Improving Resilience in Large Scale Cyber-Physical Networks

By

Saqib Hasan

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Electrical Engineering

February 28, 2019

Nashville, Tennessee

Approved:

Gabor Karsai, Ph.D.

Abhishek Dubey, Ph.D.

Richard Alan Peters, Ph.D.

D. Mitchell Wilkes, Ph.D.

Gautam Biswas, Ph.D.

Contents

Tables

Figures

This work is funded in part by the NSF under the award number CNS-1329803 and the NSF FORCES project under the award number CNS-1238959. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of NSF or FORCES.

Chapter 1

Introduction

Cyber-Physical Systems (CPS) such as power grids are systems that consists of both physical and cyber-components that are tightly coupled together to perform critical system monitoring and control functions via software programs. Therefore, it is important to build CPS systems that can anticipate change and exhibit resilience, i.e., operate under uncertain environments in the presence of faults while being dependably functional. In order to achieve resilience and reliability, power systems are slowly transforming into smart grids where it utilizes renewable energy sources and distributed monitoring/control to meet the future demands of the consumers while maintaining grid stability. This has become possible as a result of the evolution in distributed computing from small homogeneous clusters of computers to IoT technology.

In general, in addition to physical components such as transmission lines, transformers, generators, etc., power systems are equipped with a large number of cyber-devices such as Remote Terminal Units (RTUs), Phasor Measurement Units (PMUs), distance relays, circuit breakers, etc. Faults can occur in any of these physical/cyber-devices in the form of both physical faults or cyber-effects that can lead to system failure. Moreover, cyber-effects can be a result of malicious action that is performed by an adversary in the form of cyber-attack. For instance, the December 2015 Ukraine attack where, the attackers stole the credentials for the control centers and isolated several circuit breakers in addition to the Denial of Service (DoS) attack which caused a severe system damage. These attacks are becoming more prevalent [2] and smart grids are exposed to new vulnerabilities due to the increase in the potential attack surfaces [3] as a result of increase in cyber-devices. In addition, the introduction of cyber layer in smart grids further increases its complexity. Further, both pysical faults/cyber-effects often lead to a sequence of events that trigger

Table 1.1: List of Cascading Failure Events Resulting in Blackouts

Year	Place	Cause	Consequence
1996	Idaho USA	falling tree branch	18 western states in the US were blacked out due to outages of power plants and transmission lines
1997	Quebec Canada	Ice storm	A large part of New England, USA blacked out due to loss of transmission lines
1998	Minnesota	severe Lighting storm	Both Mid-Continent Area and the northwestern Ontario Hydro system of Northeast were affected due to system disturbance
1999	Sao paulo Brazil	Zone-3 relay trip	75 million people were affected due to cascading outages as a result of high voltage ac, dc line outages.
2003	Italy	Flash over due to tree trips, Unsuccessful reclosure, Miscommunication between system operators	Separation from the UCTE grids
2004	South Eastern U.S.	Hurricane Frances	Loss of 6018 MW power
2005	South Eastern U.S.	Hurricane Wilma	Loss of 10000 MW power
2012	India	Circuit breaker tripped transmission line	22 out of 28 states in India were without power and a total of 32 GW generation capacity was out of service
2015	Ukraine	Spear-phishing attack, circuit breakers opening, DoS attack	225000 people were without power
2016	South Australian	Two phases of the line were grounded	No supply to the entire South Australian region

a phenomenon called cascading failures where one or more initiating faults causes subsequent failures resulting into severe system damage commonly known as blackouts in power systems. Table 1.1 shows a list of faults that resulted into severe system blackouts as a result of cascading failures in power systems. Analyzing and understanding cascading failures in power systems to improve their resilience and stability has been a complex and challenging global problem.

Given the importance of resilience in power systems, several analysis models have already been developed in the past. Most of these models follow the traditional analysis via simulation which is considered as one of the best approaches for evaluating a system. These models analyze the vulnerabilities (physical faults/cyber-effects) by initiating them prior to the start of simulation and evaluating their effects on the system. In addition, most of these models consider vulnerabilities that are only related to physical faults in the system such as a failure in transmission line resulting from a three-phase to ground fault, etc. However, cyber-effects such as distance relay mis-operation under faulty/nominal condition is a genuine fault that could result in initiating cascading failures [4]. Therefore, analyzing smart grids require sophisticated models that include cyber-effects in addition to physical faults that can be instantiated and analyzed [5]. In addition, these models require instantiation of these faults at any instant in time in order to explore new cascading failure trajectories that remained unexplored in the previous approaches. In our context, trajectories are referred to the detailed propagation of the cascades showing the sequence of failure of the components as a result of one or more initial faults. These cascading failure trajectories, if identified, will provide directions towards building a more resilient system.

To achieve this, a power systems analysis model should focus on analyzing cascading failure that includes: 1) Detailed behavioral models for protection assemblies, i.e, distance relays, over-current relays, and circuit breakers that takes into account various cyber-effects. 2) Provides the capability to introduce physical-faults and cyber-effects at any desired time that can be used in any desired way to perform system analysis. These ca-

pabilities will allow us to solve the challenge of developing sophisticated analysis models for power systems that enable us to capture new cascading failure trajectories that would remain unexplored otherwise.

Even with the availability of sophisticated analysis models, power systems being complex networks need analysis from multiple aspects, i.e., it is necessary to perform various types of analysis such as transient analysis, time domain analysis, steady-state analysis, etc., to evaluate the overall system resilience. However, multi aspect analysis of power systems would require model building in individual platforms that would result in significant increase in modeling time and error. The modeling time, effort and error increases greatly with increase in the power system size. In addition, to achieve the overall system resilience it is necessary to perform a complex analysis on the system however, no tool provides the capability to perform a detailed multi aspect system analysis. Further, performing the analysis using individual tools require building the system model separately in each platform. This process is highly error prone and takes a lot of modeling time and effort when the size of the system increases. To address this challenge, the key is to develop a framework that can provide the capability for performing such complex analysis on power system models via the use of an appropriate Domain Specific Modeling Language (DSML). The DSML could act as the base language for modeling the power systems once and the framework would support generation of tool-specific translated system models using dedicated transformation plugins. Further, it can perform the desired analysis by selecting the appropriate tool from the tool-chain supported by the framework.

Tools that provide complex power systems analysis can be useful to evaluate only the selected single/multiple contingency and provide support towards improving the power system resilience. However, identifying system wide multiple contingencies remain a challenging problem. We define a multiple contingency as a fault in more than one component in a power system network. Moreover, a multiple contingency analysis also known as $N-k$ contingency analysis refers to the evaluation of more than one component fault in power

systems to effectively understand its effects on the system, where N is the total number of power system components and k is the total number of faults. As per the North American Electric Reliability Council (NERC), electrical power systems are designed as resilient systems and are usually $N-1$ secure [6], i.e., fault in a single component such as transmission lines, protection relays, etc., does not lead to any stability or thermal limit violations within the system. Moreover, considering the large scale of the power system networks, there are several multiple contingencies that can occur and a lot of them may cause severe cascading failures that result in blackouts [6]. Therefore, it is essential to perform such contingency analysis on the power system in order to identify the critical contingencies beforehand to improve the power system resilience by designing effective mitigation strategies. However, exhaustively performing multiple contingency analysis is a computationally challenging problem due to the combinatorial explosion of the search space, i.e., N^k as the scale of the power system increases with the increase in N and k [7, 8, 9]. Therefore, there is a need to develop effective and efficient methods that could address this problem by 1) Reducing the search space N^k considering the size of the power system and the increase in the value of k. 2) Limiting the contingency analysis to only those contingencies that qualifies for the detailed analysis via simulations using some initial pre-screening metric.

Further, as stated earlier, due to the cyber layer in the smart grids, several contingencies at present can easily be introduced in the power system networks through cyber-attacks [4] by gaining access to power grid substations. In addition, cyber-attacks have been documented as one of the major obstacles towards the reliable power system operation and are recently increasing both in number and sophistication [10, 11]. Strategic attackers can launch these attacks in the form of static attacks, i.e., attacks launched at the same time or dynamic attacks, i.e., attacks launched at different times at different points in the power system networks. Now, in order to improve the overall resilience of the power systems it is essential to identify the critical components or parts of the network (e.g., substations and protection assemblies) to protect such that when a cyber-attack is launched, the damage

to the power system remains minimum. However, because of the limited financial budget only a few components can be protected effectively in a power systems network against such attacks. To achieve this, first one needs to develop realistic static cyber-attack/defense models to analyze the system. Next, it is necessary to use such models to identify the critical components to attack that can cause severe system damage. However, due to the computational complexity it is very difficult to identify all the possible contingencies that result in large power system damage due to the attacks. So the question becomes "Which critical components to attack in order to maximize the system damage when attacks are launched simultaneously?". Moreover, identification of the most damaging attacks is not the complete solution to the problem. It can only provide the defender with the weak points in the power system models. Therefore, effective defense mechanisms are needed in order to device optimal defense strategy according to which only those critical components that can minimize the system damage when a static cyber-attack is launched are protected while utilizing the limited financial resources.

Again, cascading failures in power systems evolve slowly and takes at least a few minutes to sometimes even hours to progress [12]. Therefore, these cyber-attacks can be strategically timed and executed to cause a severe system damage. According to research discussed in [13, 14], strategically timed cyber-attacks or dynamic cyber-attacks cause significantly higher damage as compared to their static counter parts. A strategic attacker having a power system knowledge would easily be able to identify and time cyber-attacks causing severe cascading failures resulting in higher system damage. However, considering the large scale of the power system networks, the question arises "which components to attack and at what specific time instants so that the damage to the power grid is maximized?". Further, to improve the overall system reliability and resilience while considering the financial budget constraints, the question is "Which components to protect so that the system damage can be minimized when a dynamic attack is launched?". Hence, considering the financial constraints and the most damaging dynamic attacks, the available resources needs to be ef-

fectively utilized to achieve overall system stability. Therefore, there is a need for dynamic attack methods that could effectively and efficiently identify the most damaging dynamic cyber-attacks. At the same time, effective defense mechanisms are needed in order to identify the critical components that can be protected by utilizing the limited financial budget in order to minimize the damage when a dynamic cyber-attack is launched.

The remainder of this dissertation is organized as follows: Chapter 2 describes some fundamental concepts about the general terms, definitions, and some basics of power systems network. Chapter 3 describes the related research in cascading analysis models and tools, contingency analysis approaches, and cyber-physical attack based methodologies. Chapter 4 introduces the platform that takes into account both physical and cyber-faults in addition to the temporal aspect of the faults. Chapter 5 discusses the framework that integrates multiple simulation tools together to provide a mechanism for better system analysis. Chapter 6 describes an effective and efficient contingency analysis methodology. Chapter 7 introduces the techniques of limited defense resource allocation under static cyber-attacks. Chapter 8 discusses the mechanisms of limited defense resource allocation under dynamic cyber-attacks. Finally, Chapter 9 concludes the dissertation and provides some insight for the possible future research directions.

Chapter 2

Fundamentals

An electrical power system is a large complex network of power generation, delivery, monitoring, and control components. The power generation components such as generators produce the necessary power and the step-up transformers are used for increasing the voltage for transmission purposes. The delivery elements such as transmission lines, buses, etc. carry power from the generating stations to the consumption points where the voltage is stepped down using the step-down transformers and the power is distributed to the consumers. However, the control and monitoring devices such as protection assemblies, i.e., distance relays, over-current relays, etc. and circuit breakers are responsible for isolating the faulty components under abnormal system conditions. On one hand, the distance relays use a comparison of the measured value of the impedance with the actual impedance value to provide control signals to the circuit breakers. On the other hand, the over-current relays compare the measured current value with a fixed threshold value to provide the necessary control action.

A fault can occur in any component in a power network and is referred to as a *contingency*. Faults can also occur in more than one component of the power system network. These faults are referred to as $N-k$ *contingencies*, where the value of k represents the number of faults in a system that consists of N elements. Occurrence of $N-k$ *contingencies* can cause severe *cascading failures*. A *cascading failure* is an outage of one or more components in the power system network that cause subsequent failures and result in system *blackout*. A *blackout* is a system failure state in which the system cannot be operated under nominal operational conditions due to several reasons such as stability constraint violations, severe load loss, etc. $N-k$ *contingencies* often occur and some of these *contingencies* result in a system failure. These *contingencies* are referred to as critical $N-k$

contingencies. Moreover, analyzing multiple contingencies to evaluate their effect on the power system is referred to as $N-k$ *contingency analysis*.

Traditional power grids are transformed into today's smart grid due to technological advancements. As a result, smart grids are equipped with several sophisticated monitoring and control devices such as Advanced Metering Infrastructure (AMIs) that are used for collection of data and provides communication between customers and service providers [15], Phasor Measurement Units (PMUs) that are used to obtain the electrical signal measurements with respect to a common time source for synchronization [16], remotely controlled protection devices such as distance relays, over-current relays that are used to monitor the bus voltage and line current in order to take corrective actions under abnormal system conditions[17], etc. Due to the technological improvements, these devices rely on software in order to perform the desired functions. However, such software can have security flaws and thus there is an increase in the potential attack surface in the power networks that give rise to new vulnerabilities that can cause severe cascading failures. Malicious attackers take advantage of such vulnerabilities through *cyber-attacks* and cause severe damage to the system. *Cyber-attacks* can be classified into two types, i.e., *static or simultaneous cyber-attacks* and *dynamic cyber-attacks*. *Static cyber-attacks* are the attacks that take place simultaneously at multiple points in the network. However, *dynamic cyber-attacks* are those cyber-attacks that can be executed at specific instants in time at different points in the network. *Cyber-attacks* when executed can cause severe cascading failures resulting in large system damage. This damage can be considered as large when the load loss in the system is above a certain percentage, e.g., more than 40% load loss.

Chapter 3

Related Work

3.1 Design and Modeling Methods for Achieving Resilience

There are several methods and techniques to design and analyze the cascading failures in large scale power networks consisting of hundreds of nodes representing buses and edges representing the transmission lines to improve the system resilience. These methods and techniques consider distinct failures related to different components within the power systems that could cause cascading failures and can result in catastrophic damages to the system.

3.1.1 Hidden Failure Model

Distance relays that are used as a protection and control equipment to isolate the faulty components from the network under abnormal conditions tend to have hidden failures [18, 19, 20]. These hidden failures in the protection equipment are a contributing factor towards severe cascading failures that result in several small and large blackouts. Moreover, these failures are triggered only in conjunction with a normal fault isolation operation. However, they remain dormant during normal system operation, calibration and maintenance. In [18, 19], these failures are ranked based on the vulnerability indices. A vulnerability index is defined as the priority or sensitivity ranking of a region of vulnerability with respect to the other regions. In addition, a region of vulnerability is defined as a region where an occurrence of a fault will result in the incorrect operation of a distance relay leading to the exposure of any additional hidden failures. Failures with higher vulnerability index are considered to be more severe than the one's with lower vulnerability index. Similar types of faults are considered in [20] where the hidden failures in distance relays can cause unnecessary tripping of lines and these are referred to as sympathetic trippings.

A model of cascading failure has been developed in [21] to understand how cascading failures spread, to identify the key lines and to figure out some effective methods to minimize the cascading failure risk. The model considers the initial outage of transmission lines based on the Monte Carlo method. It also considers transmission line outages due to hidden failures in protection assemblies. These outages are determined using the Monte Carlo method as well. If more than two lines are determined to trip following the initial trip then the line with higher hidden failure probability is considered to be tripped next. The hidden failure probability is obtained using the statistical data in the study. In these studies, the cascade evolution paths are greatly limited as the hidden failures of only those distance relays that are connected to the same bus gets exposed under an abnormal condition. The hidden failures in protection assemblies are looked from a different perspective, i.e, with the concepts of Petri nets in [22].

3.1.2 Oak Ridge National Laboratory (ORNL)-PSerc-Alaska (OPA) Blackout Model

Considering resilience in power systems, an OPA model is developed to study the effects of the improvements in the transmission systems to reduce power system blackouts [23, 24, 25]. This model is used as it reflects the characteristics of a real power system. Three types of improvements are studied namely impact of increasing the reliability of individual components of the system, changing the operating margin of the system and impact of implementing component redundancy on the system. The effect of these improvements on the system resilience is analyzed by performing blackout studies. These blackouts are simulated by randomly removing transmission lines from the power network with a fixed probability or loss of load due to the removal of generators that exceeds their operating limits. Here, the load shedding is avoided and the process can be multi-iteration that terminates when there are no more outages or load flow divergence.

3.1.3 Small-World Model

Small-world models are defined as those models that are considered as exponential or homogeneous networks where, each node has almost the same number of links. The study discussed in [26] analyzes cascading failure using a small-world network model, where the power systems is represented as a graph with nodes and edges. It suggests that failure of nodes having higher centrality, i.e., the nodes with higher importance results in severe cascading failures that cause larger system damage. In this study, a small-world phenomenon is explained through a regular ring lattice that uses two parameters i.e., the characteristic path length and clustering coefficient that are used to explain the reason for bulk power system cascading failures. Characteristic path length in [26] is defined as the "median of the means of the shortest path lengths connecting each vertex of the graph to all other vertices". As per [27], clustering co-efficient is defined as "the average fraction of pairs of neighbors of a node that are also neighbors of each other".

3.1.4 Scale-Free System Model

It is argued in [27, 28] that most large scale complex networks such as power systems are scale-free systems. A scale-free systems are those systems whose connectivity distribution follow a power-law independent of the scale of the power networks, i.e., most nodes have few connections and a few nodes have several links [29, 30]. In this model, generator and buses are considered as nodes and an undirected graph is developed with nodes and edges. However, these graphs can grow in size if a new node is added. Thus, the topology of the graph with respect to the newly added nodes changes. In comparison with a random graph of same size and average degree, the scale-free graph has a smaller average path length and a higher clustering co-efficient. As per [27], clustering co-efficient is defined as "the average fraction of pairs of neighbors of a node that are also neighbors of each other". This could provide essential information about the critical links in the network. For exam-

ple, the big nodes with very large degrees can act as the major links that connect to various other parts of the network and are considered to be the most crucial ones. However, the limitation of the model is that it does not consider the system physics and power flows in identifying the critical links of the network. This is very important considering the highly non-linear nature of the power system models.

Another model that considers the power systems as a graphical representation to perform the cascading failure analysis is discussed in [31]. This model also considers the isolation of transmission lines due to overloads and do not consider any other types of system faults. The model also do not consider faults at different instants in time which is necessary in order to improve system resilience.

3.1.5 Dual Graph Model

Conventional studies considering topological network models considers power systems as graphs where nodes represent buses and edges represent transmission lines. However, in the dual graph model [32] the authors argue that such assumptions can cause misleading results. Hence, they provide an alternative way of representing the power grid as a graph where the vertices of the graph are considered as the transmission lines and the edges represent an interaction between power lines. The dual graph is then used to perform cascade failure analysis and identify severe critical contingencies that are based on the outage of the overloaded transmission lines.

3.1.6 CASCADE Model

The importance of the electrical transmission and distribution system towards the society is a motivation for the development and analysis of the models for cascading failures. A loading-dependent model of probabilistic cascading failure is studied in [33, 34]. Here, a simple cascading failure model is developed where each component in the power system is assumed to be loaded at some initial value below its maximum threshold level. When a

component fails, its fixed amount of load is distributed to each of the remaining components in the power system network. This could lead to overloading of subsequent components and further cascading failures. The cascade may progress in stages and the process stops when there are no further overloads in the system. Finally, the number of components failed can be obtained and the severity of the cascade is evaluated. This model is far too simple to reflect the realistic aspects of the power system and provides only an understanding on how cascade progresses. Moreover, the extent of the cascade depends on the initial loading of the system components. In addition, the developed model neglects the interaction of the various faults among different components and the time between adjacent failures.

3.1.7 Branching Process Model

A branching process model for analyzing cascading failures is studied in [35, 36, 37] that extends the previous works and improves the cascading failure analysis models by considering time. Using these models the likelihood of a blackout occurring can be explored and the cascade progression can be identified. The risk associated with the blackouts caused by the cascading failures using branching process model is effectively studied and quantified in [38]. However, these methods still ignore the interaction of different types of faults associated with various components in power systems and the effect of these faults when they are triggered at different instants in time.

3.1.8 The Manchester Model

The Manchester model described in [39] is based on the ac power flow. It represents various cascading failure interactions such as sympathetic tripping of lines, generator instability, under-frequency load shedding, post-contingency power re-dispatch, and emergency load curtailment. It employs Monte Carlo simulation to evaluate the expected blackout cost. Later it was used in [40] to develop a scale for the system stress. The scale relates system loading to blackout size and a new vulnerability index named overload risk index

(ORI) was developed. The approach utilizes the line outage distribution factor (LODFs) to perform security analysis. Based on the current system loading levels and component failure rates the ORI provides a measure of the severity of transmission line overloads. The limitation of this method is that the LODFs can only be used to solve up to N-3 contingencies in the power system. However, higher order contingencies do occur and needs to be simulated to perform an in-depth analysis.

3.1.9 Stochastic Model

The stochastic model performs the evaluation of both isolation and connection of components in a power network [41]. It computes ac and dc power flows and considers the transmission line outages due to unforeseen stochastic events. The line outages are a result of the overheating caused by excessive power flows. In this model the temperature of the transmission lines is monitored and it uses Poisson distribution to identify the transmission lines to be removed from the power network. This model can simulate the slow evolution of the cascade and has the capability to perform various network analysis such as obtaining the shortest-path, determining electrical islands, etc. However, various other faults with respect to system components are excluded from the model without which an overall system wide perceptive of resilience cannot be obtained.

3.1.10 Other Models

To effectively provide countermeasures to avoid cascading failures that could lead to blackouts it is essential to perform an extensive study on the power system for the occurrence of the widespread blackouts. One such study is discussed in [42] where the authors have developed methods to create various cascading failure scenarios to perform an in-depth analysis on the system. To develop such methods system pre-condition, post-contingency condition and the availability of control actions are the minimum information that is needed to be considered. Then the sequence of events leading to blackout should

be determined in order to get an entire idea of the cascade progression which can help in effectively designing the mitigation strategies.

Some important insights about cascading failures is provided in [43]. It is observed that if the power systems are operated below the critical point, i.e., the system loading limit below a threshold value, it experiences fewer blackouts. Operating power systems above critical point results in more blackouts that cause system upgrades. Another important insight is that the power systems tend to have a critical loading point. If the power systems are operated above the critical loading point then the number of components failing increases significantly. Another insight from the paper suggests that failure of a highly loaded transmission line causes a large disturbance to the system that could result in severe cascading failures. Moreover, according to the study certain methods and upgrades that are used to suppress small blackouts can ultimately increase the risk of large blackouts.

3.2 Fundamental Analysis Tools for Achieving Resilience

Resilience in electrical power systems is the primary goal of the system operators. In order to understand resilience, a resilience metrics is needed to be defined. There are different resilience metrics that are already available [44], where engineers can look into a system wide perspective, i.e., overall loss associated with the system or they can look into other aspects such as the transient stability, voltage stability etc. in order to evaluate a power system network.

To better understand the cascading failures and identify the critical components of the power systems for improving the reliability and resilience, it is necessary to include different aspects such as steady state analysis, transient analysis, time based analysis, time independent analysis, cyber failures in power system components, etc. in the power system analysis while performing cascading failure studies. This necessitates the need for the simulation platforms that can provide the capability to analyze the power systems depending up on these aspects. However, such simulation platforms either do not exist or are typically

very expensive. In addition, most of these platforms do not provide the flexibility to control the modeling of the system as per desired. Also, they do not allow the variation in their simulation environment according to the needs of the user to augment the functionality of these tools to perform different analysis if needed. As a result, we use multiple tools either open source or industrial tools to perform various types of analysis on individual platforms. Each of these tools are limited in their capabilities and can provide the analysis as per their functionality. Moreover, they have their own semantics and specifications depending on which a power system needs to be modeled in their modeling environment.

3.2.1 Modeling Languages

Various modeling languages are currently available for system modeling and analysis. One such language is Modelica [45] which is a multi-domain modeling language and is not a conventional programming language. It is mainly used for component-oriented modeling of complex systems such as electrical, mechanical, electronic, control, etc. Even though Modelica is similar to object-oriented programming languages, its classes are translated into objects rather than compiled in the usual sense. These objects are then used by the simulation engine. There are both commercial and free simulation environments of Modelica such as Dymola [46], MapleSim [47] and OpenModelica [48] that are available. Domain-specific models can be developed using the Editor in any of the free or commercial versions of Modelica and the developed model can be compiled by translating into a C code which can then be simulated and analyzed.

3.2.2 InterPSS

InterPSS [49] is an open source simulation platforms that performs the AC load flow analysis on the power system networks. It provides the capability for the user to develop their own objects and integrate them with the existing software platform to augment its functionality. Moreover, it also provides the capability of integrating the components of

InterPSS with other software systems.

3.2.3 Power System Analysis Toolbox

The power system analysis toolbox (PSAT) [50] is an open source Matlab [51] and GNU/ Octave-based software package. It includes the capability to perform continuous power flow, time-domain simulation, power flow, etc. It provides a user-friendly graphical interface and a Simulink-based editor for the users to easily model single line diagrams of the model. The graphical user interface (GUI) allows the user to easily interact with the developed models to perform the desired simulations. The limitation of PSAT is that it is only suitable for design and analysis of small to medium scale electrical power systems.

3.2.4 Voltage Stability Toolbox

Voltage stability toolbox (VST) [52] is another Matlab-based voltage stability toolbox. It is developed to perform bifurcation and voltage stability analysis on electrical power systems. This tool is mainly developed for enhancing the educational courses in order to easily demonstrate the fundamental concepts of voltage stability phenomenon to the user. It provides the user an ability to understand how different loading conditions affect the system stability and what corrective actions can be taken to prevent instabilities.

3.2.5 MATPOWER

Another modeling and simulation platform, i.e., MATPOWER [53] provides optimal power flow (OPF) solutions for electrical power systems. This tool is targeted towards researchers, educators and students. Its OPF architecture provides the capability to add user-friendly variables, costs, and constraints. This enables the user to design the problem as per needed. However, it does not consider other important aspect of simulation such as including cyber failures in the components, time domain simulations, voltage stability, etc.

3.2.6 GridLab-D

GridLab-D [54] is a simulation platform developed by the US Department of Energy at Pacific Northwest National Laboratory. It includes models for appliances and equipments, consumer models, etc. Capabilities such as load shedding, distributed generation, storage models, retail market modeling tools, SCADA control models, metering technologies, etc. are available in this tool. GridLab-D can be easily linked to external platforms such as Matlab, MySQL, Microsoft Excel, Microsoft Access, and other text-based tools. The analysis results from GridLab-D provide important system statistics such as profitability etc.

3.2.7 PowerFactory and PSCAD

PowerFactory [55] and PSCAD [56] are the standard simulation and analysis platform used for studying large interconnected power systems. PowerFactory can perform both AC and DC load flow analysis. It supports the simulation of FACTS, HVDC cables, etc. PowerFactory can be easily integrated with other existing platforms through interfaces such as API, DGS, CIM, etc. and is perfectly suited for transmission system operation planning. PSCAD is a very popular simulation and analysis platform and is used for performing transient analysis on electrical power systems.

3.2.8 PowerWorld Simulator

PowerWorld Simulator [57] is a user-friendly and highly interactive power systems analysis and simulation package. It is capable of solving power flow for systems with very large number of buses and provides graphical visualization as well. It also provides full-color animated one line diagrams of the power system models. These models can be easily modified on the fly or can be built from scratch using the graphical editor. Other features such as transmission lines switching, addition of generators, etc. are also available in the package.

3.3 Contingency Analysis Methods

Contingency analysis is an important aspect for achieving power system resilience. Reliable operation of power grids is the primary goal of the system operators. According to the North American Electric Reliability Corporation (NERC) standards [6], power systems are usually operated according to the $N-1$ security criterion. The $N-1$ security criterion indicates that the failure of any single component would not cause any system violations such as branch flows, violation of bus voltage or stability limits, etc. Operators are able to manage these contingencies on a day-to-day basis. However, dealing with multiple simultaneous contingencies, i.e., $N-k$ contingencies (where $k \geq 2$) becomes very challenging considering the scale of large power system networks.

In order to deal with the multiple $N-k$ contingencies, at first these contingencies needs to be identified. However, finding all the possible critical $N-k$ contingencies for a large power systems becomes computationally infeasible especially for higher values of k. The reason behind the computational in feasibility is the combinatorial explosion of the search space. For instance, ignoring the sequence of a power system, it requires $\frac{N!}{k!(N-k)!}$ number of simulations to identify all the critical $N-k$ contingencies while performing the contingency analysis. The number of simulations required increases exponentially (N^k) with increasing values of N and k. For example, let's consider a power system with $N = 5000$, where N is the total number of components in a power grid. In order to identify all the critical $N-4$ contingencies resulting in sever cascading failures causing blackouts, it requires approximately 26×10^{12} simulations. It is impossible to perform such simulations even with distributed computing platforms. Thus, the exhaustive search becomes infeasible while finding all the possible critical $N-k$ contingencies. This number even grows drastically if the sequence of the contingencies is taken into account.

Multiple $N-k$ contingencies, however, do occur and cause sever cascading failures that result in large blackouts. A few examples of such blackout cases are August 2003 North America [58], July 2012 India [59], Dec 2015 Ukraine [60], and Feb 2016 South Australia

[61] blackouts. Therefore, as per the NERC standards the system operators today are required to operate power grids against cascading failures with multiple $N-k$ contingencies [6]. Several contingency analysis methods have been developed in order to identify the critical multiple $N-k$ contingencies.

3.3.1 Ranking and Selection Methods

Various contingency ranking and selection methods have been developed and studied for identifying critical $N-k$ contingencies. In [62], a list of contingencies consisting of transmission line and generator outages are obtained that are ranked according to their severity. The severity directly relates to the effect of these contingencies on the bus voltage and transmission line flows. The ranking of these contingencies is obtained by ordering the normalized sensitivities from the highest to the lowest that are obtained using Tellegen's theorem [63] with respect to individual outages. This method does not provide the details whether the contingency will cause any system operational limit violations. However, it provides a comparison of the severity among different contingencies. In this approach, contingency mis-ordering is the main drawback due to incorrect computation of the performance index.

An extension to the contingency ranking and selection method described in [62] is discussed in [64]. Here, the drawback of the previous study is eliminated to a great degree by considering better methods toward contingency selection by designing improved expressions to obtain the performance index. Unlike the previous approach that uses only the linear terms in the expression for computing the performance index, this method includes all the terms in the infinite Taylor's series expansion to obtain a more accurate performance index that can be used to rank the contingencies more precisely. However, in some cases, the value of performance index reduces as a result of a single line overload that decreases the loading on other lines. This is referred to as masking in [64] and avoids the overloads from getting recognized that would lead to further in accuracy in the ranking process. Cer-

tain approaches are described in detail in [65] to address the problem of masking.

The method discussed in [66] proposes another way for identifying critical $N-k$ contingencies that considers some concepts from the above mentioned studies. The study is motivated to develop better contingency selection strategies. According to the study, the first strategy is based on simulating the entire single and various multiple $N-k$ contingencies using a fast approximate technique. From the results, the contingencies that cause the worst system insecurities can then be analyzed. The second strategy is based on a similar mechanism of severity index as discussed in above approaches. However, it has drawbacks related to computing times, stability of the sensitivity index, and reliability of the contingency list. Therefore, the former strategy is suggested to be more reliable and faster. But, it has drawbacks related to re-computation of the distribution factors whenever the system topology changes which will increase the computational burden.

A fast contingency screening and evaluation for voltage security analysis is discussed in [67], where a subset of voltage sensitive buses with potential voltage problems are identified. Further, these buses are then screened for voltage-reactive power. The method uses the compensation techniques [68] to simulate the component outages and utilize the sparse vector method [69] to obtain fast solutions. Finally, it uses the adaptive reduction [70] to reduce the solution time for the necessary cases.

3.3.2 Distributed Computing Based Methods

A distributed architecture for online power systems security analysis is developed in [71]. According to [71], "The real-time assessment of the systems security and reliability levels, especially under unforeseen contingencies, is known as online power system security analysis". Online contingency analysis usually requires large computational efforts for solving the power flows, checking the thermal limit violations of the components, dynamic component loadability calculations, etc. In order for the analysis methods to be useful they need to be computationally inexpensive. Therefore, a distributed architecture is developed

to address the complexity issues. Based on the architecture, a networks of remotely controlled units are distributed in the most critical sections of the electrical network. These units are responsible for field data acquisition and dynamic components loadability assessment. The data acquired from these units can be used directly by a highly scalable solution engine as it uses distributed computing to perform the online security analysis, where the jobs can be divided for obtaining a fast solution. Further, for providing the resulting system assessment a web-based interface is also available within the architecture.

3.3.3 Graph Based Methods

A two-stage screening and analysis approach for identifying severe multiple $N-k$ contingencies in power networks are discussed in [72], where an optimization problem is formulated that utilizes spectral graph theory as initial screening method for identification of the set of transmission lines of interest in stage 1 and then performs a detailed analysis of the selected transmission lines to evaluate against the system security criterion in stage 2.

In this approach, the entire power network is considered as a graph and as part of the initial screening the graph is sub-divided into subgraphs by removing the selected transmission lines from the network. The selection process for the lines to be removed is based on the system operating point and the feasibility boundary as discussed in [72].

The main drawback of this model is that the optimization problem developed is non-convex and the screening model neglects the voltage variations and the reactive power considerations while identifying the initial set of transmission lines for detailed analysis.

The above drawbacks are further addressed in [73] by formulating the problem as a mixed integer non-linear optimization problem. However, the optimization problem in the study is still non-convex. The two stage analysis approach is similar to the one discussed in [72]. Similar problem is studied by formulating it as a bilevel optimization problem in [74].

Another approach discussed in [75] uses event trees to identify high risk critical $N-k$

contingencies for online security assessment. Here, failures in protection assemblies such as inadvertent tripping, failure of protective relay to trip, and breaker failure resulting in contingencies are considered as the basis for analysis. Event trees are generated based on the probability of these rare events and a graph search method is used to identify the critical contingencies. The identified contingencies are considered to cause higher damage than the usual $N-1$ contingency. Hence, these contingencies are then further evaluated to obtain the overall risk associated with the system.

The method in [76] presents a graph theoretical approach for identifying up to $N-3$ critical contingencies, where the nodes represent the power plants or substations, however, the edges represent transmission lines. The performance of the graph is measured using an expression for computing the efficiency of the overall graph. Now, considering the set of possible damages, each graph is evaluated for its efficiency and the damages that minimize the efficiency of the graph are considered to be the most critical damages. The method also provides a mechanism to choose a single edge that can be added to the graph in order to improve its efficiency. The resulting graph that maximizes the efficiency of the overall system is considered to be the optimum solution for improving the system resilience.

The drawback associated with this approach is that with increase in the value of k generation of the graphical topologies will become computationally expensive. Additionally, considering the non-linearity of the power system networks the analysis ignore the power flow equations that play an essential role in accurately obtaining the state of the system.

Complex network theory is used in [77] to identify the vulnerable lines in a power system. The theory is based on the same principles as discussed in [26]. However, in this model a new vulnerability index named weighted line betweenness is proposed. The new vulnerability index provides the capability to identify the most critical lines and also identify other transmission lines that are not heavily loaded but are critical due to their position in the network. Another study based on the small world model in [78] uses similar betweenness index that is based on the reactance of the power transmission lines.

A concept of delta centrality is used in [79] to identify critical $N-k$ contingencies. First, the power network is modeled as a graph with nodes and edges. A transmission line that is connected between two nodes is represented by a matrix containing a unique number at its appropriate location in the matrix. This number represents the loading of the transmission line. Now, based on the obtained graph the node and line centralities of the graph are computed that identifies the critical nodes and lines that can cause severe cascading failures when removed from the network. These node and line centralities are computed using the expressions described in [79]. However, these centralities need to be computed every time the topology change occurs for the network.

The main drawback of the approaches based on the small world model can be their computational complexity with increase in the system size as it will require modeling the power network into graphical network and obtaining the weighted line betweenness, i.e., the distribution of the number of transmission lines passing through a given node, shortest electric path lengths, etc.

3.3.4 Load Outage Distribution Factors (LODFs) Based Methods

In LODF based methods the line outage distribution factors are used as a metric to identify the severity of a line outage on the rest of the network. Formally, as discussed in [80], "Line outage distribution factors (LODFs) are the linear sensitivities of line flows to a line outage". This method has the advantage of bypassing the load flow solution to identify the critical contingencies; however, it uses the LODFs of the lines that provide a metric about the impact of removal of a transmission line on the other lines in the power network.

In this study, the LODFs metric for single line and double outage contingencies are studied. The LODFs metric seems to provide unreasonable results for double line outages due to the is-landing phenomenon. Hence, a new metric for double line outage contingencies is discussed in [80] that takes into account the islanding effect and provides an effective solution.

The main drawback of the study is that the line outage distribution factors (LODFs) reduces its value when the distance increases with respect to the outaged line. Moreover, the LODFs expressions are only developed for single and double line outages. Hence, for all higher order contingencies a new LODF metric need to be developed which might become computationally expensive as 'k' increases.

Another study considered pre-screening contingencies in order to tackle the problem of combinatorial search for identifying critical $N-k$ contingencies and is discussed in [81]. Here, two algorithms are developed for identifying critical $N-2$ contingencies. The algorithms are called impact tracking structure (ITS) algorithm and overload tracking structure (OTS) algorithm. These algorithms have their own strengths but both of them make use of the linear line sensitivities. These sensitivities are obtained using the same concept of line outage distribution factors (LODFs). Due to the linear approximation the study have made several assumptions that might not be entirely effective. Moreover, the approach only provide expressions to screen and identify critical $N-2$ contingencies. However, there are no expressions for identifying higher order contingencies that do occur and cause severe cascading failures.

For $N-2$ contingency analysis, a fast algorithm is described in [9]. It selects the contingencies that causes thermal limit violations and uses DC approximation. The approach is based on the iterative pruning of the candidate contingency set which means that the contingency pairs that are considered to be safe are already pruned from the total list of contingencies that needs to be evaluated. This pruning is based on calculating the line outage distribution factors (LODFs) in the $N-1$ contingency analysis stage and using it for $N-2$ analysis. The resulting list of contingencies that needs to be evaluated for the cascading failure can be analyzed with minimum computational effort.

Further, the study in [82] builds on top of the work discussed in [9]. The approach identify the frequencies of the lines that cause constraint violations either as initiating pair or overloaded lines and statistically categorize them. In addition, it provides the correlation

in power flows between the initially removed and overloaded lines. The main drawback of the two studies is that contingency screening expressions are developed for only up to $N-2$ contingency analysis, however, higher order contingencies do occur and need to be identified too in order to improve overall system resilience.

3.3.5 Random Chemistry Method

Several methods are based on the screening and optimization methods that are constructed on the limit violations, however, random chemistry method in [7] suggests that these methods alone are not sufficient for analyzing cascading failures.

An approach is discussed in [7], where a large random set of components are selected from the entire set of components comprising the power system. This set of components also known as large, non-minimal set is obtained if the elements of the set cause a system failure when outaged from the entire power system network. The identified non-minimal set is then reduced by a constant fraction and some random subsets of the non-minimal set are obtained with size not less than the desired $N-k$ contingency. These subsets are evaluated for the system failure criterion. The subsets that cause the system failure now become the new target set. Finally, the new target set is evaluated and pruned with individual element outages and the minimal $N-k$ contingency set is obtained. The process is repeated several times to obtain large collections of critical $N-k$ contingencies.

The approach was able to identify all the critical $N-K$ contingencies up to a 'k' value of 3. However, for higher 'k' values the number of identified critical contingencies did not reach saturation. Hence, experimental results show that many higher order $N-k$ contingencies were not obtained using the approach.

3.4 Cyber-Physical Attacks Based Methods

Reliable operation of power systems is of paramount importance for the socio-economic welfare of the society. Due to the technological transformation of the traditional power

grids into smart grids, power systems employ a large number of sophisticated and autonomous components such as protection devices (distance relays, over-current relays and circuit breakers), phasor measurement units (PMUs), advanced metering instruments (AMIs) etc., that can be remotely monitored and controlled [4]. Although, necessary for meeting the future demands of the electric power, this technological advancement increases the potential attack surfaces by giving rise to new vulnerabilities [3]. For example, substations within the power system network are equipped with the remote terminal units (RTUs) that are responsible for the remote monitoring and control of the power grid. These RTUs communicate with the power system components over a local network connection which can be accessed through external means by compromising the network firewalls. This further increases the system vulnerabilities. Once an RTU is compromised the attacker may gain complete control of the substation and can cause severe damage. Hence, malicious attackers take advantage of these vulnerabilities and launch catastrophic attacks on the power networks. For instance, the recent 2015 cyber-attack in Ukraine, where the attackers stole the credentials of operators, gained complete access to the substations of the utility companies and tripped several power lines that resulted in a severe load loss [60].

Recent studies by the National Research Council documented that malicious attacks on the power grid are much more devastating than the destructions caused by natural calamities such as hurricanes etc., [83]. According to the study, cyber-attacks could result in large blackouts that can render a significant portion of the country without power even for months [83]. Moreover, these attacks can be initiated through cyber penetration [84] or physical sabotages [85]. In the recent years, cyber-attacks have been increasing both in number as well as sophistication and are considered as one of the major obstacles towards the reliable power system operations [86, 10, 11, 87, 88]. As a result, improving power system resilience considering cyber-security has gained significant attention [89]. Additionally, cyber-attacks raise new challenges for the power system reliability [90]. Several cyber-physical attack based methods have been developed in order to improve the power system

resilience. These attacks can be of two types namely; static or simultaneous attacks and dynamic attacks.

3.4.1 False Data Injection Based Models for Static or Simultaneous Cyber-Attacks

Static or simultaneous cyber-attacks are referred to as those types of attacks that take place at the same time. Several cyber-attack models are developed to formulate and analyze the simultaneous cyber-attacks in power systems.

A data integrity attack model is discussed in [91] and its impact on the voltage control loop is analyzed by identifying the Flexible AC Transmission Systems (FACTS) devices that can be targeted for a larger damage in the system. The attack model is based on [92], however, it is extended to consider the voltage control loop in the power system network. Two types of attacks are modeled in [92], namely integrity attacks and denial of service attacks. The former attack means when the values of either the measurement data or the control signal is manipulated. However, the latter means when the control signal is not allowed to reach the required destination so that the appropriate device can perform the necessary operation.

The FACTS devices are responsible for controlling the bus voltages by either inducing or supplying the reactive power. Based on the discussed model in [91], the attackers can manipulate the measured bus voltage or the control signals to the FACTS device to cause damage to the power network. The study provides an analysis on how the attack on a FACTS device can affect the nearby buses and to cause a significant impact on the load bus, the FACTS device connected to that particular bus needs to be affected.

Another study in [93] models cyber-attacks in SVC and STATCOM devices, where the attacks on the communication link between these devices are launched. These attacks can manipulate the sensor and controller values to cause system instability. The model performs transient stability on power system post cyber-attack and developed two indices, namely; a voltage and an angle stability index to evaluate power system stability.

In both these studies [91, 93], the main drawback is that the mechanisms for mitigation along with a system wide identification of highly damaging attacks are not presented in the study that could improve the overall power system resilience.

Based on several techniques on detecting bad data measurement, it was suggested that the difference between the bad data measurement and the nominal data is significant and can be easily detected. However, there is a new class of vulnerability known as the false data injection attack discussed in [94], where the authors proposed that if the measurements from the meters are tweaked a little bit then the malicious data remains undetected using the detection algorithms. In [94], false data injection attacks are modeled against state estimation in power systems. According to the study, these attacks cause arbitrary errors in the state variables that lead to erroneous state estimation by the operators. This can lead to undesired operator actions that can cause severe system damage. The study does not provide any anomaly detection mechanisms to defend against the proposed vulnerability and performs the analysis based on DC power flow models that may not yield the exact power flow solution.

A special type of false data injection attacks known as Load Redistribution (LR) attacks are discussed in [95, 96], where the line flows and power flow at only the load buses can be manipulated. This leads to an incorrect state estimation that could cause the security constrained economic dispatch (SCED) to take inappropriate generator re-dispatch actions to reduce the overall system operational cost. This could result in either immediate or delayed load shedding. The work in [96] models it as an LR attack problem and solves it as a KKT-based method and duality-based methods. The approach identifies the most damaging attacks in order to employ defense strategies to improve the power system resilience. The main drawback of the approach is that it does not use the power flow solutions and do not employ the defense strategies for the attack models.

An approach is presented in [97] that considers two types of sparse malicious false data injection attacks, i.e., random and targeted attacks. These False Data Injection Attacks

(FDIAs) are considered only in the measurement equipments that are used for state estimation. In random attacks the data from any measurement is manipulated randomly. However, in targeted attacks the data from a selected subset of measurement equipments are modified. An attack construction model is presented that generates the malicious attacks and a greedy algorithm is provided to identify the subset of measurements to be protected to defend against these attacks. This algorithm is much more effective considering the time complexity of the brute force search algorithm. Moreover, an attack detection model is developed to identify the bad data to prevent the incorrect state estimation of the power grid. The main drawback of the approach is that it considers attacks on only the measurement equipments and all the attacks take place at the same time. However, sequential cyber-attacks can be more damaging [14].

3.4.2 Topology Based Models for Static or Simultaneous Cyber-Attacks

A topology based approach, where the load distribution vector (LDV) based attack strategies are modeled is discussed in [98]. Here, based on the load distribution vector two attack strategies namely; load distribution vector based multi-node attack and load distribution vector based multi-link are proposed. The multi-node attack refers to the attacks on the nodes such as buses etc., that takes place at the same time but at different points in the network. However, the multi-link attack refers to the attacks on the transmission lines that takes place at the same time but at different points in the network. The LDV is constructed based on the loads on the nodes and links after the removal of the selected nodes or links. In this work, the failure of links and nodes are analyzed separately, however, in reality both these can be attacked simultaneously and should be analyzed together. It is necessary to include the power flow models along with topological analysis in order to perform a concrete vulnerability analysis.

Topology based approaches have also been studied with respect to cyber-attacks causing cascading failures in power systems. The approach discussed in [99] proposes a *risk graph*

based technique to identify critical nodes, i.e., substations to attack. In this study, a *risk graph* is a metric that reflects the hidden relationship between substations with respect to vulnerability and nodes/links represents substations/transmission lines respectively. The approach proposes two node selection strategies namely; sub-optimal node attack strategy and risk graph based node attack strategy. The nodes with the top most largest nodes are selected for the attack according to the sub-optimal node attack strategy. However, this strategy is not suitable for real-time attack and needs to compute the system tolerances which are infeasible. Therefore, a risk graph based attack node attack strategy is developed. It constructs the risk graphs and obtains the average risk graphs (ARG) based on different system tolerances. The ARG is then used to identify the critical nodes or substations to be attacked simultaneously to cause severe cascading failures resulting in blackouts.

An extension to the work in [99] is presented in [100]. This model incorporates the topological as well as the electrical properties of the power system model. First, a node attack strategy known as the reduced search node attack strategy is developed to identify the target nodes. These target nodes are then used by the *risk graph* method to study the attack strategies, where it suggests that the load and degree based attacks are not the strongest attack strategies. However, using risk graph based attack strategy yields stronger attacks which do not require the attacker to have any knowledge about the system tolerance parameter.

Previous studies considered attacks either on substations or transmission lines. However, cyber-attacks in reality can occur both on substations and transmission lines. Hence, another model is developed in [101] that considers the joint substation-transmission line vulnerability assessment against the power grid. The approach can capture several new vulnerabilities related to joint attacks on substations and transmission line which were not studied before. A new attack strategy based on the component interdependency graph (CIG) is developed that claims to be more effective than the node or degree based attack strategy. This metric is similar to the *risk graph* based strategy discussed in [100]. How-

ever, RGs cannot describe the joint relationship between the substations and transmission lines. The CIG on the other hand is used to obtain this joint relationship between them and the resultant graph can be used to identify the attacks that can cause severe system damage.

The main drawbacks of the above studies are that they do not propose the defense mechanism for the proposed attack methods. Moreover, they use the DC power flow approximation rather than the AC load flow that provides a more realistic system state. Other than the study in [101], only substation attacks are considered in the study. However, transmission line attacks are much more frequent. Combining the two attacks can ultimately provide better insight on improving system resilience as indicated in [101]. Further, as the number of attack nodes increases the performance of the *risk graph* decreases and it does not provide the correct relationship between the node groups. The construction of CIG in [101] can be computationally expensive for larger power systems.

3.4.3 Time Synchronization Based Models for Static or Simultaneous Cyber-Attacks

In order to obtain an accurate system state, the collected measurements needs to be aligned in time domain. This process is known as time synchronized monitoring. To achieve this type of monitoring in power systems, GPS based time synchronization monitoring devices are deployed in the power grid monitoring system. However, the attackers can manipulate the timing information and cause the operators to take incorrect decisions. The work in [102] discusses Time Synchronization Attack (TSA), where the attackers modify the measurement data by spoofing GPS. A cross-layer TSA detection scheme to identify such attacks is also proposed.

3.4.4 Real-Time Models for Static or Simultaneous Cyber-Attacks

Real time models for cyber-attacks are necessary to analyze their exact impact on the power systems. In [103], a real-time cyber-attack model is developed that considers Denial of Service (DoS) attacks and Man-In-The-Middle (MITM) attacks. The approach uses Real

Time Digital Simulator (RTDS) and RSCAD for modeling the power systems and DeterLab for providing the communication capabilities in order to simulate such attacks. The system can be easily evaluated in real-time post attacks.

The model presented in [103] is extended in [104], where an additional type of cyber-attack, i.e., communication line outage attack along with the Dos and MITM attacks are considered. In this model the voltage stability of the power system is evaluated consider these attacks. The model uses a more realistic real-time simulation models as it incorporate various other simulators such as Network Simulator-3 (NS-3), etc. for analyzing the effect on power systems. Another online model for cyber-attacks is presented in [105] that focuses more on understanding the cyber-attacks on power systems.

A real-time cyber physical system testbed for power system security and control is presented in [106], where the attack model consists of removing transmission lines and evaluating its impact in terms of voltage stability and generation loss in the power network. The approach uses a real time system simulator, i.e., OPAL-RT [107]. It also uses the SEL 351S [108] as the protection system for power networks. In addition two mitigation strategies are developed for the attack model. The first strategy is to conduct offline analysis to restore the system or bring the system to its next steady state condition. The other mitigation strategy is based on the real-time adaptive control mechanism.

The main drawback of the above models [103, 104, 5, 106] is that it can only perform evaluation of the user-specified cyber-attacks and cannot perform an overall contingency analysis. In addition, the mitigation scheme considering the attack models are not discussed in these studies [103, 104].

3.4.5 Game Theoretic Based Models for Simultaneous or Static Attacks

The problem of identifying and protecting critical resources in power systems is modeled as a defender-attacker-defender problem also known as the min-max-min problem in [109]. Here, the attacks on transmission lines are the only attacks that are considered. The

defender first protects the required transmission lines before the attack is executed. Then, the attacker launches the attack by selecting a set of transmission lines to be removed from the network. Finally, depending upon the disruption the defender takes corrective actions to minimize the damage. However, the model does not consider compromising the substations in the power grid as that is the first step in accessing the control for switching the transmission lines. This will lead to a more realistic model and can provide more exact solutions. In addition, the approach uses a DC approximation solution rather than AC load flow solution that might affect the correctness of the solution.

Another approach towards the game-theoretic modeling of attacks is presented in [110], where data injection attacks on the automatic generation control (AGC) of the power systems are considered. The approach combines the quantitative risk management techniques with decision making on protective measures. The defender's loss post data injection attacks are estimated using a measure known as the conditional value-at-risk (CVaR) that is a measure of defender's loss due to load shed. The game is designed as an attacker-defender stochastic (Markov) security game. The defender obtains the solution by solving it using the dynamic programming techniques considering budget constraints. The model considers the system state, attacker's and defender's action spaces, their payoffs and the transition state after the attacker or defender actions. The approach suggests that further improvement in complexity can be made in order to achieve efficient convergence.

Further, coordinated attacks on power systems are studied in [111]. In this approach two attack models are presented. The first attack method considers the coordinated attack between the load redistribution (LR) attack and attacks on generators. However, the second attack model considers the coordination between LR attack and attacks on transmission lines. The attack problem is formulated as the bi-level optimization problem where the attacker tries to maximize the damage; however, the defender aims at minimizing the damage. The damage is measured in terms of load lost by the power system network after the attack is launched. The approach claims that the coordinated attacks causes more dam-

age than individual attacks. The main drawback of the approach is that it does not consider the temporal nature of the attacks where these attacks can take place at different instants in time and can causes higher damage as opposed to the simultaneous attacks.

3.4.6 Other Models for Static or Simultaneous Cyber-Attacks

A man-in-the-middle attack model is discussed in [112]. The study provides a platform where real-time attacks can be launched and their effects can be analyzed on the power network. The study provides the cyber-security vulnerabilities of Supervisory Control and Data Acquisition (SCADA) systems in power grid. The address resolution protocol (ARP) based man-in-the-middle attack is studied in this platform, where the attacker can spoof into the communication layer of the SCADA system and modify the measurements or control commands that would misguide the system operators to take wrong decisions. This would result in severe cascading failures resulting in blackouts. The main drawback of the approach is that it considers only one type of cyber-attack scenario. Moreover, the attack cases provided by the user can only be simulated and analyzed. However, the maximum damage causing attacks cannot be identified automatically in order to improve system resilience.

A power system consists of components that operate both in continuous and discrete time. The physical components such as the transmission lines, generators, etc. represent analog characteristics. However, cyber components such as monitoring and control represent discrete time characteristics. Therefore, to study the exact response of the cyber-attacks on the power system it is needed to have models that include such characteristics. In [113], one such model is developed that interfaces both these characteristics in power systems. A variable structure system theory [114] approach is used to model switching cyber-control and gain insight on the cyber-physical interaction in power systems. Moreover, the new vulnerabilities that are obvious based on such interactions are evaluated.

The work in [115] suggests that it is necessary to consider the temporal features of

attacks while analyzing cascading failures in power systems. The reason behind this approach is that cascading failure propagates in stages. The number of failed components increases drastically beyond a critical point. Therefore, from the total number of stages of the cascade progression, when the critical intermediate stages are identified the cascade can be prevented to cause a severe blackout. The approach can be applied to both random and targeted cyber-attacks. The method does not provide an overall metric to improve the system resilience nor does it provide the cyber-attack models.

3.4.7 Variable Structure Systems Theory Based Model for Dynamic Cyber-Attacks

The cyber-attacks that take place at different instants in time or follow an attack sequence in time are referred to as dynamic attacks. Dynamic attacks can be more devastating than the simultaneous attacks if scheduled strategically. There is not a lot of research available on these types of attacks. However, these attacks can be strategically realized and executed in real time. Moreover, some studies have shown the effectiveness of these attacks.

In [116], a method for the coordinated multi-switch attack for cascading failures in power grid is developed. The method employs a variable structure systems theory to model the attacks. Here, rather than targeting a single circuit breaker, based on the system state information an attacker can design a strategic sequence for targeting multiple circuit breakers that control the synchronous generator switching. The switching of these breakers can lead to the transient instability and cause loss of generation that can eventually result in severe cascading failures leading to power loss. The model provides the capability for single switch attack, concurrent switch attack, and progressive switching attacks. The main drawback of the method is that it does not provide an approach to identify the most damaging switching scenarios and rather provides a model to execute the attacks. Moreover, the defense method with respect to the attack strategies is not developed for minimizing the system damage.

3.4.8 Topology Based Models for Dynamic Cyber-Attacks

Another model in [14, 13] developed a sequential attack model to launch cyber-attack on power system networks. The model considers removal of transmission lines in sequences, i.e., at different instants in time and argues that these sequential attacks can be more damaging than the corresponding simultaneous attacks on the same transmission lines. The model uses the similar approach discussed in [99] to select the links to attack and remove from the power network. It provides the capability to unravel those vulnerabilities of the power system that are not obvious via the simultaneous attack models. The results demonstrated in the study supports that sequential attacks are stronger than the simultaneous attacks.

The model discussed in [14, 13] uses a sequential attack strategy for nodes and identifies the critical nodes using the construction of the *sequential attack graph* (SAG). A *sequential attack graph* (SAG) is a metric that identifies the combinations of vulnerable nodes and their failure order that can result in higher damage. This SAG provides the attacks that are stronger than the load and node based attack strategies. Moreover, the computational complexity is reduced while identifying critical nodes. The drawback associated with the approach is that there is no defense mechanism for minimizing the damage. In addition, construction of the SAG for larger power systems can be computationally infeasible. Hence, better mechanisms for constructing SAG are needed.

Another model analyzes cascading failure caused by node overloading in power system in [117], where the node or nodes of the power system are targeted in a way that it causes overloading of the other nodes. The attack model proposes an algorithm to find the nodes maximizing the total number of failed nodes. The attack takes place sequentially rather than simultaneously. The main drawback of the approach is that only node overloading is considered and no mitigation strategy is proposed against the developed attack model.

Chapter 4

Cyber Induced Fault based Modeling and Analysis Methodology

4.1 Problem Statement

Consider a power system model that consists of generators, buses, transmission lines, measurement and monitoring devices, transformers, loads, protection equipments such as distance relays, etc. Fault in the power carrying components such as transmission lines, etc. in the power system network can cause cascading failures. These failures are the secondary effects of the load redistribution caused by line outages. Therefore, analyzing these cascading failures is very essential to improve the robustness of the entire power system. There are two ways one can perform the analysis:

- By considering models without cyber-effects that can provide the contingency to the power system cascade analysis model *prior to the start of the simulation* and evaluate its effect on the system to develop an effective mitigation strategy or to perform system upgrades to increase the robustness of the system.

- By considering models with cyber-effects that can provide the contingencies to the power system cascade analysis model *during the simulation* and evaluate its effect on the system for improving the resilience of the system.

Considering only physical contingencies *prior to the start of simulation* greatly limits the cascade analysis process. It greatly restricts the search space and a large number of cascading traces can not be identified and analyzed which in turn results in developing a weaker resilience metric. Therefore, we have developed an analysis model that provides the capability for introducing both physical faults and cyber-effects *prior to the start of simulation* and even *during the simulation* at any instant in time. The reason behind the need for

the latter capability in the analysis model is that when faults occur at different instants in time the entire cascade progression changes due to the non-linear nature of the power system networks. It could lead to a far worse cascading failure trace that can result in severe system damage.

To achieve this, we have designed detailed behavioral models of the protection devices, i.e., distance relays, over-current relays, and circuit breakers. These models represent the exact behaviors of the protection devices under *nominal* and *faulty modes of operation*. The models are developed using state machines and takes into account cyber-effects and time causality of the events. The developed protection assembly behavioral models are integrated with our cascade analysis model to provide a framework for more complex analysis. These behavioral models are used as part of a simple cascade simulation and contingency analysis model to study the evolution of cascades in the presence of physical faults and cyber-effects.

We have evaluated our developed solution by:

- First showing that the developed framework and protection assembly models provide realistic behavior under both nominal and faulty modes of system operation.
- Next, we used simulations to show how our model can take into account both physical faults and cyber-effects in order to support more complex power system analysis.
- Finally, we provide simulation results that clearly reveal how we are able to find new cascade progression traces using our framework by initiating physical faults and cyber-effects at different time instants.

In addition, we have also utilized the techniques of distributed computing to optimize the process of contingency analysis. While applying distributed computing mechanisms, each core will use the base model of the power system to perform the analysis depending upon the desired fault configuration and store the results in a separate location. Once

the entire set of contingencies is evaluated the results can be gathered and analyzed for improving system resilience.

This is based on the accepted paper in PES-Innovative Smart Grid and Technology (ISGT) Conference. The details of the publication is as below:

Hasan, Saqib, Ajay Chhokra, Abhishek Dubey, Nagabhushan Mahadevan, Gabor Karsai, Rishabh Jain, and Srdjan Lukic. "A simulation testbed for cascade analysis." In Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), 2017 IEEE, pp. 1-5. IEEE, 2017.

4.2 Introduction

Electrical power systems are heavily instrumented with protection devices whose primary responsibility is to identify and isolate faulty physical components from the power system network as per deterministic protection schemes. While these devices act on local information i.e. branch power flows and bus voltages to quickly arrest the fault propagation, the lack of a system-wide perspective could lead to cascading failures. Additionally, failures or mis-operations in the protection devices (referred to *cyber faults* in this paper) can affect the nominal behavior of the relay and/or breakers and can contribute towards cascade progression leading to blackouts as seen in Aug 2003 USA[58], 2003 Italian[118] blackouts. For instance, in the IEEE 14 bus system shown in Figure 4.1, outage of line L1_5 due to physical fault (three phase to ground fault) may not cause any further failures in the system. However, presence of an additional fault in an associated protection device (stuck breaker fault in circuit breaker PA12) will lead to a cascading failure tripping all current carrying paths to the affected line. Each protection device consists of a distance relay, an over-current relay, and a circuit breaker. The formal description is given in Section 4.5. This can cause further disturbance to the system in the form of overloads and can contribute towards cascade progression. Hence it is important to understand the unintended consequence of protection assembly failures and include these in cascading failure studies.

In order to diagnose and predict cascade evolution in a better way and to perform contingency analysis, its important for the simulation models to consider the behavior models of these discrete devices with reasonable timing accuracy. These models should be able to emulate the behavior of actual hardware in both nominal and faulty modes and allow the ability to alter the model parameters, injection of missed or spurious detection faults, modification of response delays and threshold values.

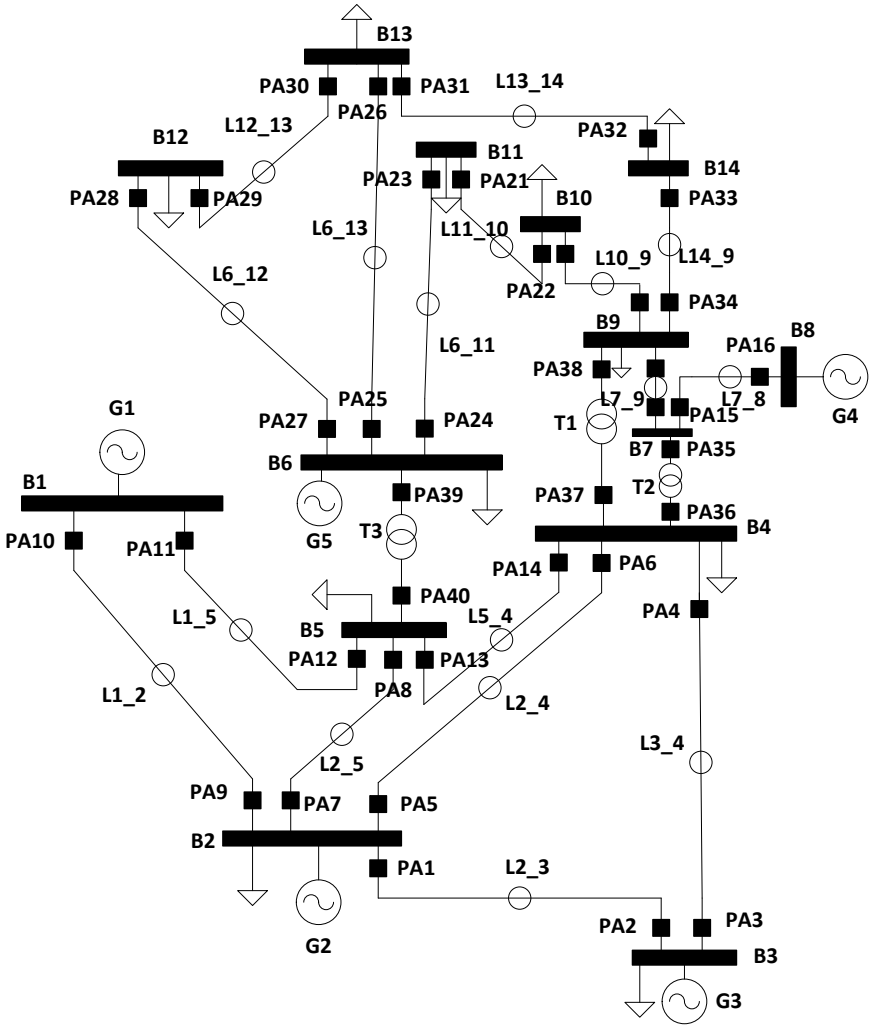


Figure 4.1: IEEE 14 Bus System[1]

Existing approaches for cascading failure analysis are to perform off-line simulations to assess the current state of power system and study its evolution using different cascade

simulation models [34],[19],[8],[31],[119],[20]. Models referenced in [34],[8],[31], [119], are based on initiating faults that cause line overloads leading to cascading failures in the system but they do not consider the interaction of cyber failures in protection devices. Models in [19],[20] considers faults in protection assembly in the form of hidden failures or sympathetic tripping. But this greatly limits the cascade evolution paths as this tripping is possible only in the lines which are connected to the same bus as the line outage fault. Moreover in all these models time causality of the events is not considered. This can be very useful in initiating a failure at any desired instant, that can change the cascade evolution path as well as in analyzing the effect of a particular fault in cascade progression. Time is also helpful for the operators in detailed cascade analysis and designing better mitigation strategies. Taking these cyber failures and time causality of events into account cascade progression will evolve in a different way, which cannot be studied based on above models but is possible via this approach.

The approach presented in this paper uses detailed behavioral model of the protection devices (distance relays, over current relays and breakers) in nominal and faulty modes of operation, taking into account *cyber faults* and time causality of the events. The behavioral models are used as part of a simple cascade simulation and contingency analysis framework to study the evolution of cascades in the presence of *cyber faults*. The results of such an analysis presents new new cascade evolution trajectory leading to blackout, which are otherwise not obvious. An example is shown with a case-study of IEEE 14 bus system.

The paper is organized as follows: Section II discusses the detailed explanation of distance relay, over current relay and circuit breaker behavioral models. Section III describes cascade simulation model and proposes a new approach of contingency analysis that involves behavioral models. Experimental setup and system under test is discussed in Section IV. The results are listed in Section V followed by the conclusion in Section VI.

Table 4.1: Protection Assembly- Parameters Description

Parameter Name	Description
Distance Relay	*Common parameters for over-current relay
F_de1*/∼F_de1*	Presence/Absence of Missed Detection Fault
F_de2_zX/∼F_de2_zX (X =1,2,3)	Presence/Absence of a zone1, zone2, zone3-Spurious Detection Fault
V, I*	3 phase bus voltages and line currents
R, L, Len	Resistance, inductance and length of the transmission line
RelayTrip	POTT scheme relay trip command reception
c_reset	Resets the relay to 'idle' state
Trip*	Relay status to disconnect the branch
Z1, Z2, Z3	Presence of zone1, zone2, zone3 faults
RelayTrip_	POTT scheme relay trip command issue
cmd_open*/cmd_close*	Open/Close command to circuit breaker
ZxWT(x=2,3)	zone2, zone3 wait times
Circuit Breaker	
F_stuck_open, F_stuck_close	Presence/Absence of Stuck open and Stuck close Faults (Stuck Faults).
cmd_open/cmd_close	Open/Close command to physical breaker
PhysicalStatus	Open/Close status of physical circuit breaker
Trip	Circuit breaker Open/Close command
st_open/st_close	Open/Close status of the circuit breaker
Over-Current Relay	
F_de2_Px/∼F_de2_Px (x=1,2,3)	Presence/Absence of high, medium and low overloads-Spurious Detection Fault
P1_OL, P2_OL, P3_OL	Presence of High, Medium, Low overloads
CThres	Max. loading value of the branch
ZoneWaitTime	Wait time for the relay

4.3 Protection Assembly Behavioral Model

The devices considered as part of the protection assembly in this work include distance relays, over-current relays and circuit breakers. The relays detect the fault conditions (reduction in impedance, increase in current) and command the breaker to open. The breakers respond to the command and open the circuit, thereby arresting the failure propagation. This nominal operation of the protection devices is affected in the presence of *cyber faults*. The behavioral models consider three types of *cyber faults* namely *Missed Detection Faults*,

Spurious Detection Faults and *Stuck Breaker Faults*. As the names suggest, in the presence of a *Missed Detection Fault*, a relay fails to detect the anomaly. As a result, the breaker is not commaded to open and arrest the failure propagation. In case of a *Spurious Detection Fault*, a relay incorrectly reports the presence of an anomaly (under nominal conditions) and subsequently commands the breaker to open. With a *Stuck Breaker Fault*, a breaker does not operate as commanded i.e. to open or close and continue to remain in their current state.

4.3.1 Distance Relay:

A distance relay is used as the primary protection in electrical power transmission systems. Its behavioral model (Figure 4.2) is designed using Matlab/Stateflow [51]. Table 4.1 shows the details about the parameters used in its modeling. Three zone reaches (zone1, zone2, zone3) are modeled in the distance relay behavioral model (Figure 4.2), which are represented by states 'chkZx' (where x=1, 2, 3 for zone1, zone2 and zone3 respectively). These zones mark the protection zones of the transmission line as per reference[120].

4.3.1.1 Normal mode operation:

During normal operation, the distance relay remains in 'idle' state when the load impedance seen by the relay is nominal. The load impedance seen by the relay is computed based on a simple detection algorithm (dl(V,I,R,L,Len)) referenced in [120],[121]. When the relay sees a drop in impedance (probably due to a physical fault such as three phase to ground fault in a transmission line), it transitions out of 'idle' state.

When the impedance falls in the zone1 reach, the relay transitions immediately from 'idle' to 'Tripped' state and sends a 'cmd_open' to its associated circuit breaker. However, if the impedance falls in zone2 or zone3 regions, the relay transitions from its 'chkZx'(x=2, 3) state to the 'waitingX' (X= 1, 2) state after the wait time for its respective zone is elapsed. These wait times are external parameters, which can be set by the user. If fault gets cleared

while the distance relay is in the 'waitingX' (X= 1, 2) state, it transitions back to the 'idle' state. However, if fault persists, the relay transitions to the 'Tripped' state and sends the 'cmd_open' to the circuit breaker.

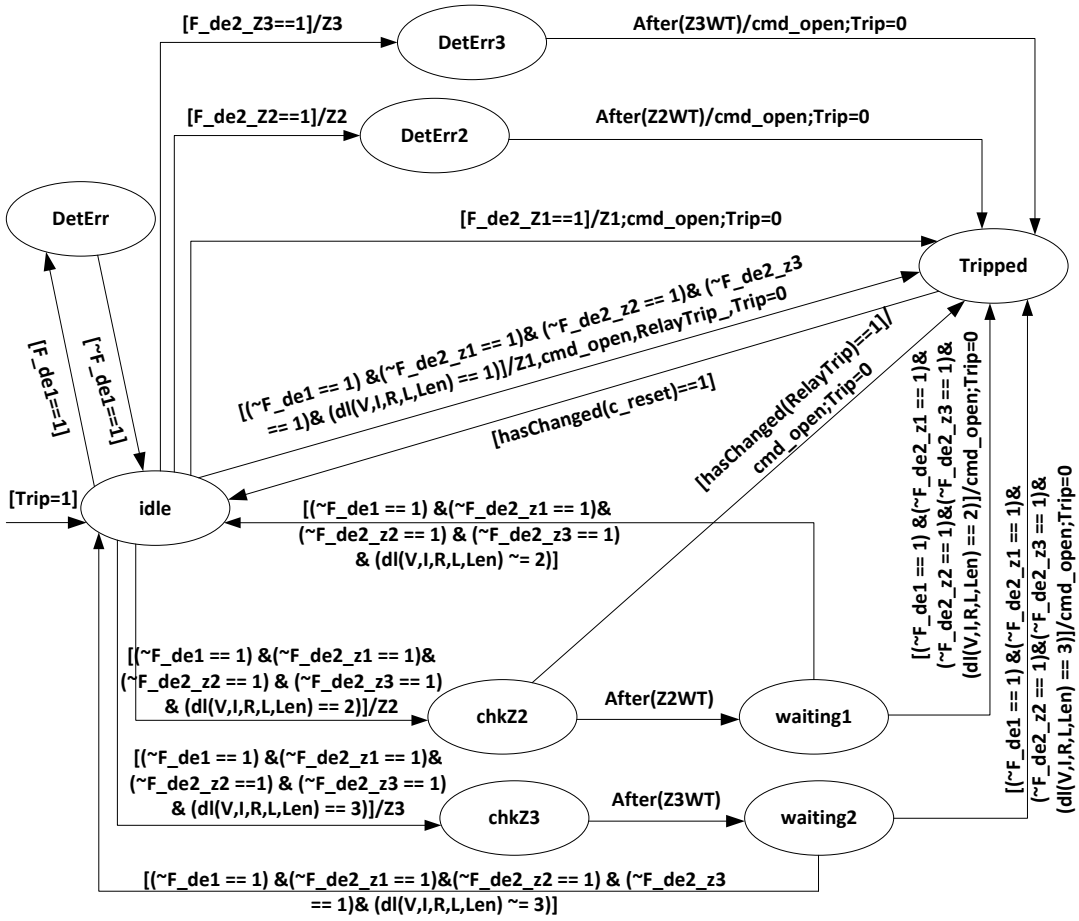


Figure 4.2: Distance Relay Stateflow Behavioral Model.

4.3.1.2 Operation under cyber faults:

In case there is a *Missed detection Fault* while the relay is in 'idle' state (Figure 4.2), it transitions to the 'DetErr' state resulting in no detection even though there might be an active zone fault. The relay will transition back to its 'idle' state once the fault is cleared. In the presence of *Spurious Detection Fault*, the relay incorrectly detects a fault and transitions from 'idle' state to the 'DetErrX'(where X=2,3) state and then transitions to the 'Tripped'

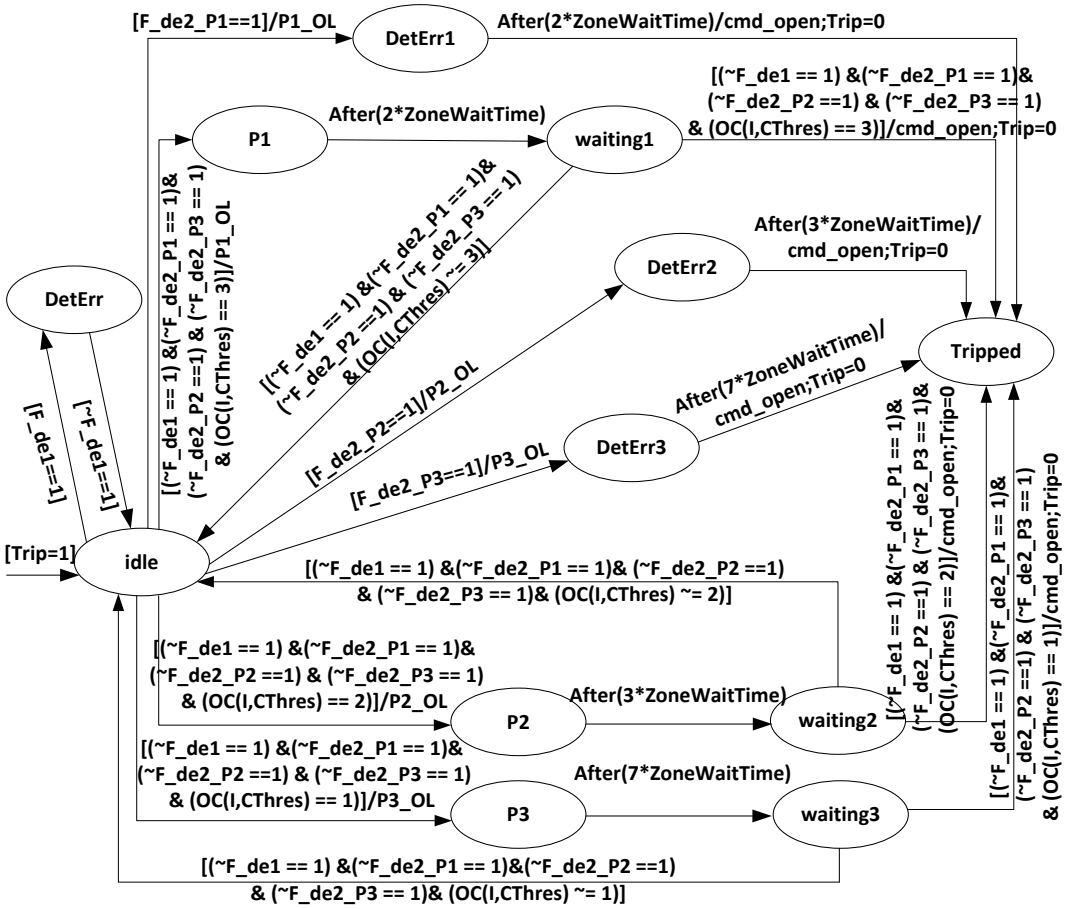


Figure 4.3: Over-Current Relay Stateflow Behavioral Model.

state based on the zone2 and zone 3 wait times. In case of a zone 1 *Spurious Detection Fault*, the relay immediately transitions from 'idle' state to the 'Tripped' state.

4.3.2 Over-Current Relay:

An over-Current relay is used as a backup protection in electrical power systems. Its behavioral model is shown in Figure 4.3 and parameters used for modeling are listed in Table 4.1. An inverse-time over-current relay is modeled for handling different amounts of overloads. These overloads are classified as high, medium and low overloads represented by states 'Px' (where x=1,2,3). There is a wait time associated with each overload, high overload having the least wait time and low overload having the longest wait time.

4.3.2.1 Normal mode operation:

During normal operation, the relay remains in the 'idle' state (Figure 4.3). However, if there is an overload condition, the relay transitions from 'idle' state to its 'Px' state (where x=1 to 3), depending on the amount of overload. These transitions are based on a simple detection algorithm (OC(I,CThres)) used for sensing overloads [122]. Being in one of the 'Px' states, the relay transitions to its 'waitingX' (X =1 to 3) state after the wait time associated with the overload elapses. If overload persists, the relay transitions to the 'Tripped' state sending a 'cmd_open' to the circuit breaker. Otherwise, the relay transitions to the 'idle' state.

4.3.2.2 Operation under cyber faults:

In case of *Spurious Detection Fault* and *Missed Detection Fault*, the over-current relay behavior is similar to the distance relay.

4.3.3 Circuit Breaker:

The circuit breaker behavioral model is designed using Matlab/Stateflow (Figure 4.4) and Table 4.1 shows the details about the parameters in its modeling.

4.3.3.1 Normal mode operation:

Under normal operation, the circuit breaker remains in 'closed' state. However, if it receives a 'cmd_open', the circuit breaker transitions from 'closed' state to the 'opening' state. Circuit breaker being a mechanical device takes time to open/close. Hence, we introduced a delay in the opening/closing operations of the circuit breaker for more realistic behavior. This delay is provided by the variables tto/ttc in the model. After the delay has elapsed it transitions from the 'opening' state to the 'wait_open' state and then transitions to the 'open' state indicating the status of the circuit breaker (as 'open') using the event

'st_open'. Similar transitions takes place if the circuit breaker receives a 'cmd_close' while being in the 'open' state.

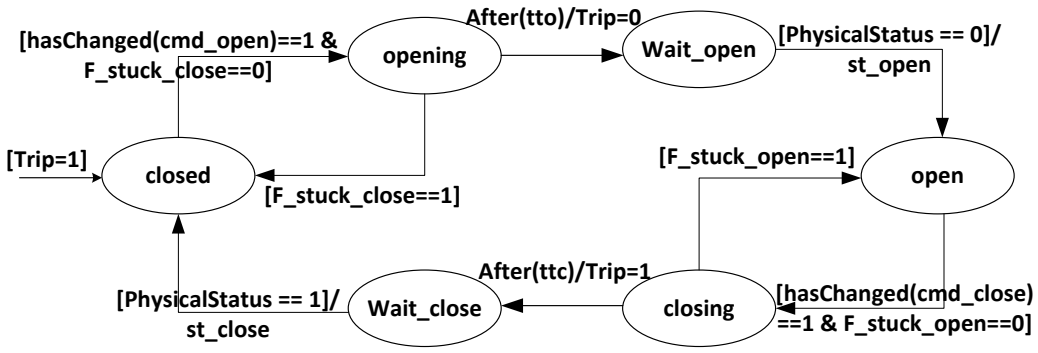


Figure 4.4: Circuit Breaker Stateflow Behavioral Model.

4.3.3.2 Operation under cyber faults:

If the circuit breaker is in 'closed' state and there is a *Stuck Close Fault* then it remains in the 'closed' state. However, if the same fault occurs while the circuit breaker is in the 'opening' state then it transitions back to the 'closed' state. Similar behavior is observed for the *Stuck Open Fault* as shown in Figure 4.4.

4.4 Towards Contingency Analysis

Contingency analysis in electrical power transmission systems is necessary to identify those critical sets, which can cause cascading failures and eventually lead to blackout. By critical set, we mean outage of those components that initiate the cascading failure. Tools such as MATCASC [123], CASCADE model [34] perform cascade analysis but they do not consider details about the time between contingencies and cyber faults in the protection equipments.

In our simulation and contingency analysis framework, we integrate the power transmission system simulation models in Matlab/ Simulink with detailed behavioral models of

protection assembly. In the phasor mode of simulation, we are able to capture the time between occurrences of different events in a contingency and also trigger cyber fault(s) in specific protection devices at specified time(s). The analysis allows us to identify contingencies which can possibly result in severe cascading outages or blackouts.

The proposed contingency analysis model is shown in Figure 4.5(a). The inputs to the analysis framework include the initial components outage (k-components outage) set, cascade simulation model and the protection assembly blocks. The initial component outage set is a initial list of components that are supposed to fail or have faults. An initial contingency can be a combination of physical and cyber faults. The protection assembly blocks will contain information about the cyber faults based on the initial component outage set. A simscape model (described later) of the power transmission system is executed taking into account the faults associated with the initial contingency set and the simulation is executed to evaluate the cascade progression through cascade simulation model. This simulation model is based on a simple cascade progression algorithm as shown in Figure 4.5(b). After the initial contingency, the system is checked for overloads and if it exists, the overloaded branches (transmission lines and transformers) are identified and tripped. The simulation is repeated to identify and trip new sets of overloaded branches. The process is repeated until there are no more overloads to trip or a blackout criteria is reached. If the blackout criteria is satisfied then the contingency is marked as the one causing 'Blackout'. Otherwise, if the blackout criteria is not satisfied and there is no further overload then the contingency is considered as 'Safe'. Currently, amount of load loss is considered as the blackout criteria in this model as referenced in [44] but it can be extended by taking into account other blackout criterion as well. At the end of the contingency analysis, a N-k ($k \in \mathbb{N}$) contingency set that can cause blackouts is identified and reported. The N-k contingency set contains the individual combinations of those initial component outages which can lead to a blackout.

The cascade analysis framework also has a feature of introducing random outages at specific times during the simulation. This could be of interest as it could reveal differ-

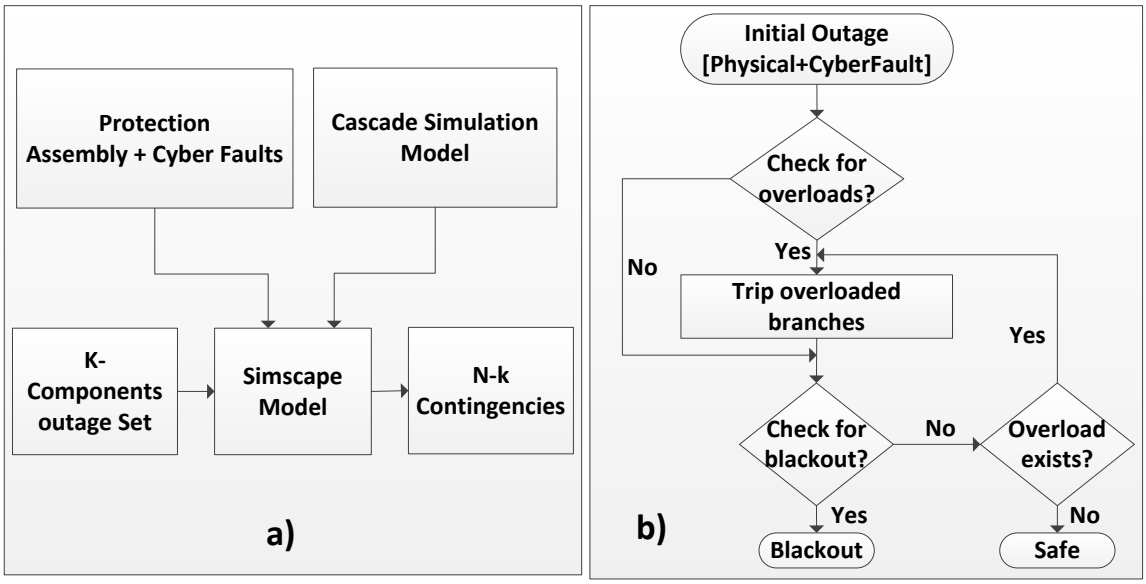


Figure 4.5: a) Contingency Analysis Model b) Cascade Flowchart

ent cascade evolution trajectory and possible blackouts due to changes in system topology. Also, the same outage when triggered at different times during the progression can contribute in finding those specific points where it is highly disruptive. This type of analysis is not possible with tools where outage can be specified only as part of the initial outage set. In our tool set, currently the random injection of faults is triggered manually. Automating it is left for future work.

4.5 System Under test and Experimental Setup

The proposed contingency analysis has been performed on an exemplar IEEE-14 Bus System[1] shown in Figure 4.1. The base voltage is 138 kV and length of each line is 16 km. The system is modeled in Matlab/Simscape using Simscape library blocks. Figure 4.6 shows the Simulink/ Simscape model corresponding to the transmission line 'L2_3 in IEEE 14 bus system (Figure 4.1), its associated bus and protection assemblies.

As shown in Figure 4.6, the transmission line is broken down into segments in-order to introduce faults at different line lengths. It is protected by a pair of protection assembly

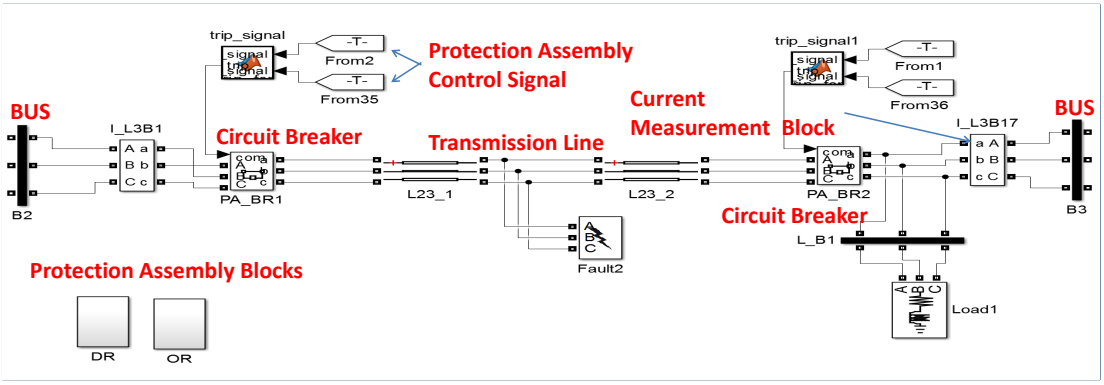


Figure 4.6: Portion of IEEE-14 Bus System- Simscape Model

on each side, which is denoted by PAn ($n \in \mathbb{N}$). Each protection assembly includes a Distance relay (PA_DRn), over-current relay (PA_ORn), and circuit breaker ((PA_BRn). The protection assembly is modeled as a separate subsystem therefore only the circuit breakers are shown at each end of the line (Figure 4.6). They receive control signals from the protection assembly subsystem. Current measurement takes place at the current measurement blocks and the voltage measurement happens at the bus. Generators are modeled as voltage sources with required base kV and MVA ratings and the loads are modeled as the constant PQ type loads. A Power GUI block is required to run the system in different modes namely phasor, discrete and continuous mode. We run the system in phasor mode for our analysis.

4.6 Results

The study is done on IEEE-14 bus system assuming that the lines are loaded at 70% of their loading capacity. It shows, how presence of cyber fault along with physical fault can lead to severe cascading failures causing blackouts and how it can be used in finding N-k contingencies which are otherwise not obvious.

- **Case 1:** At time t=0.5 sec, an initial contingency (a three phase to ground fault) occurs in the transmission line 'L3_4'(in Figure 4.1). A zone 1 fault is detected by the protection assembly 'PA_DR3', 'PA_DR4' and the fault is cleared by sending a command open

('cmd_open') to trip the circuit breakers ('PA_BR3', 'PA_BR4'). In the absence of any cyber fault, outage of transmission line 'L3_4' did not cause any further contingency and the system remained stable.

Table 4.2: Sequence of cascading events

Time(sec)	Event Description
0.500	**F:** 3ϕ-G fault- Line L3_4, Stuck close fault- PA_BR4.
0.501	**D:** Z1, Z3 in PA_DR{3,4}, PA_DR1, 'P1_OL' in PA_OR3, 'P2_OL' in PA_OR{5,1,13}, 'P3_OL' in PA_OR{9,15,21}. **CR:** 'cmd_open' in PA_BR3.
0.532	**S:** st_open-PA_BR3 is opened. **L:** Line L3_4 tripped partially.
2.000	**F:** Spurious detection fault in PA_DR27. **CS/CR:** 'cmd_open' in PA_DR27/PA_BR27.
2.031	**S:** 'st_open'-PA_BR27 is opened. **L:** Line L6_12 is removed.
3.503	**D:** 'P2_OL' in PA_OR13. **CS/CR:** 'cmd_open' in PA_OR{5,21}/PA_BR{5,21}.
3.534	**D:** 'P2_OL' in PA_OR31. **S:** 'st_open'- PA_BR{5,21} are opened. **L:** Lines L2_4, L11_10 removed.
5.505	**CS/CR:** 'cmd_open' in PA_OR13/PA_BR13.
5.536	**D:** 'P1_OL' in PA_OR{25,33}, 'P2_OL' in PA_OR {35,40}, 'P3_OL' in PA_OR{29,37}. **S:** 'st_open'-PA_BR13 is opened. **L:** Line L5_4 is disconnected.
6.536	**D:** 'P1_OL' in PA_OR31.
7.503	**CS/CR:** 'cmd_open' in PA_OR15/PA_BR15.
7.534	**S:** 'st_open'-PA_BR15 is opened. **L:** Line L7_8 is removed.
7.538	**CS/CR:** 'cmd_open' in PA_OR{25,33}/PA_BR{25,33}.
7.569	**D:** 'P3_OL' in PA_OR1. **S:** 'st_open'- PA_BR{25,33} are opened. **L:** Lines L6_13, L14_9 are removed.
14.571	**CS/CR:** 'cmd_open' in PA_OR1/PA_BR1.
14.602	**S:** 'st_open'- PA_BR1 is opened. **L:** Line L2_3 is tripped.

F: Occurrence of fault events, **D:** Detection of zone faults and overloads, **CS/CR:** Send/Receive commands from relays to circuit breakers, **S:** Status of the circuit breakers, **L:** Outage of lines.

- **Case 2:** The fault scenario in case 1 is repeated. A cyber fault (*Stuck close Fault*) is introduced in circuit breaker ('PA_BR4') of protection assembly PA4 (in Figure 4.1) in addition to the physical fault in line 'L3_4' at time t=0.5 sec. As a result of these initial faults, it is observed that a number of transmission lines gets overloaded and are eventually tripped and removed from the network. At time t=2 sec, another cyber fault (*Spurious Detection Fault*) occurred in the distance relay ('PA_DR27') of protection assembly PA27 in transmission line 'L6_12' (in Figure 4.1). This leads to overloading in other transmission lines, which gets tripped in the process.

  Occurrence of each contingency event and its impact on the system is described in detail in Table 4.2. It shows the progression of cascade with time causing multiple failures in the system. Post analysis, it is observed that transmission lines 'L12_13', 'L13_14', 'L10_9', 'L7_9' and transformers 'T1', 'T2' are also considered disconnected. This is because they do not have a current carrying path through them due to line outages listed in Table 4.2. These events eventually resulted in a load loss of 46.9% and hence caused a blackout based on the criteria referenced in [44]. Due to this, the initial contingency can be marked as a blackout causing contingency. Similar contingencies can be found based on this approach which could lead to severe cascading outages in electrical power transmission systems. Prior knowledge of such contingencies can help in designing effective mitigation strategies, which could prevent the progression of cascades.

In order to validate the generated cascade progression paths, an independent study is performed using a different simulation platform, OpenDSS[124]. WSCC 9 bus system[125] is used as the example system. The results of contingency analysis matched for all but three cases. The 3 cases where the contingency analysis results did not match can be attributed to the different solvers resulting in about $\sim$ 3% difference in the voltages and currents magnitudes computed in the two platforms.

4.7 Conclusions

In this paper detailed behavioral models of the protection assembly is presented along with the capability of introducing cyber faults at specific instants. Integration of these behavioral models with the simulation models in Matlab/Simscape helped us simulate and analyze severe cascading failures that eventually lead to blackout. The study on IEEE 14 bus system showed how introduction of cyber faults in addition to physical fault can lead to severe cascading failures causing blackout. Moreover, this approach can be applied in finding N-k contingencies as discussed in Section IV. In addition to that, the design provides the flexibility to easily understand and extend itself to incorporate more aspects, which could help improve the analysis of cascading failures. As part of the future work, more complex models need to be analyzed and the entire approach can be automated so as to find severe N-k contingencies that can result from a combination of physical and cyber faults.

Chapter 5

Component based Modeling and Analysis Approach

5.1 Problem

Power systems are complex networks that needs to be analyzed from multiple aspects such as steady state analysis, transient analysis, time independent analysis, time based analysis, cyber-faults in components, etc. All these aspects have their own advantages while performing the analysis. For instance a time independent steady state analysis is highly effective and efficient to perform an overall contingency analysis considering multiple transmission line outages. However, a transient analysis is necessary to analyze the effect of contingencies on the voltage stability of the system. Therefore, a power system needs to be analyzed through multiple facets which is referred to as a detailed analysis for improving its overall resilience. However, there is no single modeling and analysis environment that could provide such an analysis capability for the power networks. Each tool has its own limitations and capabilities. In addition, they require the system modeling according to their own semantics and specifications in individual platforms. As a result, researchers use multiple models to perform analysis and stitch together the solutions obtained to formulate an overall reliability plan. This process consumes a lot of unnecessary time while modeling the target system in multiple platforms to analyze. Moreover, it increases the risk of modeling errors.

For instance, modeling a simple IEEE-14 Bus System [1] takes approximately 2-3 hours in OpenDSS (including the calculations needed to be done before modeling) and takes nearly 5-6 hours to model it in Simscape. Considering this, one can only imagine the complexity of modeling systems on a large scale in different platforms. In addition, the developed models need to be analyzed that depends on the user requirements. Hence, it is

necessary to identify which simulation models and platforms can be required for analysis. Further, the models need to be supplied and simulated using these platforms according to the contingency analysis specification. Therefore, in order to make the entire process error free and efficient, we need a framework that could provide the capability to perform multiple analysis on system models with reduced modeling time, effort, and error.

To achieve this, we have developed a framework as shown in figure 5.2. The approach allows us to design a domain specific modeling language (DSML) for power system networks. This DSML provides the capability to capture the right abstractions for the components in power systems and allows the user to create system models depending upon the semantics and rules defined in the modeling language. Once the model is created using the DSML, appropriate simulation tools are identified from the tool-chain for performing the required analysis. Then, the models can be transformed according to the input data formats and individual specifications to the required simulation platforms such as Simulink [51], OpenDSS [124], PowerWorld [126], etc. The automatic model transformation greatly reduces system modeling time and effort. The model transformationfor a specific simulation platform is governed by the use of dedicated plug-ins/tools. These plug-ins/tools are designed specifically based on the requirements of the simulation tools that are linked with the framework in order to guarantee correct model transformation. In addition, we use mechanisms to support model correctness by employing type checking, check for duplicate nodes, etc. Post model transformation, the analysis is performed depending upon the user requirements and the results can be gathered back to the framework for analysis purposes. We provide the framework with the capability of easy extensibility to support the future user needs and requirements.

The developed framework is evaluated by creating the power system models using the developed DSML and transforming them into two different simulation platforms, i.e, Simulink and OpenDSS. These tools are integrated together and contingency analysis is performed on the transformed models. The results are obtained and are used for develop-

ing better resilience metrics.

This is based on the accepted paper in CPS WEEK workshop. The details of the publication is as below:

Hasan, Saqib, Abhishek Dubey, Ajay Chhokra, Nagabhushan Mahadevan, Gabor Karsai, and Xenofon Koutsoukos. "A modeling framework to integrate exogenous tools for identifying critical components in power systems." In Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES), 2017 Workshop on, pp. 1-6. IEEE, 2017.

5.2 Introduction

Cascading failures in electrical power systems are one of the major causes of concern for the modern society as it results in huge socio-economic loss. These failures can occur from multiple causes such as cyber-attacks, protection equipment mis-operation, system overloading, voltage collapse etc. Recent blackouts of Dec 2015 Ukraine[60], July 2012 India[59] have shown electric power grid vulnerability due to such causes and provided reasons to look deeply into the possible sources for these failures. Detailed understanding of cascading failures and identifying critical components for improving system reliability and resiliency necessitates the need to include different aspects (such as steady state vs transient analysis, time independent vs time based analysis, considering protection assembly failures etc.) of the system while performing cascading failure studies. Platforms including various aspects of the system either do not exist or are typically very expensive. Therefore, researchers tend to use multiple open source tools, which are easily available to perform disparate types of analysis on individual platforms. However, these tools have their own specifications and semantics for system modeling and are limited in their capabilities.

Comprehensive understanding for system failure requires modeling of a system in multiple tools for in-depth analysis. For instance, OpenDSS[124], an open source (steady state analysis) tool for electrical power systems can be easily used for quickly identifying critical components based on initial line outages resulting in overloads. However, it is time

independent and does not include the modeling of protection assemblies in its simulation environment. These aspects are important while studying cascading failures in detail, as any random outage can change the entire course of the cascade evolution path and can cause severe outages. Researchers interested in analyzing such failures will ultimately look for other simulation platforms (such as Matlab/Simscape[51]) with these capabilities to perform the desired analysis by modeling the same system in it. Moreover, they sometimes use multiple platforms to validate their analysis results.

System modeling in multiple simulation platforms is a tedious, error prone and time consuming process. For e.g., it takes $\sim$ 2-3 hours to model the IEEE-14 Bus System[1] in OpenDSS (including the calculations needed to be done before modeling) and $\sim$ 5-6 hours to model it in Simscape. Considering this, one can only imagine the complexity of modeling systems on a large scale in different platforms. This necessitates the need for a domain specific modeling language (DSML) which can provide the capability to capture the right abstractions for the modeling components of individual low level modeling and simulation tools in a single higher level modeling and simulation platform. System modeling errors and modeling time can be greatly reduced as this DSML is a common language from where other models can be derived.

Prior approaches for cascading failure analysis are based on determining the current state of power system and then to study its evolution using different cascade simulation models [34, 19, 8, 123, 119, 20]. These approaches can be performed using time independent platforms such as OpenDSS. While it is ideal to use such a platform for expeditious and uncomplicated analysis but performing an in-depth analysis, considering other factors such as time and protection assembly failures due to cyber-faults etc. requires system analysis in a different platform such as Simscape. This facilitates a dynamic analysis providing an advantage over the above models and helps in finding more critical components by employing a richer analysis (not possible otherwise).

A large number of modeling languages are currently available. Modelica[45] is a

multi domain modeling language and both commercial and free Modelica simulation environments such as Dymola [46], MapleSim[47] and OpenModelica[48] are available. InterPSS (AC loadflow analysis)[49], PSAT (continuation and optimal power flow)[50], VST (continuation power flow, voltage stability analysis)[52], MATPOWER (optimal power flow)[53] are some of the modeling and simulation tools for cyber physical energy systems for generation, transmission and distribution. Another modeling, simulation and analysis tool for these systems is GridLAB-D[54] and the modeling language is known as GLM. PowerFactory[55] and PSCAD[56] are some of the conventional standard solutions for simulation and analysis purposes. PowerFactory can perform both AC and DC load flow analysis. However, PSCAD is a transient simulation engine. All these modeling languages and tools provide the capability to model the system in their own specific environments with precise input data formats and can perform analysis only based on their individual capabilities. However, to the best of our knowledge most of them do not provide the ability to transform models into a different platform if needed taking into account distinct input data formats and perform the analysis based on the potentials of other tools.

This paper utilizes the concepts of model integrated computing (MIC[127]) to describe a domain specific modeling language for power systems using WebGME (Web-based Generic Modeling Environment) [128]. It identifies and captures the right abstractions for modeling components in different simulation tools namely OpenDSS and Matlab/Simscape. These tools are chosen because of the limited time but it is possible to interface more tools depending upon the requirements of the researchers. A framework is proposed that deals with system modeling using the developed DSML, identifying the type of analysis to be performed, choosing the appropriate tool(s) needed for a particular analysis from the tool-chain, transforming the model(s) based on the required specifications of a particular tool and performing the analysis. Transformed models and supporting executables are generated in order to save system modeling time and to ease the analysis process in multiple platforms. Type checking is also employed to minimize human errors during system

modeling. Modeling abstraction is validated using the transformed models of the standard WSCC-9 Bus System [125]. Since the focus of this paper is to identify critical components in electrical power systems, a case study is done on WSCC-9 Bus System, IEEE-14 Bus System and IEEE-39 Bus System[129] to demonstrate the entire workflow of the framework in identifying critical components.

The paper is organized as follows: Section II describes the modeling language. Section III discusses the system framework. Model transformation and validation is explained in Section IV. The results are demonstrated in Section V followed by the conclusion in Section VI.

5.3 Modeling Language

A domain-specific modeling language (DSML) has been developed for cyber-physical energy systems (CPS) to enable the rapid design, development and analysis on electrical power systems. A DSML is a declarative language that uses appropriate notations and abstractions to represent various facets of a system and is usually restricted to a particular domain, e.g., power systems.

The meta-model is encapsulated from the developer mode of the graphical interface (WebGME) for model specification, which allows viewing, modification and specification of the rules that administer the construction of power system models and is shown in Figure 5.1. Every object has a *name* attribute of type *string* and objects with a grayed-out name and in italics is a pure *abstract* object. These *abstract* objects cannot be instantiated in a model but they rather serve as the base class for other instantiable classes. The modeling language captures most of the relevant aspects of an electrical power transmission system and using this language engineers can create models containing instances of the objects defined in the DSML. This approach to define the semantics of the models enables a check and ensures model correctness and provides the ability to develop generic utilities called *plugins*. These *plugins* can act on the models created using the modeling language

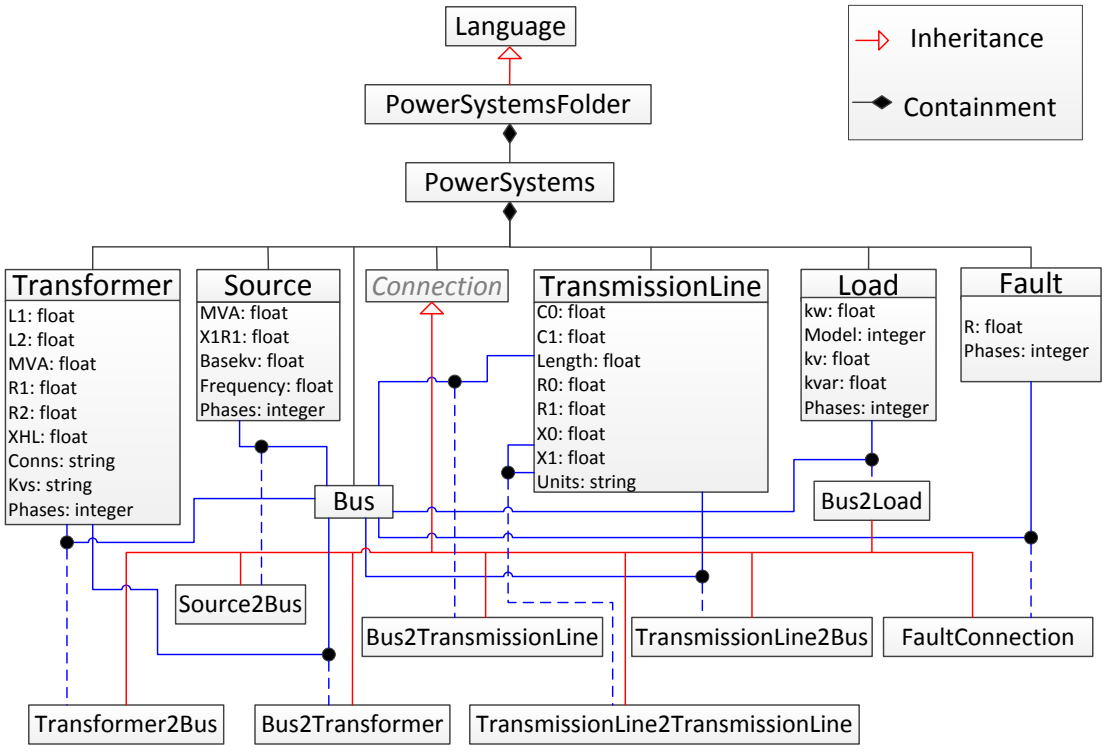


Figure 5.1: Modeling Language- UML Class Diagram.

and perform different tasks, for instance model transformation as per the requirements of a particular platform (OpenDSS or Simscape), perform the desired analysis and manage the results. It also supports code development to perform tasks as per user requirements. Although the meta-model captures many aspects of the electrical power systems but it is not a gold-standard model. It is a result of the development, deployment and analysis experiences with different tools. However, the ability to specify the meta-model and build models based on it within the same tool (WebGME) allow the users to extend or modify the meta-model based on their needs.

Figure 5.1 shows the meta-model as a UML class diagram[130] of the modeling language for power systems. *PowerSystemsFolder* is inherited from *Language*. This *PowerSystemsFolder* contains one or more *PowerSystems*. These *PowerSystems* are the models created using the developed DSML. *PowerSystems* contain one or more *Sources*, *Buses*,

Transformers, *TransmissionLines*, *Loads*, *Faults* and *Connections*. Each of these objects has a set of attributes that define their individual properties. For instance, the *Source* object has attributes that define its output power, internal source impedance, basekv, frequency of the source voltage and current, number of phases for a source. The attributes are associated with a data type, thereby enabling automatic type checking. These objects are connected together using the rules defined by the *Connection* object. *Connection* are of various types namely *Source2Bus*, *Transformer2Bus*, *Bus2Transformer*, *Bus2TransmissionLine*, *TransmissionLine2TransmissionLine*, *TransmissionLine2Bus*, *Bus2Load* and *FaultConnection*. To ensure model correctness specific *Connection* objects are used. For instance, a *Source* can be connected to a *Bus* using a *Source2Bus* connection but it cannot be connected to a *TransmissionLine* without a *Source2Bus* and *Bus2TransmissionLine* connection. The connectivity of different objects using *Connection* object is shown by solid and dotted blue lines in Figure 5.1. Once different objects are connected together a power system model is created and made available for analysis purposes.

5.4 System Framework

The proposed framework enables us to develop domain-specific modeling language (DSML) for power systems. It allows model building depending upon the semantics and rules defined in the modeling language and minimizes modeling errors through type checking. Models are transformed to different simulation platforms considering their individual specifications or input data formats thereby greatly reducing system modeling time and effort. Moreover, it identifies appropriate tool(s) from the tool-chain to perform the desired analysis on the system and manages the results post analysis.

Figure 5.2 demonstrates the proposed framework where an extendable and specialized tool WebGME is used to orchestrate the workflow. Using the developer mode of this tool a domain-specific modeling language is developed once. This language is used to create power system models within the WebGME modeling environment. Once these models are

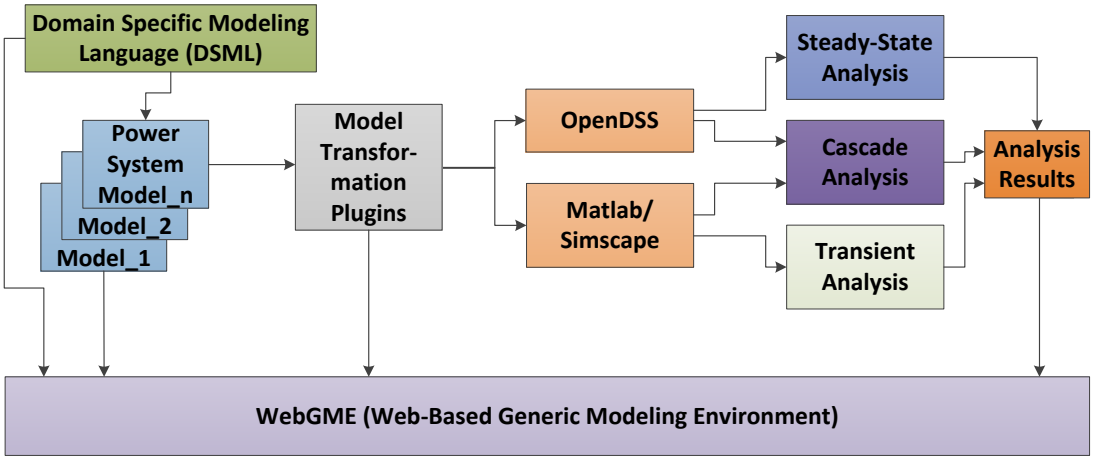


Figure 5.2: System Framework

built they can be transformed based on the requirements of different simulation tool(s), for example OpenDSS, Matlab/Simscape using the model transformation plugins which are specifically developed for model transformation. As WebGME is extendable, these dedicated plugins constitute the supporting infrastructure of the framework. Furthermore, these plugins also perform type checking on the models to ensure correct transformation, for instance they do not let duplicate named objects to be created during the transformation phase which will result in an erroneous model for the simulation tools. Other modeling transformations are also implicitly taken care of during this phase. After the model transformation, appropriate simulation tool is identified from the tool-chain and the model is automatically simulated based on the type of analysis required on the transformed model(s). The type of analysis will depend upon the needs and requirements of a user, for instance steady-state analysis, transient analysis, cascade analysis etc. Finally, post analysis results are gathered back at the WebGME environment. These results can be processed in multiple ways as WebGME is capable of facilitating graphical visualizations as well.

5.5 Model Transformation and Validation

Model transformation provides the capability of transforming the model(s) built in WebGME using the developed DSML into the required platform(s) by taking into account the

modeling semantics and specifications of individual platform. It ensures model correctness and greatly reduces the time and effort for system modeling in multiple platforms. OpenDSS and Simscape are the two tools used in this framework. Model transformation is performed on WSCC-9 Bus System created in WebGME using the DSML to the models that comply with the modeling semantics of the two tools.

5.5.1 WSCC-9 Bus System WebGME Model

Domain-specific modeling language discussed in Section II is used to model the WSCC-9 Bus system in the modeling environment of WebGME and is shown in Figure 5.3. Objects

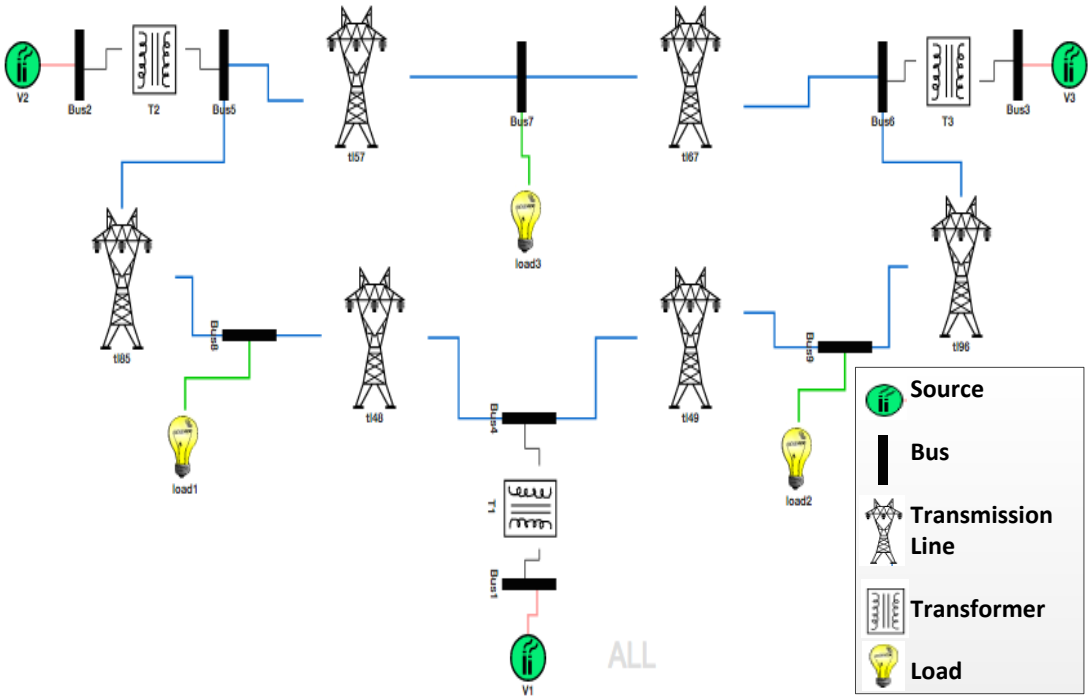


Figure 5.3: WSCC-9 Bus System WebGME Model

such as *Sources*, *TransmissionLines*, *Buses*, *Transformes* and *Loads* are selected to model the system. Attributes associated to each object are set with the appropriate data obtained from the IEEE common data format as referenced in [125].

5.5.2 WSCC-9 Bus System OpenDSS Model

OpenDSS is a time independent, script based steady-state power system modeling and simulation tool. The WSCC-9 Bus System WebGME model is automatically transformed to the OpenDSS model using dedicated *plugins* constituting the framework discussed in Section IV. As shown in Figure 5.4, the model is transformed by taking into account every object (*Sources, Buses* etc.) and its associated attributes to the appropriate semantics in OpenDSS. During the transformation, type checking is employed to ensure proper data flow for each object and to identify and remove duplicate object names which can cause compilation error during model simulation.

```
clear
New object=circuit.9bus
//Define Sources
New vsource.Source1 bus1=Bus1 phases=3 basekv=16.5 Mvasc3=247.5 r1=.0000001 x1=0.0000001
New vsource.Source2 bus1=Bus2 phases=3 basekv=18 Mvasc3=192 r1=.0000001 x1=.0000001
New vsource.Source3 bus1=Bus3 phases=3 basekv=13.8 Mvasc3=128 r1=.0000001 x1=.0000001
//Define the transmission lines and transformers
New Line.TL48 bus1=Bus4 bus2=Bus8 R1= 0.0529 R0=0.13225 X1=.4494 X0=.8972 C1=8.82  C0=5.188 length=62.1371 units=mi
New Line.TL49 bus1=Bus4 bus2=Bus9 R1=0.08993 R0=0.224825 X1=.4863 X0=1.2139 C1=7.922 C0=4.74 length=62.1371 units=mi
New Line.TL85 bus1=Bus8 bus2=Bus5 R1=0.16928 R0=0.4232 X1=.8516 X0=2.1262 C1=15.34 C0= 9.025 length=31.0686 units=mi
New Line.TL96 bus1=Bus9 bus2=Bus6 R1=0.20631 R0=0.5157 X1=.8972 X0=2.2959 C1=17.95 C0= 10.55 length=62.1371 units=mi
New Line.TL57 bus1=Bus5 bus2=Bus7 R1=0.044965 R0= 0.11241 X1=.3808 X0=.7615 C1=7.471 C0= 4.394 length=62.1371 units=mi
New Line.TL67 bus1=Bus6 bus2=Bus7 R1=0.062951 R0= 0.15737 X1=.5331 X0=1.3308 C1=10.47 C0= 6.15 length=62.1371 units=mi
New transformer.T1 phases= 3 buses= (Bus1 Bus4) Kvas=[100000 100000] conns= 'wye wye' kvs= "16.5 230" XHL=5.7147
New transformer.T2 phases= 3 buses= (Bus2 Bus5) Kvas=[100000 100000] conns= 'wye wye' kvs= "18 230" XHL=6.5619
New transformer.T3 phases= 3 buses= (Bus3 Bus6) Kvas=[100000 100000] conns= 'wye wye' kvs= "13.8 230" XHL=5.0917
//Define the loads
New Load.Load1 bus1=Bus8 phases=3 kVA=125000, 50000 Kv=230 conn= delta model=1
New Load.Load2 bus1=Bus9 phases=3 KVA=90000, 30000 Kv=230 conn= delta model=1
New Load.Load3 bus1=Bus7 phases=3 kVA=100000, 35000 Kv=230 conn= delta model=1
//Define the voltagebases
set voltagebases=[16.5, 18, 13.8, 230]
calcv
set freq=60
set mode=snapshot
solve
```

Figure 5.4: WSCC-9 Bus System OpenDSS Model

5.5.3 WSCC-9 Bus System Matlab/Simscape Model

Matlab is a time-based modeling and simulation tool. It has the capability to extend itself and perform the necessary simulation and analysis based on the users needs and requirements. Moreover, Matlab can be easily used for transient analysis in electrical power

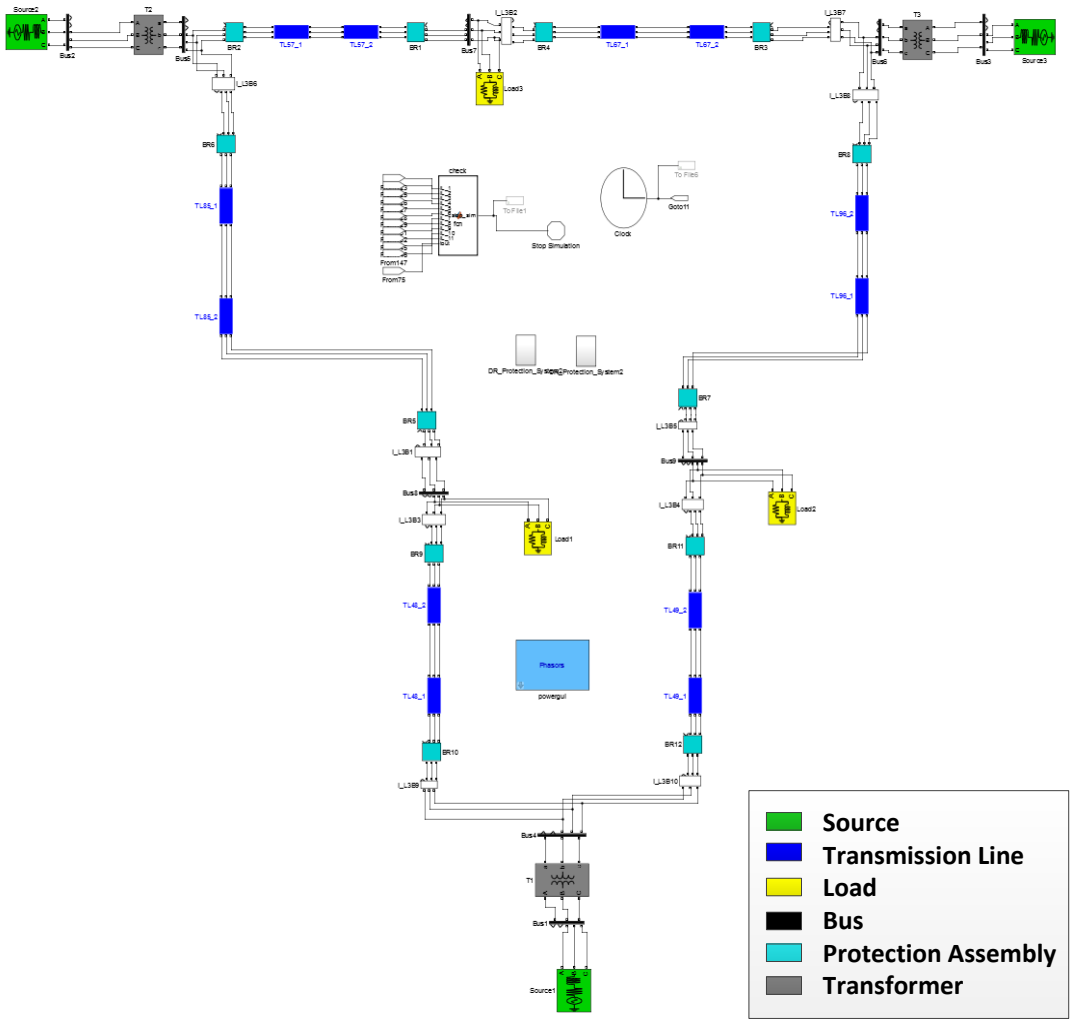


Figure 5.5: WSCC-9 Bus System Matlab/Simscape Model

systems. The WebGME model for the WSCC-9 Bus System is automatically transformed to the appropriate simulation model using dedicated *plugins* constituting the framework discussed in Section IV and is shown in Figure 5.5. The transformed model takes into account every object and its associated attributes to represent the model with correct semantics of Simscape. Certain object attributes require conversion while modeling system in multiple platforms. For e.g., the data obtained from IEEE common data format for the transmission lines in WSCC-9 Bus System has line reactances (X) but it does not contain line inductances (L), which needs conversion using the formula L = X/(2*pi*f). Such con-

versions are automatically taken into account using the model transformation plugins as OpenDSS model takes line reactances (X) as inputs to its transmission line objects and Simscape model takes line inductances (L) as inputs for its transmission line model block.

Furthermore, model transformation saves a lot of time and effort. For instance, the model transformation plugins for Simscape model automatically inserts the current and voltage measurement blocks and a protection assembly block at each end of a transmission line. These blocks are needed for the analysis but are not defined as objects to reduce the complexity and to give a higher abstraction to the DSML. The protection assembly blocks are custom designed and pre-added to the Simscape library to facilitate the model transformation process. These blocks provide the capability to introduce cyber-faults in addition to the physical faults in electrical power systems at different instants. Details about the behavior models of protection assembly blocks considering cyber-faults is referenced in [131].

5.5.4 Validation of The Transformed Models

To Validate the transformed models for the WSCC-9 Bus System, direct mapping of the objects from DSML to OpenDSS and Matlab/Simscape are listed in Table 5.1.

Table 5.1: DSML Object mapping to OpenDSS and Simscape.

DSML Object Name	**OpenDSS Object Name**	**Matlab/Simscape Block Name**
Source	Vsource	Three-Phase Source
TransmissionLine	Line	Three-Phase PI Section Line
Transformer	Transformer	Three-Phase Transformer (Two Windings)
Bus	Bus	Three-Phase VI Measurement
Load	Load	Three-Phase Parallel RLC Load
Fault	Fault	Three-Phase Fault

The transformed models of WSCC-9 Bus System are simulated in the two platforms (OpenDSS, Matlab/Simscape) under nominal mode (absence of any fault condition). These

models yield the same numerical values of bus voltages and transmission line currents with an average error of $\sim$ 1% for bus voltages and $\sim$ 3% for line currents. This variation is attributed to the different solvers in the two platforms.

5.6 Results

Using the framework discussed in Section III, critical components causing cascading failures resulting in blackouts are identified using the cascade analysis performed on OpenDSS and Matlab/Simscape models of WSCC-9 Bus System, IEEE-14 Bus System and IEEE-39 Bus System. Here, blackout criteria is considered as 40% of system load loss which is one of the criterion referenced in [44] and transmission lines are assumed to be loaded at 70% of their loading capacity for each system. Cascading analysis due to initial line outages resulting in subsequent components overloading are performed using OpenDSS (quick and easy, time-independent analysis). However, time based cascade analysis due to physical faults in transmission lines (for instance 3-phase to ground fault) and cyber-faults in protection assemblies are performed using Matlab/Simscape. Details about modeling cyber-faults in protection assembly and their integration with the Matlab/Simscape models to perform cascade analysis are presented in [131].

5.6.1 OpenDSS-Time Independent Analysis

The transformed models of WSCC-9 Bus System, IEEE-14 Bus System and IEEE-39 Bus System created using the developed DSML are used to perform the time-independent cascade analysis to identify critical components (transmission lines) causing blackout. As OpenDSS do not have an object to define protection assembly (distance relay, over-current relay and circuit breakers) and the cyber-faults associated with it, these models cannot be used to perform detailed analysis to identify critical protection assemblies causing blackout. Although, behavior of some cyber-faults in protection assemblies can be replicated in OpenDSS but it requires manually changing the OpenDSS model which is a very te-

dious process. Moreover, timing information which is useful for the operators cannot be obtained using this analysis. However, it serves as an ideal way to quickly identify critical transmission lines based on line overloading.

A simple cascade analysis framework is implemented using the COM interface in OpenDSS. N-k (N = No. of components, $k \in \mathbb{N}$) contingency analysis is performed to identify critical components based on initial line outages. These outages are a set of combinations of line outages that are iteratively removed from the network to simulate the system for possible blackouts. For instance, if k=2 then the set of initial line outages will have a total number of $\binom{N}{2}$ combinations. Each combination of outage(s) are tripped from

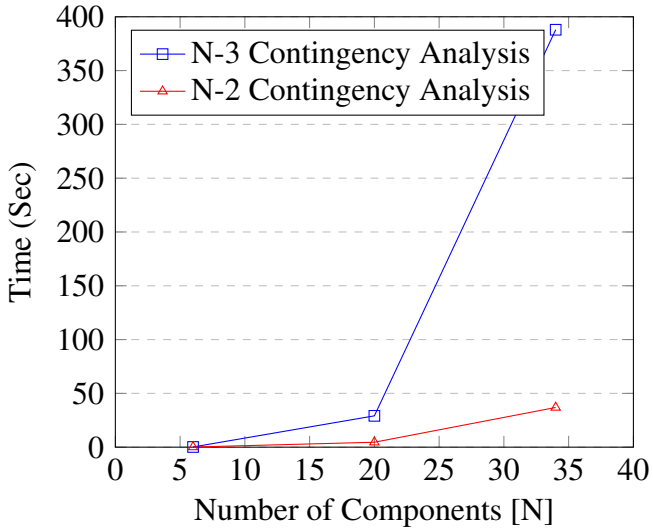


Figure 5.6: Contingency Analysis

the network and the system is checked for overloads. If it exits, the overloaded branches (transmission lines and transformers) are tripped. The system is checked for blackout criteria and if it is met the simulation is stopped and the initial outage(s) are marked as critical. If the blackout criteria is not met and there are further overloads the required branches are tripped and the check is performed again. If there are no further overloads and blackout criteria is not met then the initial outage(s) are not classified as critical. More detailed explanation is referenced in[131].

As per NERC standards, power systems are N-1 tolerant, hence N-2 and N-3 contingency analysis is performed on each system to identify combinations of critical transmission lines causing blackout. Based on the above cascade analysis framework, for N-2 contingency analysis, a total of 168 (13+40+115) combinations out of 901 (15+190+561) combinations of line outages were observed to cause blackout in WSCC-9 Bus System, IEEE-14 Bus System and IEEE-39 Bus System respectively. For N-3 contingency analysis, a total of 2515 (20+400+2095) combinations out of 7144 (20+1140+5984) combinations of line outages were observed to cause blackouts in the above mentioned systems. These combinations are marked as critical lines and can help in improving system resiliency. Figure 5.6 shows the plot of time taken to run the analysis for each system versus the number of components in each system. As 'k' increases the analysis time increases more with increase in the number of components and the plot becomes more exponential. However, this may not be an issue as it is an off-line analysis and does not take a significant amount of time. This can further be improved by employing parallel computing.

5.6.2 Matlab/Simscape-Time based Analysis

Transformed models of WSCC-9 Bus System, IEEE-14 Bus System and IEEE-39 Bus System are used to perform the time-based cascade analysis but only the results of IEEE-14 Bus System are shown due to space constraints. In this analysis cyber-faults in protection assembly (details about cyber-faults and its modeling in the protection assembly is referenced in [131]) causing cascading failures resulting in blackout are considered and critical protection assemblies are identified. It is a time-based analysis and can be useful for operators to design effective mitigation strategies as details about every failure are available with respect to time.

Analysis is performed on the IEEE-14 Bus System and every transmission line is protected using a pair of protection assembly (represented by PAn, $n \in \mathbb{N}$, as shown in Figure 5.7). Protection assembly consists of a distance relay, an over-current relay and a cir-

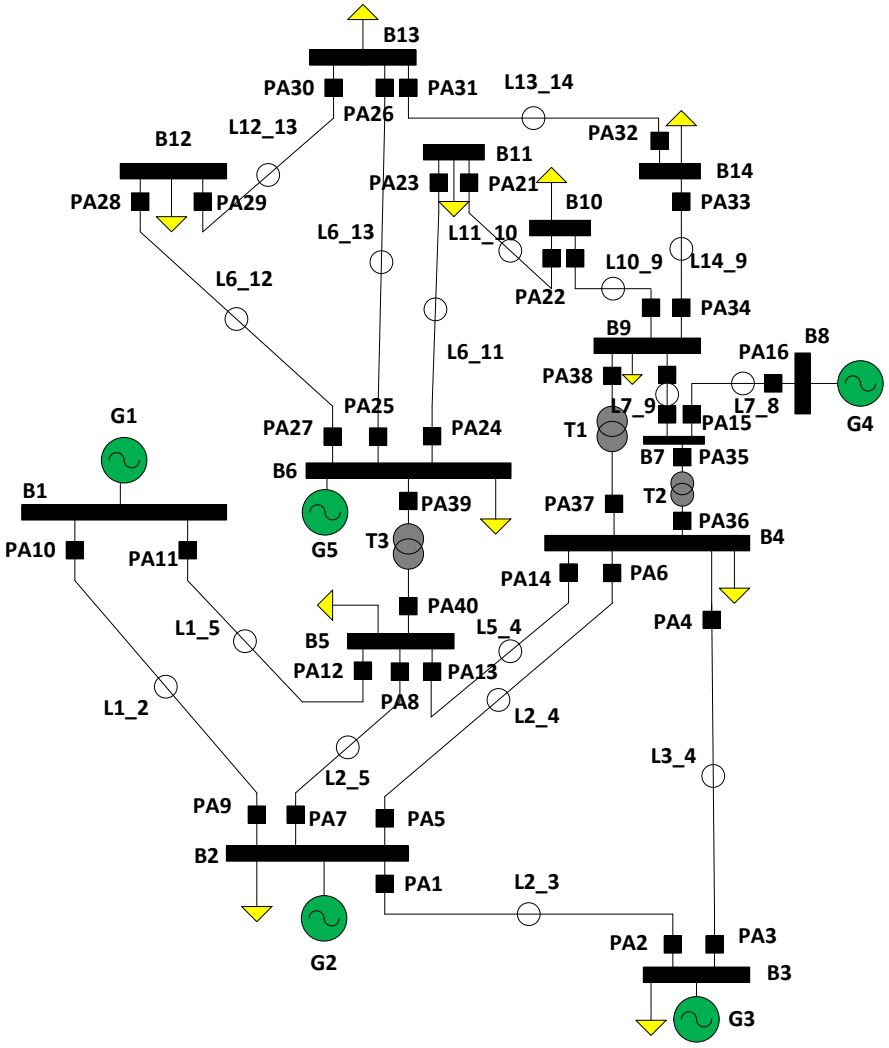


Figure 5.7: IEEE-14 Bus System[1]

cuit breaker (dentoted as PA_DRn, PA_ORn and PA_BRn respectively). Each line is given a physical fault (3-phase to ground fault) and the associated circuit breakers are given a *Stuck Close Breaker Fault* (a type of cyber-fault where the circuit breakers do not operate as desired) individually. Results of other cyber-faults are not shown due to space constraints. The simulation is run using the cascade simulation framework discussed in Section V(A). Initial outages are a combination of physical fault and a cyber-fault (referenced in [131]). As per the blackout criterion, three highly vulnerable protection assembly components (PA_BR4, PA_BR13, PA_BR14) are observed in the system with this fault combination.

Based on the study, critical components are identified and categorized in Table 5.2. Components listed in 'Category I' are the components that causes a blackout in the presence of a physical fault and a cyber-fault. However, the components in 'Category II' are likely to

Table 5.2: Critical Components Categorization

Category Name	**Component Name**	**Load Loss**
Category I	PA_BR4, PA_BR13 PA_BR14	above 40%
Category II	PA_BR6, PA_BR7	very close to 40% (39.22%)
Category III	PA_BR18, PA_BR22 PA_BR34	$>$ 25% and $<$ 35%

cause a blackout if there is any other outage that results in further load loss. These are less critical compared to 'Category I' but still should be considered while improving system resiliency. 'Category III' components are not as critical as the other two categories but can result in blackouts if drastic load loss happens due to a large number of outages.

5.7 Conclusions

In this paper, a domain-specific modeling language (DSML) for electrical power systems is described that identifies and captures the right abstractions for modeling components in different simulation tools. A framework is proposed to facilitate the development of DSML, model creation and transformation and to perform the desired analysis by choosing appropriate tool from the tool-chain. A case study is performed on WSCC-9 Bus system, IEEE-14 Bus System and IEEE-39 Bus System to show how this framework is used in identifying critical components in power systems. Moreover, the design provides flexibility to easily understand and extend the DSML and the supporting infrastructure based on the users needs and requirements. It also provides the capability to integrate more simulation tools so as to perform the desired analysis from within a single environment. As part of the future work, more complex models need to be analyzed and the entire approach can be automated, to perform the desired analysis from within the WebGME environment.

Chapter 6

Critical Contingencies Identification Methodology

6.1 Problem

Multiple $N-k$ contingencies cause cascading failures resulting into large blackouts, e.g., Feb 2016 South Australia [132], Dec 2015 Ukraine [60], July 2012 India [59], August 2003 North America [58] blackouts. Therefore, it is required by the system operators to operate the power systems against cascading failures caused by multiple contingencies based on NERC current standards. However, it becomes challenging to analyze and identify all the possible critical $N-k$ contingencies when k becomes larger ($k \geq 2$) due to the combinatorial nature of the search space. This problem becomes even more complex when the system size becomes larger.

To understand this problem, let's consider a power system consisting with N as the total number of components and k represents the simultaneous transmission line failures. Ignoring the sequence, it requires $\frac{N!}{k!(N-k)!}$ number of simulations for $N-k$ contingency analysis to identify the critical $N-k$ contingencies. Moreover the number of simulations required grows exponentially (N^k) as N and k increases. For example, let's consider a power system with $N = 5000$ and it is required to identify all the critical $N-k$ contingencies for $k = 4$ that cause severe cascading failures resulting in blackouts. This takes approximately a total of 26×10^{12} number of simulations to be performed, which is computationally infeasible.

Therefore, there is a need for effective and efficient methods to identify critical $N-k$ contingencies regardless the scale of the power system network for improving the system resilience that could meet the following objectives:

- Minimize the required number of simulations to identify the critical $N-k$ contingencies such that the non-critical contingencies should be eliminated from the target

contingency set prior to the start of the analysis process. This process can be referred to as pre-contingency screening. In addition, the pre-contingency screening process should not be based on methods that incorrectly identify critical contingencies as non-critical contingencies. This will result in loss of some critical contingencies that needs to be identified and analyzed.

- The methods should be capable enough to handle power system networks irrespective of the number of simultaneous component outages, i.e., the value of k. Moreover, it should consider AC power flow models for realistic solutions. The DC approximation methods would result in less approximate solutions due to negligence of reactive power.

To perform critical $N-k$ contingency analysis and improve the system resilience, we developed two methods that could reduce the total number of simulations and drastically reduce the computational complexity of the problem. Moreover, these methods consider the AC power flow models to identify the critical contingencies in order to obtain a more accurate solution. Both the developed methods focus on reducing the candidate contingency set. A candidate contingency set is referred to as a list of contingencies that are supposed to be evaluated for identification of critical $N-k$ contingencies. These methods are discussed as follows:

- The first method (Method 1) uses previously identified critical $N-k$ contingencies to identify the critical contingencies from the subsequent $N-k$ contingency analysis by pruning the current contingency candidate set.

- The second method (Method 2) is an improved version of the first method. This method employs the frequency distribution of the contingencies appearing in the candidate contingency set and combines it with Method 1. This enables a 2-stage pruning process that reduces the search space dramatically and identifies all the critical $N-k$ contingencies.

Unlike other approaches, our approach focuses on evaluating severe critical $N-k$ contingencies causing severe cascading failures and guarantees to capture all the possible critical contingencies regardless of the k value.

In order to evaluate our approach, we first perform the contingency analysis using both exhaustive search and the proposed methods. The obtained results provide us the details of the time that can be saved using our approach to identify the critical contingencies. Next, we record the total number of simulations that are run by both exhaustive search and the developed methods to identify all the critical $N-k$ contingencies. This provides us the details in the total number of reduction in the simulation runs. Finally, we provide the performance accuracy of the developed models (it represents the effectiveness of the developed models) where the number of identified critical $N-k$ contingencies using exhaustive search are compared with the number of identified critical $N-k$ contingencies using the developed models. This shows the effectiveness of our models.

This is based on the accepted paper in the Resilience Week (RW) conference. The details of the publication is as below:

Hasan, Saqib, Amin Ghafouri, Abhishek Dubey, Gabor Karsai, and Xenofon Koutsoukos. "Heuristics-based approach for identifying critical $N-k$ contingencies in power systems." In Resilience Week (RWS), 2017, pp. 191-197. IEEE, 2017.

6.2 Introduction

Power systems are complex electrical networks consisting of several physical (e.g., transmission lines, energy sources etc.) and computational (e.g., protection devices, PMU's etc.) components which are tightly coupled together. Reliable operation of these systems is of primary importance for the system operators. Based on the North American Electric Reliability Corporation (NERC) standards[6], these systems are generally operated according to the $N-1$ security criterion, where failure of any single component would not result in violation of branch flows, bus voltage or stability limits. Thus, system operators are rou-

tinely able to manage $N-1$ contingencies. However, it becomes challenging to deal with multiple simultaneous $N-k$ contingencies (where $k \geq 2$) which initiate severe cascading failures resulting in blackouts. Examples of such cases are Feb 2016 South Australia [132], Dec 2015 Ukraine [60], July 2012 India [59] and August 2003 North America [12] blackouts. Thus, NERC standards at present require the grid operators to operate power systems against cascading failures resulting from multiple contingencies [6].

Identifying all the possible critical $N-k$ contingencies is computationally infeasible for larger systems and higher values of k because of combinatorial explosion of the search space. For a specific power system, ignoring the sequence, $N-k$ contingency analysis requires $\frac{N!}{k!(N-k)!}$ number of simulations to identify all the critical contingencies. This number grows exponentially (N^k) as N increases. For instance, consider a power system with $N = 5000$, as the total number of components. To identify all the possible critical $N-4$ contingencies causing cascading failures resulting in blackouts, a total of approximately 26×10^{12} simulations are needed to be performed. Hence, exhaustive search is infeasible.

Various approaches have been developed to reduce the computational complexity while identifying multiple critical contingencies [62, 64, 65, 66, 67, 75, 79, 81, 133, 80, 134, 135, 9, 136]. Primitive contingency analysis techniques [62, 64, 65, 66, 67] are based on the ranking and selection of outages. As part of the ranking and selection techniques, contingencies are ranked and selected depending upon the performance index for voltage analysis, line flows, capacity, and power flow analysis. Event trees are used in identifying critical contingencies in [75]. A concept of delta centrality in [79] and line outage distribution factors in [80, 81] are used to identify groups of multiple $N-k$ contingencies. The work in [9] presented a fast $N-2$ contingency analysis algorithm based on [80, 81], which performs pruning of the contingency set. The work in [133] provides a method based on iteratively selecting random subsets which are pruned to obtain collections of multiple contingencies causing system failure. In addition, small groups of severe multiple contingencies can be identified using the optimization algorithms proposed in [134, 135, 136].

In order to improve system reliability and resilience, efficient and effective ways to identify severe cascade causing contingencies are necessary. This paper presents a new approach towards identifying all possible critical $N-k$ contingencies causing cascading failures resulting in blackouts. The approach focuses on evaluating contingencies causing severe cascading failures. The contributions from this paper are:

- We present an algorithm (Algorithm I) that uses previously identified critical $N-k$ contingencies to identify the critical contingencies from the subsequent $N-k$ contingency analysis by pruning the current contingency candidate set. This reduces the computational burden accompanied for ranking contingencies based on the above mentioned approaches.

- We present an improved algorithm (Algorithm II) that uses the frequency distribution of the contingencies appearing in the candidate contingency set and combines it with Algorithm I to employ a 2-stage pruning process identifying all the critical $N-k$ contingencies. According to this distribution, most of the critical contingencies tend to fall within a specific region of the frequency distribution curve as shown in our evaluation section.

- We evaluate our approach using case studies on the standard IEEE-14 bus system[1], IEEE-39 bus system[129], and IEEE-57 bus system[137]. Our results show that the algorithms are able to capture all the critical $N-k$ contingencies without missing any dangerous system failure causing contingency. The approach largely reduces the computational effort and takes significantly less time to identify these critical contingencies as compared to the exhaustive search. Moreover, these algorithms are based on the iterative pruning of the search space which results in very few simulations.

The remainder of the paper is organized as follows. Section II presents the heuristic algorithms to identify critical $N-k$ contingencies. Section III discusses the cascade simulation

framework used for simulating the power systems. Section IV demonstrates the results followed by the conclusions in Section V.

6.3 Contingency Analysis

In this section, we present two algorithms to identify critical $N-k$ contingencies in a power transmission system. We consider a power system $\mathcal{G}_p$ which consists of components such as buses, transmission lines, transformers, loads, and generators. The purpose of these components is to supply sufficient power from the generating stations to the loads. Failure(s) can occur in one or more component of the power system. We refer to these failures as $N-k$ contingencies, where the value of k defines the number of simultaneous multiple failures in a system consisting of N components. These $N-k$ contingencies may cause severe cascading outages resulting in a system failure, where system failure is defined by a user-supplied criterion that represents a blackout, e.g., power loss greater than or equal to 40% of the total power needed. Further, the system failure causing contingencies are referred to as critical $N-k$ contingencies. Finally, $N-k$ *contingency analysis* is defined as analyzing k simultaneous failures to understand their effects on the rest of the power network.

6.3.1 Algorithm I

First, we present Algorithm I which is based on the iterative pruning of the current candidate contingency set using the previously identified critical $N-k$ contingencies. For example, to identify critical components from $N-2$ analysis, we use the identified critical contingencies from $N-1$ analysis to prune the candidate contingency set for $N-2$ analysis. Let $\mathcal{U}$ represent the universal set of all the N possible component outages in a power system. Given a value of k, we denote by $\mathcal{S}_k$ the entire search space, defined by $\mathcal{S}_k = \{a \,|\, a \in 2^{\mathcal{U}}, |a| \leq k\}$. Further, we let C_f denote the system failure criterion.

Let $\mathcal{F} \in \mathcal{S}_{k'}$ be a contingency. If $\mathcal{F}$ causes a system failure, in the subsequent $N-k$

contingency analysis (where $k > k'$), we assume any other contingency $\mathcal{F}' \in \mathcal{S}_k$, satisfying $\mathcal{F} \subseteq \mathcal{F}'$, also causes a system failure. This assumption seems to hold true for most scenarios because, if $\mathcal{F}$ causes a system failure then intuitively $\mathcal{F}'$ will outage more number of components from the system. This will weaken the system more and result in larger damage. However, this assumption does not always hold true. That is, in some rare cases, even if $\mathcal{F}$ results in a system failure, $\mathcal{F}'$ will not cause a system failure (i.e., loss less than C_f) and eventually leads to a stable state. However, most of these $\mathcal{F}'$ still causes cascading failures that results in quite a significant loss within the system.

Algorithm 1 Algorithm for Finding $N-k$ Contingencies

Input: $\mathcal{G}_p, \mathcal{U}, C_f, k$
Initialize: $\mathcal{T} \leftarrow \emptyset, \mathcal{R} \leftarrow \emptyset$, $c_{pre} \leftarrow 0$
for all $\mathcal{F} \in \mathcal{S}_1$ **do**
    $loss \leftarrow$ `simulate_contingency`$(\mathcal{G}_p, \mathcal{F})$
    if $loss \geq C_f$ **then**
        $\mathcal{R} \leftarrow \mathcal{R} \cup \mathcal{F}$
    end if
end for
for $p = 2, \ldots, k$ **do**
    $\mathcal{P} \leftarrow \emptyset, \mathcal{R}_{cur} \leftarrow \emptyset$
    for all $\mathcal{F}' \in \mathcal{S}_p$ **do**
        for all $\mathcal{F} \in \mathcal{R}$ **do**
            if $\mathcal{F} \subseteq \mathcal{F}'$ **then**
                $\mathcal{P} \leftarrow \mathcal{P} \cup \mathcal{F}'$
            end if
        end for
    end for
    $\mathcal{T} \leftarrow \mathcal{T} \cup \mathcal{P}$
    $\hat{\mathcal{S}}_p \leftarrow \mathcal{S}_p \setminus \mathcal{P}$ ▷ prunes search space $\mathcal{S}_p$
    for all $\mathcal{F} \in \hat{\mathcal{S}}_p$ **do**
        $loss \leftarrow$ `simulate_contingency`$(\mathcal{G}_p, \mathcal{F})$
        if $loss \geq C_f$ **then**
            $\mathcal{R}_{cur} \leftarrow \mathcal{R}_{cur} \cup \mathcal{F}$
        end if
    end for
    $\mathcal{R} \leftarrow \mathcal{R} \cup \mathcal{R}_{cur}$
    if $|\mathcal{R}_{cur}| \leq c_{pre}$ **then**
        break
    end if
    $c_{pre} \leftarrow |\mathcal{R}_{cur}|$
end for
return $\mathcal{T}$

For example, consider a power system with universal set $\mathcal{U}$ containing transmission

lines $tl_1, tl_2, \ldots, tl_m$. In $N-1$ contingency analysis, if an outage $\mathcal{F} = \{tl_a\}$ satisfies C_f, then $\mathcal{F}$ is marked as a critical contingency. Next, in $N-2$ contingency analysis, any contingency $\mathcal{F}' = \{tl_a, tl_i\}$, where $i \in \{1, \ldots, m\} - \{a\}$, is assumed to cause a system failure. Therefore, the candidate pairs are pruned from the search space $\mathcal{S}_2$ and are not considered for simulation.

The algorithm takes the power system model $\mathcal{G}_p$, the N possible component outage set $\mathcal{U}$, system failure criterion C_f, and contingency range k as inputs. Further, it identifies the total number of critical $N-k$ contingencies denoted by $\mathcal{T}$. The set of critical contingencies causing system failure that are identified through simulations is denoted by $\mathcal{R}$. The set of predicted $N-k$ contingencies resulting in system failure using the set $\mathcal{R}$ is denoted by $\mathcal{P}$. The algorithm evaluates each contingency denoted by $\mathcal{F}$ using the function `simulate_contingency`$(\mathcal{G}_p, \mathcal{F})$ and adds it to $\mathcal{R}$, if the loss due to $\mathcal{F}$ is greater than or equal to the system failure criterion C_f. The function `simulate_contingency`$(\mathcal{G}_p, \mathcal{F})$ is a contingency simulator described in Section III. Given a value of p ranging from 1 to k, the algorithm identifies the search space $\mathcal{S}_p$ for the next iteration. In each iteration p, it evaluates if an element of $\mathcal{R}$ is a subset of an element in $\mathcal{S}_p$ depending upon which the elements are placed in $\mathcal{P}$. The set $\hat{\mathcal{S}_p}$ represents the pruned set of contingencies that are needed to be simulated using the function `simulate_contingency`$(\mathcal{G}_p, \mathcal{F})$ in order to identify critical $N-k$ contingencies that are not captured during the prediction stage. Using $\hat{\mathcal{S}_p}$, missed critical contingencies $\mathcal{F}$ that satisfied C_f are identified and $\mathcal{R}$ is updated accordingly. This further improves the pruning process in the subsequent iterations. The algorithm is terminated either after k iteration, or when the number of current identified critical contingencies obtained through simulations are less than or equal to the number of identified critical contingencies at the previous iteration ($|\mathcal{R}_{cur}| \leq c_{pre}$).

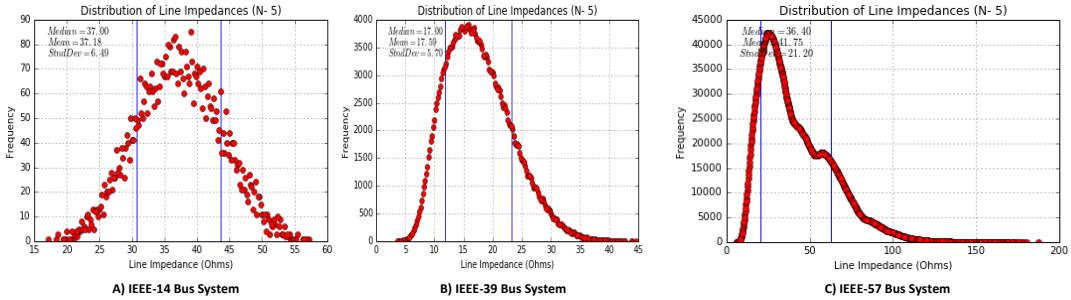


Figure 6.1: Frequency distribution curves of the candidate contingency set ($\mathcal{S}_5$) for different standard power systems.

6.3.2 Algorithm II

In Algorithm II, we use the frequency distribution curve representing the frequency distribution of the candidate contingency set and the idea from Algorithm I to employ a 2-stage pruning of the candidate contingency set $\mathcal{S}_k$. This curve represents the frequency with which a contingency $\mathcal{F}$ with impedance $\mathcal{Z}(\mathcal{F})$ appear within the search space $\mathcal{S}_k$. The frequency distribution curve of different standard IEEE systems are shown in Figure 6.1. The x-axis represents the impedance of an individual contingency. Further, the y-axis describes the frequency with which individual contingency impedances appear within the search space $\mathcal{S}_k$. For any transmission line $a \in \mathcal{U}$, let z_a denote its impedance. Given a value of k and a contingency $\mathcal{F} \in \mathcal{S}_k$, the mean impedance $\mathcal{Z}(\mathcal{F})$ of the contingency is

$$\mathcal{Z}(\mathcal{F}) = \frac{\sum_{a \in \mathcal{F}} z_a}{|\mathcal{F}|} \tag{6.1}$$

The average impedance $\bar{\mathcal{Z}}_k$ of the frequency distribution curve for a search space $\mathcal{S}_k$ is

$$\bar{\mathcal{Z}}_k = \frac{\sum_{\mathcal{F} \in \mathcal{S}_k} \mathcal{Z}(\mathcal{F})}{|\mathcal{S}_k|} \tag{6.2}$$

where $|\mathcal{S}_k|$ is the total number of contingencies.

Further, the standard deviation $\sigma_{\mathcal{Z}}$ of the frequency distribution is defined by

$$\sigma_{\mathcal{Z}} = \sqrt{\frac{\sum_{\mathcal{F} \in \mathcal{S}_k} (\mathcal{Z}(\mathcal{F}) - \bar{\mathcal{Z}}_k)^2}{|\mathcal{S}_k|}} \tag{6.3}$$

Considering a window size within the frequency distribution curve (e.g., as shown by blue

Algorithm 2 Algorithm for Finding $N-k$ Contingencies

Input: $\mathcal{G}_p, \mathcal{U}, C_f, \mathcal{Z}_w, k$
Initialize: $\mathcal{T} \leftarrow \emptyset, \mathcal{R} \leftarrow \emptyset$, $c_{pre} \leftarrow 0$
for all $\mathcal{F} \in \mathcal{S}_1$ **do**
    $loss \leftarrow$ `simulate_contingency`$(\mathcal{G}_p, \mathcal{F})$
    if $loss \geq C_f$ **then**
        $\mathcal{R} \leftarrow \mathcal{R} \cup \mathcal{F}$
    end if
end for
for $p = 2, \ldots, k$ **do**
    $\mathcal{P} \leftarrow \emptyset, \mathcal{R}_{cur} \leftarrow \emptyset, \mathcal{S}_p^{'} \leftarrow \emptyset$
    for all $\mathcal{F} \in \mathcal{S}_p$ **do**
        if $\mathcal{Z}(\mathcal{F}) \notin \mathcal{Z}_w$ **then**
            $\mathcal{S}_p^{'} \leftarrow \mathcal{S}_p^{'} \cup \mathcal{F}$ ▷ prunes search space $\mathcal{S}_p$
        end if
    end for
    $\mathcal{T} \leftarrow \mathcal{T} \cup (\mathcal{S}_p \setminus \mathcal{S}_p^{'})$
    for all $\mathcal{F}' \in \mathcal{S}_p^{'}$ **do**
        for all $\mathcal{F} \in \mathcal{R}$ **do**
            if $\mathcal{F} \subseteq \mathcal{F}'$ **then**
                $\mathcal{P} \leftarrow \mathcal{P} \cup \mathcal{F}'$
            end if
        end for
    end for
    $\mathcal{T} \leftarrow \mathcal{T} \cup \mathcal{P}$
    $\hat{\mathcal{S}}_p \leftarrow \mathcal{S}_p^{'} \setminus \mathcal{P}$ ▷ prunes search space $\mathcal{S}_p^{'}$
    for all $\mathcal{F} \in \hat{\mathcal{S}}_p$ **do**
        $loss \leftarrow$ `simulate_contingency`$(\mathcal{G}_p, \mathcal{F})$
        if $loss \geq C_f$ **then**
            $\mathcal{R}_{cur} \leftarrow \mathcal{R}_{cur} \cup \mathcal{F}$
        end if
    end for
    $\mathcal{R} \leftarrow \mathcal{R} \cup \mathcal{R}_{cur}$
    if $|\mathcal{R}_{cur}| \leq c_{pre}$ **then**
        break
    end if
    $c_{pre} \leftarrow |\mathcal{R}_{cur}|$
end for
return $\mathcal{T}$

lines in Figure 6.1), most critical $N-k$ contingencies fall in this region. Hence, there is a

higher probability of picking a critical contingency within this region. The window size is then obtained by

$$\mathcal{Z}_w = [\bar{\mathcal{Z}}_k - \sigma_{\mathcal{Z}}, \bar{\mathcal{Z}}_k + \sigma_{\mathcal{Z}}] \tag{6.4}$$

Based on the assumption from the frequency distribution curve, in Algorithm II, a contingency $\mathcal{F} \in \mathcal{S}_k$ that appears within $\mathcal{Z}_w$ is pruned from the search space $\mathcal{S}_k$. This is referred to as stage-1 prediction and pruning. After stage-1 pruning of $\mathcal{S}_k$, further pruning is done based on the same approach as Algorithm I. This provides a stage-2 prediction and pruning, which further improves the efficiency of our method.

The Algorithm takes the power system model $\mathcal{G}_p$, the N possible component outage set $\mathcal{U}$, system failure criterion C_f, frequency distribution curve window size denoted by $\mathcal{Z}_w$ and contingency range k as inputs. Further, it identifies the total number of critical $N-k$ contingencies denoted by $\mathcal{T}$. The set of critical contingencies causing system failure that are identified through simulations is denoted by $\mathcal{R}$. The set of predicted $N-k$ contingencies resulting in system failure using the set $\mathcal{R}$ is denoted by $\mathcal{P}$. The algorithm evaluates each contingency denoted by $\mathcal{F}$ using the function `simulate_contingency`$(\mathcal{G}_p, \mathcal{F})$ and adds it to $\mathcal{R}$, if the loss due to $\mathcal{F}$ is greater than or equal to the system failure criterion C_f. Given a value of p ranging from 1 to k, the algorithm identifies the search space $\mathcal{S}_p$ for the next iteration. In each iteration p, it evaluates if a contingency $\mathcal{F} \in \mathcal{S}_p$ does not exist within the specified region of the frequency distribution curve denoted by $\mathcal{Z}_w$, it is added to $\mathcal{S}_p^{'}$. This step defines the stage-1 pruning of the candidate contingency set $\mathcal{S}_p$.

Further, in the same iteration p, the algorithm evaluates if an element of $\mathcal{R}$ is a subset of an element in $\mathcal{S}_p^{'}$ depending upon which the elements are placed in $\mathcal{P}$. This step mark the stage-2 pruning of the search space $\mathcal{S}_p$. The set $\hat{\mathcal{S}_p}$ represents the pruned set of contingencies that are needed to be simulated in order to identify critical $N-k$ contingencies that are not captured during the prediction stage. Using $\hat{\mathcal{S}_p}$, missed critical contingencies $\mathcal{F}$ that satisfied C_f are identified and $\mathcal{R}$ is updated accordingly. This further improves the pruning process in the subsequent iterations. The algorithm is terminated either after k

iteration, or when the number of current identified critical contingencies obtained through simulations are less than or equal to the number of identified critical contingencies at the previous iteration ($|\mathcal{R}_{cur}| \leq c_{pre}$).

Note, when the system size becomes too large, the two algorithms can be iteratively used over the subset of the search space $\mathcal{S}_k$ to identify critical $N-k$ contingencies. Another possible solution can be to run the algorithms on the subset of the search space $\mathcal{S}_k$ over a distributed computing platform. Moreover, the approach will still be able to capture all the possible critical contingencies without missing any dangerous contingency. In addition, we use simple data structures while implementing these heuristics. However, the efficiency of these algorithms can further be improved by making use of efficient data structures, such as trees etc., if needed. This will further improve the analysis results discussed in Section IV.

6.4 Contingency Simulator

In this section, we describe our contingency simulator framework. There are various cascade simulation models and each of these models have their own assumptions, capabilities and limitations [123, 119, 20]. Among these models, there is no standard cascade simulation model for simulating cascading failures. In this work, we select a commonly used cascade simulation model, but it should be noted that the considered cascade model can be easily replaced by any other model while keeping the algorithms fixed.

We have developed a contingency simulator framework by integrating OpenDSS power system model and Cascade simulation model with the OpenDSS contingency simulator, which is a modified version of the simulator used in [131] for Simscape Models. The OpenDSS contingency simulator is an OpenDSS-based AC power flow solver for power systems[124]. The simulator allows us to capture critical $N-k$ contingencies causing severe cascading outages resulting in system failure. Further, the identified critical contingencies can help operators design effective mitigation strategies.

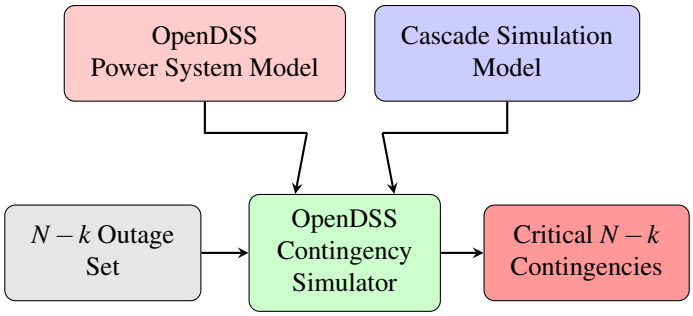


Figure 6.2: Cascade Simulator Framework

The contingency simulator framework is shown in Figure 6.2, where the inputs to the simulator are the OpenDSS power system model, cascade simulation model, and $N-k$

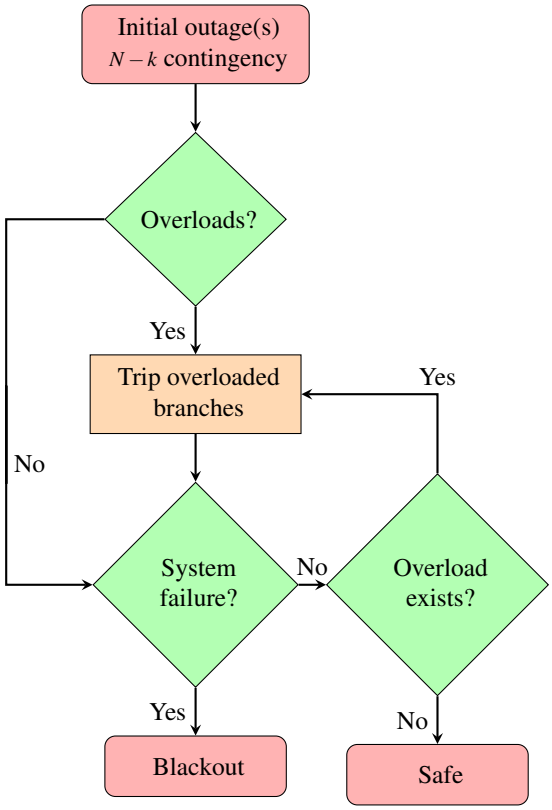


Figure 6.3: Cascade Simulation Model

outage set. The $N-k$ outage set is the set of contingencies that are needed to be simulated and analyzed. Further, the contingency simulator analyzes each contingency based on the cascade simulation model and identifies the set of critical $N-k$ contingencies.

The flowchart for the cascade simulation model is shown in Figure 6.3. Initial outages

in the form of $N-k$ contingencies are given to the simulator. After the initial outages, the OpenDSS power system model is executed by solving power flow and the system is evaluated for overloads. If an overload is observed, identified branches are tripped and the system is evaluated for the system failure (i.e., blackout) criterion. Here, the system failure criterion is considered to be a load loss of 40% or more, which is one of the criterion in [44]. If the criterion is met then the contingency is marked as a critical $N-k$ contingency. However, if the criterion is not met and the system is still overloaded, then the overloaded branches are tripped until all the overloaded branches are eliminated, or the system satisfies the system failure criterion, or the system reaches a stable state (as described by Safe state in Figure 6.3). Moreover, if there are no overloads after the initial outages then the system is directly evaluated for the system failure criterion. If the criterion is not satisfied then the contingency is marked as Safe.

6.5 Evaluation

To validate and test the developed algorithms, we apply them to the standard IEEE-14 bus system, IEEE-39 bus system, and IEEE-57 bus system. We noticed that the two algorithms are able to identify all the possible critical $N-k$ contingencies without missing any dangerous $N-k$ outage resulting in system failure. Furthermore, we observed that the heuristics are much more effective and efficient as compared to the exhaustive search and use significantly smaller number of simulations to predict the actual number of critical $N-k$ contingencies compared to exhaustive search.

6.5.1 Execution Time Analysis of the Algorithms

First, we compare the time complexity of the two algorithms with the exhaustive search. $N-9$, $N-5$, $N-4$ contingency analysis are performed on IEEE-14 bus system, IEEE-39 bus system and IEEE-57 bus system respectively. Figure 6.4 shows the time complexity results of these systems. In each figure, x-axis represents the value of k and the y-axis

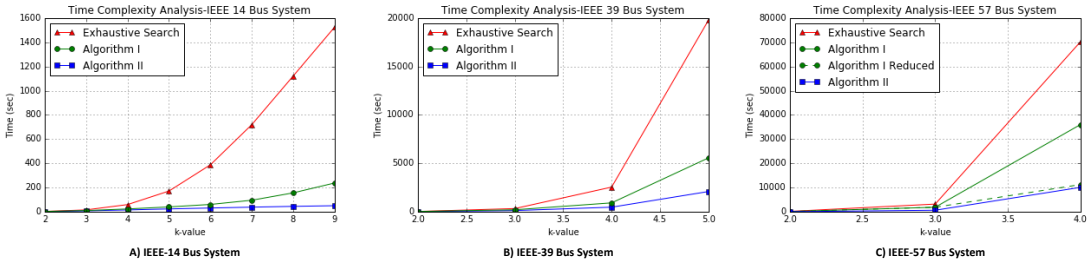


Figure 6.4: Execution Time Analysis-Time taken by Exhaustive search, Algorithm I and Algorithm II to Identify Critical $N-k$ Contingencies

represents the time taken to perform the contingency analysis given a value of k. The red, green, and blue lines represent the execution time for exhaustive search, Algorithm I, and Algorithm II, respectively. Figure 6.4 clearly shows that Algorithm I and Algorithm II are much faster than the exhaustive search for all k values. Moreover, Algorithm II seems to be even faster than Algorithm I. This is because of the 2-stage pruning performed in each iteration of Algorithm II, which reduces the search space significantly and improves the algorithm's efficiency.

The exhaustive search to perform $N-4$ contingency analysis for IEEE-57 bus system identifies a total of 346,214 system failure contingencies from a total of 722,865 contingencies. Using Algorithm I, which requires performing only 24,469 simulations, 345,662 critical $N-k$ contingencies out of the 346,214 critical contingencies are identified. To identify the remaining 552 critical $N-4$ contingencies, Algorithm I uses 259,600 simulations in its final iteration. If the computational cost of running these 259,600 simulations to identify the remaining 552 critical $N-4$ contingencies is high, the algorithm can terminate prior to these simulations. This will significantly improve the execution time of Algorithm I, as shown by the green dotted line in Figure 6.4.C. The same approach can similarly be applied to the other considered systems but it is not shown due to figure clarity.

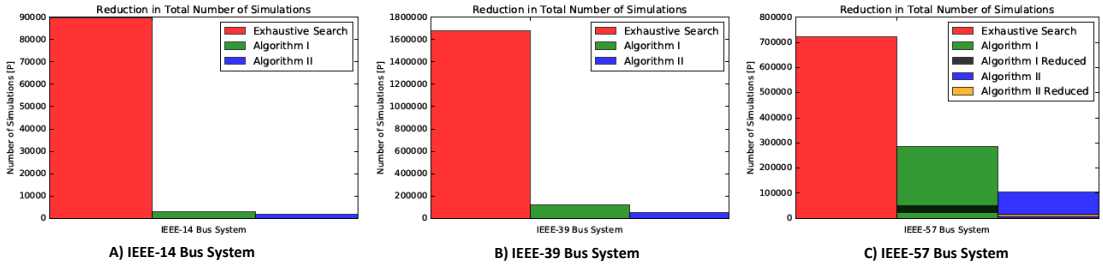


Figure 6.5: Total Number of Simulations Run using Exhaustive Search, Algorithm I and Algorithm II to Identify Critical $N-k$ Contingencies

6.5.2 Reduction in the Total Number of Simulations

Now, we compare the total number of simulations needed to identify critical $N-k$ contingencies using exhaustive search, Algorithm I, and Algorithm II. The analysis on the standard IEEE systems shows that Algorithm I and Algorithm II effectively reduce the total number of simulations as compared to the exhaustive search and still are capable of identifying all the critical $N-k$ contingencies. Figure 6.5 shows the total number of simulations used by the exhaustive search, Algorithm I, and Algorithm II to identify all the critical $N-k$ contingencies. The y-axis represents the total number of simulations used by each algorithm. The red, green, and blue bars represent the number of simulations performed by the exhaustive search, Algorithm I, and Algorithm II, respectively. Figure 6.5.A shows that in IEEE-14 bus system, there are 89845, 3095, and 1734 number of simulations performed for the exhaustive search, Algorithm I, and Algorithm II to identify all the possible critical $N-k$ contingencies (where $k=9$). Further, Figure 6.5.B shows that in IEEE-39 bus system, there are 1676115, 117536, and 48046 number of simulations carried out for the exhaustive search, Algorithm I and Algorithm II to identify the cirtical $N-k$ contingencies (where $k=6$). As show in Figure 6.5, the numbers of performed simulations are significantly reduced by Algorithm I and Algorithm II as compared to the exhaustive search. In addition, Algorithm II performs much better than Algorithm I in terms of reducing the number of simulations.

The total number of simulations can be further reduced in both of our algorithms by avoiding running simulations for scenarios where large number of simulations results in identifying only a very few critical $N-k$ contingencies. In our case, for Algorithm I, if the 259,600 simulations that will capture only 552 critical contingencies are avoided, the total number of simulations will be reduced to only 24,469 simulations (marked by the black-colored region in Figure 6.5.C). Similarly, for Algorithm II, the total number of simulations can be reduced to only 7,413 simulations (marked by the orange-colored region in Figure 6.5.C). This is a significant reduction in the number of simulations if there is a leverage to identify most but not all critical contingencies.

6.5.3 Performance Accuracy of the Algorithms

First, the effectiveness of Algorithm II is shown using Figure 6.6. During the stage-1 prediction and pruning process in Algorithm II, we identified a total of 14,968 out of 19,778 and a total of 15,272 out of 21,879 critical contingencies for IEEE-14 bus system and IEEE-39 bus system respectively. These critical contingencies are identified without any simulations. Based on our heuristics for the frequency distribution curve, most critical $N-k$ contingencies are expected and do fall within the region shown by blue lines (Figure 6.6), which represents the window size defined by $\mathcal{Z}_w$ in Section II. Furthermore, most of the remaining critical contingencies that are not captured in stage-1 prediction and pruning process are identified using the stage-2 prediction and pruning process. In addition, only a very few critical contingencies are needed to be identified using simulations. Thus, all the critical $N-k$ contingencies are identified through minimum computational effort.

Now, we compare the performance accuracy of the two algorithms. Performance accuracy is a measure of the ability of these algorithms to capture the number of critical $N-k$ contingencies when compared with the exhaustive search. In Figure 6.7, the x-axis represents the value of k and the y-axis represents the total number of identified critical $N-k$ contingencies. The red, green and blue lines represent the identified critical $N-k$

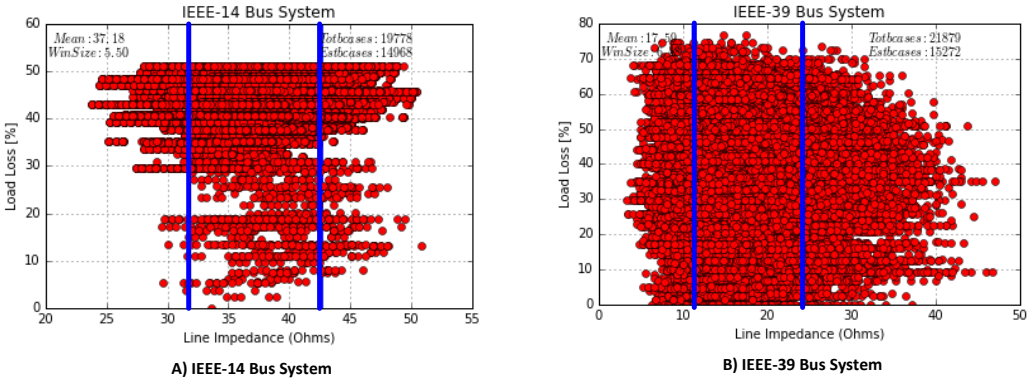


Figure 6.6: Effectiveness of Stage-1 Prediction and Pruning Process of Algorithm II

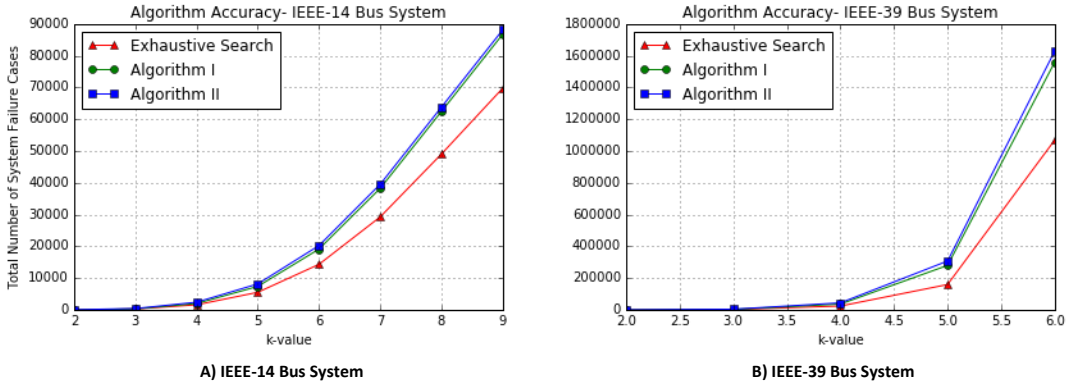


Figure 6.7: Prediction Accuracy of Algorithm I and Algorithm II

contingencies using exhaustive search, Algorithm I, and Algorithm II respectively. Figure 6.7 shows that Algorithm I and Algorithm II are nearly the same with respect to their performance accuracy, that is, they capture nearly the same number of critical contingencies. In IEEE-14 bus system, a total of 69749, 86733, and 88111 critical $N-k$ contingencies (where, $k=9$) are identified using exhaustive search, Algorithm I and, Algorithm II respectively. In IEEE-39 bus system, a total of 1068603, 1558545, and 1628069 $N-k$ contingencies (where, $k=6$) resulting in system failure are identified using exhaustive search, Algorithm I and Algorithm II respectively. Performance accuracy results for IEEE-57 bus system are not shown but can be obtained similarly.

As shown in Figure 6.7, the two algorithms are capable of capturing all the critical $N-k$ contingencies that can be identified using exhaustive search. Moreover, they also capture and classify some non-critical contingencies as critical. But, most of these non-critical contingencies still cause cascading failures that result in significant losses within the system. Figure 6.6 shows most of the non-critical contingencies within the region marked by blue lines cause significant loss. If remained not addressed, these non-critical contingencies combined with another outage can instantly put the system in a non-recoverable state. In addition, these identified non-critical contingencies together with the critical contingencies from the algorithms can further help operators improve the system reliability and resilience.

6.6 Conclusions

Two heuristic algorithms were developed for $N-k$ contingency analysis problem. The idea for both these heuristics is based on the iterative pruning of the candidate contingency set $\mathcal{S}_k$. The pruning process is based on the previously identified critical $N-k$ contingencies and the information from the frequency distribution curve of the $N-k$ candidate contingency set. Even though the approach is based on the developed heuristics, it captures all the critical $N-k$ contingencies without missing any dangerous contingency resulting in system failure. The algorithms are validated and tested on the standard IEEE-14 bus system, IEEE-39 bus system and IEEE-57 bus system. The results described in Section IV demonstrate that these heuristics perform much better than the exhaustive search by reducing the computational time and minimizing the total number of simulations.

Although the results prove the effectiveness of the two heuristics, they can further be improved by using efficient data structures, such as trees etc. that could improve the overall efficiency of the algorithms. Another approach is to use the concept of distributed computing, where the algorithms can be run in parallel on multiple cores to optimize the approach further. Additionally, for very large candidate contingency set, a smaller critical subset can be used to identify critical $N-k$ contingencies. Furthermore, the identified critical $N-k$

contingencies can be used by system operators to design effective failure detection and mitigation strategies to improve system resilience and reliability [138]. As part of the future work, larger power transmission systems can be analyzed and the approach can be applied to radial distribution power networks to perform $N-k$ contingency analysis for identifying its performance.

Chapter 7

Modeling and Analysis of Static Cyber-Physical Attacks

7.1 Problem

In order to provide reliable system operation and meet future energy demands, smart grids are equipped with several sophisticated instruments such as Advanced Metering Infrastructures (AMIs), Phasor Measurement Units (PMUs), distance relays, etc. However, this advancement in technology increases the potential attack surfaces due to the introduction of the cyber components. 'Cyber' here means the software associated with protection assemblies such as distance relay, over-current relay, etc. Malicious attackers are able to take advantage of the technological advancements and compromise the control centers by gaining access into the control stations. The attackers can then launch sophisticated attacks such as remotely operating circuit breakers to disconnect transmission lines, false data injection attacks for incorrect state estimation, etc. that could cause severe cascading failures resulting in large system damages.

Cyber-attacks can be easily modeled and analyzed, however, it becomes really challenging when the defenders are resource bounded and tries to identify the critical components to protect in order to minimize the system damage as a result of large cascading failures. This is mainly due to the fact that power systems are very large systems with several hundreds of buses, transmission lines and identifying multiple critical components to protect from the entire power systems is computationally infeasible. In addition, similar analogy is true for identifying critical components to attack when the attacker is resource bounded while attacking a power system. Hence, there is a need for a *cyber-attack model* that considers the following:

- Realistic attack models that could analyze and simulate realistic attack behaviors

when multiple attacks are executed simultaneously. The model should use AC load flow model for more accurate solutions.

- It is necessary to develop optimization methods that could identify critical components to attack simultaneously in order to maximize the system damage as a result of cascading failures. The optimization method should be capable enough to identify efficient and effective solutions irrespective of the power system size.

- It is necessary to develop optimization methods that could identify critical components to protect considering the most damaging attacks that could improve the resilience of the system. The methods should not be dependent on the scale of the power system and provide effective and efficient solution such that after deploying the defense resources the damage should be minimized when a static cyber-attack is launched.

To improve the power system resilience, it is necessary to identify the critical components to protect considering the financial budget constraints. In our work, we considers that an attacker is resource bounded, i.e., the adversary can attack only a limited number of components due to various reasons such as cost of attack, etc. We follow a *game-theoretic approach* to design an attacker/defender model in power systems. The approach effectively addresses the above mentioned challenges by considering the following:

- A formal model for an attacker is defined, where the cost of attacking any substation is uniform. In this model, the attacker is able to identify the critical substations and its components, i.e., protection assemblies that can be manipulated to disconnect transmission lines from the network to cause cascading failures that maximize system damage based on the attacker's budget.

- An efficient polynomial-time algorithm is presented to identify the *worst-case attack*, i.e., to identify the critical substations and protection assemblies to attack based on the attacker model.

- A formal model for a defender is defined, where the cost of protecting any substation is uniform. In this model, given a defense budget, a defender is able to identify the critical substations to protect in order to minimize system damage from cascading failures when a cyber-attack is launched.

- An efficient polynomial-time algorithm is presented to identify the critical substations to prioritize and defend to minimize damage under *static cyber-attacks*.

Note that in our approach, from the physical systems perspective we do not consider the transient instabilities due to generator out of sync, etc., since the focus is to show the applicability of the approach which can be easily demonstrated using the steady state analysis. From the cyber systems perspective, all the components, i.e., substations and protection assemblies comprising the power system has a uniform cost to attack/defend and the success probability of any attack/defense action is always 1, i.e., any attack or defense on the identified component is considered to be 100% feasible. Further, from the cyber network perspective, we do not consider the actual delay associated with the control commands, however, we are more interested in the final effect of these commands based on an attackers/defenders action.

In order to evaluate the our approach, we first use simulations to demonstrate an attack-/defense scenario that shows how the approach works in general. Next, we use simulations to identify the attacks/defense with different budgets to show how the approach is able to minimize the system damage by intelligently identifying and protecting the critical substations in a power systems network. In addition, we also show that the proposed approach is significantly better than the exhaustive search by comparing the attack/defense execution time and the obtained solution against the exhaustive search.

This is based on the accepted paper in the PES-Innovative Smart Grid and Technology (ISGT) conference. The details of the publication is as below:

Hasan, Saqib, Amin Ghafouri, Abhishek Dubey, Gabor Karsai, and Xenofon Koutsoukos. “Vulnerability analysis of power systems based on cyber-attack and defense mod-

els." In 2018 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), pp. 1-5. IEEE, 2018.

7.2 Introduction

Smart grids are a result of increasing demand for reliable electric energy. The advancement in the grid's technology is responsible for expanding the capabilities of the traditional power grids generation, transmission, and distribution systems. Technologies such as substation automation, phasor measurement units (PMUs), and advanced metering infrastructures (AMIs) are currently deployed to achieve reliable supply for electric power. However, it increases the cyber component in a smart grid, which potentially increases the attack surface. Furthermore, cyber-attacks are documented as one of the major obstacles towards the reliable power system operation[10], [11]. Attackers take advantage of these technological advancements and launch sophisticated attacks causing severe damage to the systems, e.g., recent blackout of Dec 2015 Ukraine[60].

A power network consists of substations, control centers, AMIs etc. The substations have remote terminal units (RTUs) to monitor and control the field devices such as relays, and circuit breakers. These devices can be remotely manipulated to isolate transmission lines from the network during maintenance or faulty conditions[4] that can result in cascading failures. Therefore, RTUs become the primary target for cyber-attacks. The adversary aims to gain complete control of the RTUs and cause severe power system damage by modifying the relay settings, remotely opening circuit breakers, changing measurement data etc. However, the time and effort required in compromising a RTU ensure that an attacker can only access a few RTUs before they are detected[94]. Consequently, strategic attackers try to identify the critical substations to launch a successful attack that maximizes the damage[139].

In order to minimize the damage, defending all the substations simultaneously against cyber-attacks is very difficult given financial budget constraints. Besides, the defense mech-

anisms become more expensive since dedicated IT professionals are required to continuously monitor the system to identify and patch the vulnerabilities for a stable system operation. Therefore, it becomes necessary to intelligently identify a critical set of substations that can be prioritized and protected to minimize the system damage during a cyber-attack.

Previous works [113, 112, 140, 91, 93, 97, 96, 117, 131] have explored the frameworks and cyber-attack models that could simulate and analyze specific type of cyber-physical vulnerabilities in the power system. A framework for modeling cyber-physical switching attacks and man-in-the-middle attack is presented in [113, 112]. A class of attacks impacting physical processes excluding anomalies in the cyber-domain is referenced in [140]. Data integrity attacks and load redistribution attacks are discussed in [91], [96]. In [93], the impact of cyber-attacks on the transient stability of the system is described. An approach to identify and protect a subset of measurements from the adversaries considering false data injection attacks is presented in [97]. Node overloading attacks due to increasing load resulting in cascading failures are studied in [117]. The above approaches emphasize on specific vulnerabilities, however, they do not focus on system-wide identification of critical components to attack in a power system. This can provide important insight on prioritizing and protecting the substations and its components for improving the overall system resilience. Moreover, unlike our approach they do not consider a system-wide view on defending against these vulnerabilities especially with limited resource availability.

In this paper, we consider a game-theoretic approach to design an attacker / defender model for power systems. A strategic attacker tries to maximize the damage by identifying the worst-case attack whereas the defender tries to minimize the damage by protecting the critical substations. A worst-case attack refers to an attack on a subset of substations that cause maximum system damage. Here, we consider the cyber-attack on substations to gain access to the RTUs where the adversary can open the circuit breakers by manipulating protection assembly control signals resulting in severe cascading failures. Identifying all the possible attack and defense scenarios is computationally infeasible. Therefore, it

is necessary to identify the worst-case attack and defense in an efficient way. The main contributions of the paper are:

- A formal model for an attacker is described, where the cost of attacking any substation is uniform. In this model, the attacker can identify the critical substations and its components that can be manipulated to disconnect transmission lines from the network that maximize system damage based on the attackers budget.
- An efficient polynomial-time algorithm is presented to identify the worst-case attack based on the attacker model.
- A formal model for a defender is described, where the cost of protecting any substation is uniform. In this model, given a defense budget, a defender can identify the critical substations to protect in order to minimize system damage during a cyber-attack.
- An efficient polynomial-time algorithm is presented to identify the critical substations to defend to minimize damage.
- The case study is performed on the standard IEEE-14, 39, and 57 bus systems[141]. Our results show that the approach captures the worst-case attacks on the power network and effectively uses the defense model to minimize the damage.

The remainder of the paper is organized as follows. Section II presents the attacker model followed by the defender model in Section III. Section IV demonstrates the results. The conclusions are provided in Section V.

7.3 System Model

A power system is a complex network of power generation, delivery, monitoring and control components. The power delivery elements such as transmission lines, buses, and

transformers supply power from the generation points to the loads. However, the monitoring and control devices such as protection assemblies and circuit breakers are responsible for isolating faulty elements from the network during abnormal conditions. Due to the advancement in technology, power networks can be remotely controlled using RTUs and SCADA systems. Attackers may compromise these systems and isolate components from the power network causing cascading outages resulting in severe load loss [60].

We consider a power system G_p, where U is a set of buses, G is a set of generators, T is a set of transformers, L is a set of loads, and P is a set of protection assemblies. The power system is divided into substations. Each substation has its own monitoring and control units referred to as RTUs. Let $S = \{S^i\}_{i=1}^m$ be the set of substations. Each substation consists of a set of protection assemblies from P. We define $F(S^i)$ as a function that returns the set of protection assemblies in a substation S^i. Clearly, the union of all the protection assemblies in every substation represents the set of protection assemblies in the power network, that is, $\cup_{i=1}^m F(S^i) = P$.

7.4 Attacker Model

In this section, we provide the attacker model that could result in maximum load loss in a power network.

7.4.1 Worst-Case Attack

The goal of the malicious attacker is to destabilize the power system and maximize the load loss. The attacker achieves this by gaining access to a subset of substations $S^{'} \subseteq S$. The adversary is resource bounded, i.e., it can compromise at most B_S substations. Then, the attacker identifies the protection assemblies $P^{'} \subseteq F(S^{'})$ that will be manipulated to isolate transmission lines from the power network. Here, the attacker manipulates at most B_P protection assemblies. The budget B_P can represent the maximum number of protection assemblies that the attacker can attack due to a stealthiness criterion. Note that a non-strategic

attacker may choose a large B_P and potentially attack all the protection assemblies within the compromised substations, however, a more strategic attacker may favor a small B_P as the attack may remain undetected for a longer period of time, which could potentially cause more damage. Also, note that manipulating all the protection assemblies of a substation to isolate power lines may not lead to cascading failures resulting in severe load loss due to the reduction in overall system load. We define the attack on a set of substations $S^{'}$ and protection assemblies $P^{'}$ by A_P.

Let the loads in the power network be defined by L_j and current flowing through each load is given by I_j, where $j = 1$ to $n, n \in \mathbb{N}$. Now, the load loss function is computed as below:

$$J(A_P) = \frac{\sum_{j=1}^{n} L_j}{L_T} \times 100, \ \ \forall I_j = 0 \tag{7.1}$$

where L_T is the total system load and A_P is the attack. The problem is formally defined below.

Problem 1 (Worst-Case Attack) *Given a power system network G_p, a substation budget B_S, and a protection assembly budget B_P, find a worst-case attack A_P that maximizes the load loss in the power system network. Formally,*

$$\begin{aligned} &\underset{S^{'}}{\text{argmax}} \ \max_{P^{'} \subseteq F(S^{'})} J(A_P) \\ s.t. \quad &|S^{'}| \leq B_S, \quad |P^{'}| \leq B_P \end{aligned} \tag{7.2}$$

7.4.2 Algorithm for Finding Worst-Case Attack

Using exhaustive search to identify worst-case attack is computationally infeasible due to the combinatorial nature of search space [142]. Hence, we present an efficient Algorithm 8 to find the worst-case attack. The algorithm starts with an empty set and intelligently selects the critical substations one-by-one that cause maximum system damage. Next, from the selected substations, the algorithm iteratively identifies the protection as-

semblies to manipulate. It takes as input the power system model G_p, the substation budget B_S, the protection assembly budget B_P, and the substation and its component information S_P^{info}. Here, substation components refer to the protection assemblies of the substation. Then, it finds the worst-case attack by identifying the critical substations S_w to compromise, the transmission lines T_w corresponding to the protection assemblies that are manipulated to be removed from the network, and the resulting load loss L_w. The substation and

Algorithm 3 Algorithm for Finding Worst-Case Attack

Input: $G_p, B_S, B_P, S_P^{info}$
Initialize: $L_w \leftarrow 0, T_w \leftarrow \emptyset, S_w \leftarrow \emptyset, L_g \leftarrow 0$
for $j = 1, \ldots, B_S$ **do**
    if $S_w = \emptyset$ **then**
        $\hat{S} \leftarrow$ `Substation_comps`$(S_P^{info}, \emptyset)$
    else
        $\hat{S} \leftarrow$ `Substation_comps`(S_P^{info}, S_w)
    end if
    for all $s \in \hat{S}$ **do**
        $P_t \leftarrow F(s)$
        $T_P, L_P \leftarrow$ `Worst_Attack`(G_p, P_t, B_P)
        if $L_P > L_w$ **then**
            $L_w \leftarrow L_P$, $T_w \leftarrow T_P$, $S_w \leftarrow s$
        end if
    end for
    if $(L_g - L_w) \leq \varepsilon$ **then**
        break
    else
        $L_g \leftarrow L_w$
    end if
end for
return S_w, T_w, L_w

its components i.e., protection assemblies is denoted by $\hat{S}$, which represents a hash table. At each iteration j, based on S_w, $\hat{S}$ is obtained by using `Substation_comps()`. If S_w is non-empty, the function selects the substations S_w from the power system that cause maximum load loss in the previous iteration and uses it to obtain a new set of substations to select from that may result in maximum damage in the current iteration j. For each $s \in \hat{S}$, `Worst_Attack`(G_p, P_t, B_P) computes and returns the transmission line outages corresponding to the selected protection assemblies and load loss denoted by T_P, L_P respectively that cause maximum damage. In each iteration j, if $L_P > L_w$ then the solution is updated. The

loop terminates if no further improvement $L_g - L_w$ is observed.

The function `Worst_attack()` is described as Algorithm 6. The algorithm intelligently selects the critical protection assemblies one-by-one to isolate transmission lines that cause maximum load loss (equation 8.1). It takes as input the power system model G_p, substation components P_t, and the protection assembly budget B_P. Further, it identifies the maximum load loss $L_w^{'}$ and the outages $T_w^{'}$. The algorithm starts with an empty set and uses $\texttt{Max\_loss}(G_p, P_t, \emptyset)$ to identify the component outages resulting in maximum damage. This function simulates a set of contingencies, i.e., outages of components and returns the one that cause maximum load loss. The components and corresponding load loss is represented by $T_P^{'}$, $L_P^{'}$ respectively. For each iteration i, $\texttt{Updated\_comps}(P_t, T_P^{'})$ uses $T_P^{'}$ and returns a new set of components $\hat{P}_t$ to be removed from G_p depending upon B_P. $\hat{P}_t$ represents a list of contingencies that are needed to be simulated. Next, $\texttt{Max\_loss}(G_p, P_t, \hat{P}_t)$

Algorithm 4 Algorithm for Worst_Attack() Function

Input: G_p, P_t, B_P
Initialize: $L_w^{'} \leftarrow 0, T_w^{'} \leftarrow \emptyset$
$T_P^{'}, L_P^{'} \leftarrow \texttt{Max\_loss}(G_p, P_t, \emptyset)$
$L_w^{'} \leftarrow L_P^{'}$, $T_w^{'} \leftarrow T_P^{'}$
for $i = 1, \ldots, B_P$ **do**
    $\hat{P}_t \leftarrow \texttt{Updated\_comps}(P_t, T_P^{'})$
    $T_P^{'}, L_P^{'} \leftarrow \texttt{Max\_loss}(G_p, P_t, \hat{P}_t)$
    if $L_P^{'} > L_w^{'}$ **then**
        $L_w^{'} \leftarrow L_P^{'}$, $T_w^{'} \leftarrow T_P^{'}$
    end if
end for
return $T_w^{'}, L_w^{'}$

uses the updated component list $\hat{P}_t$ to identify the maximum load loss causing components in the current iteration i. In each iteration i, if the load loss $L_P^{'}$ is greater than the maximum load loss $L_w^{'}$ then the solution is updated. The worst-case running time of Algorithm 8 is $O(|S| \times |B_S| \times |P| \times |B_P|)$, which is non-exponential.

7.5 Defender Model

In this section, we provide the defender model to improve the power system resilience by minimizing the load loss. Here, based on the attack on the substations and its components, a set of critical substations to be protected is identified.

7.5.1 Defender's Problem

The goal of the defender is to improve the system resilience and minimize the load loss possible. A defender achieves this by protecting a subset of substations D_S from the total number of substations S, i.e., $D_S \subseteq S$. The defender is resource bounded, i.e., it can protect at most B_D substations. The substations can be protected using various methods such as better firewall protection against intrusion, application whitelisting, network segmentation[143]. Note that this model can provide important insight upon which substations can be upgraded first considering financial budget constraints and the worst-case attack. The problem is formally defined below.

Problem 2 (Defender's Problem) *Given a power system G_p and a defense budget B_D, find a defense strategy D_P that minimizes the load loss in the power network. Formally,*

$$\begin{aligned} &\underset{D_S}{\operatorname{argmin}} \max_{S' \subseteq S - D_S} \max_{P' \subseteq F(S')} J(A_P) \\ s.t. \quad &|D_S| \leq B_D, \quad |S^{'}| \leq B_S, \quad |P^{'}| \leq B_P \end{aligned} \tag{7.3}$$

7.5.2 Algorithm for Finding the Critical Substations to Protect

Using exhaustive search to identify the critical substations to protect is computationally infeasible due to the combinatorial nature of search space[142]. Hence, we present an efficient Algorithm 9 to find the set of critical substations to protect. The algorithm starts with an empty set and intelligently selects the critical substations one-by-one to protect that minimizes system damage. It takes the power system model G_p, the substation budget

B_S, the protection assembly budget B_P, and the defense budget B_D as inputs. Further, it identifies the critical substations S_D to be protected to minimize the load loss during an attack.

Algorithm 5 Algorithm to Find Critical Substations to Protect

Input: G_p, B_S, B_P, B_D
Initialize: $S_d^{'} \leftarrow \emptyset, S_D \leftarrow \emptyset, L_w \leftarrow 100$
$\hat{T}_w, \hat{L}_w, \hat{S}_w \leftarrow$ `Get_Attack`$(G_p, B_S, B_P, \emptyset, \emptyset)$
for $i = 1, \ldots, B_D$ **do**
    $L_w \leftarrow 100$
    if $S_D \neq \emptyset$ **then**
        $\hat{S}_w \leftarrow$ `Get_Attack`$(G_p, B_S, B_P, S_D, \emptyset)$
    end if
    for all $s \in \hat{S}_w$ **do**
        $\hat{T}_w, \hat{L}_w, S_{sub} \leftarrow$ `Get_Attack`(G_p, B_S, B_P, S_D, s)
        if $\hat{L}_w < L_w$ **then**
            $L_w \leftarrow \hat{L}_w$, $S_d^{'} \leftarrow s$
        end if
    end for
    $S_D \leftarrow S_D \cup S_d^{'}$
end for
return S_D

First, the worst-case attack with no defense is obtained using `Get_Attack`$(G_p, B_S, B_P, \emptyset, \emptyset)$, which is same as Algorithm 8. It provides the substations $\hat{S}_w$ to compromise that maximizes the damage. From the identified worst-case attack, the substation to be protected is identified using `Get_Attack`(G_p, B_S, B_P, S_D, s). This function is similar to Algorithm 8, however, it computes the worst-case attack after removing the substations to be protected S_D and the substation s that belongs to $\hat{S}_w$ from the attackers list of attackable substations. For each iteration i and for each $s \in \hat{S}_w$, the substation to be protected that minimizes the load loss, i.e., if $\hat{L}_w < L_w$ is identified and the solution is updated, i.e., $S_D \leftarrow S_D \cup S_d^{'}$. Next, depending upon B_D, for each iteration i, the new set of critical substations to compromise $\hat{S}_w$ is obtained using `Get_Attack`$(G_p, B_S, B_P, S_D, \emptyset)$ based on S_D. This function returns the new worst-case attack by considering only the substations that are not protected. The worst-case running time of Algorithm 9 is $O(|S| \times |B_D| \times |S| \times |B_S| \times |P| \times |B_P|)$, which is non-exponential.

Table 7.1: IEEE-14 Bus System Attack-Defense Scenario

Attack Budget (B_S)	B_P	Defense Budget (B_D)	Pre-Defense Load Loss	Post-Defense Load Loss	Substations Attacked	Substations Defended	Improvement (%)
2	2	3	51.17	48.30	S7	S4, S3, S2	5.61
2	2	4	51.17	43.46	S1, S6	S4, S3, S2, S7	15.07
2	2	5	51.17	29.55	S8, S9	S4, S3, S2, S7, S6	42.25
2	2	6	51.17	21.84	S5, S10	S4, S3, S2, S7, S6, S9	57.31

7.6 Evaluation

To evaluate the developed algorithms, we apply them to the standard IEEE-14, 39, and 57 bus systems. We used a steady state simulator discussed in [142] for our analysis.

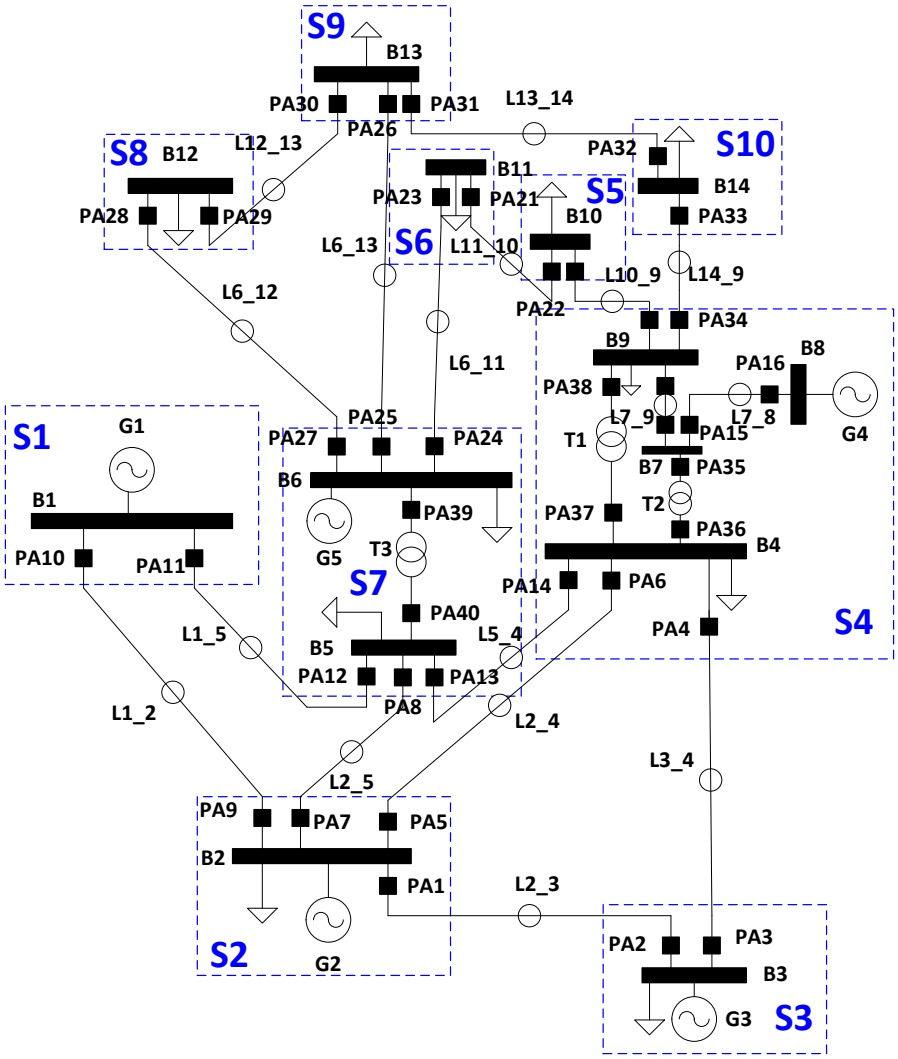


Figure 7.1: IEEE-14 Bus System[1]

First, we discuss an attack-defense scenario for the standard IEEE-14 bus system shown in Figure 7.1. The blue colored dotted lines represent the substations denoted by S1 , . . . , S10. Each transmission line is protected by a pair of protection assembly denoted by PAn, where n $\in \mathbb{N}$. These protection assemblies within the substations can be manipulated to open the circuit breakers that can disconnect the transmission lines from the network to initiate the cascading failures causing severe damage to the power network. Table 7.1 shows the details of the performed case study. The attack budget for the system is assumed to be 2. However, the defense budget is increased in steps up to a total of 6 substations. From Table 7.1, it is clear that the load loss for the IEEE-14 bus system is significantly minimized by intelligently selecting the substations to be protected. Moreover, with an in-

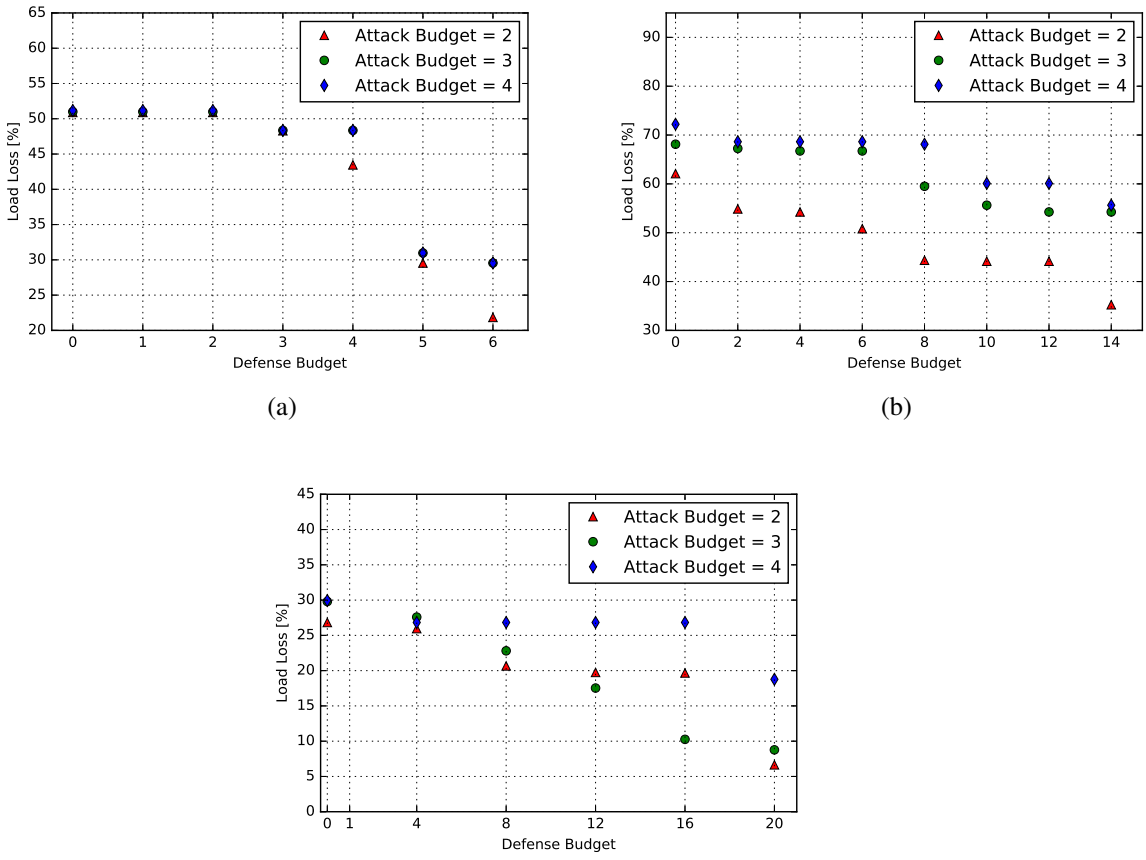


Figure 7.2: Load loss as a function of various attack and defense budgets for (a) IEEE-14 bus system, (b) IEEE-39 bus system, (c) IEEE-57 bus system.

crease in the defense budget, a total of 57.31% improvement in load loss is observed. The substations that are attacked and defended are mentioned in Table 7.1. Similar results for IEEE-39, 57 bus systems can be obtained using the developed models.

Now, we identify the worst-case attack and defense for the three standard IEEE systems. Figure 7.2 represents the load loss as a function of various attack and defense budgets. In each figure the x-axis represents the defense budget and the y-axis represents the overall system load loss. Red, green and blue colored markers represent attack budget 2, 3 and 4 respectively. The respective colored markers at defense budget '0' corresponds to the load loss with no defense. From Figure 7.2, it is clear that by carefully selecting the substations to be protected and with increase in the defense budget the overall system loss is significantly minimized and the adversary is unable to maximize the damage even with increase in the attack budget. In our analysis, we choose a defense budget of 0-50% of the total number of substations for each system.

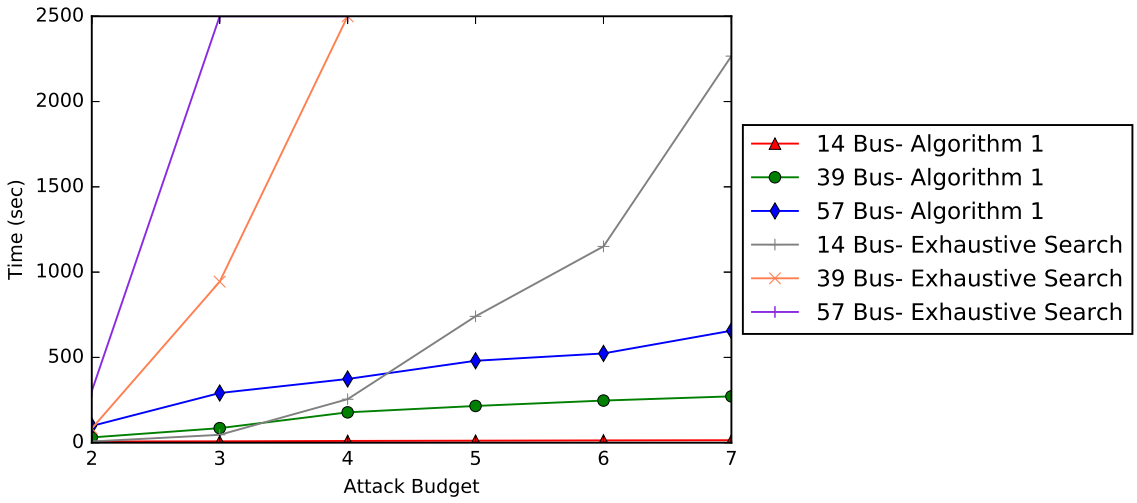


Figure 7.3: Attack analysis execution time

Further, we discuss the time taken to identify the worst-case attack and defense shown in Figures 7.3 and 7.4 respectively. In each figure, the x-axis represents the attack or defense budget whereas the y-axis represents the time. Red, green and blue lines represent the worst-case attack or defense identification time for IEEE-14, 39, and 57 bus systems

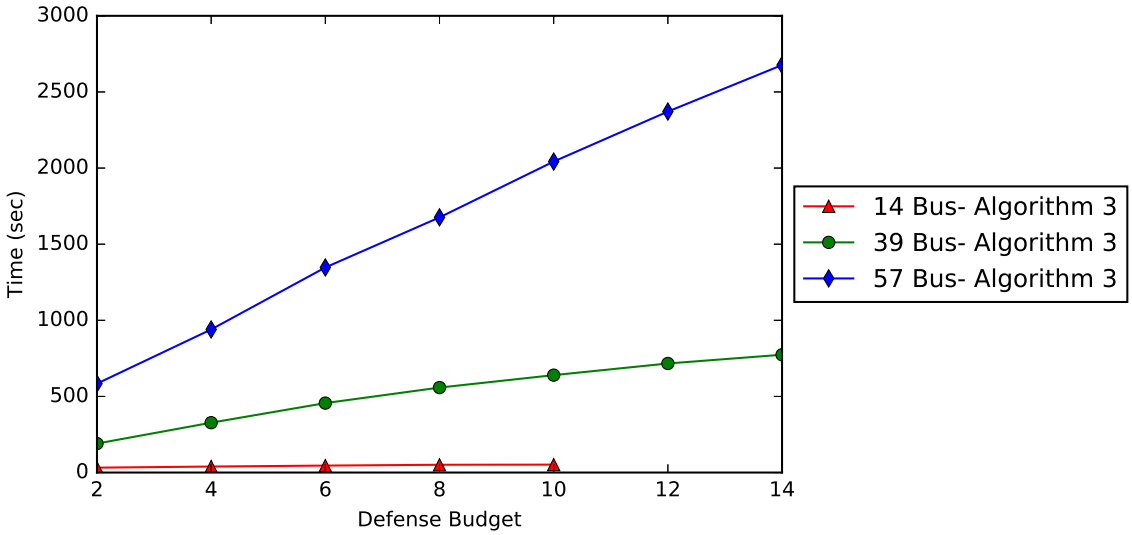


Figure 7.4: Defense execution time

respectively. From Figure 7.3, it is clear that as the attack budget increases, the time taken to obtain the worst-case attack increases marginally for the three systems which is insignificant when compared with the time taken by exhaustive search for 14, 39 and 57 bus systems shown by gray, coral, and violet colored lines respectively. The algorithm also provides mostly the same solution for these systems as the exhaustive search. Similar analysis can be performed for defense scenario. Figure 7.4 shows that as the defense budget increases, the time taken to identify the critical set of substations to be protected increases slightly which is again inconsiderable if compared with the exhaustive search. It clearly shows that our algorithms perform much better than the exhaustive search. This is mainly because of the fact that in each iteration of these algorithms, our search is guided intelligently to significantly reduce the search space in order to obtain an efficient and effective solution.

7.7 Conclusions

The attacker and defender models along with the algorithms to obtain the worst-case attack and defense were developed. The main idea of the attacker model is to select the subset of substations causing maximum system damage when compromised by an adversary.

However, the defender model operates to protect the subset of substations that minimize the system damage. The case study on IEEE systems showed how the damage to the power network can be significantly reduced by intelligently selecting a subset of substations to protect, given a defense budget. Under financial budget constraints, prioritizing and protecting the critical substations can greatly increase system resilience. Moreover, these algorithms can be easily applied to larger systems with higher attack and defense budgets. As part of the future work, these models can be applied to changing network topologies to provide online solutions for prioritizing defense resources.

Chapter 8

Modeling and Analysis of Dynamic Cyber-Physical Attacks

8.1 Problem

With technological advancements, the *cyber-attacks* are increasing both in number and sophistication. *Dynamic cyber-attacks* are a great example of a sophisticated attack, where the attacker is assumed to have knowledge about the power system so that they can dynamically launch an attack for causing severe system damage. The key idea is that cascading failures usually takes at least a few minutes and sometimes up to hours to progress and cause catastrophic damage. Attackers can take advantage of this mechanism and schedule the attacks dynamically to cause severe system damage as against their static counter parts. A motivating example is demonstrated to explain the concept in detail in Section 8.3. There are several challenges that needs to be addressed while considering the analysis of *dynamic attacks*.

- First, realistic *dynamic cyber-attacks models* needs to be developed for analyzing cascading failures that can cause severe system damage which can be very challenging. The model should use AC load flow model for more accurate solutions.
- Identifying all the high impact strategically timed *dynamic attacks* becomes computationally infeasible for large power networks. Hence, it is necessary to develop optimization methods that could identify critical components to attack that can be scheduled appropriately in order to maximize the system damage as a result of cascading failures. The optimization method should be capable enough to identify efficient and effective solutions irrespective of the power system size.
- It is necessary to develop optimization methods that could identify critical compo-

nents to protect considering the most damaging *dynamic attacks* that could improve the resilience of the system. The methods should not be dependent on the scale of the power system and provide effective and efficient solution such that after deploying the defense resources the damage should be minimized when a *dynamic cyber-attack* is launched.

In order to improve the power system resilience, we have developed a *game-theoretic approach* for designing *cyber-attack* and *defense* models in power systems. Here the attacker is resource bounded, i.e., the adversary can attack only a limited number of components due to various reasons such as cost of attack, etc. and can schedule the attacks to maximize the system damage. On the other hand, the defender is resource bounded, i.e., he can only protect a limited number of nodes/components due to reasons such as limited financial budget, etc. Our approach includes the following that effectively addresses the challenges mentioned in the problem statement:

- A formal dynamic attacker model is developed, where the attacking cost of any substation and its components, i.e., protection assemblies is uniform. In this model, the attacker will be able to strategically identify the critical substations and its components that can be intelligently attacked at different time instants in order to isolate the transmission lines from the power network to maximize the system damage depending upon the attacker's budget.

- An efficient polynomial-time algorithm is designed to intelligently identify the most damaging strategically timed *dynamic attack* by selecting critical substations and its components to attack based on the *dynamic attack* model.

- A formal dynamic defense model is developed, where the protection cost of any substation is uniform. In this model, given a defense budget, a defender can strategically identify the most critical substations to prioritize and protect so as to minimize the overall system damage when an attacker launches a *dynamic attack*.

- An efficient polynomial-time algorithm is introduced to strategically identify the critical substations and its components to defend against the *dynamic attack* in order to minimize the overall system damage when an attacker launches a dynamic attack.

In order to evaluate our approach, we first use simulations to demonstrate how random *static attacks* can be scheduled optimally to cause higher system damage. Next, we use simulations to identify the *worst-case static attacks* and find optimal *dynamic attack* sequence to maximize the system damage at different attack budgets. In addition, we demonstrate a static/dynamic attack scenario to show how the attacks can be optimized by intelligently initiating them at different time instants. Further, we show how the optimal use of the limited defense resources can be effective in minimizing the system damage by selecting the substations to protect. Finally, we also show that the developed approach is significantly better than the exhaustive search by comparing the attack/defense execution time against the exhaustive search.

This is based on the submitted paper in the International Journal of Electrical Power & Energy Systems. The details of the publication is as below:

Hasan, Saqib, Abhishek Dubey, Gabor Karsai, and Xenofon Koutsoukos. "A Game-Theoretic Approach for Power Systems Defense Against Dynamic Cyber-Attacks." 'Submitted, In Review' In Elsevier-International Journal of Electrical Power & Energy Systems, 2019.

8.2 Introduction

Recent studies by the National Electric Research Council (NERC) documented that malicious attacks on power grids are much more devastating than the destruction caused by natural calamities [83] and can be instigated through cyber penetration [84] or physical obstruction [85] resulting in large blackouts. Today, power system resilience considering cyber-security has gained significant attention [89] as cyber-attacks are increasing both in number as well as sophistication and are considered as one of the major obstacles towards

the reliable system operations [88, 11, 10, 86]. For instance, due to the technological transformation of the traditional power grids into smart grids, power systems employ a large number of sophisticated and autonomous components such as protection devices, phasor measurement units (PMUs), remote terminal units (RTUs), etc. This increases the potential attack surface by giving rise to new vulnerabilities [3].

The attackers take advantage of such cyber components and gain access to the network by compromising the firewall and can launch catastrophic attacks, compromising system reliability [90] e.g., the recent Ukraine 2015 cyber-attack[60]. What makes the problem worse is the fact that most operators follow the guidelines from NERC [6] requiring only $N-2$ reliability criterion [144], since analysis of higher order contingencies is computationally hard [7, 142], however, a cyber-attack is not limited to only two component failures.

Given such challenges, it is crucial to not only analyze a power system topology for reliability failures but it is also important to analyze the effect of cyber-attacks. In principle this can be approached by considering static attacks, where the devices are affected simultaneously or by dynamically sequenced attacks, which as shown in this paper, can cause significantly higher damage as compared to their static counterparts. Therefore, methods to study dynamic attack are important.

Several frameworks and attack models have been developed to study security vulnerabilities [113, 112, 140, 91, 97, 96, 117, 93, 131, 106, 145]. A man-in-the middle attack and modeling of cyber-physical switching attacks are presented in [113, 112]. Several data integrity attack studies the effect of manipulating control messages, measurement data in[91, 97]. A special type of false data injection attack, i.e., load redistribution (LR) attack is presented in [96, 117]. The effect of cyber-attack on the voltage stability of support devices is provided in [93]. The work in [131] considers cyber-failures in protection assemblies and provides a platform to obtain new cascading traces. A real-time cyber-physical system testbed that provides mitigation strategies against attacks is discussed in

[106]. Additionally, a number of game-theoretic approach based studies exist. For example, an efficient algorithm to solve the defender-attacker-defender problem for system protection is discussed in[109]. In [146], the authors formulate the problem as a minmax non-cooperative game and solved it using genetic algorithm. Moreover, the work in[111] formulates the coordinated attacks on power systems as a bi-level optimization problem. The authors in [116] consider coordinated multi-switch attacks that leads to cascading failures in smart grid. In [101], the authors studied a joint substation-transmission line vulnerability and proposed a component interdependency graph based attack strategy. Based on false data injection attacks, a Markov security game for attacks on automatic generation control is formulated in [110] and a time synchronization based attack is presented in [102]. Further, in [94] the effect of false data injection attacks against state estimation in power grids are studied. Finally, the work in [115, 14] studies the temporal features of attacks in power systems.

However, there are several limitations in these approaches. The frameworks in [113, 112, 106] do not consider a system-wide identification of critical components to compromise. Attack models and strategies referenced in [140, 97, 96, 117, 93, 109, 146, 111, 101, 110, 102, 116, 94] focus on simultaneous attacks on different aspects of the system such as opening of circuit breakers, false data injection attacks in monitoring components, etc. However, none of these approaches consider cyber-attacks from the perspective of time domain, which is a vital facet in cascading failures since the progression of such failures takes at least minutes[12] or at times hours[147]. An attacker can easily and realistically sequence these attacks in a stealthy manner such that the attack mimics the trace of a normal cascading failure that could easily misguide the system operators. Moreover, considering strategically timed cyber-attacks reveal new system vulnerabilities which can not be found using previous approaches and their identification can enhance the overall power system resilience. Further, the attack model in[14] is based on the constructed sequential attack graph (SAG) which can be computationally infeasible for large power networks and most

of them do not provide any defense model.

In this paper, we consider a game-theoretic approach to design attacker-defender cyber-attack and defense models for power systems to identify the worst-case dynamic attack. This work proposes a much simpler approach which does not require the construction of complex SAG as required by [14]. Further, we do not choose attacks based on node degree or load which enables us to explore a wider attack area. The specific contributions are:

- A formal dynamic attack model is described, where the attacking cost of any substation and their components is uniform. In this model, the attacker can strategically identify the critical substations and its components to attack at different time instants in order to maximize the system damage constrained by the attacker's budget.
- A formal dynamic defense model is described, where the protection cost of any substation is uniform. In this model, given a defense budget, a defender can strategically identify the most critical substations to prioritize and protect so as to minimize the overall system damage.
- Two efficient polynomial-time algorithms are introduced to identify both the worst-case dynamic attack and a defense strategy which minimizes overall system damage.

Our results (shown using IEEE 39 and 57 bus examples) demonstrate that the approach captures the worst-case dynamic attacks on the power system networks and effectively uses the dynamic defense model to minimize the overall system damage. It also proves the effectiveness and efficiency of our algorithms. Moreover, the attack algorithm is able to maximize the system damage for both static and random attacks.

The remainder of this paper is organized as follows. The system model along with a motivating example is discussed in Section II. Section III and IV give a detailed formal description of the static attack and defense models. The dynamic attack and defense models along with their algorithms are formally presented in Section V and VI. Results are discussed in Section VII followed by the conclusions in Section VIII.

Table 8.1: List of Commonly Used Symbols

Symbol	Description
General Symbols	
S	set of substations
P	set of protection assemblies in a power system
S^i	i^{th} substation in S
$F(S^i)$	function that returns the set of protection assemblies in substation S^i
B_S	substation attack budget
B_P	protection assemblies attack budget
B_D	substation defense budget
Static, Dynamic Attack and Defense Model	
$S^{'}$	set of substations selected from S for static attack
$P^{'}$	set of protection assemblies selected from $P^{'}$ for static attack
$A_{P'}$	static attack on substations $S^{'}$ and protection assemblies $P^{'}$
k	time instant in $\{1,\ldots,T\}$
$S^{'}(k)$	set of substations selected from S for dynamic attack
$P^{'}(k)$	set of protection assemblies selected from P for dynamic attack
$A_{P'}(k)$	dynamic attack on substations $S^{'}(k)$ and protection assemblies $P^{'}(k)$
$x(k)$	state of the system at k^{th} time instant
$H(k)$	attack history of the system $\mathcal{G}_{\mathcal{P}}$
$G(H(k))$	function representing the power system state under the presence of attack history $H(k)$ at time step k
$g(H(k))$	function representing nominal system state with no attack history
D_S	set of protected substations

Table 8.2: List of Methods

Method Name	Use
`Gen_Contin`$(S_{info}, \hat{P}_a)$	Returns the set of contingencies based on the protection assemblies in S_{info}, and $\hat{P}_a$
`Simulate_Model`$(\mathcal{G}_\mathcal{P})$	Simulates the nominal state of the power system model $\mathcal{G}_\mathcal{P}$
`Isolate_Branches`$(\mathcal{G}_\mathcal{P}, p)$	Removes branch(es) from the power system model $\mathcal{G}_\mathcal{P}$ associated with the attacked protection assemblies p
`Simulate_Contin`$(\mathcal{G}_\mathcal{P}, p, k)$	Simulates the power system model $\mathcal{G}_\mathcal{P}$ with branch(es) removal at specific time instants k
`Get_Branches`$(\mathcal{G}_\mathcal{P}, p)$	Returns the overloaded branches in the power system model $\mathcal{G}_\mathcal{P}$ post attack
`Get_Loads`$(\mathcal{G}_\mathcal{P}, p, k)$	Returns the load names l that are disconnected in the power system model $\mathcal{G}_\mathcal{P}$ post attack
`Get_Damage`$(\mathcal{G}_\mathcal{P}, l)$	Returns the overall damage in the power system model $\mathcal{G}_\mathcal{P}$ post attack
`Obtain_Subs`(S_{info}, p)	Returns the substation(s) corresponding to the attacked protection assemblies p in the power system model $\mathcal{G}_\mathcal{P}$

8.3 System Model and Motivating Example

We consider a power system $\mathcal{G}_\mathcal{P}$, where U is a set of buses, G is a set of generators, R is a set of transmission lines, L is a set of loads, and P is a set of protection assemblies.

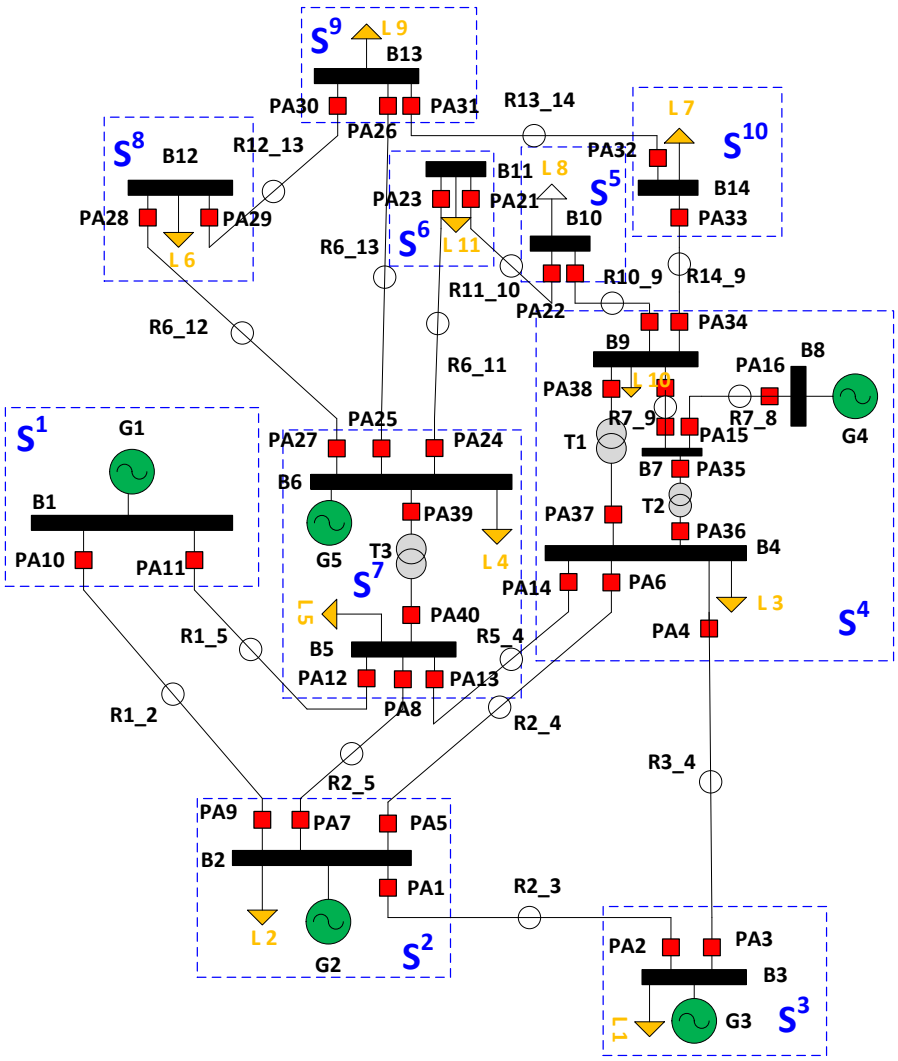


Figure 8.1: IEEE-14 Bus System[1]

The power system is divided into substations. Each substation has its own monitoring and control units referred to as RTUs. Let $S = \{S^i\}_{i=1}^{m}$ be the set of substations. Each substation consists of a set of protection assemblies from P. We define $F(S^i)$ as a function that returns the set of protection assemblies in a substation S^i. Clearly, the union of all the protection assemblies in every substation represents the set of P in the power network, that is, $\bigcup_{i=1}^{m} F(S^i) = P$. The symbols used have been summarized in Table 8.1. Table 8.2 describe the main subroutines referred later in the algorithm sections.

Let us consider an IEEE-14 bus system [1] as shown in Figure 8.1 to demonstrate the

concept of static and dynamic attack. The system is divided into substations shown by blue dotted rectangles labeled as S^n, where $n \in \mathbb{N}$. The protection assemblies within the substations are labeled as PAn. The transmission lines labeled as 'Rn_m' can be isolated by manipulating the protection assemblies associated with it. Now consider the static attack scenario where the protection assemblies associated with the transmission lines 'R6_13' and 'R7_8' are manipulated to isolate them from the power network simultaneously. This led to removal of lines 'R9_14', 'R6_12', 'R9_10', 'R12_13' and loads 'L 5, L9, L4, and L7' from the power network due to subsequent system overloading. Now, in case of dynamic attack, only transmission line 'R6_13' is isolated initially which causes a cascading failure in lines 'R12_13', 'R9_14' and 'R6_12' due to overloading of these lines. The transmission lines are isolated and at this time another attack is executed, i.e., transmission line 'R7_8' is isolated which results in the outage of lines 'R10_11', and 'R9_10' in the subsequent cascading stage. Post dynamic attack, the system lost a total of five loads namely 'L 5, L8, L9, L4, and L 7' as opposed to 'L 5, L9, L4, and L7' in the static attack scenario. This is obviously a higher damage as compared to the static attack considering the same attacks are executed with a difference in the attack execution time and provides the motivation to the problem.

8.4 Static Attack Model

In this section, we first formulate the static attack model that aims to maximize the load loss in the power system network. Then, we provide an efficient algorithm to identify the worst-case static attack.

8.4.1 Worst-Case Static Attack

The objective of the malevolent attacker is to maximize the load loss and destabilize the power system network. In order to achieve this, first the attacker may gain access to a subset of substations $S^{'} \subseteq S$ where the attacker is resource bounded, i.e., the attacker can

compromise at most B_S substations. Now, the adversary can identify the protection assemblies $P^{'} \subseteq F(S^{'})$ to manipulate in order to isolate the transmission lines from the power network. Note that these protection assemblies belong to the selected substations $S^{'}$. The attacker is again resource bounded and can attack at most B_P protection assemblies. The budget B_P can represent the maximum number of protection assemblies that the attacker can compromise due to a stealthiness criterion. It is important to note that a naive attacker may select a large B_P and probably attack all the protection assemblies within the compromised substations, whereas, a strategic attacker may favor a small B_P as it would enable the attacker to remain undetected for a considerably longer period of time. Remaining undetected for a longer time could provide the attacker with an opportunity to potentially cause more damage to the power system network.

Additionally, note that transmission lines are rated to carry a maximum amount of power in order to satisfy the thermal limits criterion[148]. In the presence of power flow violations, the protection assembly associated with the transmission line isolates it from the rest of the power network to avoid abnormal system conditions. This action often results in cascading failures causing severe load loss. Manipulating all the protection assemblies of a substation to disconnect power transmission lines may reduce the overall power flowing through the rest of the active transmission lines due to the reduction in overall system load. As a result, this may not lead to severe cascading failures causing higher load loss. Note that, similar assumptions hold true for the dynamic attack model. Next, the attack on a set of substations $S^{'}$ and protection assemblies $P^{'}$ is denoted by $A_{P^{'}}$.

Let L_j denote the j^{th} load in the power system network $\mathcal{G}_{\mathcal{P}}$. The current flowing through each load L_j is given by I_j, where $j = 1$ to $n, n \in \mathbb{N}$. Now, we compute the damage/load loss function for the static attack model as below:

$$J(A_{P^{'}}) = \frac{\sum_{j=1}^{n} L_j}{L_T} \times 100, \;\; \forall I_j = 0 \tag{8.1}$$

where L_T represents the total system load. The problem is formally defined below.

Problem 3 (Worst-Case Static Attack) *Given a power system network $\mathcal{G}_\mathcal{P}$, a substation budget B_S, and a protection assembly budget B_P, find a worst-case static attack $A_{P'}$ that maximizes the damage/load loss in the power system network. Formally,*

$$\underset{S'}{\text{argmax}} \max_{P' \subseteq F(S')} J(A_{P'}) \tag{8.2}$$

$$\begin{gathered} |S^{'}| \leq B_S \\ \forall S^{'}, S^{''} \in S : S^{'} \cap S^{''} = \emptyset \end{gathered} \tag{8.3}$$

$$\begin{gathered} |P^{'}| \leq B_P \\ \forall P^{'}, P^{''} \in P : P^{'} \cap P^{''} = \emptyset \end{gathered} \tag{8.4}$$

$$B_S \leq B_P \tag{8.5}$$

8.4.2 Algorithm for Finding Worst-Case Static Attack

This section describes the algorithm for finding the worst-case static attack in detail. `Get_WSA`$(\mathcal{G}_\mathcal{P}, B_P, S_{info})$: Algorithm 6 is based on iteratively identifying attacks that maximize the system damage depending upon the budget constraints, i.e., B_S and B_P. Here, the algorithm intelligently selects the protection assemblies to manipulate one-by-one that maximizes power system damage and maps it back to their respective substations. This approach reduces the overall run time of the algorithm. It takes as inputs the power system model $\mathcal{G}_\mathcal{P}$, protection assemblies budget B_P, and power system substation configuration information S_{info}. Further, it identifies the worst-case static attack by identifying a set of critical substations to compromise $S^{'}$, the protection assemblies to manipulate $P^{'}$ and the damage caused by the attack L_w.

As a first step, the algorithm identifies the maximum damage causing protection assemblies that can be manipulated from the entire set of protection assemblies using the method

Algorithm 6 Algorithm for Finding Worst-Case Static Attack: $\texttt{Get\_WSA}(\mathcal{G}_{\mathcal{P}}, B_P, S_{info})$

Input: $\mathcal{G}_{\mathcal{P}}, B_P, S_{info}$
Initialize: $L_w \leftarrow 0, P^{'} \leftarrow \emptyset, S^{'} \leftarrow \emptyset, L_g \leftarrow 0$
$P_t \leftarrow F(s)$
$\hat{P}, L_P \leftarrow \texttt{Get\_Static\_Attack}(\mathcal{G}_{\mathcal{P}}, P_t)$
$L_w \leftarrow L_P, P^{'} \leftarrow \hat{P}$
for $k = 2, \ldots, B_P$ **do**
    $\hat{P}_t \leftarrow \texttt{Get\_Contin}(S_{info}, \hat{P})$
    $\hat{P}, L_P \leftarrow \texttt{Get\_Static\_Attack}(\mathcal{G}_{\mathcal{P}}, \hat{P}_t)$
    if $L_P > L_w$ **then**
        $L_w \leftarrow L_P$, $P^{'} \leftarrow \hat{P}$
    end if
    if $(L_g - L_w) \leq \varepsilon$ **then**
        break
    else
        $L_g \leftarrow L_w$
    end if
end for
$S^{'} \leftarrow \texttt{Obtain\_subs}(S_{info}, P^{'})$
return $S^{'}, P^{'}, L_w$

$\texttt{Get\_Static\_Attack}(\mathcal{G}_{\mathcal{P}}, P_t)$. The set of entire protection assemblies can be obtained by using the function $F(S)$. $\texttt{Get\_Static\_Attack}(\mathcal{G}_{\mathcal{P}}, P_t)$ is similar to Algorithm 9, however, we do not consider the attack execution at different time instants in this algorithm, i.e., all the attacks take place at the same time. Further, for every following iteration, the algorithm identifies the new set of attackable protection assemblies. For instance, let S_{info} be the set that represents the information about the substations and its protection assemblies. If an attacker has attacked a protection assembly $\hat{P}$ from the set of substations S_{info} then in the next iteration, $\texttt{Get\_Contin}(S_{info}, \hat{P})$ uses the $\hat{P}$ to return a new attackable set of protection assemblies such that the attacker can choose only one new protection assembly from the total number of protection assemblies P in S_{info} that has not been previously attacked. Here, the function $\texttt{Get\_Static\_Attack}(\mathcal{G}_{\mathcal{P}}, \hat{P}_t)$ identifies the protection assemblies that cause maximum damage and updates the solution if the damage L_P caused by the selected protection assemblies is greater than the worst-case static damage L_w, where $\hat{P}_t$ represents the set of protection assemblies that are available for the attack. The function $\texttt{Get\_Static\_Attack}(\mathcal{G}_{\mathcal{P}}, \hat{P}_t)$ is similar to Algorithm 9, however, it does not consider the

time vector for scheduling attacks. The algorithm terminates if no further improvement in system damage is observed. At the end, the substations $S^{'}$ that should be compromised in order to maximize system damage corresponding to the attacked protection assemblies are identified through direct mapping using the method `Obtain_subs`$(S_{info}, P^{'})$. The worst-case running time of Algorithm 6 is non-exponential and is given by $O(|P| \times |B_P|)$.

8.5 Static Defense Model

In this section, first we provide the formulation of the defender model to improve the power system resilience by minimizing the damage/load loss. Then, we provide an efficient algorithm for identifying the critical substations to be protected in order to minimize the system damage considering the static attack model. Here, based on the substations and its components i.e. protection assemblies targeted by the attack, a set of critical substations to be protected is identified.

8.5.1 Defender's Problem

The primary goal of a defender is to improve the power system resilience by protecting the critical substations in order to minimize the possible load loss when an attack is launched. To achieve this, the defender can protect a subset of substations D_S from the total number of substations S, i.e., $D_S \subseteq S$. The defender is resource bounded and it can prioritize and protect up to B_D substations due to financial budget constraints. The substation protection from the attacks can be achieved using various methods such as better firewall protection against intrusion, application whitelisting and network segmentation[143]. More importantly while considering financial budget constraints it is impossible to protect and upgrade all the substations simultaneously. Also, a strategic attacker would aim at maximizing the system damage by attacking the most critical substations. Hence, this model can provide important insight upon which substations can be prioritized for the upgrade and protected first against the malicious adversarial attack considering financial budget

constraints and the worst-case static attack. Note that similar assumptions hold true for the dynamic defense model. The problem is formally described below.

Problem 4 (Defender's Problem) *Given a power system network $\mathcal{G}_\mathcal{P}$, a defense budget B_D, a substation budget B_S, a protection assembly budget B_P, find a defense strategy D_P to minimize the load loss when an attacker launches a static attack. Formally,*

$$\underset{D_S}{\operatorname{argmin}} \max_{S' \subseteq S \setminus D_S} \max_{P' \subseteq F(S')} J(A_{P'}) \tag{8.6}$$

$$|D_S| \leq B_D \tag{8.7}$$

$$\begin{gathered} |S^{'}| \leq B_S \\ \forall S^{'}, S^{''} \in S : S^{'} \cap S^{''} = \emptyset \end{gathered} \tag{8.8}$$

$$\begin{gathered} |P^{'}| \leq B_P \\ \forall P^{'}, P^{''} \in P : P^{'} \cap P^{''} = \emptyset \end{gathered} \tag{8.9}$$

$$B_S \leq B_P \tag{8.10}$$

8.5.2 Algorithm for Finding the Critical Substations to Protect

Algorithm 7 starts with an empty set and strategically identifies the critical substations to protect one-by-one such that when an attacker launches an attack the overall system damage can be minimized. The algorithm takes the same inputs as Algorithm 6 with the defense budget B_D as an additional input. It then identifies the critical substations D_S to prioritize and protect to minimize the system damage when a static attack is launched.

First, the worst-case static attack is identified using `Get_WSA`$(\mathcal{G}_\mathcal{P}, B_P, S_{info})$ which is explained in Algorithm 6. Next, for the first iteration when there are no critical substations in D_S^t to protect, we use the critical substations $S_t^{'}$ identified from the worst-case attack to identify the first substation to protect. D_S^t represents the intermediate solution set for substations to be protected in order to obtain a better solution. We iteratively pro-

tect each substation in $S_t^{'}$ and evaluate the overall system damage post static attack using $\texttt{Get\_WSA2}(\mathcal{G}_{\mathcal{P}}, B_P, S_{info}, D_S^t, s)$. The computed system damage in each iteration is used to

Algorithm 7 Algorithm to Find Critical Substations to Protect: $\texttt{Get\_Static\_Defense}(\mathcal{G}_{\mathcal{P}}, B_P, B_D, S_{info})$

Input: $\mathcal{G}_{\mathcal{P}}, B_P, B_D, S_{info}$
Initialize: $D_s^{'} \leftarrow \emptyset, D_S \leftarrow \emptyset, D_S^t \leftarrow \emptyset, \hat{L}_w \leftarrow 100, L_{Prev} \leftarrow 100, L_H \leftarrow \emptyset$
$S_t^{'}, P^{'}, L_w \leftarrow \texttt{Get\_WSA}(\mathcal{G}_{\mathcal{P}}, B_P, S_{info})$
$L_H \cup L_{Prev}$
for $i = 1, \ldots, B_D$ **do**
    $\hat{L}_w \leftarrow 100, flag \leftarrow 0$
    if $D_S^t \neq \emptyset$ **then**
        $S_t^{'}, L_{Prev} \leftarrow \texttt{Get\_WSA1}(\mathcal{G}_{\mathcal{P}}, B_P, S_{info}, D_S^t, \emptyset)$
        $L_H \cup L_{Prev}$
    end if
    for all $s \in S_t^{'}$ **do**
        $S_s^{'}, P_s^{'}, L_s^{'} \leftarrow \texttt{Get\_WSA2}(\mathcal{G}_{\mathcal{P}}, B_P, S_{info}, D_S^t, s)$
        if $L_s^{'} < \hat{L}_w$ **then**
            $\hat{L}_w \leftarrow L_s^{'}, D_s^{'} \leftarrow s, flag \leftarrow 1$
        end if
    end for
    $D_S \leftarrow D_S \cup D_s^{'}, D_S^t \leftarrow D_S^t \cup D_s^{'}$
    if $\hat{L}_w > min(L_H)$ AND $flag == 1$ **then**
        $D_S \leftarrow D_S \setminus D_s^{'}$
    else
        $D_S \leftarrow D_S^t$
    end if
end for
return D_S

select the substation to protect, i.e., $D_S \leftarrow D_S \cup D_s^{'}$, $D_S^t \leftarrow D_S^t \cup D_s^{'}$, where $D_s^{'}$ is the substation that is to be protected and is obtained in the i^{th} iteration. Note that, the function $\texttt{Get\_WSA2}(\mathcal{G}_{\mathcal{P}}, B_P, S_{info}, D_S^t, s)$ is same as Algorithm 6, however, here the worst-case static attack is computed by eliminating the protected substations D_S^t and the substation s from the attackable list of substations, i.e., $S \setminus (D_S^t \cup s)$. If the computed damage $L_s^{'}$ is smaller than the maximum damage $\hat{L}_w$, the solution is updated. Additionally, in each iteration, if the protected substations set D_S^t is non-empty then a new set of critical substations are identified using worst-case static attack function, i.e., $\texttt{Get\_WSA1}(\mathcal{G}_{\mathcal{P}}, B_P, S_{info}, D_S^t, \emptyset)$. This function is also same as Algorithm 6, however, the protected substations D_S^t are removed from the attackable list of substations while executing the worst-case static attack on the power sys-

tem model $\mathcal{G}_{\mathcal{P}}$. It ensures that once the substations are protected, the attacker can only launch the static attack on the remaining substations depending on the attack budget. The obtained attack can further be utilized to identify the substation to protect considering the defense budget constraints. In the algorithm L_H keeps a track of all the previous load losses obtained after protecting the substations in D_S^t and updates the final solution D_S depending upon the comparison of the obtained damage with the previous system damages. This ensures a better protection mechanism that provides an effective solution. The worst-case running time of Algorithm 7 is non-exponential and is given by $O(|S| \times |B_D| \times |P| \times |B_P|)$.

8.6 Dynamic Attack Model

In this section, we first formulate the dynamic attack model that could result in maximum damage/load loss in a power system network. Then, we provide an efficient algorithm for identifying the worst-case dynamic attack that maximizes the system damage.

8.6.1 Worst-Case Dynamic Attack

The objective of the malicious attacker is to destabilize the power system by maximizing the load loss. In order to achieve this, first the attacker can gain access to a subset of substations $S^{'}(k) \subseteq S$ at different time instants k, where $k \in \{1,\ldots,T\}$. The attacker is resource bounded and can compromise up to B_S substations. Next, the adversary can identify the protection assemblies $P^{'}(k) \subseteq F(S^{'}(k))$ to manipulate within the selected substations in order to disconnect transmission lines from the power system network at different time instants k. Here, the attacker is again resource bounded, i.e., it can manipulate at most B_P protection assemblies. Finally, the dynamic attack on a set of substations $S^{'}$ and protection assemblies $P^{'}$ at time step k is denoted by $A_{P^{'}}(k)$.

Let L_j denote the loads in the power system network. The current flowing through each load L_j is given by I_j, where $j = 1$ to $n, n \in \mathbb{N}$. Now, we compute the dynamic attack

damage/load loss function as below:

$$J(A_{P'}(k), x(k)) = \frac{\sum_{j=1}^{n} L_j(k)}{L_T} \times 100, \;\; \forall I_j(k) = 0 \tag{8.11}$$

where $k \in \{1, \ldots, T\}$, $x(k)$, L_T, and $A_{P'}(k)$ represents the time step, system state, total system load, and the attack at time step k respectively. The problem is formally defined below.

Problem 5 (Worst-Case Dynamic Attack) *Given a power system network $\mathcal{G}_{\mathcal{P}}$, a substation budget B_S, and a protection assembly budget B_P, find a worst-case dynamic attack $A_{P'}(k)$ that maximizes the damage/load loss in the power system network. Formally,*

$$\underset{\{S'(k)\}_{k=1}^{T}}{\operatorname{argmax}} \; \max_{(\{P'(k) \subseteq F(S'(k))\}_{k=1}^{T})} \sum_{k=1}^{T} J(A_{P'}(k), x(k)) \tag{8.12}$$

$$x(k) = \begin{cases} G(H(k)), & \text{if} \quad H(k) = \{A_{P'}(i)\}_{i=1}^{k-1} \\ g(H(k)), & \text{if} \quad H(k) = \emptyset \end{cases} \tag{8.13}$$

$$\begin{gathered} \sum_{k=1}^{T} |S^{'}(k)| \leq B_S \\ \forall k, k' \in \{1, ..., T\} : S^{'}(k) \cap S^{'}(k') = \emptyset, k \neq k' \end{gathered} \tag{8.14}$$

$$\begin{gathered} \sum_{k=1}^{T} |P^{'}(k)| \leq B_P \\ \forall k, k' \in \{1, ..., T\} : P^{'}(k) \cap P^{'}(k') = \emptyset, k \neq k' \end{gathered} \tag{8.15}$$

$$B_S \leq B_P \tag{8.16}$$

where, $x(k)$ represents the state of the system at time step k and $H(k)$ represents the attack history of the system. $G(H(k))$ is a function that represents the power system state at time step k when there is a history of attack present in $H(k)$ at any time step $k-1$. However, the

function $g(H(k))$ represents the system state under no attack history $H(k) = \emptyset$, i.e., nominal system operation.

8.6.2 Algorithm for Finding Worst-Case Dynamic Attack

This section describes the algorithm for finding the worst-case dynamic attack and the supporting algorithms, i.e., Algorithms 6 and 9 in detail.

- `Get_WDA`$(\mathcal{G}_{\mathcal{P}}, B_P, S_{info}, a_k)$: Algorithm 8 is based on iteratively identifying the attacks at specific instants in time depending upon the budget constraints, i.e., B_S and B_P. It takes as inputs the power system model $\mathcal{G}_{\mathcal{P}}$, protection assemblies budget B_P, power system substation configuration information S_{info}, and the time instant vector a_k at which the attacks can be initiated. The algorithm then identifies the worst-case dynamic damage L_w^d causing attack by finding a set of critical substations to compromise $S^{'}(k)$, a set of protection assemblies to attack $P^{'}(k)$, and the attack vector a_k^d that represents the time instants at which the attack needs to be executed.

  First, we use `Get_WSA`$(\mathcal{G}_{\mathcal{P}}, S_{info}, B_P)$ (explained as Algorithm 6) to identify the worst-case static attack. Here, we identify the maximum damage causing attack that provides the substations to compromise $S^{'}$, and the protection assemblies $P^{'}$ within the substations to manipulate in order to isolate the transmission lines from the power network without considering different time instants, i.e., assuming the attacks take place at the same time. The set of $P^{'}$ is iteratively used to generate a new set of contingencies C using `Gen_Contin`$(P^{'}, P^d)$. The contingencies C are used by `Get_Dynamic_Attack`$(\mathcal{G}_{\mathcal{P}}, C, a_{temp}^d, a_k)$ (explained as Algorithm 9) which returns the maximum damage L^* causing attack consisting of substations and associated protection assemblies P^* and the attack time vector a^*. In each iteration one attack is intelligently identified along with its time instant vector a_{temp}^d and added to the solution. Note that during the contingency generation process, P^* is utilized in

such a way that the search space remain much smaller than the exhaustive search but still effective. In each iteration, if the maximum damage L^* obtained from $\texttt{Get\_Dynamic\_Attack}(\mathcal{G}_\mathcal{P}, C, a^d_{temp}, a_k)$ is larger than the worst-case dynamic damage L^d_w then the solution is updated. At the end, the method $\texttt{Obtain\_subs}(S^{'}, P^{'}(k))$ is used to obtain the direct mapping of the substations to be attacked. This is possible because the corresponding protection assemblies belong to the respective substations. This process reduces unnecessary algorithm run time and still provides effective solution.

Algorithm 8 Algorithm for Finding Worst-Case Dynamic Attack: $\texttt{Get\_WDA}(\mathcal{G}_\mathcal{P}, B_P, S_{info}, a_k)$

Input: $\mathcal{G}_\mathcal{P}, B_P, S_{info}, a_k$
Initialize: $L^d_w \leftarrow 0, P^{'}(k) \leftarrow \emptyset, S^{'}(k) \leftarrow \emptyset, a^d_k \leftarrow \emptyset, a^{'}_k \leftarrow 0$
$S^{'}, P^{'}, L_w \leftarrow \texttt{Get\_WSA}(\mathcal{G}_\mathcal{P}, S_{info}, B_P)$
$S^{'}(k) \leftarrow S^{'}, P^{'}(k) \leftarrow P^{'}, L^d_w \leftarrow L_w$
for all $p \in P^{'}$ **do**
    $a^d_k \leftarrow a^d_k \cup a^{'}_k$
end for
for all $p \in P^{'}$ **do**
    $P^d \leftarrow \emptyset, a^d \leftarrow a^{'}_k, a^d_{temp} \leftarrow a^d$
    $P^d \leftarrow P^d \cup p$
    for $i = 1, \ldots, (|P^{'}|)$ **do**
        $C \leftarrow \texttt{Gen\_Contin}(P^{'}, P^d)$
        $P^*, L^*, a^* \leftarrow \texttt{Get\_Dynamic\_Attack}(\mathcal{G}_\mathcal{P}, C, a^d_{temp}, a_k)$
        $P^d \leftarrow P^*, a^d_{temp} \leftarrow a^*$
        if $L^* \geq L^d_w$ **then**
            $L^d_w \leftarrow L^*, P^{'}(k) \leftarrow P^*, a^d_k \leftarrow a^*$
        end if
    end for
end for
$S^{'}(k) \leftarrow \texttt{Obtain\_subs}(S^{'}, P^{'}(k))$
return $S^{'}(k), P^{'}(k), L^d_w, a^d_k$

- $\texttt{Get\_Dynamic\_Attack}(\mathcal{G}_\mathcal{P}, C, a^d_{temp}, a_k)$: Given a set of contingencies, Algorithm 9, i.e., $\texttt{Get\_Dynamic\_Attack}(\mathcal{G}_\mathcal{P}, C, a^d_{temp}, a_k)$ identifies the protection assemblies one-by-one and the best sequence in which the attack can be executed to maximize the power system damage. The inputs to the algorithm are the power system model $\mathcal{G}_\mathcal{P}$, the set of contingencies C, the time instants of the attack vector a^d_{temp} of the set of

contingencies in C and the time instants a_k at which the next attack needs to be placed that maximizes the damage. The set of contingencies C is a two dimensional matrix that represents the protection assemblies at each (i, j), where i, j denotes the row and column of a 2D matrix respectively that needs to be manipulated in order to isolate transmission lines from the system model $\mathcal{G}_{\mathcal{P}}$. Note that, the attack vector a^d_{temp} of any contingency $C(i, j)$ represents the time instants of the previously attacked protection assemblies in $C(i, j)$. Since protection assemblies are identified one-by-one and added to the solution, the maximum damage causing protection assembly that needs to be identified in any iteration will have an empty time instant ([]) in $C(i, j)$ before the algorithm is executed. Further, for any iteration in Algorithm 8, Algorithm 9 computes the maximum damage causing attack identifying the set of protection assemblies P^* to manipulate within the identified substations $S^{'}$, the system damage L^* caused by the attack, and the time instants a^* at which the attacks at the substations $S^{'}$ needs to be executed.

For each contingency, the algorithm first simulates the power system model in its nominal state, i.e., without any attack. Then, depending upon a contingency $C(i, j)$ and the attack vector a^d_{temp}, all the transmission lines associated with $C(i, j)$ are removed from the power network for which the time instants are '0', i.e., initial attack. The power system $\mathcal{G}_{\mathcal{P}}$ is then simulated with the initial attack and is further evaluated for the secondary effects of this attack, i.e., additional system overloads. If there are any overloaded transmission lines they are identified and removed from the power network. Additionally, if there are any other attacks in $C(i, j)$ that are available to be executed using the attack vector a^d_{temp} at any other time instants they are also identified and executed. Next, the algorithm uses the time instant vector a_k to manipulate the protection assembly with empty time instant to isolate the associated transmission line such that it maximizes the system damage. The power system model is then simulated with the contingencies $(P_{C(i,j)} \cup P_{a_k})$ and its associated attack vector

Algorithm 9 Algorithm for Finding Dynamic Attack: `Get_Dynamic_Attack`$(\mathcal{G}_{\mathcal{P}}, C, a^d_{temp}, a_k)$

Input: $\mathcal{G}_{\mathcal{P}}, C, a^d_{temp}, a_k$
Initialize: $L^* \leftarrow 0, P^* \leftarrow \emptyset, a^* \leftarrow \emptyset, a^*_k \leftarrow \emptyset, P_{a_k} \leftarrow \emptyset$
for $i = 1, \ldots, |C|$ **do**
  `Simulate_Model`$(\mathcal{G}_{\mathcal{P}})$
  for $k = 1, \ldots, |a_k|$ **do**
    $P_{C(i,j)} \leftarrow \emptyset, k_c \leftarrow 0, a_{C(i,j)} \leftarrow \emptyset$
    for $j = 1, \ldots, |a^d_{temp}|$ **do**
      if $a^d_{temp}(j) = 0$ **then**
        `Isolate_Branches`$(\mathcal{G}_{\mathcal{P}}, C(i,j))$
        $a_{C(i,j)} \leftarrow a_{C(i,j)} \cup a^d_{temp}(j)$
        $P_{C(i,j)} \leftarrow P_{C(i,j)} \cup C(i,j)$
      end if
    end for
    `Simulate_Contin`$(\mathcal{G}_{\mathcal{P}}, P_{C(i,j)}, a_{C(i,j)})$
    $e \leftarrow 1$
    while $e = 1$ **do**
      $e \leftarrow 0, k_c \leftarrow k_c + 1$
      $c \leftarrow$ `Get_Branches`$(\mathcal{G}_{\mathcal{P}}, C(i,j))$
      if $|c| \neq 0$ **then**
        for $y = 1, \ldots, |c|$ **do**
          `Isolate_Branches`$(\mathcal{G}_{\mathcal{P}}, c(y))$
        end for
        $e \leftarrow 1$
      end if
      for $j = 1, \ldots, |a^d_{temp}|$ **do**
        if $k_c = a^d_{temp}(j)$ **then**
          `Isolate_Branches`$(\mathcal{G}_{\mathcal{P}}, C(i,j))$
          $P_{C(i,j)} \leftarrow P_{C(i,j)} \cup C(i,j)$
          $a_{C(i,j)} \leftarrow a_{C(i,j)} \cup a^d_{temp}(j)$
        end if
      end for
      if $k_c = a_k(k)$ **then**
        `Isolate_Branches`$(\mathcal{G}_{\mathcal{P}}, C(i, |C(i)| - 1))$
        $P_{a_k} \leftarrow C(i, |C(i)| - 1), a^*_k \leftarrow k_c$
      end if
      `Simulate_Contin`$(\mathcal{G}_{\mathcal{P}}, P_{C(i,j)} \cup P_{a_k}, a_{C(i,j)} \cup a^*_k)$
      $L_l \leftarrow$ `Get_Loads`$(\mathcal{G}_{\mathcal{P}}, P_{C(i,j)} \cup P_{a_k}, a_{C(i,j)} \cup a^*_k)$
      $L_C \leftarrow$ `Get_Damage`$(\mathcal{G}_{\mathcal{P}}, L_l)$
    end while
    if $L_C > L^*$ **then**
      $L^* \leftarrow L_C, P_t \leftarrow P_{C(i,j)}, P_i \leftarrow P_{a_k}$
      $a^{'}_C \leftarrow a^*_k, a_C \leftarrow a_{C(i,j)}$
    end if
    `Simulate_Model`$(\mathcal{G}_{\mathcal{P}})$
  end for
end for
$P^* \leftarrow P_t, a^* \leftarrow a_C$
$P^* \leftarrow P^* \cup P_i, a^* \leftarrow a^* \cup a^{'}_C$
return P^*, L^*, a^*

$(a_{C(i,j)} \cup a_k^*)$. Next, the amount of system damage caused by the attack is computed for every contingency set in C. If the computed load L_C is larger than the maximum damage L^* in any iteration, the solution is updated. Note that, after evaluating each contingency set in C, the power system model is set back to its nominal state.

8.7 Dynamic Defense Model

In this section, first we provide the formulation of the defender model to improve the power system resilience by minimizing the load loss. Then, we provide an efficient algorithm for identifying the critical substations to be protected to minimize the system damage considering the dynamic attack. Here, based on the attack on the substations and its components, a set of critical substations to be protected is identified.

8.7.1 Defender's Problem

The objective of the defender is to improve the power system resilience by minimizing the damage/load loss possible. In order to achieve this, defender can protect a subset of substations D_S from the total number of substations S in the power system network, i.e., $D_S \subseteq S$. Due to financial budget constraints, the defender is resource bounded and can prioritize and protect at most B_D substations. The problem is formally defined below.

Problem 6 (Defender's Problem) *Given a power system network $\mathcal{G}_\mathcal{P}$, a defense budget B_D, a substation budget B_S, a protection assembly budget B_P, find a defense strategy $D_P^{'}$ to minimize the damage/load loss when an attacker launches a dynamic attack at different time instants k. Formally,*

$$\underset{D_S}{\operatorname{argmin}} \max_{\{(S'(k)\subseteq S\setminus D_S)\,(P'(k)\subseteq F(S'(k)))\}_{k=1}^{T}} \sum_{k=1}^{T} J(A_{P'}(k), x(k)) \tag{8.17}$$

$$x(k) = \begin{cases} G(H(k)), & if \quad H(k) = \{A_{P'}(i)\}_{i=1}^{k-1} \\ g(H(k)), & if \quad H(k) = \emptyset \end{cases} \tag{8.18}$$

$$|D_S| \leq B_D \tag{8.19}$$

$$\begin{gathered} \sum_{k=1}^{T} |S^{'}(k)| \leq B_S \\ \forall k, k' \in \{1, ..., T\} : S^{'}(k) \cap S^{'}(k') = \emptyset, k \neq k' \end{gathered} \tag{8.20}$$

$$\begin{gathered} \sum_{k=1}^{T} |P^{'}(k)| \leq B_P \\ \forall k, k' \in \{1, ..., T\} : P^{'}(k) \cap P^{'}(k') = \emptyset, k \neq k' \end{gathered} \tag{8.21}$$

$$B_S \leq B_P \tag{8.22}$$

where, $x(k)$ represents the state of the system at time step k and $H(k)$ represents the attack history of the system.

8.7.2 Algorithm for Finding the Critical Substations to Protect

Algorithm 10 starts with an empty set and intelligently identifies the critical substations to protect one-by-one such that when the attack is launched the overall system damage can be minimized. The algorithm takes the same inputs as Algorithm 8 with the defense budget B_D as an additional input. It then identifies the critical substations D_S to prioritize and protect so as to minimize the system damage when a dynamic attack is executed.

First, the worst-case dynamic attack is identified by using $\texttt{Get\_WDA}(\mathcal{G}_{\mathcal{P}}, B_P, S_{info}, a_k)$ which is described as Algorithm 8. Next, if there are no critical substations in D_S, We use the critical substations $S_t^{'}(k)$ identified from the worst-case dynamic attack to identify the first substation to protect. We iteratively protect each substation in $S_t^{'}(k)$ and evaluate the overall system damage post dynamic attack using $\texttt{Get\_WDA2}(\mathcal{G}_{\mathcal{P}}, B_P, S_{info}, a_k, D_S^t, s)$. The computed system damage in each iteration is used to select the substation to protect,

i.e., $D_S \leftarrow D_S \cup D_s^{'}$. A track of intermediate solution $D_S^t \leftarrow D_S^t \cup D_s^{'}$ is kept in order to

Algorithm 10 Algorithm for Finding Critical Substations to Protect: `Get_Dynamic_Defense`$(\mathcal{G}_{\mathcal{P}}, B_P, B_D, S_{info}, a_k)$

Input: $\mathcal{G}_{\mathcal{P}}, B_P, B_D, S_{info}, a_k$
Initialize: $D_s^{'} \leftarrow \emptyset, D_S \leftarrow \emptyset, \hat{L}_w \leftarrow 100, L_{Prev} \leftarrow 100, L_H \leftarrow \emptyset$
$S_t^{'}(k), P^{'}(k), L_w^d, a_k^d \leftarrow$ `Get_WDA`$(\mathcal{G}_{\mathcal{P}}, B_P, S_{info}, a_k)$
$L_H \cup L_{Prev}$
for $i = 1, \ldots, B_D$ **do**
    $\hat{L}_w \leftarrow 100, flag \leftarrow 0$
    if $D_S^t \neq \emptyset$ **then**
        $\hat{S}_t^{'}(k), L_{Prev} \leftarrow$ `Get_WDA1`$(\mathcal{G}_{\mathcal{P}}, B_P, S_{info}, a_k, D_S^t, \emptyset)$
        $L_H \cup L_{Prev}$
    end if
    for all $s \in S_t^{'}(k)$ **do**
        $S_s^{'}(k), P_s^{'}(k), L_s^{'} \leftarrow$ `Get_WDA2`$(\mathcal{G}_{\mathcal{P}}, B_P, S_{info}, a_k, D_S^t, s)$
        if $L_s^{'} < \hat{L}_w$ **then**
            $\hat{L}_w \leftarrow L_s^{'}$, $D_s^{'} \leftarrow s, flag \leftarrow 1$
        end if
    end for
    $D_S \leftarrow D_S \cup D_s^{'}, D_S^t \leftarrow D_S^t \cup D_s^{'}$
    if $\hat{L}_w > min(L_H)$ AND $flag == 1$ **then**
        $D_S \leftarrow D_S \setminus D_s^{'}$
    else
        $D_S \leftarrow D_S^t$
    end if
end for
return D_S

obtain a better solution. Note that, the function `Get_WDA2`$(\mathcal{G}_{\mathcal{P}}, B_P, S_{info}, a_k, D_S^t, s)$ is same as Algorithm 8, however, here the worst-case dynamic attack is computed by eliminating the protected substations D_S^t and the substation s from the attackable list of substations, i.e., $S \setminus (D_S^t \cup s)$. If the computed damage $L_s^{'}$ is smaller than the maximum damage $\hat{L}_w$, the solution is updated. Additionally, in each iteration, if the protected substations set D_S^t is non-empty then a new set of critical substations are identified using the worst-case dynamic attack function, i.e., `Get_WDA1`$(\mathcal{G}_{\mathcal{P}}, B_P, S_{info}, a_k, D_S^t, \emptyset)$. This function is also same as Algorithm 8, however, the protected substations D_S^t are removed from the attackable list of substations while executing the worst-case dynamic attack on the power system model $\mathcal{G}_{\mathcal{P}}$. It ensures that once the substations are protected, the attacker can only launch the dynamic attack on the remaining substations depending on the attack budget. The obtained

attack can further be utilized to identify the substation to protect considering the defense budget constraints. In the algorithm L_H keeps a track of all the previous load losses obtained after protecting the substations in D_S^t and updates the final solution D_S depending upon the comparison of the obtained damage with the previous system damages. This ensures a better protection mechanism that provides an effective solution.

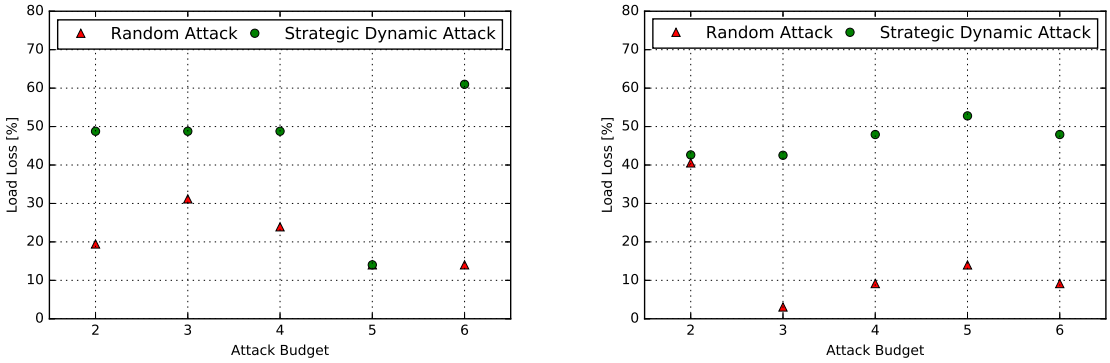
8.8 Evaluation

We considered two standard IEEE systems, the 39 bus and 57 bus systems to evaluate our approach. We used a modified version of the steady state simulator discussed in [142] to perform the analysis. First, we discuss how randomly chosen attacks can be optimized using our dynamic attack model. Next, we show the optimization of the worst-case static attacks using the dynamic attack model. Then, we present the dynamic defense results that show the reduction in the overall system damage/load loss. Finally, we evaluate the performance of our algorithm's execution time for the dynamic attack and defense algorithms in comparison with the naive exhaustive search algorithm.

8.8.1 Optimizing Random Attacks

Figure 8.2 shows the optimization of the random attacks using the dynamic attack model discussed in Section V. Here, depending upon the attack budget (up to 6), we randomly picked the components to attack from the power system model. Then, we used these attacks as inputs to our dynamic attack algorithm to obtain a strategic sequence in which the attacks can be executed so as to maximize the system damage. We performed our evaluation on the IEEE 39, 57 bus system and the results are shown in Figure 8.2. The x-axis represents the attack budget whereas the y-axis represents the system damage, i.e., load loss. Red, green color markers represent the random and strategic dynamic attacks respectively.

For both the standard IEEE systems, we can clearly see from Figures 8.2a and 8.2b that



(a) IEEE-39 bus system

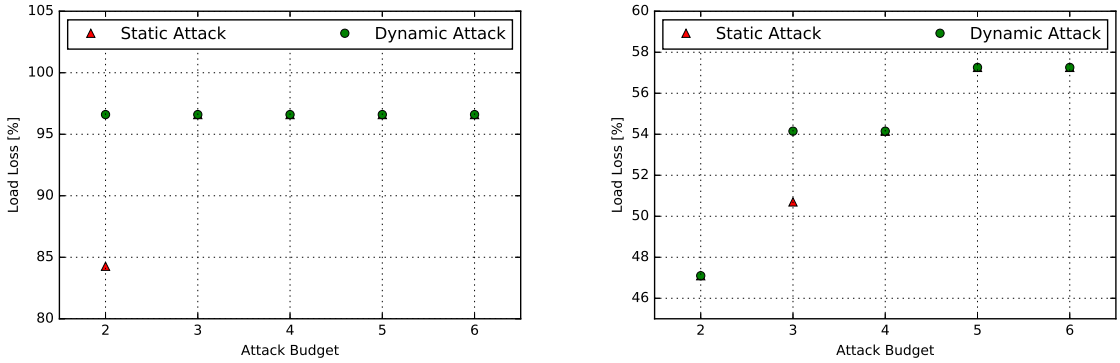
(b) IEEE-57 bus system

Figure 8.2: Random Attacks Vs Dynamic Attacks: Load loss as a function of various attack budgets for different standard IEEE systems.

our dynamic attack algorithms described in this paper are able to strategically identify the specific instants (or sequences) at which different attacks can be executed and maximize the system damage for a randomly identified set of components to attack. From Figure 8.2a, for an attack budget of 6 in IEEE-39 bus system the random attack caused a load loss of 14.03%, however, the same attack when executed at different instants in time, i.e., dynamic attack resulted in a total load loss of 60.99%. The dynamic attack on the same components caused a 334.71% higher load loss than the static attack. For the same attack budget in the IEEE-57 bus system the random attack caused a load loss of 9.16%, whereas, the dynamic attack resulted in a load loss of 47.93% as shown in Figure 8.2b. This dynamic attack load loss is 423.25% higher than the random attack.

8.8.2 Optimizing Static Attacks

We perform the analysis on the same IEEE systems. First, we identified the worst-case static attack and then we use it to identify the worst-case dynamic attack in order to further maximize the system damage. Figure 8.3 shows the results for the optimization of the worst-case static attack using our dynamic attack model and algorithm. The x-axis represents the attack budget, whereas, the y-axis represents the system damage. Red, green



(a) IEEE-39 bus system

(b) IEEE-57 bus system

Figure 8.3: Static Attacks Vs Dynamic Attacks: Load loss as a function of various attack budgets for different standard IEEE systems.

colored markers represent the worst-case static and dynamic attacks respectively. Here, we consider an attack budget of up to 6 components/substations.

From Figures 8.3a and 8.3b it is clear that the dynamic attack causes higher damage with different attack budgets. As shown in Figure 8.3a, for an attack budget of 2 in IEEE-39 bus system the worst-case static attack caused a load loss of 84.27%, however, the optimized worst-case dynamic attack resulted in a load loss of 96.60%. Here, the dynamic attack on the same components caused a 14.63% higher load loss. Similarly, for the IEEE-59 bus system in Figure 8.3b, the worst-case static attack caused a load loss of 50.70%, whereas, the optimized worst-case dynamic attack resulted in a load loss of 54.15% for an attack budget of 3. The dynamic attack caused a higher load loss by 6.80%. Note that, the worst-case static attacks are already identified as the attacks that cause maximum damage, however our dynamic attack algorithms are still able to optimize them for obtaining even higher system damage if there is a possibility for optimization. The dynamic attack algorithm results from Figure 8.3 clearly show that the dynamic attacks on the same components that are identified from the static attack scenario when scheduled and executed strategically resulted in a higher system damage. Note that, in Figure 8.3, the static and dynamic attack load loss becomes equal for some attack budgets because there is no additional load loss

Table 8.3: Scenario representing the maximization of system damage using dynamic attack model

Static Attack		Dynamic Attack	
Initial Attack	Attack time vector: [0, 0] Substations compromised: [S^{13}, S^{24}] Protection assemblies attacked: [PA10, PA16] Transmission lines Isolated due to the attacked protection assemblies: [‘R16_19’, ‘R2_3’] Load loss: ‘0%’	Initial Attack	Attack time vector: [0] Substations compromised: [S^{24}] Protection assemblies attacked: [PA16] Transmission lines Isolated due to the attacked protection assemblies: [‘R2_3’] Load loss: ‘0%’
Stage 1 Outages	Isolation of transmission lines due to the secondary effect of the outages from the initial attack: [‘R2_25, R25_26, R18_17, R27_26’], Load loss: ‘0%’	Stage 1 Outages	Isolation of transmission lines due to the secondary effect of the outages from the initial attack: [‘R2_25, R18_17’], Load loss: ‘0%’
Stage 2 Outages	Isolation of transmission lines due to the seco of the outages from the stage 1: [‘R6_5, R14_15, R14_13, R10_13, R26_28, R21_22] Load loss: ‘35.48%’	Additional Attack	Attack time vector: [1] Substations compromised: [S^{13}] Protection assemblies attacked: [PA10] Transmission lines Isolated due to the attacked protection assemblies: [‘R16_19’] Load loss: ‘0%’
Stage 3 Outages	Isolation of transmission lines due to the secondary effect of the outages from the stage 2: [‘R8_7, R6_7, R10_11, R6_11] Load loss: ‘64.80%’	Stage 2 Outages	Isolation of transmission lines due to the secondary effect of the outages from the stage 1: [‘R6_5’], Load loss: ‘0%’
Stage 4 Outages	Isolation of transmission lines due to the secondary effect of the outages from the stage 3: [‘R9_39, R8_9’], Load loss: ‘84.27%’	Stage 3 Outages	Isolation of transmission lines due to the secondary effect of the outages from the stage 2: [‘R8_7, R6_7, R4_14, R14_13, R10_13’], Load loss: ‘7.25%’
		Stage 4 Outages	Isolation of transmission lines due to the secondary effect of the outages from the stage 3: [‘R9_39, R8_9, R21_22, R24_23’], Load loss: ‘56.42%’
		Stage 5 Outages	Isolation of transmission lines due to the secondary effect of the outages from the stage 3: [‘R25_26, R17_27, R27_26, R16_17, R26_28, R28_29, R26_29, R16_21’], Load loss: ‘96.60%’

possible within the system. Also note that, for some attack budgets the difference in the load loss between the static attack and the dynamic attack can remain very small because the additional loads that gets disconnected during the dynamic attack maybe smaller in magnitude as compared to the total load loss. However, if the additional load loss is larger in magnitude, then this difference can be significantly larger as shown by attack budget 2,

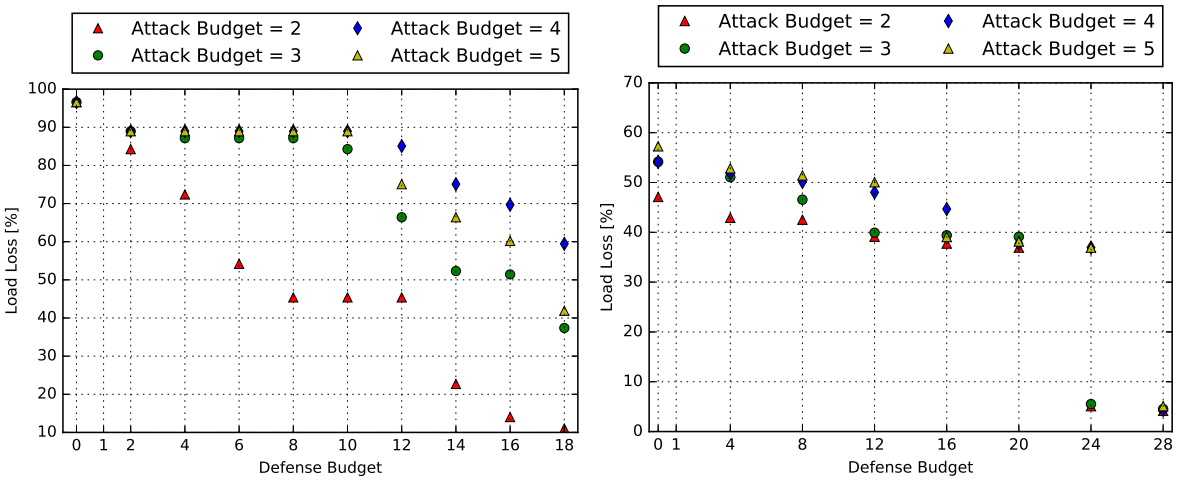
3 in Figures 8.3a and 8.3b respectively.

We have shown the exact cascade progression for one of the static and dynamic attack scenarios in Table 8.3 that can easily answer the question of 'how dynamic attacks can have higher impact?'. For both the attack scenarios, we consider the same substations and its components to attack, but the only difference is the attack time. For the static attack scenario with an attack budget of 2, Table 8.3 shows that both the attacks are launched at the same time [0, 0] ([0, 0] indicates simultaneous attack or static attack). As a result of the static attack the transmission lines associated with the attacked protection assemblies are isolated. This resulted in a sequence of cascading failures as shown by the 'Stage 1 Outages' through 'Stage 4 Outages' in Table 8.3 and the total system load loss was observed to be 84.27%.

Now, we consider the same substations and protection assemblies for the dynamic attack scenario. Here, the initial attack takes place at time instant 0 that initiated a cascading event causing subsequent failures (Stage 1 Outages in Table 8.3). At time instant 1, another attack was launched that further weakened the system causing Outages through Stage 2 to Stage 5 resulting in a significant damage to the system. The overall system load loss was observed to be 96.60% (Stage 2 and Stage 5 Outages in Table 8.3) which is considerably higher than the static attack. Note that the specific time at which these attacks can be executed are computed using the algorithms described in Section V.

8.8.3 Minimizing System Damage Using Dynamic Defense

We evaluate our defense model and algorithm using the standard IEEE-39 and 57 bus systems. Figure 8.4 shows the load losses in the power system at different attack budgets when a dynamic attack is launched after the critical substations are intelligently identified and protected depending upon the defense budget. In each figure, the x-axis represents the defense budget and the y-axis represent the total system damage. Red, green, blue, and yellow colored markers represents the attack budgets 2, 3, 4 and 5 respectively. The



(a) IEEE-39 bus system

(b) IEEE-57 bus system

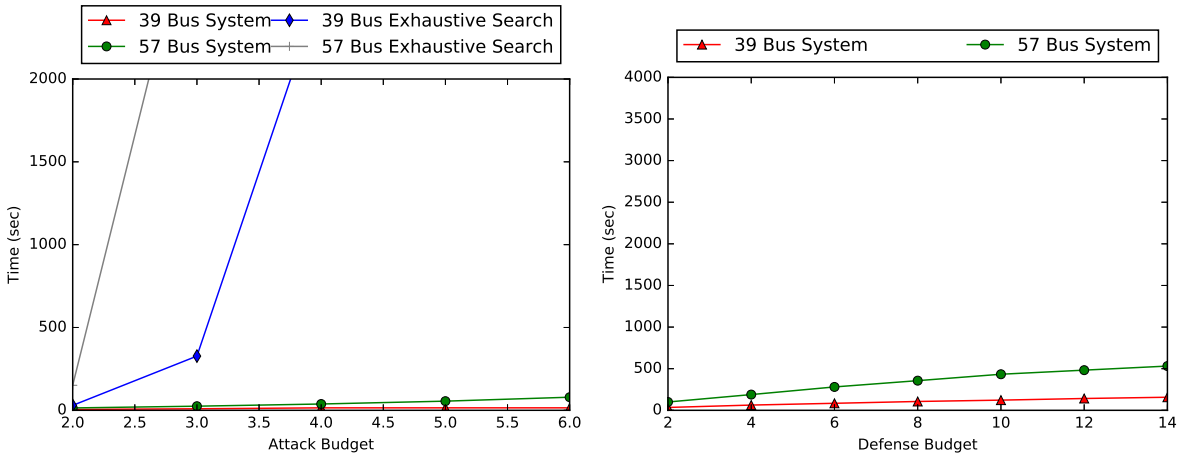
Figure 8.4: Dynamic Defense: Load loss as a function of various defense budgets for different standard IEEE systems.

respective color markers at the defense budget 0 represent the total system damage without any defense.

From Figure 8.4, we can clearly see that by intelligently selecting and protecting the critical substations of the power network, the system damage can be significantly reduced for IEEE-39 bus system (Figure 8.4a) and 57 bus system (Figure 8.4b) when a dynamic attack is launched. In Figure 8.4a, for an attack and a defense budget of 2, the load loss is reduced from 96.60% to 84.27%, that is, a total of 12.76% reduction in load loss. Moreover, for the same attack budget and a defense budget of 18, a total of 88.58% reduction in load loss is observed. For other attack budgets, as the defense budget increases we can see significant improvement in the reduction of total system load loss.

8.8.4 Performance of the Dynamic Attack and Defense Algorithms

Here, we compare the execution time of our dynamic attack and defense algorithms with the naive exhaustive search algorithms. We use the same standard IEEE systems to perform our analysis. Figure 8.5 shows the dynamic attack and defense execution time



(a) Execution time of attack analysis

(b) Execution time of defense analysis

Figure 8.5: Analysis execution time for attack and defense for different standard IEEE systems

with respect to the exhaustive search. In each figure, the x-axis represents either the attack budget or the defense budget and the y-axis represents the time taken by the algorithm to identify the attack or defense. The details of the markers are shown in the legend box of Figure 8.5.

From Figure 8.5a, we can clearly see that the time taken to identify the dynamic attack for IEEE-39, 57 bus system increases very slightly with increase in the attack budget. However, the time taken to identify the attack using the exhaustive search algorithm is observed to be exponential even at smaller attack budgets. The exhaustive search execution time in Figure 8.5a represents the time taken to identify the maximum damage causing static attack. Moreover, the exhaustive search execution time for identifying the maximum damage causing dynamic attack will be much larger than the time taken to identify the static attack. Similarly, it is clear from Figure 8.5b that the time taken to identify the defense increases slowly with the increase in the defense budget. We know that dynamic defense via exhaustive analysis will take much longer than the exhaustive attack since it will have to first identify the attack and then identify the defense. Hence, if we compare only against the attack time, it still shows that the developed approach is much faster than the exhaus-

tive search. Therefore, as demonstrated in Figure 8.5, our algorithms prove to be far more efficient than the naive exhaustive search.

8.9 Conclusions and Future Works

We described the static and dynamic cyber-attack and defense models for electrical power systems using game-theoretic approach. From the attacker's perspective, we provide an efficient and effective algorithm that is able to strategically identify the dynamic attacks that maximizes the system damage by considering both random attacks as well as worst-case static attacks. We also provide an efficient algorithm from defenders perspective that identifies the critical substations to protect in order to minimize the overall system damage. Our results shows that, under financial budget constraints, intelligently selecting the substations to prioritize and protect can significantly improve the power system resilience. In addition, these algorithms are efficient and perform significantly better than the exhaustive search even with the complex dynamic attack and defense models. As part of the future work, the attacker-defender models can be easily extended to consider randomness, i.e., a success probability can be associated with an attack and a defense that can give us more insight to improve the power system resilience under probabilistic scenarios. Further, under unknown circumstances where the defender has no idea whether an attacker follows a static attack model or a dynamic attack model, a defense strategy that could improve the overall power system resilience irrespective of the attack model can be an interesting direction to explore.

Chapter 9

Summary and Future Work

In this thesis, we have described the importance of resilience in large scale cyber-physical systems. We provided important insight on the importance of contingency analysis models, identification of critical contingencies, *cyber-attack* and *defense models* with respect to *static* and *dynamic attacks* in electrical power systems. The following list summarizes the contributions from this thesis:

- Developed the framework for introducing both physical and cyber-faults at any desired time instant and study their effects on the system. The framework provides the capability to identify new critical vulnerabilities that can not be identified otherwise.

- Designed a component based modeling and analysis approach. The approach provides a common DSML for system design and integrates multiple simulation tools together to provide the capability of performing richer analysis by significantly reducing the modeling time and error.

- Developed methods for identifying higher order critical $N-k$ contingencies. The developed methods effectively prune the search space and result in only a limited number of simulation runs that reduces the analysis time significantly without compromising the performance, i.e., the methods are able to identify every single critical contingency.

- Designed the game-theoretic approach for modeling static cyber-attack and defense models. The approach identifies the effective deployment of the limited defense resources under financial budget constraints. This is achieved by first identifying the worst-case static attack. The developed algorithms effectively improves the system

resilience and prove to be significantly faster than the exhaustive search mechanism.

- Developed the game-theoretic based cyber-attack and defense models, algorithms for dynamic cyber-attacks in large scale cyber-physical systems. The algorithms effectively identify the worst-case attacks which are then used to identify the effective deployment of the limited defense resources. The evaluation results demonstrate that the solutions obtained using our algorithms are significantly faster than the exhaustive search mechanism and effectively improves the overall system resilience.

Although, the work presented in this thesis provides a concrete mechanism to improve the overall resilience of CPS. However, there can be several extensions to this work. The analysis approach used in this work is based on the off-line analysis mechanism. It is important to identify ways to apply this approach to perform online analysis which will further improve the effectiveness of the developed mechanisms. Further, various modeling and analysis methods described in this work can be integrated into a single analysis tool. This tool can serve as a base to perform the desired type of analysis by the engineers to identify the critical points in the system and obtain useful solutions. Moreover, block-chain based approach can be applied to the resilience problem discussed in this thesis. The block-chain based approach will improve the security of the overall system by recording each transaction in a distributed ledger. The transactions will be hard to manipulate by an attacker since it will need a significantly high computing power which is next to impossible for a single attacker. In addition, machine learning (ML) based approaches can be designed where the ML model can be trained based on the problem using the required data set and then the predictions can be obtained by using this model for the unknown events, contingencies, etc. Similarly, attack and defense models can also be applied to the transactive energy domain to identify various vulnerabilities and their solutions. The above mentioned extensions can further improve the overall resilience of the large-scale cyber-physical systems.

Chapter 10

List of Publications

The published papers are reviewed by at least 3 reviewers. Below is the list of all the accepted and under review publications:

- Hasan, Saqib, Abhishek Dubey, Gabor Karsai, and Xenofon Koutsoukos. "A Game-Theoretic Approach for Power Systems Defense Against Dynamic Cyber-Attacks." 'Submitted, In Review' In Elsevier-International Journal of Electrical Power & Energy Systems, 2019.

- Hasan, Saqib, Amin Ghafouri, Abhishek Dubey, Gabor Karsai, and Xenofon Koutsoukos. "Vulnerability Analysis of Power Systems Based on Cyber-Attack and Defense Models." In Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), 2018 IEEE, 2018.

- Hasan, Saqib, Ajay Chhokra, Abhishek Dubey, Nagabhushan Mahadevan, Gabor Karsai, Rishabh Jain, and Srdjan Lukic. "A simulation testbed for cascade analysis." In Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), 2017 IEEE, pp. 1-5. IEEE, 2017.

- Hasan, Saqib, Abhishek Dubey, Ajay Chhokra, Nagabhushan Mahadevan, Gabor Karsai, and Xenofon Koutsoukos. "A modeling framework to integrate exogenous tools for identifying critical components in power systems." In Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES), 2017 Workshop on, pp. 1-6. IEEE, 2017.

- Hasan, Saqib, Amin Ghafouri, Abhishek Dubey, Gabor Karsai, and Xenofon Koutsoukos. "Heuristics-based approach for identifying critical Nk contingencies in power

systems." In Resilience Week (RWS), 2017, pp. 191-197. IEEE, 2017.

- Hasan, Saqib, Ajay Chhokra, Nagabhushan Mahadevan, Abhishek Dubey, and Gabor Karsai. "Technical Report: Cyber-physical vulnerability analysis, Institute for Software-Integrated Systems." ISIS 17 (2017): 101.

- Chhokra, Ajay, Amogh Kulkarni, Saqib Hasan, Abhishek Dubey, Nagabhushan Mahadevan, and Gabor Karsai. "A Systematic Approach of Identifying Optimal Load Control Actions for Arresting Cascading Failures in Power Systems." In Proceedings of the 2nd Workshop on Cyber-Physical Security and Resilience in Smart Grids, pp. 41-46. ACM, 2017.

- Chhokra Ajay, Abhishek Dubey, Saqib Hasan, Nagabhushan Mahadevan, and Gabor Karsai. "Diagnostics and prognostics using temporal causal models for cyber physical energy systems." In CPS Week, 2017.

- Chhokra, Ajay, Abhishek Dubey, Nagabhushan Mahadevan, Saqib Hasan, and Gabor Karsai. "Diagnosis in Cyber-Physical Systems with Fault Protection Assemblies." In Diagnosability, Security and Safety of Hybrid Dynamic and Cyber-Physical Systems, pp. 201-225. Springer, Cham, 2018.

Bibliography

[1] http://icseg.iti.illinois.edu/ieee-14-bus-system/.

[2] USNEWS. Cyberattacks surge on energy companies, electric grid. *US-NEWS,[Online] Available at: https://www.usnews.com/news/blogs/data-mine/2016/04/08/cyberattacks-surge-on-energy-companies-electric-grid*.

[3] US DOE. Enabling modernization of the electric power system. *Quadrennial technology review*, 22, 2015.

[4] P Oman, A Risley, Jeff Roberts, and E Schweitzer. Attack and defend tools for remotely accessible control and protection equipment in electric power systems. In *Fifty-Fifth Annual Conference for Protective Relay Engineers*, 2002.

[5] Chih-Che Sun, Chen-Ching Liu, and Jing Xie. Cyber-physical system security of a power grid: State-of-the-art. *Electronics*, 5(3):40, 2016.

[6] Standard NERC. Top-004–2: Transmission operations. *North American Electric Reliability Corporation*, 2007.

[7] Margaret J Eppstein and Paul DH Hines. A random chemistry algorithm for identifying collections of multiple contingencies that initiate cascading failure. *IEEE Transactions on Power Systems*, 27(3):1698–1705, 2012.

[8] Paul DH Hines, Ian Dobson, Eduardo Cotilla-Sanchez, and Margaret Eppstein. ” dual graph” and” random chemistry” methods for cascading failure analysis. In *System Sciences (HICSS), 2013 46th Hawaii International Conference on*, pages 2141–2150. IEEE, 2013.

[9] Konstantin S Turitsyn and PA Kaplunovich. Fast algorithm for n-2 contingency

problem. In *System Sciences (HICSS), 2013 46th Hawaii International Conference on*, pages 2161–2166. IEEE, 2013.

[10] Wenye Wang and Zhuo Lu. Cyber security in the smart grid: Survey and challenges. *Computer Networks*, 57(5):1344–1371, 2013.

[11] Danda B Rawat and Chandra Bajracharya. Cyber security for smart grid systems: Status, challenges and perspectives. In *SoutheastCon 2015*, pages 1–6. IEEE, 2015.

[12] B Liscouski and W Elliot. Final report on the august 14, 2003 blackout in the united states and canada: Causes and recommendations. *A report to US Department of Energy*, 40(4), 2004.

[13] Yihai Zhu, Jun Yan, Yufei Tang, Yan Sun, and Haibo He. The sequential attack against power grid networks. In *Communications (ICC), 2014 IEEE International Conference on*, pages 616–621. IEEE, 2014.

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

[15] Advanced metering infrastructure (ami). https://www.ferc.gov/CalendarFiles/20070423091846-EPRI%20-%20Advanced%20Metering.pdf/.

[16] Abhishek Gupta, Smriti Jain, Vishal Kumar Agrawal, Sohail Ahamed, Saksham Aggrawal, and Rishabh Kumar. Phasor measurement unit. *International Journal of Engineering and Management Research (IJEMR)*, 6(2):221–224, 2016.

[17] B Ravindranath and M Chander. *Power system protection and switchgear*. New Age International, 1977.

[18] AG Phadke and James S Thorp. Expose hidden failures to prevent cascading outages [in power systems]. *IEEE Computer Applications in Power*, 9(3):20–23, 1996.

[19] J Chen and JS Thorp. A reliability study of transmission system protection via a hidden failure dc load flow model. In *Power System Management and Control, 2002. Fifth International Conference on (Conf. Publ. No. 488)*, pages 384–389. IET, 2002.

[20] Dusko P Nedic, Ian Dobson, Daniel S Kirschen, Benjamin A Carreras, and Vickie E Lynch. Criticality in a cascading failure blackout model. *International Journal of Electrical Power & Energy Systems*, 28(9):627–633, 2006.

[21] Yi Jun, Zhou Xiaoxin, and Xiao Yunan. Model of cascading failures in power systems. In *Power System Technology, 2006. PowerCon 2006. International Conference on*, pages 1–7. IEEE, 2006.

[22] JL Sanchez, G Ramos, and MA Rios. Modeling of operative sequences of protections in power transmission systems using petri nets. In *Transmission and Distribution Conference and Exposition: Latin America, 2008 IEEE/PES*, pages 1–6. IEEE, 2008.

[23] Benjamin A Carreras, Vickie E Lynch, ML Sachtjen, Ian Dobson, and David E Newman. Modeling blackout dynamics in power transmission networks with simple structure. In *System Sciences, 2001. Proceedings of the 34th Annual Hawaii International Conference on*, pages 719–727. IEEE, 2001.

[24] Ian Dobson, BA Carreras, V Lynch, and D Newman. An initial model for complex dynamics in electric power system blackouts. In *hicss*, page 2017. IEEE, 2001.

[25] David E Newman, Benjamin A Carreras, Vickie E Lynch, and Ian Dobson. The impact of various upgrade strategies on the long-term dynamics and robustness of the transmission grid. In *Proc. Conf. Electricity Transmission in Deregulated Markets*, page 2004, 2004.

[26] Lu Zongxiang, Meng Zhongwei, and Zhou Shuangxi. Cascading failure analysis of bulk power system using small-world network model. In *Probabilistic Methods Applied to Power Systems, 2004 International Conference on*, pages 635–640. IEEE, 2004.

[27] Xiao Fan Wang and Guanrong Chen. Complex networks: small-world, scale-free and beyond. *IEEE circuits and systems magazine*, 3(1):6–20, 2003.

[28] Paul Hines and Seth Blumsack. A centrality measure for electrical networks. In *Hawaii International Conference on System Sciences, Proceedings of the 41st Annual*, pages 185–185. IEEE, 2008.

[29] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.

[30] Albert-László Barabási, Réka Albert, and Hawoong Jeong. Mean-field theory for scale-free random networks. *Physica A: Statistical Mechanics and its Applications*, 272(1-2):173–187, 1999.

[31] Yakup Koc, Trivik Verma, Nuno AM Araujo, and Martijn Warnier. Matcasc: A tool to analyse cascading line outages in power grids. In *Intelligent Energy Systems (IWIES), 2013 IEEE International Workshop on*, pages 143–148. IEEE, 2013.

[32] Paul DH Hines, Ian Dobson, Eduardo Cotilla-Sanchez, and Margaret Eppstein. " dual graph" and" random chemistry" methods for cascading failure analysis. In *System Sciences (HICSS), 2013 46th Hawaii International Conference on*, pages 2141–2150. IEEE, 2013.

[33] Ian Dobson, Benjamin A Carreras, and David E Newman. A probabilistic loading-dependent model of cascading failure and possible implications for blackouts. In *System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on*, pages 10–pp. IEEE, 2003.

[34] Ian Dobson, Benjamin A Carreras, and David E Newman. A loading-dependent model of probabilistic cascading failure. *Probability in the Engineering and Informational Sciences*, 19(1):15–32, 2005.

[35] Ian Dobson, Benjamin A Carreras, and David E Newman. Branching process models for the exponentially increasing portions of cascading failure blackouts. In *System Sciences, 2005. HICSS'05. Proceedings of the 38th Annual Hawaii International Conference on*, pages 64a–64a. IEEE, 2005.

[36] Ian Dobson, Kevin R Wierzbicki, Benjamin A Carreras, Vickie E Lynch, and David E Newman. An estimator of propagation of cascading failure. In *System Sciences, 2006. HICSS'06. Proceedings of the 39th Annual Hawaii International Conference on*, volume 10, pages 245c–245c. IEEE, 2006.

[37] Kevin R Wierzbicki and Ian Dobson. An approach to statistical estimation of cascading failure propagation in blackouts. In *CRIS, Third International Conference on Critical Infrastructures*, pages 1–7, 2006.

[38] Ian Dobson, Kevin R Wierzbicki, Janghoon Kim, and Hui Ren. Towards quantifying cascading blackout risk. In *Bulk Power System Dynamics and Control-VII. Revitalizing Operational Reliability, 2007 iREP Symposium*, pages 1–12. IEEE, 2007.

[39] Mario A Rios, Daniel S Kirschen, Dilan Jayaweera, Dusko P Nedic, and Ron N Allan. Value of security: modeling time-dependent phenomena and weather conditions. *IEEE Transactions on Power Systems*, 17(3):543–548, 2002.

[40] Daniel S Kirschen, Dilan Jayaweera, Dusko P Nedic, and Ron N Allan. A probabilistic indicator of system stress. *IEEE Transactions on Power Systems*, 19(3):1650–1657, 2004.

[41] Marian Anghel, Kenneth A Werley, and Adilson E Motter. Stochastic model for

power grid dynamics. In *System Sciences, 2007. HICSS 2007. 40th Annual Hawaii International Conference on*, pages 113–113. IEEE, 2007.

[42] Badrul H Chowdhury and Sushant Baravc. Creating cascading failure scenarios in interconnected power systems. In *Power Engineering Society General Meeting, 2006. IEEE*, pages 8–pp. IEEE, 2006.

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

[44] Ieee cascading failure working group. http://sites.ieee.org/pes-cascading/presentations/.

[45] Peter Fritzson. *Principles of object-oriented modeling and simulation with Modelica 2.1*. John Wiley & Sons, 2010.

[46] http://www.3ds.com/.

[47] https://www.maplesoft.com/products/maplesim/.

[48] Peter Fritzson, Peter Aronsson, Håkan Lundvall, Kaj Nyström, Adrian Pop, Levon Saldamli, and David Broman. The openmodelica modeling, simulation, and software development environment. *Simulation News Europe*, 44, 2005.

[49] Michael Zhou and Shizhao Zhou. Internet, open-source and power system simulation. In *Power Engineering Society General Meeting, 2007. IEEE*, pages 1–5. IEEE, 2007.

[50] Federico Milano. An open source power system analysis toolbox. *IEEE Transactions on Power systems*, 20(3):1199–1206, 2005.

[51] http://www.mathworks.com/.

[52] Saffet Ayasun, Chika O Nwankpa, and Harry G Kwatny. Voltage stability toolbox for power system education and research. *IEEE Transactions on Education*, 49(4):432–442, 2006.

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

[54] David P Chassin, K Schneider, and C Gerkensmeyer. Gridlab-d: An open-source power systems modeling and simulation environment. In *Transmission and distribution conference and exposition, 2008. t&d. IEEE/PES*, pages 1–5. IEEE, 2008.

[55] http://www.digsilent.de/.

[56] https://hvdc.ca/pscad/.

[57] PowerWorld Simulator. Powerworld corporation, 2005.

[58] U-CPSOT Force. Final report on the august 14th blackout in the united states and canada. *Department of Energy and National Resources Canada*, 2004.

[59] Joshua J Romero. Blackouts illuminate india's power problems. *IEEE spectrum*, 49(10), 2012.

[60] Tereza Pultarova. News briefing: Cyber security-ukraine grid hack is wake-up call for network operators. *Engineering & Technology*, 11(1):12–13, 2016.

[61] April Reese. Blackouts cast australia's green energy in dim light, 2017.

[62] GC Ejebe and BF Wollenberg. Automatic contingency selection. *IEEE Transactions on Power Apparatus and Systems*, (1):97–109, 1979.

[63] SW Director and R Rohrer. The generalized adjoint network and network sensitivities. *IEEE Transactions on Circuit Theory*, 16(3):318–323, 1969.

[64] TA Mikolinnas and BF Wollenberg. An advanced contingency selection algorithm. *IEEE Transactions on Power Apparatus and Systems*, (2):608–617, 1981.

[65] GD Irisarri and AM Sasson. An automatic contingency selection method for on-line security analysis. *IEEE transactions on power apparatus and systems*, (4):1838–1844, 1981.

[66] Mark K Enns, John J Quada, and Bert Sackett. Fast linear contingency analysis. *IEEE Transactions on Power Apparatus and Systems*, (4):783–791, 1982.

[67] GC Ejebe, HP Van Meeteren, and BF Wollenberg. Fast contingency screening and evaluation for voltage security analysis. *IEEE Transactions on Power Systems*, 3(4):1582–1590, 1988.

[68] O Alsac, B Stott, and WF Tinney. Sparsity-oriented compensation methods for modified network solutions. *IEEE Transactions on Power Apparatus and Systems*, (5):1050–1060, 1983.

[69] WF Tinney, Vladimir Brandwajn, and SM Chan. Sparse vector methods. *IEEE transactions on power apparatus and systems*, (2):295–301, 1985.

[70] William F Tinney and Joseph M Bright. Adaptive reductions for power flow equivalents. *IEEE transactions on power systems*, 2(2):351–359, 1987.

[71] Michele Di Santo, Alfredo Vaccaro, Domenico Villacci, and Eugenio Zimeo. A distributed architecture for online power systems security analysis. *IEEE Transactions on Industrial Electronics*, 51(6):1238–1248, 2004.

[72] Vaibhav Donde, Vanessa Lopez, Bernard Lesieutre, Ali Pinar, Chao Yang, and Juan Meza. Identification of severe multiple contingencies in electric power networks.

In *Power Symposium, 2005. Proceedings of the 37th Annual North American*, pages 59–66. IEEE, 2005.

[73] Vaibhav Donde, Vanessa López, Bernard Lesieutre, Ali Pinar, Chao Yang, and Juan Meza. Severe multiple contingency screening in electric power systems. *IEEE Transactions on Power Systems*, 23(2):406–417, 2008.

[74] Javier Salmeron, Kevin Wood, and Ross Baldick. Analysis of electric grid security under terrorist threat. *IEEE Transactions on power systems*, 19(2):905–912, 2004.

[75] Qiming Chen and James D McCalley. Identifying high risk nk contingencies for online security assessment. *IEEE Transactions on Power Systems*, 20(2):823–834, 2005.

[76] Paolo Crucitti, Vito Latora, and Massimo Marchiori. Locating critical lines in high-voltage electrical power grids. *Fluctuation and Noise Letters*, 5(02):L201–L208, 2005.

[77] Xiaogang Chen, Ke Sun, Yijia Cao, and Shaobu Wang. Identification of vulnerable lines in power grid based on complex network theory. In *Power Engineering Society General Meeting, 2007. IEEE*, pages 1–6. IEEE, 2007.

[78] Ajendra Dwivedi, Xinghuo Yu, and Peter Sokolowski. Identifying vulnerable lines in a power network using complex network theory. In *Industrial Electronics, 2009. ISIE 2009. IEEE International Symposium on*, pages 18–23. IEEE, 2009.

[79] Rodel D Dosano, Hwachang Song, and Byongjun Lee. Network centrality based nk contingency scenario generation. In *Transmission & Distribution Conference & Exposition: Asia and Pacific, 2009*, pages 1–4. IEEE, 2009.

[80] Charles Davis and Thomas Overbye. Linear analysis of multiple outage interaction.

In *System Sciences, 2009. HICSS'09. 42nd Hawaii International Conference on*, pages 1–8. IEEE, 2009.

[81] C Matthew Davis and Thomas J Overbye. Multiple element contingency screening. *IEEE Transactions on Power Systems*, 26(3):1294–1301, 2011.

[82] PA Kaplunovich and Konstantin S Turitsyn. Statistical properties and classification of n-2 contingencies in large scale power grids. In *System Sciences (HICSS), 2014 47th Hawaii International Conference on*, pages 2517–2526. IEEE, 2014.

[83] M Granger Morgan, Massoud Amin, Edward V Badolato, WO Ball, AB Nae, and C Gellings. Terrorism and the electric power delivery system. *National Research Council of the National Academies [Online], Washington DC*, 2012.

[84] Peter J Hawrylak, Michael Haney, Mauricio Papa, and John Hale. Using hybrid attack graphs to model cyber-physical attacks in the smart grid. In *Resilient Control Systems (ISRCS), 2012 5th International Symposium on*, pages 161–164. IEEE, 2012.

[85] JE David. Double threat: Us grid vulnerable on two fronts. *CNBC. Retrieved from http://www. cnbc. com/id/101306145*, 2014.

[86] Siobhan Gorman. Electricity grid in us penetrated by spies. *The Wall Street Journal*, 8, 2009.

[87] Cyberattacks surge on energy companies, electric grid. https://www.usnews.com/news/blogs/data-mine/2016/04/08/cyberattacks-surge-on-energy-companies-electric-grid, 2016.

[88] Robert K Knake. A cyberattack on the us power grid. *Council on Foreign Relations*, 2017.

[89] Cyril W Draffin Jr. Cybersecurity white paper. 2016.

[90] Siddharth Sridhar, Adam Hahn, and Manimaran Govindarasu. Cyber–physical system security for the electric power grid. *Proceedings of the IEEE*, 100(1):210–224, 2012.

[91] Siddharth Sridhar and G Manimaran. Data integrity attack and its impacts on voltage control loop in power grid. In *Power and Energy Society General Meeting, 2011 IEEE*, pages 1–6. IEEE, 2011.

[92] Yu-Lun Huang, Alvaro A Cárdenas, Saurabh Amin, Zong-Syun Lin, Hsin-Yi Tsai, and Shankar Sastry. Understanding the physical and economic consequences of attacks on control systems. *International Journal of Critical Infrastructure Protection*, 2(3):73–83, 2009.

[93] Bo Chen, Salman Mashayekh, Karen L Butler-Purry, and Deepa Kundur. Impact of cyber attacks on transient stability of smart grids with voltage support devices. In *Power and Energy Society General Meeting (PES), 2013 IEEE*, pages 1–5. IEEE, 2013.

[94] Yao Liu, Peng Ning, and Michael K Reiter. False data injection attacks against state estimation in electric power grids. *ACM Transactions on Information and System Security (TISSEC)*, 14(1):13, 2011.

[95] Yanling Yuan, Zuyi Li, and Kui Ren. Modeling load redistribution attacks in power systems. *IEEE Transactions on Smart Grid*, 2(2):382–390, 2011.

[96] Yanling Yuan, Zuyi Li, and Kui Ren. Quantitative analysis of load redistribution attacks in power systems. *IEEE Transactions on Parallel and Distributed Systems*, 23(9):1731–1738, 2012.

[97] Jinping Hao, Robert J Piechocki, Dritan Kaleshi, Woon Hau Chin, and Zhong Fan. Sparse malicious false data injection attacks and defense mechanisms in smart grids. *IEEE Transactions on Industrial Informatics*, 11(5):1–12, 2015.

[98] Yihai Zhu, Yan Sun, and Haibo He. Load distribution vector based attack strategies against power grid systems. In *Global Communications Conference (GLOBECOM), 2012 IEEE*, pages 935–941. IEEE, 2012.

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

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

[101] Yihai Zhu, Jun Yan, Yufei Tang, Yan Lindsay Sun, and Haibo He. Joint substation-transmission line vulnerability assessment against the smart grid. *IEEE Transactions on Information Forensics and Security*, 10(5):1010–1024, 2015.

[102] Zhenghao Zhang, Shuping Gong, Aleksandar D Dimitrovski, and Husheng Li. Time synchronization attack in smart grid: Impact and analysis. *IEEE Transactions on Smart Grid*, 4(1):87–98, 2013.

[103] R Liu and A Srivastava. Integrated simulation to analyze the impact of cyber-attacks on the power grid. In *Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES), 2015 Workshop on*, pages 1–6. IEEE, 2015.

[104] Ren Liu, Ceeman Vellaithurai, Saugata S Biswas, Thoshitha T Gamage, and Anurag K Srivastava. Analyzing the cyber-physical impact of cyber events on the power grid. *IEEE Transactions on Smart Grid*, 6(5):2444–2453, 2015.

[105] Katherine R Davis, Charles M Davis, Saman A Zonouz, Rakesh B Bobba, Robin Berthier, Luis Garcia, and Peter W Sauer. A cyber-physical modeling and assessment framework for power grid infrastructures. *IEEE Transactions on Smart Grid*, 6(5):2464–2475, 2015.

[106] Shiva Poudel, Zhen Ni, and Naresh Malla. Real-time cyber physical system testbed for power system security and control. *International Journal of Electrical Power & Energy Systems*, 90:124–133, 2017.

[107] Opal-rt technologies. https://www.opal-rt.com/.

[108] Sherwin Paul Transfiguracion Edra. *Electrical Engineering Laboratory 444 System Protection Laboratory Experiment: Schweitzer Engineering Laboratories, Inc. SEL-351S Protection and Breaker Control Relay*. PhD thesis, California Polytechnic State University, 2004.

[109] Wei Yuan, Long Zhao, and Bo Zeng. Optimal power grid protection through a defender–attacker–defender model. *Reliability Engineering & System Safety*, 121:83–89, 2014.

[110] Yee Wei Law, Tansu Alpcan, and Marimuthu Palaniswami. Security games for risk minimization in automatic generation control. *IEEE Transactions on Power Systems*, 30(1):223–232, 2015.

[111] Yingmeng Xiang, Lingfeng Wang, and Nian Liu. Coordinated attacks on electric power systems in a cyber-physical environment. *Electric Power Systems Research*, 149:156–168, 2017.

[112] Yi Yang, Kieran McLaughlin, Timothy Littler, Sakir Sezer, Eul Gyu Im, ZQ Yao, B Pranggono, and HF Wang. Man-in-the-middle attack test-bed investigating cyber-security vulnerabilities in smart grid scada systems. 2012.

[113] Shan Liu, Salman Mashayekh, Deepa Kundur, Takis Zourntos, and Karen Butler-Purry. A framework for modeling cyber-physical switching attacks in smart grid. *IEEE Transactions on Emerging Topics in Computing*, 1(2):273–285, 2013.

[114] Zhendong Sun. *Switched linear systems: control and design*. Springer Science & Business Media, 2006.

[115] Jun Yan, Yihai Zhu, Haibo He, and Yan Sun. Revealing temporal features of attacks against smart grid. In *Innovative Smart Grid Technologies (ISGT), 2013 IEEE PES*, pages 1–6. IEEE, 2013.

[116] Shan Liu, Bo Chen, Takis Zourntos, Deepa Kundur, and Karen Butler-Purry. A coordinated multi-switch attack for cascading failures in smart grid. *IEEE Transactions on Smart Grid*, 5(3):1183–1195, 2014.

[117] Volker Turau and Christoph Weyer. Cascading failures caused by node overloading in complex networks. In *Cyber-Physical Security and Resilience in Smart Grids (CPSR-SG), Joint Workshop on*, pages 1–6. IEEE, 2016.

[118] A Berizzi. The italian 2003 blackout. In *Power Engineering Society General Meeting, 2004. IEEE*, pages 1673–1679. IEEE, 2004.

[119] Benjamin A Carreras, David E Newman, Ian Dobson, and Naga S Degala. Validating opa with wecc data. In *System Sciences (HICSS), 2013 46th Hawaii International Conference on*, pages 2197–2204. IEEE, 2013.

[120] J Lewis Blackburn and Thomas J Domin. *Protective relaying: principles and applications*. CRC press, 2015.

[121] Jeff Roberts and Armando Guzman. Directional element design and evaluation. In *proceedings of the 21st Annual Western Protective Relay Conference, Spokane, WA*, 1994.

[122] NPTEL. National program on technology enhanced learning. http://www.nptel.ac.in/courses/108101039/download/lecture-15.pdf.

[123] Yakup Koc, Trivik Verma, Nuno AM Araujo, and Martijn Warnier. Matcasc: A tool to analyse cascading line outages in power grids. In *Intelligent Energy Systems (IWIES), 2013 IEEE International Workshop on*, pages 143–148. IEEE, 2013.

[124] OpenDSS PVSystem Model and OpenDSS Storage Element. Opendss manual. *EPRI,[Online] Available at: http://sourceforge. net/apps/mediawiki/electricdss/index. php*.

[125] ICSEG. Illinois center for a smarter electric grid(icseg). http://http://icseg.iti.illinois.edu/wscc-9-bus-system/.

[126] PowerWorld Simulator. Version 10.0 scopf. *PVQV, PowerWorld Corporation, Champaign, IL*, 61820, 2005.

[127] Janos Sztipanovits and Gabor Karsai. Model-integrated computing. *Computer*, 30(4):110–111, 1997.

[128] https://webgme.org/.

[129] http://icseg.iti.illinois.edu/ieee-39-bus-system/.

[130] Grady Booch. *The unified modeling language user guide*. 2005.

[131] Saqib Hasan, Ajay Chhokra, Abhishek Dubey, Nagabhushan Mahadevan, and Gabor Karsai. A simulation testbed for cascade analysis. *IEEE PES Innovative Smart Grid Technologies*, 2017.

[132] April Reese. Blackouts cast Australia's green energy in dim light, 2017.

[133] Margaret J Eppstein and Paul DH Hines. A random chemistry algorithm for identifying collections of multiple contingencies that initiate cascading failure. *IEEE Transactions on Power Systems*, 27(3), 2012.

[134] Vaibhav Donde, Vanessa López, Bernard Lesieutre, Ali Pinar, Chao Yang, and Juan Meza. Severe multiple contingency screening in electric power systems. *IEEE Transactions on Power Systems*, 23(2):406–417, 2008.

[135] Daniel Bienstock and Abhinav Verma. The nk problem in power grids: New models, formulations, and numerical experiments. *SIAM Journal on Optimization*, 20(5):2352–2380, 2010.

[136] Claudio M Rocco, Jose Emmanuel Ramirez-Marquez, Daniel E Salazar, and Cesar Yajure. Assessing the vulnerability of a power system through a multiple objective contingency screening approach. *IEEE Transactions on Reliability*, 60(2):394–403, 2011.

[137] http://icseg.iti.illinois.edu/ieee-57-bus-system/.

[138] Amin Ghafouri, Aron Laszka, Abhishek Dubey, and Xenofon Koutsoukos. Optimal detection of faulty sensors used in route planning. In *Science of Smart City Operations and Platforms Engineering (SCOPE)*, 2017.

[139] Destructive cyber attacks increase in frequency, sophistication. *AFCEA,[Online] Available at: https://www.afcea.org/content/Article-destructive-cyber-attacks-increase-frequency-sophistication*.

[140] Hui Lin, Homa Alemzadeh, Daniel Chen, Zbigniew Kalbarczyk, and Ravishankar K Iyer. Safety-critical cyber-physical attacks: Analysis, detection, and mitigation. In *Proceedings of the Symposium and Bootcamp on the Science of Security*, pages 82–89. ACM, 2016.

[141] http://icseg.iti.illinois.edu/power-cases/.

[142] Saqib Hasan, Amin Ghafouri, Abhishek Dubey, Gabor Karsai, and Xenofon Kout-

soukos. Heuristics-based approach for identifying critical n- k contingencies in power systems. *Resilience Week*, 2017.

[143] TC Robert M Lee, Michael J Assante, and Tim Conway. Analysis of the cyber attack on the ukranian power grid. defense use case. *SANS ICS*, 2016.

[144] J Duncan Glover, Mulukutla S Sarma, and Thomas Overbye. *Power System Analysis & Design, SI Version*. Cengage Learning, 2012.

[145] Jun Yan, Yihai Zhu, Haibo He, and Yan Sun. Multi-contingency cascading analysis of smart grid based on self-organizing map. *IEEE Transactions on Information Forensics and Security*, 8(4):646–656, 2013.

[146] Kjell Hausken and Gregory Levitin. Minmax defense strategy for complex multi-state systems. *Reliability Engineering & System Safety*, 94(2):577–587, 2009.

[147] Helen Pidd. India blackouts leave 700 million without power. *The guardian*, 31, 2012.

[148] NERC Planning Standard. North american electric reliability council, sep. 1997.