark system .................. 65
Table 4.1. Generator rescheduled after OPF ........................................ 86
ABSTRACT

CASCADING FAILURES AND CONTINGENCY ANALYSIS FOR SMART GRID SECURITY

SHIVA POUDEL

2016

The modern electric power grid has become highly integrated in order to increase the reliability of power transmission from the generating units to end consumers. In addition, today’s power system are facing a rising appeal for the upgrade to a highly intelligent generation of electricity networks commonly known as Smart Grid. However, the growing integration of power system with communication network also brings increasing challenges to the security of modern power grid from both physical and cyber space. Malicious attackers can take advantage of the increased access to the monitoring and control of the system and exploit some of the inherent structural vulnerability of power grids. Therefore, determining the most vulnerable components (e.g., buses or generators or transmission lines) is critically important for power grid defense. This dissertation introduces three different approaches to enhance the security of the smart grid.

Motivated by the security challenges of the smart grid, the first goal of this thesis is to facilitate the understanding of cascading failure and blackouts triggered by multi-component attacks, and to support the decision making in the protection of a reliable and secure smart grid. In this work, a new definition of load is proposed by taking power flow into consideration in comparison with the load definition based on degree or network connectivity. Unsupervised learning techniques (e.g., K-means algorithm and
self-organizing map (SOM)) are introduced to find the vulnerable nodes and performance comparison is done with traditional load based attack strategy.

Second, an electrical distance approach is introduced to find the vulnerable branches during contingencies. A new network structure different than the original topological structure is formed based on impedance matrix which is referred as electrical structure. This structure is pruned to make it size compatible with the topological structure and the common branches between the two different structures are observed during contingency analysis experiments. Simulation results for single and multiple contingencies have been reported and the violation of line limits during single and multiple outages are observed for vulnerability analysis.

Finally, a cyber-physical power system (CPS) testbed is introduced as an accurate cyber-physical environment in order to observe the system behavior during malicious attacks and different disturbance scenarios. The application areas and architecture of proposed CPS testbed have been discussed in details. The testbed’s efficacy is then evaluated by conducting real-time cyber attacks and exploring the impact in a physical system. The possible mitigation strategies are suggested for defense against the attack and protect the system from being unstable.
CHAPTER 1 INTRODUCTION

1.1 Background of Smart Grid

The power system has been evolved for more than one hundred years. With its development and extension from the 20th century, at present, U.S. interconnected system include about 2,000 electric distribution utilities, more than 300,000 miles of transmission lines and more than 7,200 power plants [1]. In addition, there is not a single national electric grid in the United States and it is fragmented. Electric infrastructure today is aging, outmoded and overstressed. Right now, we are using the 19th-century system from the days of Edison and Westinghouse that uses 20th-century equipment in an effort to keep up with a 21st-century economy [2]. The existing electric grid is a strictly hierarchical system in which power plants at the top of chain ensure power delivery to the consumers at the bottom of the chain. Simple, it is a single way pipeline where the sources have no real-time information about their end points which is, therefore, over-engineered to withstand maximum anticipated peak demand. The power industry is facing great challenges at present in terms of the economy, reliability and efficiency issues.

This section introduces the smart grid and its advantages over the traditional power grid. The next-generation electricity grid, commonly known as “smart grid” is expected to address the major shortcomings of the existing electric grid. Smart grid is a modern electric power grid that uses optimized control algorithm, smart meters, bi-directional communication between the source and destination of the utility. The U.S Department of Energy (DOE) has identified seven principles characteristics of the modern electric grid which includes self-healing, consumer participation, attack resistance, power quality for
21st century needs, accommodation of generation and storage, enabling of markets and optimization of assets [3]. Fig. 1.1 shows a basic view of the smart grid that uses communication and information technology embedded with the power system, IEDs for sensing and monitoring purposes, integration of renewable energy and demand response. These technologies are used to enhance all the areas of the electric grid, like generation, transmission, distribution and even electric markets [4].

With the help of advanced metering infrastructure (AMI), smart grid makes the power system a two-way interactive grid. And, with this interactivity, power companies are able to regulate the supply in real time depending on the energy demand. So, there is no need to generate excessive power anymore, which reduces the power waste in the grid and maximizes the power supply efficiency. Also, customers can adjust their power usage according to the real-time pricing with the help of smart meters and time-based pricing.
This power usage habit will reduce the peak demand by shifting some part of power utilizing time. In addition, with the application of advanced control algorithms in smart grid, it is possible to coordinate renewable energies and electric vehicles into the grid. This makes the power grid more robust during the fluctuate demand from costumers. The wide area measurement system (WAMS) based on phasor measurement units (PMUs) and accurate GPS-based timing are being deployed to accurately analyze the flow of electricity through bulk power system. Similarly, in the substation layer, the advanced communications paradigms and improved field devices like intelligent electronic devices (IEDs) are being used. This allows multicast transmission of device status and helps perform sophisticated operations such as grid protection.

1.2 Smart Grid Security Challenges

Smart grid is expected to be more economical, reliable and efficient as compared to the traditional grid because of the technologies like AMI, WAMS, and automated substation. However, these technologies provide some target point for attack and increase the security issues. For example, a smart meter, PMU data, communication protocols in a substation can be the point of interest for an attacker to compromise the power grid elements (buses or transmission lines). Hence, the coupling of the power infrastructure with communication and information technology will introduce some new problems of which the most imminent is the cyber security issues and physical vulnerabilities growing in the smart grid.

Cascading failure in power transmission usually begins when one part of the system fails and shifts its load to nearby elements. Each component has a loading failure
threshold above which it fails. After any disturbance in the system, the loading of the component is increased and if it is above the threshold value, it fails and the load is transferred to remaining system components. It is a common effect seen in high voltage system where the failure process cascades through the power grid components like a ripple on a pond and continues until all of the system’s components are compromised or none of the components have a loading value above their threshold. North American Electric Reliability Council (NERC) standard [5] requires the power system to demonstrate transient, dynamic and voltage stability, and there should not occur any cascading or uncontrolled separation following the single contingency. This means it is not possible to cause cascading failure if any of the single components is compromised. However, because of openness and complexity of the envisioned smart grid, it is possible for intelligent attackers to carry out well-coordinated cyber, physical or cyber-physical attacks to initiate cascading failures which ultimately leads to large scale blackouts. With the use of sophisticated communications and control schemes in the power system, network complexity is increasing which enables new flexibility in operation in terms of reliability and efficiency. But, this might also contribute new ways that the system could fail or change in system behavior. Northeast blackout in 2003 [6], European blackout in 2006 [7], Northern India blackout in 2012 [8] are some examples of large blackouts caused by cascading failures. These examples show how the power grids are operated under stressed conditions and hence more effort is needed to understand and mitigate the risk of cascading failures.

Contingency analysis is a traditional approach to testing all contingencies sequentially to evaluate system performance and reliability. CA simulates the outage of
particular grid components and evaluates the consequences following the outage [9].

North American Reliability corporation requires the system operator to maintain $N - 1$ contingency criterion. However, multiple outage contingencies are becoming increasingly relevant because of the way the power system is being operated [10] and the growing threats from cyberspace that attackers are gaining more useful information to knock down multiple grid components. Usually, a power system is guaranteed to be $N - 1$ secure due to computational complexity in evaluating multiple contingencies for a large power system. In particular, the list size or possible number of events, $C$ for $k$ contingency is given by,

$$C = \binom{N}{k} = \frac{N!}{k!(N-k)!} \quad (1.1)$$

where $N$ is the total number of components (nodes or branches) in power system and $k$ is the number of outaged/failed components. Hence, even for a modestly sized system with $N = 5000$, the number of combinations for the double outage is around 12.5 million. So, it is required to model the power system efficiently and effectively for simulating its behavior during multiple components failure.

Although the smart grid network introduces enhancements and improved capabilities to the conventional power grid, it is becoming more complex and vulnerable to different kinds of cyber attacks. It is also considered as a typical cyber-physical system due to tight coupling between ICT and physical power system. Vulnerabilities with ICT allows attackers to access the network and break confidentiality and integrity of data for interrupting the service. There are numerous access points within the smart grid which facilitate the attackers to compromise the system through potential cyber attacks. Recent
research has proved that an intentional cyber-attack can cause a significant impact on the power grid in terms of stability, efficiency and economic operation [11], [12], [13]. So, the interdependency between cyber and physical domains must be understood in order to implement and evaluate realistic cyber attack-defense experiments.

1.3 Motivations and Contributions

Determining the most vulnerable components (e.g., buses or generators) is critically important for power grid defense. Since it is infeasible to handle multi-contingency cases in large scale power grid due to computational complexity, researchers have been looking for optimal approaches to have a balance between cost of power grid modeling and efficiency of security analysis. In this work, a new definition of load is proposed by taking power flow into consideration in comparison with the load definition based on degree or network connectivity. Unsupervised learning techniques (e.g., K-means algorithm and self-organizing map (SOM)) are introduced to cluster the nodes (i.e., buses) in the IEEE-39 bus and IEEE-57 bus benchmarks. Then most vulnerable node in each cluster is determined based on their load information to form initial victim set. percentage of failure (PoF) is used to compare the performance of clustering based approach and traditional load-based approach during cascading failure process. With the simulation results, the unsupervised learning (clustering based) approaches are more efficient in finding the most vulnerable nodes and our proposed definition of load is relatively useful in studying power grid security.

The Large-scale power system outage is one of the most catastrophic disaster in modern society that results in enormous damage of billions per year for US economy
alone. So, system operators are required to maintain plans for any unforeseen events in the power system that forces power system to operate without reliability. Contingency analysis is one of the well-known methods to paint the future scenarios for any contingencies in the power system. However, a large number of possible $N-k$ combinations make their assessment computationally prohibitive. A new method to search most vulnerable transmission lines efficiently based on electrical distance is introduced. Specifically, a new electrical network is first built based on the impedance matrix (by inverting admittance matrix). Then, this impedance matrix is pruned based on the number of connection in topology network. Next, the common connections in two different structures (i.e., electrical network and topology network) will be observed for contingency experiments. Our results verify that violations of transmission lines limit due to contingencies are mostly associated with those common branches. In addition, voltage profile is further studied to validate that the vulnerable transmission lines found above are critical in power system stability.

The next-generation electricity grid, commonly known as “smart grid” is expected to address the major shortcomings of the existing electric grid. The existing power grid is upgraded into a smart grid through an intelligent communication infrastructures, layers of information, extensive computing and sensing technologies. Thus, these cyber and power components of grid together constitute a complex cyber-physical system and it is critical to understand the interdependencies among these domains. In addition, this integration also increases the risk of cyber attacks and introduces new vulnerabilities to the power system. Hence, it is always necessary to understand power system phenomenon during any kind of disturbances. Researchers need a power system test bed which can provide a
real-time platform for simulating different power system events and attacks. In this work, a real-time cyber-physical testbed is introduced to study about the power system security experiments. Various research applications supported by the proposed testbed has been presented. The impact of the possible cyber attack on physical power grid has been analyzed and possible mitigation strategy is suggested.

1.4 The Structure of Thesis

The rest of the thesis is organized as follows. In Chapter 2, the application of computational intelligence for cascading failures in a power system is discussed where unsupervised learning is adopted for multi-contingency analysis. A detailed description of the simulated model and performance comparison of attack strategies will be provided. Chapter 3 discusses the new approach for searching vulnerable branches in the power system. The proposed algorithm is discussed and applied to various standard test cases. The contingency experiment and transient stability results are presented in order to verify the proposed approach. Similarly, Chapter 4 will describe the architecture and application of proposed testbed developed at South Dakota State University where power system security experiments are investigated in real time. Finally, conclusions of the thesis and possible future works are presented in Chapter 5.
CHAPTER 2 UNSUPERVISED LEARNING FOR POWER GRID SECURITY

2.1 Introduction

The electric power grid interconnect generating units and load over large geographical areas into an entity to achieve efficient and reliable power supply. With thousands of substations and interconnected transmission lines, the modern electric power grid is regarded as one of the most complex networks. U.S. power grid as of today consists of more than 9,200 electric generating units with more than 1 million MW of generating capacity which is connected to more than 300,000 miles of transmission lines [14]. The electric power grid is regarded as an engineering marvel, however, for meeting the nation’s energy demand efficiently and reliably, aging infrastructures need extensive upgrades. A smart electric power grid, with intelligent use of information, increases the connectivity, automation, and coordination among generating plants, networks within the grid and consumers. With the purpose of adding resiliency to the existing electric power system, information and communication technologies (ICTs) are introduced and electric grid is becoming smarter these days. With the help of ICT network, smart grid provides better situational awareness during emergencies such as storms, earthquakes, and terrorist attacks. One of the significant features of the smart grid, in contrast to the traditional power system, is the large scale deployment of a two-way communication network connecting both the power plants and the end consumers. This interactive system will enhance the delivery of quality power, optimize the efficiency and stability at a lower cost with the help of computer-based automation.
2.1.1 Attack threats in Smart Grid

The smart grid boosts not only the economic benefits but also a growing number of potential threats from the cyber-space [15]. Because of the huge volume of data flowing through the power transmission network, they are more vulnerable to data interception and unauthorized modification, which can be utilized by an attacker to disrupt power system operations. Due to the large amount of information being exchanged in the network, an attacker with enough knowledge of the power grid can penetrate into the network and find the vulnerable components for initiating an attack. These smart attacks if wisely designed and launched successfully, can compromise some critical components causing a disastrous impact on the power grid. Recently, malicious attacks against power grids have drawn growing attention from many aspects, e.g., power industry, educational institutions, government and even the public. Since the smart grid generally referred as next-generation power transmission system are relying on communication networks and smart meters, great concerns regarding cyber intrusions are raised. Such intrusions and physical sabotage can be controlled by the attackers to target on critical substations and transmission lines to cause a large-scale power outage. Attackers are referred to as those people who have a strong will to carry out attacks in order to disable the power grid and they might be individuals, terrorists or any hostile countries. Individual attacks and terrorist attacks are highly possible, and some of them have already happened in past. In 2013, a 37-year old Arkansas man launched three attacks on local power grid [16]. Specifically, the attacker was interested in high voltage transmission lines and substation. In addition, on April 16, 2013, group of snipers assaulted an electrical substation near San
Jose, California [17]. They fired 17 giant transformers which then took a nearly a month to repair causing loss of 15 million USD. This incident served as a warning sign to power industry in the community. However, attacks from hostile countries are less often than individual and terrorist attacks. But, once they happen can take U.S. power grid down even for months [18].

One of the vulnerabilities of the smart grid that could be taken advantage of by the attackers is the possibility of cascading failure events events in power systems. During the cascading failure, a few failed components can trigger the collapse of normal power transmission and consequently results in a blackout. Because of the interconnected structure of modern grid, the result of an attack can trigger a cascading failure or a blackout as it allows the local failures to propagate throughout the system. These cases often occur in an unpredictable manner. Although US government claims that the electric power grid in the U.S. is 99.97 % reliable, still allows room for frequent power interruptions that cost $150 million each year [14]. Although many works have been done to enhance the security and reliability of the U.S. grid, major blackouts are still inevitable. Table 2.1 shows the notable blackouts in U.S. history for past 50 years.

2.1.2 Vulnerability of power grids to cascading failure

Power sector vulnerability has been a key issue in society for over a decade. A component failure may trigger a cascade of failures across the grid and lead to a large blackout. Cascading failures in large-scale electric power systems are a succession of transmission and generation outages, one precipitating another [19]. In order to provide continuously and high-quality electricity service to end consumers, the power system is
### Table 2.1. History of Major United States Blackouts

<table>
<thead>
<tr>
<th>Date</th>
<th>Location</th>
<th>Cause</th>
<th>Consequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nov. 1965</td>
<td>Northeastern of U.S., Ontario in Canada</td>
<td>Human errors</td>
<td>30 million people</td>
</tr>
<tr>
<td>Jul.1977</td>
<td>New York City</td>
<td>An electrical substation stroke by lighting</td>
<td>9 million people without power</td>
</tr>
<tr>
<td>Jan.1981</td>
<td>Utah</td>
<td>Knocking out transmission lines</td>
<td>1.5 million people lost power</td>
</tr>
<tr>
<td>Oct.1989</td>
<td>Northern California</td>
<td>Substations damaged by earthquakes</td>
<td>1.4 million people lost power</td>
</tr>
<tr>
<td>Jan.1989</td>
<td>Northeast of North America</td>
<td>Transmission towers destroyed by ice</td>
<td>3.5 million people affected</td>
</tr>
<tr>
<td>Aug. 2003</td>
<td>Northeast of U.S., Ontario in Canada</td>
<td>Transmission lines tripped by trees</td>
<td>55 million people without power</td>
</tr>
<tr>
<td>Sept. 2011</td>
<td>California</td>
<td>Technical error</td>
<td>7 million people without power</td>
</tr>
<tr>
<td>Jul. 2012</td>
<td>New York, New Jersey</td>
<td>Hurricane Sandy</td>
<td>10 million people without power</td>
</tr>
</tbody>
</table>

designed to operate securely during any single critical component failure, which is called the $N - 1$ criterion. However, an unexpected rare disturbance in unfavorable circumstances may trigger a series of device outages (cascading failures), or even a system breakdown resulting in a widespread power blackout. Thus, cascading failure is considered to be the leading reason of large-scale power outages. As reported in [20], failure to perform a critical redistribution of power for an overloaded transmission line resulted in the major blackout that affected more than 55 million people in the Northeast American region. Cascading failure refers to a sequence of dependent failures of individual components that successively weakens the power system and it includes the initial failure/s and dependent failures.

The initial failure/s can occur on substations, transmission lines, or other components. The cause of initial failures can be random like natural disasters (e.g.,
earthquakes), operational errors, equipment failures or can be malicious attacks like cyber intrusions or physical sabotages from the attackers perspective. Malicious attacks are more powerful and harmful than random cause since the attack can be controlled in terms of initiating events and a different number of targets. Dependent failures are triggered by initial failure/s. Many grid components are failed subsequently after initial failure/s which is referred to as failure propagation and the sequence of cascading event is uncertain [21]. This is because the power system always requires a balance between supply and demand. The initial failure of critical components could cause large-scale power redistribution which causes subsequent failures. A typical cascading-failure is divided into three stages: the pre-cascading stage, the cascading stage, and the post-cascading stage.

The pre-cascading stage refers to all events and preconditions before the cascading process happens. This stage is characterized by the gradual weakening of the system as different events occur. Typical events are bad weather, an unanticipated demand increase, unscheduled outage of generators or transmission lines, insufficient right-of-way maintenance, serious transmission congestion, hidden failures of protection systems. At the pre-cascading stage, a power system is stressed and the probability of cascading outages increases rapidly.

During cascading outage, a system-wide outage sequence is usually triggered by outages occurring at key locations. There will be series of uncontrollable tripping of system elements spreading from initial location to entire power grid network. The final phase of the cascading failures is the breakdown of the entire or the major portion of the system.

The restorative process of the power systems after the system breakdown is referred
as post cascading stage. The restorative process might take few hours to days and
sometimes even weeks depending on its location, the degree of damage, and propagation
size. In this thesis, the focus is on the understanding of the pre-cascading and cascading
stage. Post cascading stage is beyond the scope of our current work.

The electric substation is the major components in any power grids. The reliability
of power grid is strongly related to the security of the substation. In reality, substations are
confronting various risks, e.g., natural disaster [22], cyber attack [23], physical attacks
[16] [17], and so forth. Hence it is an urgent task to investigate the cascading failure from
substation perspective. An attacker’s goal during the attack is to identify the set of target
nodes whose simultaneous failure causes maximum damage to the electric power grid. So,
an attacker always targets certain substations to initiate a cascading failure that maximizes
the number of customers without electricity and causes tremendous economic loss. If an
attacker wants to launch successful and powerful attacks, they need to have prior
knowledge about following questions.

- In what ways can attackers initially attack the targets?

- Which cascading model is the best one to predict the attack performance and
  quantify the damage of attack?

- Which component(s) should be identified as targets and compromised initially?

Answering any of aforementioned questions needs a significant amount of research.
Since in the current literature, the electric power grid is considered as a cyber-physical
system, and there is growing attention against cyber-physical attacks [24], [25]. Also,
there are different models which have been developed previously by using different
information of power grids [26], [27]. However, in this dissertation two different
cascading model is used to evaluate the attack performance and it mainly focuses on the
third question from the attackers’ perspective. It is also assumed that an attacker has
enough knowledge about the power grid and cascading models.

2.1.3 Related work

This section will briefly review some of the existing researches related to smart grid
security issues. Smart grid, as an integration of power transmission networks and
communication networks, can be vulnerable in both physical and cyber space [24]. These
include challenges in accurate measurements and monitoring of power system states,
power transmission reliability against disruptive events, detection of malicious events as
well as control of access and authentication.

A significant amount of research work has been done and numerous issues have
been satisfactorily resolved for smart grid security. But, there are still many emerging
challenges in the applications of smart grid especially ensuring its security from various
power system attacks causing cascading failures and even blackout. This slows down the
effort of its installation and replacement of traditional power grids. According to [21],
uncertainties of blackouts are associated with three sources: (a) initiating or triggering
events, (b) sequence of dependent events, and (c) ultimate cost. A probabilistic approach
to cascading failure was proposed in [28] to predict the “next” event in cascading
phenomenon. However, the authors have neglected the uncertainties with the initial
triggering event. An extended topological approach was proposed for assessing the
vulnerability of power grid components during cascading failures with limited knowledge
about dynamics of power system information [29]. Transient stability and voltage security assessment using machine learning was studied in [30]. Random sampling techniques were considered for screening all the operating situations of the power system. The use of modular neural networks and multilayer perceptrons (MLP) helps to detect and recognize intrusions in computer networks with higher accuracy [31]. A distributed smart grid attack strategies to destabilize power system was provided to create cascading failure within multiple targets of the system [32]. Optimal node attack strategy (NAS) based on DC power flow analysis was used to investigate the vulnerability of nodes [33], where the authors proposed a new metric called risk graph for showing the hidden relationship in nodes. In [34] combination of unsupervised and supervised learning was used for the online security evaluation of $N - 1$ contingency. Generally, higher order contingency is one of the major contributions to cascading failures and the identification of the contingency set is the foremost step for studying power grid security [35]. The application of well-known higher order contingency ($N - k$) used for searching $k$ critical components has the drawback of high computational and analysis cost. To overcome the complexity of assessing all $N - k$ contingency combinations, risk analysis methodologies like cluster-based approach, enumeration of likely cascade paths, uniform sampling, and bulk analysis methods are discussed [21]. During the study of cascading failure in the power grid, it has been recognized that topological structure of power grid has a key impact on the propagation of cascading events [36], [37]. Utilization of spatial features of electric power grid can help to analyze the electrical system behavior in cascading failure scenarios [38].

In order to study the cascading process for power grid security, one needs to define
load associated with buses (nodes) so that it helps to keep track of which nodes is knocked down in successive time steps. Load of a particular node in the complex network can be defined as overall transmission capabilities, which is also referred to as extended betweenness of that node [39]. A new model was proposed to define a load of a node as a product of its own degree and sum of the degree of its immediate neighbors [38], [40]. A degradation model based on degree was proposed to evaluate the stress on a particular node due to the failure of neighboring nodes. This stress on node was supposed to decrease its life expectancy analogous to “wear out” process [41]. However, these definitions of load are based on network connectivity and ignore the power flow governed by basic circuit laws in the power system. So, these definitions may not be realistic and useful from power system viewpoint as they do not take power flow analysis into consideration.

2.2 Power Grid Modeling

In this section, the two different models of power grid used in the cascading failure experiment analysis are described.

2.2.1 System model

Among the power system model available for cascading analysis, in this work the topology-based model will be compared with proposed model during the cascading failure analysis. In order to represent the power grid as a topological network, it is assumed that a substation in the power grid is referred as a node and a transmission line connecting substation will be regarded as a branch. A substation may consist of a generator, load or it may be simply a pass-through transmission substation. In addition, a load of a particular
node is also defined, on which the process of cascading failure depends on. The two
different definition of load used in this work will be explained in the following section.

2.2.1.1 Definition of load based on network connectivity

First, a load of a particular node is defined based on degree (connectivity). Previous
work on cascading failures of high-level power grid structure has suggested that a load of
a particular node is related to the connectivity with/of its immediate neighbors [27], [42].
It suggests that a node, either connecting to a number of immediate neighbors or whose
direct neighbors have a greater connection, will take a greater portion of the load during
power delivery. Hence, in this model, a load of a particular node is defined as a product of
its degree and sum of the degree of its neighboring nodes. Let $k_v$ be the degree of a
particular node, the initial load, $L_v$ is defined by,

$$L_v = k_v \sum_{m \in Nbr(v)} k_m$$  \hspace{1cm} (2.1)

where $Nbr(v)$ is the set of neighbouring nodes of a particular node $v$. If one or more nodes
are knocked down, they are assumed to be out of service and the load flowing through that
node gets proportionally redistributed to the neighboring nodes. So, the nearby nodes
must take up the slack for the failed node. Hence the load gets redistributed according to
the model proposed in [27] and the load of immediate neighbors is updated with some
additional load as defined by following equations,

$$\delta_m = \frac{L_m}{\sum L_m} * L_v$$  \hspace{1cm} (2.2)
where $L_v$ is the initial load of failure node, $\delta_m$ is the additional load assigned to neighbouring nodes due to failure of node $v$. In order to explain the definition and distribution of load, a numeric example for equations 2.1 and 2.2 is shown in Fig. 2.1. It represents a small portion of IEEE 39 bus system where buses (nodes) are represented by circle and number inside the circle represent the bus number. The number next to each node in black is the initial load of that particular node. For example, the load of node 16 is given by,

$L(16) = k(16) \times \{k(17) + k(15) + k(19) + k(24) + k(21)\}$

$L(16) = 5 \times \{3 + 2 + 3 + 2 + 2\} = 60$

When node 16 is failed either by fault or some direct attacks, it will be disconnected from the power grid. This means the connection between this node with its neighboring nodes will also be removed since no power could be delivered through it. So, based on equation (2.3), the load of the failed node will be redistributed to its immediate neighbors.
The number in blue is the extra load that will be added on each of the neighboring nodes of node 16. The newly added load is proportional to the initial load of neighboring nodes in accordance with defined equation.

2.2.1.2 Definition of load based on Power Flow

The first model is a pure topological model that neglects the fundamental electric laws behind power flow. Since these models typically assume that cascades propagate locally, they can be misleading [43]. For instance, if any component $x$ fails, the next component to be overloaded and fail is one that is topologically connected to $x$. However, as suggested by [44], real cascading failure propagates non-locally which means the next component to fail after $x$ may be hundred of miles away from it. In this model, a load of a particular node is defined as the burden of total MW that it carries during its healthy state. When electric power from the generating units is delivered to end consumers, it can travel through different routes through different buses (nodes). The node may have some MW of real load, or a generator, or it may be simply a transmission path delivering power to other substations. Assume $D$ is the demand present at a particular node $v$, $I_1, I_2, ..., I_k$ be the amount of power flowing into node $v$ from $k$ neighbouring nodes and $O_1, O_2, ..., O_m$ be the amount of power flowing out from node $v$ to $m$ neighbouring nodes. An initial load that a particular node carries during normal operation of the grid is defined by equation (2.4). Information regarding power flowing into the node $I$, flowing out of node $O$, and demand in that node $D$ can be obtained by running power flow experiments. The power flow is
presented in Fig. 2.2 and the load definition is provided as,

\[ L(v) = \sum_{i=1}^{k} I_i = \sum_{j=1}^{m} O_j + D, \quad v \in n \]  

(2.4)

where \( v \) is a given node of a system with \( n \) nodes.

Figure 2.2. Load burden of a node using power flow information.

The attack model for this case is similar to that as described earlier in section 2.2.1.1. But, the redistribution of load after failure of particular node is based on circuit laws and independent of the initial load of neighboring nodes as described in equation (2.2). Also, this redistribution is not only limited to adjacent network components [45] and despite overloading, there may be a situation in cascading analysis where a neighboring node may get functionally disconnected from the system causing load
shedding at those nodes or form isolated networks [46]. After each node is taken off from the system because of an attack or cascading failure process, the load redistribution in the remaining network takes place according to power flow method. A well-known Full-Newton method is used for analyzing the system behavior after the loss of a particular node. The relationship between node current \( I \) and node voltage \( E \) for a particular node \( i \) in a network of \( n \) nodes is given by the following linear equation,

\[
I_i = \sum_{k=1}^{n} Y_{ik} E_k
\]  

(2.5)

where \( Y_{ik} \) is an element of the admittance matrix joining nodes \( i \) and \( k \). Following equation (4.6), complex power at node \( i \) is given by,

\[
S_i = E_i I_i^* 
\]  

(2.6)

\[
P_i + jQ_i = E_i \sum_{k=1}^{n} Y_{ik}^* E_k^* 
\]  

(2.7)

Equation (3.3) represents real and reactive power flowing in any branch and it is used to update the branch flow according to \( Y_{bus} \) after any outage of node(s). So, the survived nodes will have new value of load assigned to them. In this work, only real power is taken into account for the calculation of load of the nodes as shown in Fig. 2.2.

2.2.2 Cascading tree

As the load is redistributed in the network according to the model described in previous sections, this will cause some of the surviving nodes in the grid to be overloaded. When a node is overloaded beyond some predefined capacity, it is also taken off from the
system and all the branches or transmission lines connected with that node will be disconnected. The capacity $C(v)$ of each node is defined which is directly proportional to its initial load $L(v)$ that it carries in a healthy network as,

$$C(v) = \alpha L(v), \ v \in n$$

(2.8)

where $\alpha \geq 1$ is the threshold of overloading ratio, above which a node is considered failed and referred as system capacity. A Higher value of $\alpha$ means the higher capability of a node to resist perturbations. The failure propagation will continue in the system as long as new overloaded node appears in the grid leading to cascading process. A universal system tolerance is assigned to define capacity of all nodes within the network as there is no ground truth for the practical value of system tolerance. So, during the simulation for cascading failure analysis different values of $\alpha$ will be used.

Finally, when a number of nodes are failed, the concept of “round” is used to help to describe the successive propagation of cascading failure. The definition of a round is illustrated in Fig. 2.3. The very first set of failed nodes are the victims in the initial attack as chosen by an attacker. Then the nodes knocked down by the cascading failure of initial victims will be regarded as the victims of the second round, so on and so forth. In this way, failed nodes at different rounds of a cascading process form a tree-like structure where the “child” nodes are the direct victims of their parent node’s failure.

In summary, the overall cascading failure simulator and concept of round can be generalized in following steps:

1. Trigger a multi-node cascading failure by knocking down some victims in the grid.
2. Calculate the load redistribution because of failures and mark fatally overloaded nodes as failed in the next round.

3. Disconnect failed nodes and branches from the grid.

4. Repeat step 2 and 3 until the process reaches a final stabilized state.

2.2.3 Evaluation metric

As it is desired identify the most critical power grid components from the cascading failure perspective, the damage of the attack needs to be quantified. The final percentage/fraction of failure is used in the power grid with respect to system tolerance $\alpha$, denoted as PoF, as the assessment metric:

$$PoF = 1 - \frac{N'}{N}$$  \hspace{1cm} (2.9)
where $N$ is a total number of nodes in the system before the attack and $N'$ is a number of nodes survived the attack. For each attack simulated, the value of PoF is measured after the cascading failure stops at the final stabilized state. The physical meaning of PoF is related to the size of blackout as several components are knocked down during the cascading failure process. According to the previous definition of “round”, a cascading “tree” with more “leaves”, i.e., a higher value of PoF at the final round, indicates that the initial victims have compromised a large number of grid components during cascading failure causing a larger blackout. By using this measurement, it is able to illustrate the effectiveness of the proposed approach and identify the critical components in the power grid during multi-victim attack scenarios.

2.3 Multi-component Attack Strategies

The primary goal of this research is to analyze the cascading failures during various multi-victim attack strategies. Traditional load based attack and clustering based attack are compared for both definitions of load, i.e., load based on degree or connectivity and load based on power flow. After the cascading failure experiment, the performance of two different attack strategies is compared for two different power grid models separately as shown in Fig. 2.4.

2.3.1 Traditional load based attack strategy

The attack on nodes with the highest load is a common attack strategy and is based on the fact that the failure of a node with the largest value of load causes a significant amount of load to be redistributed among its neighbors. The optimal strategy in this attack is to select the desired number of victim nodes in descending order of loads and to remove
them from the network. Authors in [47] studied cascading failures in North American power grid using information from its network structure. Loss of single high-load or high-degree substation reduced the efficiency of the power grid by 25%. It seems reasonable to choose victim nodes according to the load, but one should always remember that network topology plays a crucial role in complex network failures.

2.3.2 Unsupervised learning based attack strategy

Unsupervised learning method draws inferences from the datasets without labeled responses or without external help. By contrast with supervised learning, there are no explicit target outputs associated with each input. Since clustering or partitioning of data in the absence of class labels is often a requirement, unsupervised learning tries to build a model from data without any external help [48]. Among different approaches to
unsupervised learning, it is used for clustering analysis. The commonly used unsupervised
learning algorithms in data clustering are k-means and self-organizing map (SOM).

2.3.2.1 K-means Clustering

K-means clustering (MacQueen, 1967) is a method commonly used to partition n
data set into k groups in which each data belongs to the cluster with nearest mean [49]. It
is one of the simplest unsupervised learning algorithms for solving the clustering problem.
The procedure follows a simple way to classify given a set of data to a certain number of
clusters fixed apriori. K-means algorithm attempts to find the cluster centers so as to
minimize the within-cluster sum of square (WCSS) distance. Simply, its objective is to
minimize the following objective function,

\[ J = \sum_{j=1}^{k} \sum_{i=1}^{n_j} |V_{ij}^t - C_j|^2 \]  

where \( V_{ij}^t \) is the \( i^{th} \) input of \( j^{th} \) cluster, \( C_j \) is the centroid of a particular cluster, \( k \) is the
total number of clusters, and \( n_j \) is the number of data set belonging to \( j^{th} \) cluster. The
algorithm behind this objective can be summarized as:

Step 1. Initialization: Choose \( C_1, C_2, \ldots, C_k \) as initial cluster centers. They represent
the “temporary” means of the clusters. The dimension of initial centers is same as that of
the input vector.

Step 2. Assignment: Each observation (data set) is assigned to a closest cluster by
calculating the squared Euclidean distance.

\[ D = |V(t) - C(t)| \]  

(2.11)
where \( V(t) \) is the sample input vector at \( t^{th} \) iteration, \( C(t) \) is the weight vector of cluster center.

Step 3. Update: The new centroid of the clusters are calculated using (2.12)

\[
C_j = \frac{1}{n_j} \sum_{V_i \in j} V_i
\]  
(2.12)

where \( n_j \) is the number of data set \( (V_i) \) belonging to a particular cluster \( j \) after assignment. Since the arithmetic mean is a least square estimator, this will minimize the WCSS objective in equation (2.10).

Step 4. Convergence: Repeat step 2 and 3 until there is no further change in position of centroid of each cluster.

2.3.2.2 Self Organizing Map

Self-organizing map commonly known as Kohonen network is a class of artificial neural network (ANN) in an unsupervised learning category [50]. It is an effective platform for visualizing and analyzing high dimensional data. The two modes of operation of SOM are training and mapping. During the training process, the neurons are settled at different locations in the lattice. During the mapping process, the input vector is clustered. The number of neurons depends on the number of initial victims targeted for multi-victim attack. The training of neurons occurs in several steps as described below.

Step 1. Initialization: Before training, each neuron’s weight needs to be initialized. Among various initialization processes, random initialization approach is used for its simplicity. So, the weight of the neurons is set to small standardized random values such that \( 0 < W < 1 \), where \( W \) is the weight vector of the neuron. The dimension of weight
vector is same as that of the input vector.

Step 2. Best Matching Unit (BMU): After initialization, a sample input vector is selected from the set of training data and presented as input to SOM. Each neuron is now examined to find which one’s weight is closest to the current input. The winner is called the BMU. To determine this unit, Euclidean distance between each neurons’ weight vector and current input vector being presented to lattice is calculated as,

\[
\text{Dist} = |V(t) - W(t)|
\]  

(2.13)

where \(V(t)\) is the sample input vector at \(t^{th}\) iteration and \(W(t)\) is the weight vector of neuron. The neuron that minimizes this distance is tagged as BMU.

Step 3. Parameters Update: Now, a neighborhood function is defined to calculate which neurons are within the neighborhood of BMU. The Gaussian kernel function is used as neighborhood function which adjusts the weights of neurons based on its distance to BMU. It is defined by,

\[
H(\sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\delta^2/2\sigma^2}
\]  

(2.14)

where \(\delta\) is the distance between the neuron and BMU.

A unique feature of this learning is that the area of the neighborhood and learning rate decreases over time. The neighborhood size (i.e., \(\sigma_o\)) and SOM learning rate (i.e., \(L_o\)) are updated using exponential decay function defined by,

\[
L(t) = L_o e^{-t/\lambda_t}
\]  

(2.15)
\[ \sigma(t) = \sigma_0 e^{-t/\lambda_2} \]  
\[ \text{(2.16)} \]

where \( t \) is number of iterations and \( \lambda_1 \) is the time constant whose value depends on \( \sigma_0, \lambda_1 = 100/log(\sigma_o) \).

Step 4. Convergence: After the parameters are updated, neurons are dragged towards input vector to adjust their weight. Since Gaussian kernel is used as neighborhood function, the weight of the neurons closer to BMU will be updated as,

\[ W(t+1) = W(t) + H(\sigma)L(t)(V(t) - W(t)) \]  
\[ \text{(2.17)} \]

At this stage, the training of neurons is finished and they will find their own positions in the lattice. Each of them represents the centroid of clusters. Finally, the mapping process starts in which the input vector will be clustered according to the Euclidean distance.

2.4 Simulation Results

The IEEE-39 bus and IEEE-57 bus system are studied for analyzing the multi-victim attack from two different strategies. A simulator is built in MATLAB 2014a environment for simulating load redistribution process for different attack strategies. For preserving the topological information in both benchmark system, X and Y coordinates for different nodes are extracted. A system case with branch and bus state for 39 bus system is obtained from Illinois Center for a Smarter Electric Grid (ICSEG) [51] and 57 bus system from Matpower [52].
2.4.1 Parameters and environment setup

The coordinates of buses (nodes) for both test cases are normalized by using the following equation as,

\[ z = \frac{z - \text{min}(Z)}{\text{max}(Z) - \text{min}(Z)}, \quad z \in Z \] (2.18)

where \( z \) stands for either \( X \) or \( Y \) coordinate.

We plan to attack two different nodes in IEEE-39 bus system so the number of neurons in SOM training are 2. 1-D SOM lattice is used for training of neurons and the dimension of input vector in our case is 2, i.e., \( X \) and \( Y \) coordinate, each neuron carries a 1-by-2 weight vector. The weight of two neurons is set to uniform random values such that \( 0 < W < 1 \). In our simulation, the initial values of neighborhood size (i.e., \( \sigma_0 \)), SOM learning rate (i.e., \( L_0 \)), and length of rough training (\( \lambda_2 \)) are selected as 1.2, 0.01, and 15 respectively.

Since we plan to attack three different nodes in the IEEE-57 bus system, we have three different initial cluster centers. The dimension of the sample input vector in our case is 2, i.e., \( X \) and \( Y \) coordinate and hence these cluster centers also carry a 1-by-2 weight vector. The weight of three cluster centers is set to uniform random values such that \( 0 < C < 1 \).

2.4.2 SOM and K-means clustering

From Fig. 2.5, it can be seen that 39 buses (nodes) of IEEE-39 bus system benchmark are grouped into two different clusters using SOM. Similarly, Fig. 2.6 shows that the 57 buses (nodes) of the IEEE-57 bus system are grouped into three different clusters using K-means algorithm. Centroid of each cluster is represented by a solid
Figure 2.5. Two clusters of IEEE 39 bus system using SOM

Figure 2.6. Three clusters of IEEE 57 bus system using K-means
(green) circle in both the clustering results.

2.4.3 Performance comparison

2.4.3.1 Power system model based on network connectivity

In this model, a load of a particular node is calculated using equation (2.1). Based on the loading information of all nodes, an attack was initiated on power system benchmark from both traditional load-based and proposed clustering-based approach. For traditional load based strategy, the nodes are sorted according to their initial load and find the most loaded ones to form the initial victim set. In IEEE-39 bus system, it is found that nodes 16 and 26 are the most loaded ones and in 57 bus system nodes 15, 13 and 9 are most loaded nodes. These set of nodes are selected to form the initial victim set for traditional load based attack strategy. In order to study the cascading failure process, the initial set of nodes are knocked down in both benchmark system. This causes the redistribution of the load in remaining nodes according to equation (2.2) and it is observed that some of the nodes are overloaded beyond their capacities. Next, these overloaded nodes are also knocked down. This process is repeated several times until the system achieved a final steady state. Finally, the number of survived components in steady state is evaluated and PoF is calculated using equation (2.9).

To further explain the cascading phenomenon, failure of different nodes are presented in successive time steps for IEEE-39 bus system at system tolerance of 1.6 as shown in table 2.2. The benchmark is simulated in Matlab for calculation and redistribution of load. Nodes 16 and 26 in gray colored cells are knocked down to initiate an attack. These set of nodes corresponds to the load based attack strategy. As these
Table 2.2. Cascading events for IEEE 39 bus system at tolerance of $\alpha =1.6$.

<table>
<thead>
<tr>
<th>Bus No.</th>
<th>Initial load $L_i$ (C)</th>
<th>Capacity $C$</th>
<th>Round of cascading events</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
<th>5th</th>
<th>6th</th>
<th>7th</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12</td>
<td>19.2</td>
<td></td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>24.1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>36</td>
<td>57.6</td>
<td></td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>58.0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>27</td>
<td>43.2</td>
<td></td>
<td>27</td>
<td>27</td>
<td>54.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>27</td>
<td>43.2</td>
<td></td>
<td>27</td>
<td>27</td>
<td>53.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>30</td>
<td>48</td>
<td></td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>60.4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>36</td>
<td>57.6</td>
<td></td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>74.1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>22.4</td>
<td></td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>52.6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>21</td>
<td>33.6</td>
<td></td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>43.2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>16</td>
<td></td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>23.2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>21</td>
<td>33.6</td>
<td></td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>51.15</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>27</td>
<td>43.2</td>
<td></td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>102.2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>19.2</td>
<td></td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>29.2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>24</td>
<td>38.4</td>
<td></td>
<td>24</td>
<td>24</td>
<td>47.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>24</td>
<td>38.4</td>
<td></td>
<td>24</td>
<td>24</td>
<td>49.6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>16</td>
<td>25.6</td>
<td>25.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>60</td>
<td>96</td>
<td></td>
<td>96</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>27</td>
<td>43.2</td>
<td>25.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>12</td>
<td>19.2</td>
<td>27.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>24</td>
<td>38.4</td>
<td>25.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>8</td>
<td>12.8</td>
<td>36.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>21</td>
<td>16</td>
<td>25.6</td>
<td>25.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>22</td>
<td>18</td>
<td>28.8</td>
<td>36.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>23</td>
<td>18</td>
<td>28.8</td>
<td>25.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>24</td>
<td>25.6</td>
<td>25.7</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>25</td>
<td>27</td>
<td>43.2</td>
<td></td>
<td>41.2</td>
<td>41.2</td>
<td>41.2</td>
<td>41.2</td>
<td>82.9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>26</td>
<td>40</td>
<td>64</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>27</td>
<td>14</td>
<td>22.4</td>
<td>49.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>28</td>
<td>14</td>
<td>22.4</td>
<td></td>
<td>21.3</td>
<td>21.3</td>
<td>21.3</td>
<td>21.3</td>
<td>21.3</td>
<td>21.3</td>
<td>21.3</td>
</tr>
<tr>
<td>29</td>
<td>21</td>
<td>33.6</td>
<td></td>
<td>32.0</td>
<td>32.0</td>
<td>32.0</td>
<td>32.0</td>
<td>32.0</td>
<td>32.0</td>
<td>32.0</td>
</tr>
<tr>
<td>30</td>
<td>4</td>
<td>6.4</td>
<td></td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>8.0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>31</td>
<td>4</td>
<td>6.4</td>
<td></td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>6.4</td>
<td>0</td>
</tr>
<tr>
<td>32</td>
<td>3</td>
<td>4.8</td>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>8.1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>33</td>
<td>3</td>
<td>4.8</td>
<td>13.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>34</td>
<td>2</td>
<td>3.2</td>
<td></td>
<td>2</td>
<td>2</td>
<td>38.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>35</td>
<td>3</td>
<td>4.8</td>
<td></td>
<td>3</td>
<td>3</td>
<td>5.8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>36</td>
<td>3</td>
<td>4.8</td>
<td></td>
<td>3</td>
<td>3</td>
<td>8.4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>37</td>
<td>3</td>
<td>4.8</td>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>85.9</td>
</tr>
<tr>
<td>38</td>
<td>3</td>
<td>4.8</td>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>39</td>
<td>8</td>
<td>12.8</td>
<td></td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>32.1</td>
</tr>
</tbody>
</table>
victim nodes are knocked down to initiate an attack, the load carried by those victim
nodes gets redistributed to their immediate neighbors. Now the updated load is calculated
using equation (2.2) and compared with their capacity. If their updated load is higher than
capacity, i.e., if they are overloaded then these nodes are again knocked down in next time
steps. The overloaded nodes in different time steps are represented by bold figures in the
table.

For clustering based approach, the mostly loaded nodes from each cluster are
selected to form the initial victim set. For 39 bus system it is found that node 16 and 2 are
the most loaded ones and in 57 bus system nodes 9, 13 and 12 are the most loaded. The
load redistribution process and the cascading phenomenon are similar as described for a
traditional load-based attack strategy. PoF for the clustering-based attack is calculated and

![Graph showing performance comparison](image_url)
compared with traditional approach for both benchmark system. From Figs. 2.7 and 2.8, it is observed that the clustering-based attack is able to find the victim set with nodes carrying less load but with greater impact than the traditional load based strategy.

2.4.3.2 Power system model based on actual power flow

In this approach, the victim nodes are selected based on the load defined by equation (2.4). The power flow information is obtained from PowerWorld and Matpower for 39 bus system and 57 bus system respectively. After calculating the load of each bus, the attack is initiated by selecting victim set for both traditional load-based and clustering-based strategy. For traditional load-based strategy, nodes 6 and 39 are found to be most loaded for 39 bus system. Whereas for 57 bus system, nodes 8, 1 and 12 are most loaded. These set of nodes are the initial victims which are knocked down to trigger a cascading event.
Table 2.4. Cascading events for IEEE 39 bus system at tolerance of $\alpha =1.25$.

<table>
<thead>
<tr>
<th>Bus No.</th>
<th>Initial load $L_i$</th>
<th>Capacity C</th>
<th>Round of cascading events</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1st</td>
</tr>
<tr>
<td>1.0</td>
<td>116.3</td>
<td>145.4</td>
<td>0.0</td>
</tr>
<tr>
<td>2.0</td>
<td>481.5</td>
<td>601.9</td>
<td>872.2</td>
</tr>
<tr>
<td>3.0</td>
<td>398.8</td>
<td>498.5</td>
<td>897.5</td>
</tr>
<tr>
<td>4.0</td>
<td>500.0</td>
<td>625.0</td>
<td>1274.2</td>
</tr>
<tr>
<td>5.0</td>
<td>482.8</td>
<td>603.5</td>
<td>769.0</td>
</tr>
<tr>
<td>6.0</td>
<td>910.0</td>
<td>1137.5</td>
<td>0.0</td>
</tr>
<tr>
<td>7.0</td>
<td>426.8</td>
<td>533.5</td>
<td>233.8</td>
</tr>
<tr>
<td>8.0</td>
<td>523.0</td>
<td>653.8</td>
<td>763.0</td>
</tr>
<tr>
<td>9.0</td>
<td>12.1</td>
<td>15.1</td>
<td>0.0</td>
</tr>
<tr>
<td>10.0</td>
<td>650.0</td>
<td>812.5</td>
<td>650.0</td>
</tr>
<tr>
<td>11.0</td>
<td>350.3</td>
<td>437.9</td>
<td>31.9</td>
</tr>
<tr>
<td>12.0</td>
<td>7.5</td>
<td>9.4</td>
<td>31.8</td>
</tr>
<tr>
<td>13.0</td>
<td>299.7</td>
<td>374.6</td>
<td>642.4</td>
</tr>
<tr>
<td>14.0</td>
<td>292.9</td>
<td>366.1</td>
<td>717.8</td>
</tr>
<tr>
<td>15.0</td>
<td>320.1</td>
<td>400.1</td>
<td>397.4</td>
</tr>
<tr>
<td>16.0</td>
<td>823.9</td>
<td>1029.9</td>
<td>822.7</td>
</tr>
<tr>
<td>17.0</td>
<td>204.1</td>
<td>255.1</td>
<td>183.8</td>
</tr>
<tr>
<td>18.0</td>
<td>192.3</td>
<td>240.4</td>
<td>183.7</td>
</tr>
<tr>
<td>19.0</td>
<td>629.2</td>
<td>786.5</td>
<td>629.0</td>
</tr>
<tr>
<td>20.0</td>
<td>680.2</td>
<td>850.3</td>
<td>680.2</td>
</tr>
<tr>
<td>21.0</td>
<td>604.6</td>
<td>755.8</td>
<td>604.2</td>
</tr>
<tr>
<td>22.0</td>
<td>650.0</td>
<td>812.5</td>
<td>650.0</td>
</tr>
<tr>
<td>23.0</td>
<td>601.3</td>
<td>751.6</td>
<td>601.4</td>
</tr>
<tr>
<td>24.0</td>
<td>353.8</td>
<td>442.3</td>
<td>353.9</td>
</tr>
<tr>
<td>25.0</td>
<td>538.3</td>
<td>672.9</td>
<td>538.3</td>
</tr>
<tr>
<td>26.0</td>
<td>409.6</td>
<td>512.0</td>
<td>512.8</td>
</tr>
<tr>
<td>27.0</td>
<td>281.9</td>
<td>352.4</td>
<td>372.9</td>
</tr>
<tr>
<td>28.0</td>
<td>347.5</td>
<td>434.4</td>
<td>347.8</td>
</tr>
<tr>
<td>29.0</td>
<td>824.8</td>
<td>1031.0</td>
<td>824.7</td>
</tr>
<tr>
<td>30.0</td>
<td>250.0</td>
<td>312.5</td>
<td>741.0</td>
</tr>
<tr>
<td>31.0</td>
<td>571.3</td>
<td>714.1</td>
<td>9.2</td>
</tr>
<tr>
<td>32.0</td>
<td>650.0</td>
<td>812.5</td>
<td>650.0</td>
</tr>
<tr>
<td>33.0</td>
<td>632.0</td>
<td>790.0</td>
<td>632.0</td>
</tr>
<tr>
<td>34.0</td>
<td>508.0</td>
<td>635.0</td>
<td>508.0</td>
</tr>
<tr>
<td>35.0</td>
<td>650.0</td>
<td>812.5</td>
<td>650.0</td>
</tr>
<tr>
<td>36.0</td>
<td>560.0</td>
<td>700.0</td>
<td>560.0</td>
</tr>
<tr>
<td>37.0</td>
<td>540.0</td>
<td>675.0</td>
<td>540.0</td>
</tr>
<tr>
<td>38.0</td>
<td>830.0</td>
<td>1037.5</td>
<td>830.0</td>
</tr>
<tr>
<td>39.0</td>
<td>1116.3</td>
<td>1395.4</td>
<td>0.0</td>
</tr>
</tbody>
</table>
For clustering based approach, the mostly loaded nodes from each cluster are selected to form the initial victim set. Based on Fig. 2.5, nodes 38 and 39 are selected from two different clusters for 39 bus system. Similarly for 57 bus system, nodes 8, 1 and 13 are selected as initial victim set as shown in Fig. 2.6. These nodes are marked with a red circle in clustering results of both benchmark system.

To explain the cascading process in this model, load based attack in IEEE-39 bus system is considered. Nodes 6 and 39 are knocked down initially to initiate an attack. As these victim nodes are knocked down, the load carried by those victim nodes is redistributed according to basic circuit laws. Unlike the previous model, it is observed that the load redistribution is not only limited to immediate neighbors and independent of the initial load of nodes. The updated load of nodes are calculated using branch flow information from PowerWorld. Next, the overloaded nodes are identified and knocked down in successive time steps. The whole process is summarized in table 2.4.

Finally, the number of survived components in steady state is evaluated and PoF is calculated using equation (2.9). Performance comparison in Figs. 2.9 and 2.10 suggests that clustering-based attack is more efficient than load-based attack in searching of vulnerable components which can cause greater damage in power system benchmark. In addition, Fig. 2.9 shows different behavior of SOM based attack from system tolerance of 1.75 to 2.5 which represents system blackout. In this simulation, when system tolerance is increased from 1.75 to 2.5 in SOM based attack, nodes 4, 26, and 27 (nodes with some MW of demand) survived due to increased capacity, whereas nodes 5, 13, 14, and 17 failed, which represented a pathway for power to flow.
Figure 2.9. Performance comparison of traditional load based and clustering based attack for IEEE-39 bus system based on power flow.

Figure 2.10. Performance comparison of traditional load based and clustering based attack for IEEE-57 bus system based on power flow.
2.4.4 Transient stability

In this case, the transient stability analysis with a large disturbance is investigated. Our benchmark IEEE-39 bus system consists of ten synchronous generators. The transients in rotor angle of all generators after a large disturbance is observed. A balanced three-phase fault is created at $t = 1$ s and cleared at $t = 1.2$ s in two different pair of nodes, i.e., nodes 6 & 39 (load based) and nodes 38 & 39 (SOM-based). Following this disturbance, the frequency of synchronous machines undergo transient deviations and the power angle of these machines will oscillate over time. The oscillations, unless damped, can lead to grid failure causing the system to collapse. The main objective is to study whether the rotor angles return to a steady state value to maintain synchronism following the clearing of disturbance.

![Figure 2.11. Rotor angle of generators for load based attack.](image-url)
Figs. 2.11 and 2.12 shows the rotor angles of all ten generators for a load based attack (i.e., nodes 6 and 39) and the SOM-based attack (i.e., nodes 38 and 39) respectively. For the load based attack, it is observed the rotor angles of all ten generators are able to achieve a new steady state after the damping process. Thus, the system is able to achieve synchronism with the transition from one equilibrium state to another. For the SOM-based attack, it is observed that the rotor angles diverge over time and most of the generator’s rotor angles are overlapped as shown in Figure 2.12. Synchronous generators are not able to return to a new equilibrium state with new steady state rotor angles. This situation is not expected in power systems and is referred to as “going out of step” or “loss of synchronism”. This supports our investigation that an attack at nodes 6 and 39 is less vulnerable than at nodes 38 and 39.
2.5 Summary

The simulation of cascading failure for power system benchmark using the concept of network connectivity ignores the actual power flowing through the buses. This concept assumes that in any interconnected network, the nodes having a greater number of connection with its neighbors should take a higher amount of load for delivery. Also, the load redistribution after the failure of a parent node is limited to its immediate neighbors only. But this is not always true in power system network because the power flowing in the electric grid depends on demand at a particular node, generating units in a node, and transmission line parameters. For simplicity, the cascading phenomenon in the electric power grid can be explained based on network connectivity. However, the results are not supportive to study the real cascading event in power system as it ignores load loss, the formation of isolated networks and system blackout.

The concept of load flow is introduced in addition to network connectivity to calculate the actual burden of a node by taking power flow into considerations. During simulating the cascading process using real power flow information, the formation of isolated networks as well as load loss is taken into account to calculate the severity of damage. The performance comparison of two different attack strategies is made based on the failure of a number of grid components. It is observed that preserving both the electrical features and topological information of grid can assist in selecting most vulnerable components in cascading failures. Thus, for both test cases, the clustering-based strategy is more effective than traditional load based strategy in selecting the victim sets for initiating an attack to cause the cascading failure.
Not only substations, transmission lines are also vulnerable to different kinds of failures in the power system. Because of the loss of transmission lines, the imposed limits of the power system components gets violated and it goes into emergency condition. So, it is required to study the effect of outages in terms of severity. However, it is challenging to deal with each and every combination of the outage to evaluate its consequences. In the following chapter, the detailed explanation of proposed approach for finding vulnerable branches during contingency experiments efficiently will be provided.
CHAPTER 3 ELECTRICAL DISTANCE APPROACH FOR CONTINGENCY ANALYSIS

3.1 Introduction

The electric power grid is regarded as one of the critical infrastructures in the world. It is a complex network consisting of numerous components like generators, transformers, transmission lines, circuit breakers, etc. As the power systems continue to increase in size and complexity because of grid modernization, cascading events leading to blackout are more likely to happen. U.S. electricity consumers incur a loss of almost $80 billion annually because of power interruptions [53]. The operation of the power grid has also changed dramatically due to change in the way the system is being owned and operated because of deregulation. With the aim of increasing reliability and efficiency of the power grid, the use of advanced information and communication technology is incorporated in addition to the physical infrastructure [54], [55]. However, security is a critical issue in the operation of such cyber-physical system because of increasing malware from cyber side to compromise the physical components. Since many large scale electrical disruptions have occurred in the past [56], [57], it has become necessary to ensure the operation of power systems economically and reliably.

3.1.1 Power system security

A detailed security assessment is essential in dealing with all credible outages in the system, its consequences and the remedial actions for them. For a power system to be secure, it must have a continuous supply of power without loss of load. With this aim,
security analysis is performed to develop various control strategies to guarantee the avoidance and survival of any emergency conditions. Whenever the imposed limits of the power system get violated, the system is said to be in emergency condition. The limits can be line limit, voltage limit or generation limit and these violations of the limit occurs due to contingencies occurring in the system. Thus an important part of the security analysis revolves around the power system ability to withstand the effect of any contingencies. An important factor in the operation of the power grid is the desire to operate it robustly because any kind of unplanned outages could lead to cascading events or even costly blackouts. One of the major agenda of power system planning and its operation is to study the effect of outage in terms of severity. Power system security also involves the contingency analysis where the simulation is conducted for the list of possible outages so as to give system operators a symptom of how a system will behave in an event of unscheduled failure of power grid components.

3.1.2 Contingency analysis

Contingency analysis (CA) is a well-known function in modern energy management system (EMS) to ensure power system security during equipment failure. It assists engineers to operate power system at a secure operating point, where transmission lines are loaded within their safe limit and consumers are provided power with acceptable quality standards [58]. In general, an outage of one transmission line or transformer or combination of different outages may lead to overload in other branches and/or sudden system voltage rise or drop. CA is used to calculate violations and analyse those violations for maintaining system security. It executes a power flow analysis for each
credible contingency event defined on the contingency list [54], which in turn helps to identify the thermal and voltage violations. Results after each power flow study are compared with limits of each element in the power system to identify the violations. For instance, a transmission line that was loaded at 80% of its MVA rating before any event might be overloaded above 100% after some outages in the system.

Introduction of new North American Reliability Corporation (NERC) Standards necessitate the system operators to ensure that the performance of power system is within the operating limits such that all single and multiple contingencies do not result in cascading outages [59]. While these standards mandate the power industry to consider multiple contingencies, it is still challenging to solve the problem due to a high number of possible events. Multiple approaches have been proposed previously to address the complexity problem of $N - k$ contingency. Because of the way the power system is designed and operated, not all the outages will actually cause trouble. Hence, most of the time and effort spent while running power flow experiment will go for solutions which discover that there are no any violations in the system. In fact, only a few of the power flow solutions will conclude an overload or voltage violation in the system. The solution to this situation is to find an efficient way to select only those contingencies that are likely to result in an overload of branches or voltage limit violations.

3.1.3 Related work

Contingency screening or contingency selection is an essential task in contingency analysis which helps to reduce the numerous computations. Contingency selection criterion based on the calculation of performance indices has been first introduced by
Ejebe and Wollenberg [60], where the contingencies are sorted in descending order of the values of performance index (PI) reflecting the severity. However, the results were not reliable and to improve the reliability, authors in [61] proposed the use of higher-order sensitivities. One way to gain speed of solution in a contingency analysis procedure is to use an approximate model of the power system. Linear sensitivities using the DC power flow solution in contingency analysis is computationally much faster, however, this approach will not catch all of the contingency violations due to the underlying assumptions [62]. The $1P - 1Q$ (one $P - \theta$ calculation and one $Q - V$ calculation) method for contingency selection using fast decoupled load flow has been presented in [63] where solution is interrupted after one iteration. The application of Genetic Algorithm for contingency ranking has been studied in [64] where the ranking problem is formulated as an optimization problem with an objective of finding the critical cases. Ranking of contingency based on risk index along with the likelihood of each contingency and severity is computationally efficient and could be used for power system security assessment [65]. Inspired from [66], multi-element contingency screening algorithms were able to detect nearly all contingencies resulting violations by solving small number fraction of possible contingencies [10]. The screening algorithms considered linear sensitivities like line outage distribution factor (LODF) as well as flow and line limit information. Non-heuristic selection procedure for $N - 2$ contingency problem based on the idea of iterative pruning of the possible candidate sets was proposed in [67]. Eppstein et al. [68] developed a randomized algorithm based approach for identifying a minimum $N - k$ contingency that initiates a failure in the system as a sequence of cascading outages. A cyber-physical environment for security assessment was introduced in [69] by including
both accidental contingencies and malicious compromise of cyber assets with the aim of identifying the possible contingencies through cyber attacks. A novel heuristic algorithm for identifying critical $N - 2$ contingencies resulting overload during post-contingency condition was proposed in [70] where the authors claimed its zero missing rates in DC approximation model.

Despite the fact that plenty of algorithms for $N - k$ contingency screening were proposed in last 3 decades, most of the aforementioned literature still require running of power flow for various combination of contingencies to identify the vulnerable branches for security assessment. With this consideration, an electrical distance approach is proposed for searching of vulnerable branches efficiently during a single and double outage in the system. A new connection of power system is defined based on the electrical distance between nodes or buses using the inverse of an admittance matrix (i.e., an impedance matrix). The entries of non-sparse impedance matrix are pruned in order to obtain a comparable structure in consistent with original topological structure. Both $N - 1$ and $N - 2$ contingency analysis are performed to analyze the post-contingency results in vulnerable branches. This method is more efficient to search vulnerable branches and reduce computational complexity during contingency screening. Also, it is not required to conduct power flow experiment for each possible events. In addition, the effect of loss of vulnerable branches on bus voltage profile is studied to validate the proposed approach.

3.2 Mathematical Formulation of Contingency Analysis

A well-known Full-Newton method is used for analysing the system behavior during pre and post contingencies. This method will determine the overloads and voltage
limit violations more accurately than approximate DC power flow. The relationship between node current $I$ and node voltage $V$ for a particular node $i$ in a network of $n$ nodes is given by the linear equation,

$$I_i = \sum_{k=1}^{n} Y_{ik} V_k$$  \hspace{1cm} (3.1)

where $Y_{ik}$ is an element of the admittance matrix connecting nodes $i$ and $k$. Following equation (4.6), complex power at node $i$ is given by,

$$S_i = V_i I_i^*$$  \hspace{1cm} (3.2)

$$P_i + jQ_i = V_i \sum_{k=1}^{n} Y_{ik}^* V_k^*$$  \hspace{1cm} (3.3)

Equation (3.3) represents real and reactive power flowing in $i_{th}$ branch. Tripping of some set of transmission lines changes the topology of grid which is reflected in admittance matrix, $Y_{bus}$ matrix. The updated $Y_{bus}$ matrix is then used to redistribute the branch flows after any outage of transmission line/s. The algorithm for AC power flow security analysis with contingency case selection is shown in Algorithm I. Each of the possible event is simulated by removing the elements defined in contingency list by updating the model. The post contingency flow in branches are tested for overloads and all the limit violations are reported. The impact of line outages is illustrated in Fig. 3.1.

![Figure 3.1](image-url)  

Figure 3.1. Contingency result for outage in power system. The lines identified in circle are overloaded due to outage of line/s identified in rectangle.
This figure illustrates the post contingency flow due to outage of particular line/s which are identified in rectangle. Only those lines which are carrying power beyond their limit are presented in the list (shown in circle) and their percentage loading due to outage is represented as \( f_{a,b} \), where \( f \) is the percentage loading of line \( a \) due to outage of line \( b \).

For example, during \( N-1 \) contingency shown in first row, \( f_{3,1} \) represents the percentage loading of line 3 due to outage of line 1. Similarly, during \( N-2 \) contingency shown in second row, \( f_{3,1&2} \) represents the percentage loading of line 3 due to combined outage of lines 1 and 2.

**Algorithm 1: Traditional contingency analysis**

<table>
<thead>
<tr>
<th>Input:</th>
<th>List of possible outages, Thermal limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Result:</td>
<td>Alarm list containing overload branches</td>
</tr>
<tr>
<td>for all list of possible outages do</td>
<td></td>
</tr>
<tr>
<td>1. pick outage ( i ) from the list and remove that component from the model</td>
<td></td>
</tr>
<tr>
<td>2. run AC power flow in the updated model</td>
<td></td>
</tr>
<tr>
<td>if post contingent flow &gt; thermal limit, then</td>
<td></td>
</tr>
<tr>
<td>create list for vulnerable branch;</td>
<td></td>
</tr>
<tr>
<td>else</td>
<td></td>
</tr>
<tr>
<td>identify system as secure for that outage;</td>
<td></td>
</tr>
<tr>
<td>end</td>
<td></td>
</tr>
<tr>
<td>end</td>
<td></td>
</tr>
</tbody>
</table>

Although NERC requires maintaining power grid security against \( N-1 \) contingency, they are still vulnerable to events which involve multiple component failures, i.e., \( N-k \) contingency. So, as multiple outages are taken into considerations the number of events to be simulated grows rapidly and screening is intractable when \( k > 1 \). The number of total contingencies to handle for \( k \) outages with \( N \) number of branches in a system is given by,

\[
Total = \binom{N}{k} = \frac{N!}{k!(N-k)!} \tag{3.4}
\]
For $k = 1$ the total number is simply $N$ which corresponds to $N - 1$ contingency and for $k = 2$, the total number of combination is given as,

$$\binom{N}{2} = \frac{N!}{2!(N-2)!} = \frac{N(N-1)}{2} = \frac{(N^2 - N)}{2}$$  \hspace{1cm} (3.5)

Equation (3.5) suggests that for maintaining system security against $N - 2$ contingency requires analyzing events in the order of $N^2$. In general, for simulating a $k$ number of outages, $O(N^k)$ power flow solutions are required to process. The number of transmission lines in the power system can be linearized with a number of buses ($n$) in the system, i.e., $N \approx 1.5n$ [10]. Based on this the computational complexity can be expressed as function of $n$ as,

$$O(N^k) = O((1.5n)^k) = O((1.5)^k(n)^k) = O(n^k)$$  \hspace{1cm} (3.6)

The use of Newton’s method for power flow adds some computational effort in solving multiple outage contingency analysis. The elements of Jacobian matrix are $2 \times 2$ blocks of real numbers and computational effort for required factorization for $n$ number of buses is $O(n^{1.4})$ [71]. In this thesis, third order outage is considered as a highest order, and following equation (3.6), the computational complexity for $k = 2$ is given as

$$CE = O(n^{1.4}).O(n^k) = O(n^{k+1.4}) = O(n^{3.4}).$$  \hspace{1cm} (3.7)

Equation (3.7) suggests that finding the critical combinations of contingency is challenging even for modest values for $N$ and $k$. So, to determine the vulnerable branches
in the system for unseen events without any qualitative assumptions can be the remedy of this complexity. The proposed approach is discussed in the following section.

3.3 Proposed Electrical Distance Method

3.3.1 Different structure of power grid

This section discusses the proposed method for finding vulnerable branches for any probable outages in the system. A new network structure is proposed based on electrical distance which is referred as electrical structure of the network. It will have a same number of the node to node connections as that of topological structure. The topological structure is an electrical power network formed from the admittance matrix \( Y_{bus} \). This new electrical structure is then compared with an original topological structure in order to explore the vulnerable branches. Here, “compare” means to find the branches that are common in both topological and electrical structure. In order to study the structure of power grids from a complex networks perspective, the electrical structure, as well as its topology, needs to be studied. Since flow in electrical networks is governed by Kirchoff’s law, this results in unique patterns of interaction between nodes in a network.

The bus admittance matrix defined by equation (3.8) captures the topological structure of power system network.

\[
Y_{bus}^{ls} = \begin{cases} 
G_{ls} + jB_{ls}, & \text{if } l \neq s \\
- \sum (G_{ls} + jB_{ls}) & \text{if } l = s \\
0, & \text{if no connection exists}
\end{cases}
\]  

(3.8)

Per definition, the \( Y_{bus} \) matrix tends to be sparse as the value of some entries is 0 if nodes \( l \)
and s do not have a direct physical connection. The admittance matrix, $Y_{bus}$ for W&W 6 bus network [72] in Fig. 3.2 is shown in Table 3.1. This matrix shows all 11 branches of the system which can be represented by non-zero entries in the upper or lower triangular matrix.

### Table 3.1. $Y_{bus}$ for W&W 6 bus system.

<table>
<thead>
<tr>
<th>$Y_{bus}$</th>
<th>bus 1</th>
<th>bus 2</th>
<th>bus 3</th>
<th>bus 4</th>
<th>bus 5</th>
<th>bus 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>bus 1</td>
<td>4.0 - j11.8</td>
<td>-2.0 + j4.0</td>
<td>0</td>
<td>-1.2 + j4.7</td>
<td>-0.8 + j3.1</td>
<td>0</td>
</tr>
<tr>
<td>bus 2</td>
<td>-2.0 + j4.0</td>
<td>9.3 - j23.2</td>
<td>-0.7 + j3.8</td>
<td>-4.0 + j8.0</td>
<td>-1.0 + j3.0</td>
<td>-1.5 + j4.5</td>
</tr>
<tr>
<td>bus 3</td>
<td>0</td>
<td>-0.7 + j3.8</td>
<td>4.2 - j16.6</td>
<td>0</td>
<td>-1.5 + j3.2</td>
<td>-1.9 + j9.6</td>
</tr>
<tr>
<td>bus 4</td>
<td>-1.2 + j4.7</td>
<td>-4.0 + j8.0</td>
<td>0</td>
<td>6.2 - j14.7</td>
<td>-1.0 + j2.0</td>
<td>0</td>
</tr>
<tr>
<td>bus 5</td>
<td>-0.8 + j3.1</td>
<td>-1.0 + j3.0</td>
<td>-1.5 + j3.2</td>
<td>-1.0 + j2.0</td>
<td>5.3 - j14.2</td>
<td>-1.0 + j3.0</td>
</tr>
<tr>
<td>bus 6</td>
<td>0</td>
<td>-1.6 + j4.5</td>
<td>-1.9 + j9.6</td>
<td>0</td>
<td>-1.0 + j3.0</td>
<td>4.5 - j17.0</td>
</tr>
</tbody>
</table>

#### 3.3.2 Measure of electrical distance

There are variant measures of electrical distance for a power network ([73], [74]), but one of the efficient ways is the absolute value of the inverse of the system admittance matrix. To define the electrical distance in our work, the absolute value of the inverse of the $Y_{bus}$ matrix is used which is a non-sparse (dense) matrix, i.e., $Z_{bus} = |Y_{bus}^{-1}|$. The power system network is coupled by the following equation,

![Figure 3.2. Single line diagram of W&W 6 bus system [72].](image-url)
\[ YV = I \]  

where \( V \) and \( I \) represent the bus voltage and injected current vectors respectively; and \( Y \) is the admittance matrix. Suppose a network has \( n \) nodes and \( m \) branches or links, each link has an impedance \( z = r + jx \) where \( r \) is resistance and \( x \) is reactance. The line admittance can be written as,

\[ y = g + jb = 1/z \]

where \( g \) is conductance and \( b \) is susceptance of any branch. Assume that a unit current flows along the link from node \( l \) to \( s \) which causes the voltage difference between ends of link equal to \( \delta v = V(l) - V(s) = Z_{ls} \). Therefore \( Z_{ls} \) can be interpreted as the electrical distance between two nodes. It is also important to note that, electrical distance (\( Z_{bus} = |Y_{bus}^{-1}| \)) does not perfectly represent all of the ways in which components in a grid connect, it is a useful starting point for structural analysis [43]. More mathematical justification of the proposed approach can be found in [75] and [76]. The distance matrix, \( Z_{bus} \) is a full matrix where each element \( Z_{ls} \) reflect the propagation of the voltage variation following a current injection in a given node pair throughout the system [77]. Since Kirchhoff’s and Ohm’s laws provide connectivity among all nodes pair in the system, the graph defined from electrical distance matrix \( Z_{bus} \) is fully connected. For the system shown in Fig. 3.2, there will be 15 links in the system as defined by,

\[ p = \frac{n^2 - n}{2} \]
where \( n \) is number of nodes in the network and \( p \) is the total number of connecting links between each node pair. These connections are represented by entries of electrical distance matrix, \( Z_{bus} \) given in Table 3.2. Each term of the matrix (\( Z_{ls} \) or \( e_{ls} \)) describes the amount of connectivity between node pairs in the system.

Table 3.2. Equivalent \( Z_{bus} \) for W&W 6 bus system. Each entry represents the amount of coupling between pair of nodes.

<table>
<thead>
<tr>
<th>( Z_{bus} )</th>
<th>bus 1</th>
<th>bus 2</th>
<th>bus 3</th>
<th>bus 4</th>
<th>bus 5</th>
<th>bus 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>bus 1</td>
<td>3.6404</td>
<td>3.7041</td>
<td>3.7251</td>
<td>3.6963</td>
<td>3.7149</td>
<td>3.7236</td>
</tr>
<tr>
<td>bus 2</td>
<td>3.7041</td>
<td>3.6744</td>
<td>3.7089</td>
<td>3.6970</td>
<td>3.7146</td>
<td>3.7066</td>
</tr>
<tr>
<td>bus 3</td>
<td>3.7251</td>
<td>3.7089</td>
<td>3.6446</td>
<td>3.7221</td>
<td>3.7092</td>
<td>3.6800</td>
</tr>
<tr>
<td>bus 4</td>
<td>3.6963</td>
<td>3.6970</td>
<td>3.7221</td>
<td>3.6515</td>
<td>3.7166</td>
<td>3.7203</td>
</tr>
<tr>
<td>bus 5</td>
<td>3.7149</td>
<td>3.7146</td>
<td>3.7029</td>
<td>3.7166</td>
<td>3.6678</td>
<td>3.7091</td>
</tr>
</tbody>
</table>

The equivalent electrical distance between nodes \( s \) and \( l \) is thus given by the magnitude of the relevant entry of the \( Z_{bus} \). A small value of \( e_{ls} \) corresponds to a shorter electrical distance but a stronger coupling between these nodes (\( s \) and \( l \)). This reflects a larger propensity for power to flow between these nodes.

3.3.3 Formation of new electrical structure

A graph representation of the electrical structure of the system from the electrical distance matrix, \( Z_{bus} \) can be generated. The algorithm for representation of electrical structure graphically is given in Algorithm 2. Since \( Z_{bus} \) is a non sparse matrix, it will reflect all possible links between the nodes present in network. In the electrical structure, the number of nodes \( n \) will be same as a topological network. But the existing \( m \) links in the topological network will be replaced by \( m \) smallest entries from the upper or lower triangle of \( Z_{bus} \) matrix (since \( Z_{bus} \) is symmetric). This means the new electrical structure will have a same number of the node to node connections as that of topological structure.
Algorithm 2: Graphical representation of $Z_{bus}$ matrix

**Input:** $Y_{bus}$, Threshold ($t$)

**Result:** Adjacency matrix with elements $a_{ls}$

Graphical representation of electrical network

$Z_{bus} = |Y_{bus}|$;

for every element of ($Z_{bus}$) do

    if $e_{ls} < t$, then
        $a_{ls} = 1$;
        draw connection between node $l$ and $s$;
    else
        $a_{ls} = 0$;
    end

    if $a_{ls} \in Y_{bus}$, then
        branch $a_{ls}$ exist;
    else
        branch $a_{ls}$ does not exist;
    end

end

Hence, the total $p$ connection links of $Z_{bus}$ matrix will be reduced to $m$ by calculating the proper threshold value. With these newly selected connection links, there will be a different topology than the original network which is electrical structure of the system.

The elements of the adjacency matrix, $A$ of this new network will be defined as follows,

$$a_{ls} = \begin{cases} 
1, & \text{if } e_{ls} < t \\
0, & \text{if } e_{ls} \geq t
\end{cases}$$

(3.12)

where $e_{ls}$ is any term of $Z_{bus}$ matrix and $t$ is the calculated electrical distance threshold in order to capture the same number of links as the previous network. Next, the branches in an electrical network are compared with topological connections and it is observed that some branches are common in both structures. So, using the adjacency matrix $A$ and $Y_{bus}$ matrix, the common connections under the set $n(T \cap E)$ are observed, where $T$ represents...
the set of topological connections and \( E \) represents the set of electrical connections obtained from pruned \( Z_{bus} \) matrix. The remaining connections are those which do not have a physical connection between the nodes. The steps to represent the electrical structure graphically and to compare it with topological structure is shown in Algorithm 2. After finding the most important electrical branches from an electrical distance, contingency analysis is performed in power system test cases. With the aim of finding the most critical or vulnerable branches, attention is paid to those branches that are important according to our proposed model and also have physical connections between the nodes.

3.3.4 Discussion on the proposed method

The computational complexity for \( N - 2 \) contingency as given by equation (3.7) is \( O(n^{3.4}) \) and this complexity increases to \( O(n^{4.4}) \) if \( N - 3 \) contingency \((k = 3)\) is considered. In our proposed work, the only computational cost is the calculation of inverse of \( n \times n \) matrix and this cost is independent of multiple outages taken into account. According to [78] and [79], the computational cost for obtaining inverse of matrix is \( O(n^{2.373}) \) which is less than traditional contingency analysis.

Our proposed approach is discussed in finding the critical branches based on impedance matrix. In our approach, the formulation is same as that in references [43] and [80] such that smaller value of electrical distance corresponds to the larger propensity for power flow between the nodes. It is also true that loading will impact the location of critical branches. However, during the early stage of power system planning and operation, the transmission line parameters are adjusted based on the loading. This means, a line which has to carry higher MW has different line parameter (impedance) than the
Figure 3.3. IEEE-24 bus one-line diagram and node branch representation of the system exactly following the one-line diagram. The number in node are in accordance with the bus number.

line which needs to carry less MW. So, the line criticality is closely related to its parameters which are reflected in terms of electrical distance. Hence, loading will impact the location of critical branches and this will be somehow reflected in terms of electric connectedness between the nodes.
3.4 Simulation Results

In this section, the IEEE-24 bus system is used as a test case to illustrate the electrical distance approach for selecting vulnerable branches. This system is obtained from Illinois Center for Smarter Electric Grid (ICSEG) [81]. It has 38 transmission lines with 6 transformers, which are given unique numbers as shown in Fig. 3.3. The transmission lines are numbered so that it helps identify the failed and overloaded branches during contingency analysis. The node positions and system admittance matrix, $Y_{bus}$ is obtained from PowerWorld. The dense impedance matrix ($Z_{bus}$) is then calculated by inverting $Y_{bus}$ matrix. Following equation (3.11), the total number of distinct node to node connections for 24 bus system is calculated as $[(24^2 - 24)/2] = 276$.

3.4.1 Case I: Contingency analysis based on electrical distance

In addition to the one-line diagram, Fig. 3.3 also shows the topological structure (node-branch representation) of the IEEE-24 bus system using the bus (node) position and connections between branches according to admittance matrix. Algorithm 2 presented in Section 3.3 is applied to IEEE 24 bus test case system to generate the equivalent electrical structure of the network. Initially, the $Z_{bus}$ matrix is made triangular since the node pair $i \rightarrow j$ and $j \rightarrow i$ represents the same connection between the nodes. Then each entry is selected and its impedance value is compared with remaining entries. The entries with higher impedance are made zero and the ones with lower impedance are kept unchanged in $Z_{bus}$ matrix. The total nonzero elements of the matrix are counted to find the total number of connections whose impedance are less than that of the selected entry. The particular element or entry which gives a number of connections equal to that of the
topological structure is tagged as a threshold value. However, it is not always necessary that a particular entry will itself be a threshold value and give exact required number of non-zero elements in $Z_{bus}$. In such case, it is necessary to adjust the nearest impedance value by trial and error method to find exact threshold value. For IEEE-24 bus test case, among 276 entries, 228th entry is obtained as the threshold value of 0.4090 which gave a number of connections equal to 38.

![Figure 3.4](image.png)

**Figure 3.4.** Node-branch representation of different structure of power grid network; (a) Dense electrical structure showing connection among all node pairs; (b) Important electrical connections with number of connections equal to number of branches in topological structure (Fig. 3.3); (c) Electrical connections that are physically present in benchmark system.

With the calculated threshold value, out of 276 total electrical connections (Fig. 3.4a), only 38 strongest electrical connections are presented. The links shown in Fig. 3.4b represents the 38 different node-node connections and hence it is size-compatible with the topological structure. The two representation of the 24 bus network suggests different structure. From an electrical perspective (which captures the behaviour of the network, not simply the physical structure), 24 bus network seems to have a distinct group of nodes that are electrical hubs. That is, those buses have a high electrical connectivity to the rest of the network. Power flowing through the network as governed by Kirchhoff’s law is more likely to pass through these nodes than the remaining other nodes. In language of
social networks, these nodes are often referred to as nodes with “high betweenness” or “information centrality” [80]. It is not necessary that all the electrical connections defined by adjacency matrix $A$ will be physically present in the original system. So, particular connections are searched, which are electrically important according to our approach and exist in the topological structure as well. Fig. 3.4c shows particular branches in the system which have the shortest electrical distance (electrically important connections) and are common between topological structure and electrical structure. The set of branches belonging to a different structure is also represented in Venn-diagram in Fig. 3.5 where the color of marks represent branches from different network structure. The branches belonging to $n(T \cap E)$ are physically connected branches in the benchmark, and they are not selected based on their impedance value. So, it is not necessary that they have higher/lower impedance value of branches belonging to $n(E)$.

After finding the electrical structure of 24 bus system, it is simulated for $N - 1$, $N - 2$, and $N - 3$ contingency analysis to observe different possible scenarios in the post-contingency analysis. With the loss of single ($N - 1$) or a combination of transmission lines ($N - 2$ and $N - 3$), there occurs power flow redistribution in the
network. This forces some of the remaining transmission lines to carry power beyond their line limit. During contingency analysis, only real power flow is considered for checking the line limits, and voltage limit violations are not accounted in this research.

For example, loss of line 3 causes lines 1, 2, 4, 5, 8, and 10 to carry power beyond their capacity and forces them to be overloaded. An illustration of post contingency analysis for single outage is shown in Fig. 3.6. The circles contain a line identifier or name and the post contingent flow (in terms of percentage of their limits) for an outage of line identified in the rectangle.

Table 3.3. Contingency analysis for IEEE-24 bus system showing line index and number of times it is violated for \(N-1\), \(N-2\), and \(N-3\) contingencies

<table>
<thead>
<tr>
<th>S.No</th>
<th>Line Index</th>
<th>Impedance (Z_{ls})</th>
<th>Number of violations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>(N-1)</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0.3570</td>
<td>31</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0.4085</td>
<td>28</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>0.3802</td>
<td>30</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>0.3862</td>
<td>29</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>0.3979</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>0.3997</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>0.3962</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>0.3906</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>6*</td>
<td>0.4111</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>134</strong></td>
</tr>
</tbody>
</table>

* \(\rightarrow\) line not captured by electrical distance approach

Figure 3.6. \(N-1\) contingency analysis results for loss of line 3 and line 6. The line identified in circles are overloaded lines due to outage of line identified in rectangle. The number in circle represents the percentage loading of lines.
Figure 3.7. $N-2$ contingency analysis results for lines 3&6 and lines 6&8. The line identified in circles are overloaded lines due to outage of line identified in rectangle. The number in circle represents the percentage loading of lines.

Similarly, Fig. 3.7 shows the post contingent flow in branches for an outage of a combination of lines, i.e., $N-2$ contingency. One key thing to observe from this analysis is that the line which is overloaded during an outage of a particular line ($N-1$ outage) might not be overloaded for a combination of outages involving the previous line. For example, line 4 is overloaded due to loss of line 6 whereas it is not during the loss of line 6 and line 8 together. But it is obvious that $N-2$ is more severe than $N-1$ when observing their overall impact. Based on these post contingent flow for all single and higher order contingency list, the total number of violations and violated lines are identified. In IEEE-24 bus system, there are 134 line limit violations for $N-1$, 2063 line limit violations for $N-2$ and 19783 line limit violations for $N-3$ contingency respectively.

These violations are contributed by various branches in the system. Table 3.3 shows the line identifier, their impedance and the number of times it is violated for defined contingency. There are 8 branches which are loaded beyond their capacity by different line outages in the system for $N-1$ and $N-2$ contingency. There is one extra line whose line limit is violated for $N-3$ contingency and the line index suggests that except line 6, all of these lines are connections in the electrical network which are also present in the topological structure, i.e., physically present in Fig. 3.3. This means all the *sensitive* or
Table 3.4. Descriptive statistics for power system test cases

<table>
<thead>
<tr>
<th>Test cases</th>
<th>No. of Buses</th>
<th>No. of Tr. Lines</th>
<th>No. of Generators</th>
<th>Gen. MW</th>
<th>Load MW</th>
<th>Voltage level (p.u.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>W&amp;W 6</td>
<td>6</td>
<td>11</td>
<td>3</td>
<td>219.52</td>
<td>210</td>
<td>0.92-1.00</td>
</tr>
<tr>
<td>IEEE-14</td>
<td>14</td>
<td>20</td>
<td>5</td>
<td>272.39</td>
<td>259</td>
<td>1.01-1.09</td>
</tr>
<tr>
<td>IEEE-24</td>
<td>24</td>
<td>38</td>
<td>11</td>
<td>1701.58</td>
<td>1619</td>
<td>0.92-1.00</td>
</tr>
<tr>
<td>IEEE-30</td>
<td>30</td>
<td>41</td>
<td>6</td>
<td>300.95</td>
<td>283.4</td>
<td>0.99-1.08</td>
</tr>
<tr>
<td>G&amp;S-37</td>
<td>37</td>
<td>57</td>
<td>10</td>
<td>819.17</td>
<td>808.72</td>
<td>0.99-1.03</td>
</tr>
<tr>
<td>IEEE-39</td>
<td>39</td>
<td>46</td>
<td>10</td>
<td>6191.28</td>
<td>6149.5</td>
<td>0.98-1.07</td>
</tr>
</tbody>
</table>

vulnerable branches for contingencies are being captured using electrical distance approach for \( N - 1 \) and \( N - 2 \). For \( N - 3 \) contingency, 1 out of 19783 line limit violations is missing from being captured, however, the total percentage of violations captured under electrical distance is still 99.99%.

Algorithm 2 is applied and outage for different contingency combinations are simulated for several power system benchmarks. A 6-bus system is built in PowerWorld using all branch and loading information from [72]. A design project case with 37 bus and 57 transmission lines is taken from [82]. Similarly, remaining IEEE test cases are directly accessed from Illinois Center for Smarter Electric Grid (ICSEG) [81]. The descriptive statistics for the power system test cases used in this research are provided in table 3.4.

Table 3.5 summarizes the contingency analysis results. The total number of violations and line identifier of the violated lines that contributed for those recorded violations are observed. The total number of violations due to \( N - 1, N - 2, \) and \( N - 3 \) contingencies captured by important electrical branches are expressed in percentages and a total number of sensitive lines for each contingency are reported. A total number of sensitive lines means the total number of branches which are overloaded during to \( N - 1, N - 2, \) and \( N - 3 \) contingency experiments. It is observed that “most of the violations” are associated
Table 3.5. Contingency analysis in different benchmark system

<table>
<thead>
<tr>
<th>System</th>
<th>Number of Tr. Lines</th>
<th>Number of line limit violations</th>
<th>Violations captured under electrical distance</th>
<th>Number of sensitive lines</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>N-1</td>
<td>N-2</td>
<td>N-3</td>
</tr>
<tr>
<td>W&amp;W-6</td>
<td>11</td>
<td>3</td>
<td>105</td>
<td>211</td>
</tr>
<tr>
<td>IEEE-14</td>
<td>20</td>
<td>1</td>
<td>23</td>
<td>200</td>
</tr>
<tr>
<td>IEEE-24</td>
<td>38</td>
<td>134</td>
<td>2063</td>
<td>19783</td>
</tr>
<tr>
<td>IEEE-30</td>
<td>41</td>
<td>49</td>
<td>1005</td>
<td>12131</td>
</tr>
<tr>
<td>G&amp;S-37</td>
<td>57</td>
<td>3</td>
<td>100</td>
<td>2352</td>
</tr>
<tr>
<td>IEEE-39</td>
<td>46</td>
<td>20</td>
<td>557</td>
<td>9321</td>
</tr>
</tbody>
</table>
with those branches which are obtained from electrical distance approach and have a physical existence in the power system network as well. Here, the term “most of the violations” is used for those benchmarks where all the violations are not captured by electrical distance approach. As shown in table 3.5, in some test cases, the violations captured under electrical distance is not 100%. For example, in IEEE-39 bus system during $N-1$ contingency, the total number of violations captured by electrical distance approach is 85.0% which means out of total 20 violations, 17 violations are from the lines which have strongest electrical connections. Similarly, during $N-2$ contingency analysis, 86.5% of total 557 violations are captured by electrical distance approach. Remaining 75 line limit violations are from those transmission lines which are not identified as vulnerable based on our approach.

The results for contingency analysis in different power system test cases suggest that electrical distance approach for searching the vulnerable branches is more efficient and does not require power flow experiment for each probable event in the power system. With this proposed method, the computational effort in solving thousands of possible outages during higher order contingency to predict the effect of outages can be reduced.

3.4.2 Case II: Loss of sensitive branches and voltage profile

All the simulations set up for this case study are the same as that for Case I. Here, the effect in voltage profile of system is studied with loss of transmission line. In any power system, transmission lines have a limited capability for power transfer, as well known from circuit theory [83]. This limit marks the onset of voltage instability where the system is not able to maintain the steady voltage at all buses in the system after being
subjected to a disturbance from a given initial condition.

Based on our proposed approach of electrical distance, there are some branches in a system that are identified as critical during contingency analysis and those branches are among having small electrical distance. The different set of branches belonging to the electrical and topological structure are shown in Fig. 3.5. To verify the proposed approach for selection of critical branches, the bus voltage profile is studied for loss of a particular branch in the system. Two branches are selected from a set of critical lines which are under the set $n(T \cap E)$. And, two other branches are selected which are not captured in electrical connections and belong to set $n(T)-n(T \cap E)$. These branches are opened one by one at $t = 1\, sec$ during normal operation condition. Transient stability analysis is done for these disturbances to see the change in voltage level of all 24 buses.

During normal operating condition, the voltage profile of the system is between 0.92 p.u. and 1 p.u. as shown in Fig. 3.8. First, line 1 is removed and the change in voltage level of all the buses is observed. This caused the voltage of several buses to decrease around 0.8 p.u. and increase above 1.1 p.u. which are beyond the normal operating range as shown in Fig. 3.8a. Again, line 4 is opened which is also identified as critical line according to our approach. Fig. 3.8b shows the degradation in voltage profile because of loss of another critical line. This decline in voltage level is not desired in power system and can damage the load or eventually lead to voltage collapse.

Similarly, two different branches from set $n(T)-n(T \cap E)$ are selected. Lines 19 and 22 are opened at $t = 1\, sec$ and voltage profile is observed for both the cases. Fig. 3.9 suggests that these disturbances have a minimal effect in the system voltage profile and only few bus voltages are found to be declined below normal operating point of 0.92 p.u.
In addition, the voltage distribution of all 24 buses for loss of branches in the system are studied. Fig. 3.10 shows the upper and lower voltage level of buses for an outage of each branch set. The blue mark represents the upper voltage level and black mark represents the lower voltage level for a particular loss. The red marks show the voltage level of particular buses which are beyond the range specified by the green line, and the x-axis shows the corresponding index of the branch. The line index suggests that the line captured in $n(T \cap E)$ caused the voltage to distribute beyond $\pm 10\%$ limit (i.e., below 0.9
Figure 3.10. Voltage distribution for loss of branches. Red mark represents the bus voltage that is out of normal operating range and number refers to the buses whose voltage is mostly affected by loss of particular branch.

p.u. or above 1.1 p.u.). For example, when line-1 is removed from the system, the bus voltage of all 24 buses ranges from 1.1308 to 0.8094. Our result shows that the deviation of voltage level from the defined boundary is because of loss of branch which belongs to \( n(T \cap E) \) set. This suggests that branches with strongest electrical connections are more critical and hence their outage has a severe impact in the voltage.

3.5 Summary

A method of searching vulnerable or critical branches is studied to avoid the computational expense of processing higher order outage event. In our proposed approach, the vulnerable branches are screened through the electrical distance between node pairs, so that the candidate search in contingency analysis for all possible events can be avoided. This significantly reduces the computational cost while having the ability to identify
almost all vulnerable branches in the system. Our experimental results show that branches with small electrical distance are identified as vulnerable for \( N - 1 \), \( N - 2 \) and \( N - 3 \) contingency. Study of the voltage profile of system for transmission line loss validated our proposed approach for searching the vulnerable branches. Based on vulnerable branches initial flow, loading limit, and the way power system is being operated, the identification alone can help in decision making during power system outage. Correspondingly, power system security can be enhanced by making those branches more rigid or re-scheduling of generator output in order to decrease the burden on these particular branches.

Contingency analysis is a critical activity in the context of the power infrastructure because it provides a guide for resiliency and enables the grid to continue operating even in the case of failure. A critical issue with the current evolution of the power grid into a so-called smart grid is the introduction of cyber-security threats due to the pervasive deployment of communication networks and information technologies. So, in addition, to having plans for the accidental contingencies, it is also required to study the power system behaviour for malicious compromises. This requires a model of a cyber-physical system including interactions among cyber and physical components. In the following chapter, a cyber-physical power system testbed is provided which can be used to assess potential impacts of both cyber and physical contingencies as well as to investigate the mitigation efforts.
4.1 Introduction

Smart grid is a modernized electrical grid and is generally referred as the next generation power system. For the purpose of sensing, monitoring, protection and control, information and communication technology system are being deployed in the modern power system. With this integration, the smart grid is expected to greatly enhance efficiency, reliability, and economy of power production and consumption along with the integration of renewable energy resources, as well as demand response and distributed intelligence [84]. Although the current smart grid initiatives are expanding the use of information technologies to modernize the existing grid, their adoptions in cyber-physical systems (CPS) have introduced power system security issues. Attacks on either cyber or physical parts of the smart grid will possibly impact the stability of the entire system.

4.1.1 Vulnerabilities in Smart Grid

As there is growing dependency of nations infrastructure on the cyber domains, these systems become an attractive target for well-trained attackers. The combination of an increasingly interconnected power grid and cyber domains presents concerns for the grid’s current security posture. As shown in Fig. 4.1 threats could target the generation, transmission, distribution, and market domains. Unfortunately, as smart grid security concerns have grown, researchers have begun to identify the vulnerabilities within both cyber and physical domains. Analysis from Department of Homeland Security and Idaho National Laboratory have revealed that most of the serious vulnerabilities are prevalent
through ICS software platforms and network configurations [85], [86]. North American Electric Reliability Corporation (NERC) from the private sector has also acknowledged the threat to the electric power grid. NERC created a cyber attack task force (CATF) in order to explore the vulnerabilities in cyber domain and identify proper detection capabilities. The “High-Impact, Low-Frequency Event Risk to the North American Bulk Power System” report released by NERC in 2009 highlighted major threats to the grid, specifically the coordinated attacks blending both cyber and physical methods [87].

4.1.2 Research needs

Recent research in a cyber attack against smart grid has shown that these intentional attacks can have an impact on power system operation in terms of stability and economy. For example, authors in [11] commented that cyber attack in measurements of static var compensator (SVC) or static synchronous compensator (STATCOM) can degrade the system’s stability margin. Cyber attacks including false data injection attacks can mislead
the state estimating process [89] or even can impact the economic operation of electric power market operations by manipulating the nodal price [12]. Similarly, denial of service (DoS) attacks in the cyber layer of smart grids can affect the dynamic performance of physical power system [90]. It is also important to verify the device settings, algorithms, and application before they are deployed in the real power system to avoid any unfortunate incident. For example, malfunctioning of relays can lead to false tripping of breakers which can cause cascading failures. In this case, cyber-physical testbeds can serve as a tool for simulating the power system model accurately and also helps to understand the complex relation between cyber and physical domains. Although United States Department of Energy (DoE) is giving considerable attention to the security enhancements of the cyber-physical power system, the research related to cyber attack and impacts are constrained by the availability of realistic cyber-physical system testbed.

4.1.3 Related work

Suitable power system testbeds are needed in order to accurately capture the attack effects, attack impacts in a physical system, and possible mitigation strategies as well as control algorithms to ensure stability and security of power grid. Several cyber defense testbeds have been developed at various entities such as national labs, universities, and research centers for the purpose of studying the consequences associated with these cyber-physical threats and mitigate those consequences. The researchers at national SCADA testbed at Idaho National Laboratory investigated how a cyber attack can cause damage to a physical system through an aurora generator test [13]. In the experiment, the researcher used a computer program to open and close the breaker out of phase from the
grid to maximize the stress. Virtual control system environment (VSCE) was developed at Sandia National Laboratory which uses hybrid modeling and simulation architecture in order to understand the possible impact of particular cyber threats, cyber defense training and exploring power system vulnerabilities [91].

Beside national labs, universities and research institutes are also focusing research in the development of a CPS testbed for cyber security issues. A testbed has been developed at University of Arizona using PowerWorld and MODBUS protocol to detect cyber attacks on SCADA system [92]. In this work, authors presented compromised human machine interface (HMI) and denial of service as different attack scenarios. Authors in [93] suggested various applications of testbed developed at Mississippi State University. The proposed testbed was used for simulation of common power system contingencies (generator loss, transmission loss, and sudden load loss), and event detection using data mining of phasor measurement unit data. Similarly, a power system cyber-physical testbed was developed for intrusion detection and it also provides a platform for hardware in the loop (HIL) simulation, cyber-attack and generated data sets for developing and validating an intrusion detection system for monitoring power system events [94]. The SCADASim testbed has been developed at Royal Melbourne Institute of Technology University for building SCADA simulations which support a combination of network simulation and real device connectivity [95]. It helps to analyze the effects of malicious attacks like denial of service, eavesdropping, man-in-the-middle and spoofing on the devices and simulated network. Emerging smart grid distributed control algorithms were examined in the developed smart-grid cyber-physical system testbed where authors highlighted the impacts of different interactions in the operation of micro-grid [96].
cyber-physical security testbed at University College Dublin [97] provides an accurate tool for analyzing cyber-physical vulnerabilities, allows monitoring of the dynamic behavior of power system as a response to cyber attacks (impact analysis), and mitigation of cyber attacks. For analyzing the physical impact due to compromise in the cyber network, a cyber-physical contingency analysis framework was introduced in [69] for both accidental contingency and malicious compromise. The impact of three different types of real-life cyber attacks namely, communication line outage, denial of service and man-in-the-middle attack in physical power grid was studied through the proposed testbed developed at Washington State University [98]. Other related CPS power testbed research have been discussed in [99], [100], [101].

Research efforts have been made worldwide from industry, academia and national laboratories for the development of real-time cyber-physical testbed. The realistic cyber-physical environment achieved from these testbeds is helpful for investigating power system vulnerabilities, mitigation strategies, and system behavior during different kinds of conditions for our research.

4.2 Proposed Cyber-Physical System Testbed

A. Testbed Application

A CPS testbed needs to have certain capabilities to provide the realistic cyber-physical environment. This section provides the various research applications of a cyber-physical power system testbed. The comprehensive set of testbed application areas is shown in Fig. 4.2 and are elaborated below in detail.

1. Vulnerability analysis: The vulnerability is any weakness that an intruder takes
benefit of in order to compromise the security goals (confidentiality, integrity, availability or authenticity) of smart grid [102]. In the cybersecurity context, it is equally important to incorporate both cyber and physical layers of the power system for vulnerability analysis [97]. The vulnerability might be associated with communication protocols, firewall or VPNs, sensing and monitoring devices in a substation or even in control centers. However, unavailability of these technologies publicly is the main constraint for vulnerability assessment. A real-time cyber-physical testbed provides an environment to analyze the weakness as well as testing of these platforms and architectures.

2. Disturbance scenarios: Different disturbances can be simulated in a real-time CPS testbed. They might be an attack or any dynamic events within the power system. For example, an attacker can trigger a circuit breaker to isolate a transmission line during normal condition or there might be a sudden increase in demand at a particular time. Such events can be simulated in real time and the consequences can be analyzed to propose the necessary mitigation and control strategies.

3. Impact analysis: This is one of the important applications of a testbed which involves quantifying the effects of given event on the physical electrical grid and provide a degree of damage based on power delivery disruptions [103]. The physical impact from possible cyber attack, power system faults, and power system contingencies can be evaluated from the CPS testbed. The interdependency between cyber and physical part of power system makes the analysis part more complicated. In addition, a testbed helps to analyze the impact of potential cyber attacks and physical hazards in terms of power system’s stability, reliability, and economic operation so that operating personnel can have a clear idea about the potential damage that can happen because of a particular event.
4. **Stability and control:** It is always desired to have power system ability to move from one steady-state following a disturbance to another stable operating point without loss of synchronism and without having unacceptable frequency deviations [82]. CPS testbed facilitates the simulation of major disturbances like loss of generation, faults and sudden changes in load. The control strategies like adaptive dynamic programming based supplementary control [104] can be implemented within the testbed in real-time for efficient damping of the oscillations and help the multi-machine system return to the synchronous frequency with new steady state power angles.

5. **Cyber-physical assessment metrics:** Threat and damage are much easier to describe than to measure them. For evaluating the damage in the cyber-physical environment, metrics must combine both cyber and physical properties. Development of metrics helps to improve ability to understand the damage, control it and even defend against it. Physical metrics may include the quantified impact in power flow, change in system’s stability, loss of load, percentage of failure, the risk if failure and even effects in markets [12], [40]. Similarly on the cyber side, metrics can be evaluated based on probabilities of vulnerability, vulnerability criticality and different other methods like artificial intelligence (AI) assessments techniques, models of security measurements,
concrete measurement methods and so on [105].

6. Mitigation module: Depending on the impact caused by an attack or any change in system configuration, a mitigation strategy can be proposed and its effectiveness can be explored in the CPS testbed in real time. This feature helps to bring the system back to normal condition after going through some transients. It will help to protect the system from further damage that might result from a single initiating event and reduce the vulnerabilities of the power system infrastructure.

7. Training and Education: The CPS testbed is a platform for academic purpose and will be useful for gaining knowledge related to cyber security of the smart grid. Users can have the ability to interface with the power system operations and controls during different kind of disturbances and attacks being simulated in the testbed. In addition to research, with the help of data visualization and user-friendly modeling environment, testbed can be used for educational purposes to study different real scenarios in the power system.

B. Testbed Architecture

The architecture of cyber-physical testbed at South Dakota State University (SDSU), highlighting the cyber, physical, control, and communication system is shown in Fig. 4.3.

1. Physical System: Power system simulation in this testbed is performed using two different tools, OPAL-RT, and RT-lab. OPAL-RT is a platform equipped with analog and digital I/Os that provides the capability to perform real-time experiments by interfacing with real hardware devices like protection relays. RT-LAB, fully integrated with MATLAB/Simulink is the open real-time simulation software environment. The test case to be simulated is modeled using the SimPowerSystems (MATLAB/Simulink) Toolbox. The reason behind using SPS is that the models built in SPS are compatible with
Figure 4.3. Schematic architect of the proposed cyber-physical microgrid testbed.

OPAL-RT and can be executed in real time. In order to achieve hardware in loop simulation, some specific blocks from [106] are added to SPS model in order to access the analog and digital ports of the real time simulator. The power system model used in the simulation is based on the Western System Coordination Council (WSCC) 3 machine, 9 bus test case [107]. The system consists of three generating units at buses 1, 2, and 3, and three loads at buses 5, 6, and 8. Two microprocessor based relays or IEDs equipped with current and voltage input modules are used as a protection system for two different transmission lines. The algorithms or settings incorporated in their microprocessor utilize the measured signals from OPAL-RT to detect any fault or anomalies in the power system. These relays generate trip signal once they detect any abnormal condition. The generated trip signal is used to open the circuit breaker, send information to the control center or communicate with other devices and controllers for protecting the system [108]. The three-phase current of the transmission lines is sent to analog outputs of the simulator and these signals are then fed into the CT inputs of the respective relay. The output contact of the relay is connected to the digital input of simulator to open the circuit breaker in the SPS model. The current flowing through the lines is constantly monitored through
human-machine interface tool. The settings of the relay are adjusted/modified using AcSELerator Quickset through PC.

2. Cyber, Control and Communication: In this testbed, SEL-C662 based communication is used between intelligent electronic devices (IEDs) and the user interface in substations where the user interface acquire the monitored data from relays. The distributed network protocol DNP 3.0 is used for control and measurements between the control center and substation which is similar to many real-world SCADA systems. The testbed control center can support general SCADA functions like collecting measurements and field device status from substations, giving commands to an operator working with the field devices and even store the collected data on the server for future reference. In the testbed environment illustrated in Fig. 4.3, intrusions are possible from remote access connections to a substation communication network. Thus, via VPN an attacker can intrude into the substation network and perform malicious attempts like modify data packets, eavesdropping, and compromise user interface.

4.3 Experimental Study on Cyber Security

4.3.1 Parameters and environment setup

Based on the testbed architecture described in section 3.4, a real time hardware-in-loop simulation environment is developed based on OPAL-RT. Fig. 4.4 shows the hardware setup of proposed CPS testbed in Micro-grid lab of South Dakota State University.

Two SEL 351S relays are connected to PC for human machine interface and current flow through the transmission lines being protected (line 7-5 & line 7-8) is constantly
monitored. A simple triggering circuit is built to utilize the relay contact for the opening of the breaker in Simulink. The WSCC 9 bus system is simulated in RT-Lab and a PC is used to launch an attack in one of the protected transmission lines. Here, it is assumed that attacker has penetrated into the substation and has compromised the human machine interface.

A trip signal is sent by an attacker from relay #1 at $t = 5.29s$ during normal operating condition and hence the line 7-5 is opened at $t = 5.29s$. This loss of transmission line forces the line 7-8 to carry all the power generated at bus 2. Hence, the current in line 7-8 increases by a significant amount. The current in line 7-8 is constantly monitored and is being protected from over-current through relay #2. The settings in the relay are adjusted in such a way that whenever the current exceeds the limit for a predefined time, the line is taken off from the system to prevent the damage caused by over-current. The settings are implemented through overcurrent curve where U.S. very inverse curve is used for defining the operating time. Fig. 4.5 shows how the operating time is calculated based
on multiples of pickup. The equation associated with this U.S. curve is given as [109],

\[ t_p = TD(0.0963 + \frac{3.88}{M^2 - 1}) \]  

(4.1)

where \( t_p \) is operating time in seconds, \( TD \) is time-dial settings which is 2 in our case and \( M \) is applied multiples of pickup current (for \( t_p, M > 1 \)).

4.3.2 Experiment results

During this experiment, it is observed that the current flow in line 7-8 is increased from 193.68 A to 333.74 A after the loss of line 7-5. It is assumed that the transmission line 7-8 is designed to carry a maximum current of 250 A. This overcurrent in line 7-8 is detected by the relay and based on provided settings, it sends a trip signal to open the line for preventing further damage due to overload.

![Time-Overcurrent curve](image)

Figure 4.5. U.S. Very Inverse Curve [109]. This curve is drawn based on equation (4.1).

Using equation (4.1), operating time of relay is calculated as,
\[ t_p = TD \times \left( 0.0963 + \frac{3.88}{M^2 - 1} \right) \]
\[ = 2 \times \left( 0.0963 + \frac{3.88}{\left(\frac{33.74}{250}\right)^2 - 1} \right) \]
\[ = 10.114 \text{s} \]  

Figure 4.6. Current in transmission line 7-8. Current increases after 5.29s and due to relay settings, the line 7-8 is opened at 15.30s after which the current is zero.

If the over-current persists till this period, the relay reaches the maximum time for withstanding the thermal overload and it will send the command to the breaker for the opening of the line. Once the line is tripped, the current flow through it reduces to zero as shown in Fig. 4.6. Because of simultaneous tripping of both the lines, the generation station at bus 2 is isolated. This causes the remaining generators at bus 1 and bus 3 to supply the load in the system. The impact of the attack on power system is shown in Figs. 4.7 and 4.8.
Fig. 4.7 shows how the system voltages are impacted by the attack and Fig. 4.8 shows the change in real and reactive power generation as a result of the attack. Each of these figures has two major events which took place as a part of the attack. The first event represents the tripping of line 7-5 by an attack and the second event represents the tripping of line 7-8 by operating personnel to prevent it from overload. Fig. 4.7 shows that the first event did not cause much impact on the system voltage and the voltage at all buses stayed close to 1.0 p.u. However, after the tripping of the second line, the generator at bus 2 is completely isolated from the grid and this impacted the voltage at several buses significantly. From Fig. 4.8, it can be observed that the tripping of line 7-5 changes the generation in all three generators by a small amount, but tripping of line 7-8 completely isolates generator 2 from the system and therefore it results in a huge loss of generation. Such event could cause some frequency stability problems in a real power system.
4.3.3 Mitigation using optimal power flow

In order to protect the system from being unstable, some emergency control action should be taken before the loss of line 7-8. One of the mitigation strategies is to use optimal power flow (OPF). The goal of OPF is to determine the best way to instantaneously operate a power system. Usually, “best” means to minimize the operating cost or minimize the control change by taking some constraints into consideration. In our case, the only constraint is the current limit of line 7-8 as 250 A. The generator cost function of all three generators for this system are given by [52],

\[
C.F_1 = 0.11 \times P^2 + 5 \times P + 150
\]

\[
C.F_2 = 0.085 \times P^2 + 1.2 \times P + 600
\]

\[
C.F_3 = 0.1225 \times P^2 + 1 \times P + 335
\]
4.3.3.1 Solution 1

Primal linear programming (LP) OPF solution is used to optimize the system. The objective function is to minimize the operating cost by rescheduling of the generator. Solving a primal LP OPF changes system control to remove any violation. It determines the optimal solution by iterating between solving a standard power flow and then solving a linear program. So, the quadratic input/output model (non-linear cost functions) given in equation (4.3) are approximated into piecewise linear model as a series of straight line segments.

<table>
<thead>
<tr>
<th>Generator</th>
<th>Max. Output (MW)</th>
<th>Power Output (MW) Before attack</th>
<th>After OPF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1*</td>
<td>250</td>
<td>71.63</td>
<td>121.21</td>
</tr>
<tr>
<td>2</td>
<td>300</td>
<td>163</td>
<td>100.21</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>85</td>
<td>100.00</td>
</tr>
</tbody>
</table>

* → Most expensive generator

The OPF results in Table 4.1 shows how the generators are optimally dispatched while simultaneously enforcing the transmission line limit. The system is now at new operating condition and current through line 7-8 is found to be 247.28 A. It is observed that the generation from generator 2 is decreased because of the transmission line limit (line 7-8) and generator 3 is limited by its maximum capacity. So, to fulfill the demand, the most expensive generator is increasing its output. When the system is under OPF AGC control, all generators output are varied automatically by AGC in conjunction with the solutions solved by the OPF algorithm [110]. This also helps to maintain the system frequency while minimizing operating costs and satisfying all necessary constraints.
4.3.3.2 Solution 2

In addition to the rescheduling of the generator, optimal power flow is also used with dispatchable loads to relieve transmission overloading. The generator cost functions for this case are similar to that as given in equation (4.3). In this case, the objective function is to minimize the control change and OPF will seek to relieve transmission overloads with the minimum change from the initial operating point [111]. This objective function will enable load and generation control (generator controls are also needed to follow any changes in load) with OPF. The first step is to ensure that each load in the system is AGCable and OPF load dispatch is enabled in the area. The simulation result is shown in Fig. 4.9 where the OPF solution has curtailed the load at bus 5 by 44.7 % to relieve the overload in branch 7-8. With the new configuration, the current flow in line 7-8 is observed to be 246.82 A.

Figure 4.9. OPF mitigation simulation for load curtailment.

Hence, optimal power flow can be used as one of the effective mitigation strategies which help to adjust the power flows to restore a normal operating condition.
4.4 Summary

A real time cyber-physical testbed is presented to simulate different attacks and disturbances in the smart grid. The discussed testbed is a valuable tool for simulating events in the realistic cyber-physical environment. It is used to demonstrate the impact of the cyber attack on the physical power grid in terms of voltage stability and loss of generation. A mitigation strategy is also proposed using optimal power flow for preventing the system from being unstable.
CHAPTER 5 CONCLUSIONS AND FUTURE WORK

5.1 Conclusions

The smart grid is enjoying the integration of latest technology for efficient and reliable operation. Its integrated nature and inclusion of different communication paradigms is making the modern power system more intelligent and economic. However, it is facing some critical problems because of its structural vulnerability and cyber security issues. This thesis has focused on the security of modern power grid from three different perspectives.

First, the cascading failures caused by potential attacks in the power grid was studied. The vulnerability of power grids from the attackers perspective is investigated. The attackers will always want to compromise the particular grid components which will trigger a cascading failure causing blackout of large size. Two different models of power system were discussed and performance comparison of various attack strategies was presented. The selection of initial victim sets for causing cascading failure should be done preserving both topological and electrical structure of the grid since neither of them is robust enough to provide comprehensive evaluation of failure propagation alone. Clustering based attack was found to be more critical as it causes greater damage in the power system benchmarks. In addition, the model based on power flow was found to be more realistic in studying the cascading failures in the power grid.

Second, the screening of vulnerable branches was done based on electrical distance approach. It was observed that the proposed approach is computationally efficient than the traditional contingency analysis method. The vulnerability of the branches was examined
based on their impedance values. The proposed algorithm is applied in different power system test cases for \( N - 1 \), \( N - 2 \) and \( N - 3 \) contingencies. The number of line limit violations during single and multiple contingencies were found to be associated with branches under the electrical structure, i.e., branches with smallest electrical distance. The proposed approach was found to be efficient in comparison to traditional methods for screening of the vulnerable branches during contingencies. Study of the voltage profiles further supported that the branches with smallest electrical distance are more prone to vulnerability.

Finally, the development and applications of a CPS testbed were discussed. The application of a CPS testbed was provided and the architecture of developed platform was presented. The physical and cyber domains of the CPS testbed were discussed and their interdependencies was evaluated based on the potential cyber attack. The impact of the possible cyber attack in the physical power system was studied in terms of system voltages and generation loss. The effectiveness of the mitigation strategy can be evaluated using the testbed before implementing it in the real power system cases.

5.2 Future work

The future work along this direction includes the following major tasks:

1. Consider both the node (substation) capacity and line limits during cascading failure process and visualize the failure propagation in large scale power grid, which consists of thousands of substations and transmission lines. In addition, the cascading model could be improved by introducing dynamic condition of the power system.
2. Simulate the cascading failure phenomenon using the developed testbed in real time.

   The effect of multi-node attack could be simulated in proposed testbed and evaluate
   the damage, as well as defense the attack for preventing failure propagation.

3. Integration of distributed energy resources like solar, wind and storage devices
   (battery) in real-time CPS platform for investigating the systems’ reliability during
   unplanned outages and malicious attacks.

   In general, all these works are expected to enhance the smart grid security against
   potential attacks and failures.
LITERATURE CITED


[78] V. V. Williams, “Multiplying matrices in o (n2. 373) time,” 2014.


