TOWARDS THE AUTOMATION OF DATA-DRIVEN REMEDIAL ACTION SCHEME DESIGN

A Dissertation

by

HANYUE LI

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Thomas J. Overbye
Committee Members,	Katherine R. Davis
	Chao Tian
	Timothy A. Davis
Head of Department,	Miroslav M. Begovic

August 2021

Major Subject: Electrical Engineering

ABSTRACT

The emerging advancements in the electric power industry are constantly redefining the scope of power system planning and operations, where engineering decisions made are reliant on the most up-to-date knowledge of the grid. Managing and operating a conventional electric grid is a task in which engineers and operators have a century of collected experience. However, the understanding we obtained from previous experiences might not be sufficient for the grid of the future.

Traditional remedial action scheme (RAS) design takes a holistic approach that requires years of expertise in a specific power system. Building on the industrys current practice, this dissertation develops a power system tool called Auto-RAS that provides a systematic approach to create remedial action schemes in a robust, effective, and automated manner. Leveraging data-driven techniques, responses and control practices of the current power system are contextualized as statistical characteristics and mathematical expressions to guide the design of remedial action schemes.

To maintain a similar level of size and modeling complexity as the real power system while still be share the research result publicly and freely, synthetic electric grid network models are used as test cases for the development and testing of Auto-RAS. In this dissertation, a chronological power system operation simulation framework is developed to provide large amount of scenarios representing a wide spectrum of system operating conditions as input data for the Auto-RAS process.

The determination of Auto-RAS condition logic starts with operational scenario analysis, where the need of remedial action schemes are identified as a list of severe violation system elements. A two-stage linear SVM algorithm is implemented to selected features that can best represent the operational scenarios, and learns a hyperplane that can optimally divide the scenarios with and without risks of severe violations. The selected scenario features, along with the learned hyperplane, and list of violation-causing contingencies are then leveraged as Auto-RAS condition logic.

A sensitivity-based methodology is developed in this dissertation to create the corrective ac-

tions that can be deployed adaptively for remedial action scheme. Leveraging network connectivity analysis, a subset of power system elements that can be controlled as part of the corrective action scheme is selected. Sensitivity analysis such as line outage distribution factor and transmission loading relief are used to quickly determine the most effective controllable elements and the corresponding corrective actions to address specific operational violations.

To evaluate the performance of remedial action schemes developed by the Auto-RAS framework, a checklist is developed to follow typical industrial system planning standard, and to ensure that the designed remedial action schemes can operate to perform their intended functionalities, and do not introduce unintentional or unacceptable reliability risks to the bulk electric power system.

ACKNOWLEDGMENTS

First of all, I thank my advisor, Prof. Thomas Overbye, and my Ph.D. committee members, for their guidance of this work. Prof. Overbye's vision, sincerity, and motivation have deeply inspired me as a student, researcher, and power engineer.

I am immensely grateful for the friendship and camaraderie from Team Overbye, and the Texas Power and Energy Conference organizing committee over the years. I am glad to have shared the Inwood Drive house with Jessica Wert. All the times we had discussing research ideas over tea at the front porch swing will be missed dearly. I thoroughly enjoyed all the long nights spent organizing TPEC since 2017 with the power and energy community at Texas A&M University. Wei Trinh's unlimited snack provision, and Arun Karngala's dance celebration are some of my favorite moments to remember. We always joke that one day we will be invited back as seminar speakers, or TPEC keynote speakers. I am eagerly waiting for that day to come.

I am thankful for my wonderful husband, Raymund Lee, for his support ever since the very beginning of my Ph.D. journey. With him I share the greatest passion for the power and energy industry. From him I gain endless motivations to stay resilient and focused when things get difficult.

Lastly, I would like to thank my family, especially my mother and father, Shanhong Yang and Lian Li, for supporting me unconditionally in all my pursuits.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of advisor Professor Thomas Overbye, Professor Katherine Davis, and Professor Chao Tian of the Department of Electrical and Computer Engineering and Professor Timothy Davis of the Department of Computer Science.

The work of synthetic load time series development and scenario creation in Chapter 4 and 5 was developed with assistance from Ashly Smith, Juhee Yeo, Jessica Wert, who are students in the Department of Electrical and Computer Engineering working under Prof.Thomas Overbye.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

This work is funded in part by the Advanced Research Projects Agency Energy (ARPA-E), U.S. Department of Energy, under Award Number DEAR0000714.

This work is funded in part by the Sandia National Laboratories, U.S. Department of Energy, under projects Laboratory Directed Research & Development (LDRD) and Harmonized Automatic Relay Mitigation of Nefarious Intentional Events (HARMONIE) Special Protection Scheme.

This work is funded in part by the National Science Foundation (NSF), under Grant Number 1916142.

This work is funded in part by the Power Systems Engineering Research Center (PSERC), under Project Number S-91.

This work is funded in part by the Thomas W. Powell Fellowship from the Department of Electrical and Computer Engineering at Texas A&M University.

The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

1. INTRODUCTION

At the beginning of this century, the U.S. National Academy of Engineering (NAE) looked back at the past 100 years, and named the development of the electric grid as one of the greatest engineering achievements of the 20th century [1]. Over time, the power grid became the lifeblood of society by delivering power from generating plants to the end-use customers, and supporting the economy and daily function of the world.

To adapt to the changing need of the modern society, the first two decades of the 21st century are experiencing an unprecedented rate of reform and development in power systems. In the early 2000s, to mitigate the monopoly and create a good business environment, part of the United States electric power industry was deregulated and competitive markets were introduced [2]. With increasing electricity production and lower market prices, natural gas displaced coal as the top fuel type of electricity in the United States for the first time in 2016 [3, 4]. Following rising environmental concerns, the installation of renewable generation also increased rapidly. Over the past 20 years, the total wind and solar capacity have increased by more than 90,000 and 50,000 MW respectively in the United States [5]. By 2020, renewable energy sources provided more than 20% of the total U.S. electricity generation [6].

Incentivized by the monetary opportunities in the deregulated electricity market and benefits of using renewable energy resources, though developed at a slower rate, transmission projects are constantly being proposed and constructed [7]. The grain belt express project, for example, is a 800-mile overhead DC transmission line. This project is designed to deliver 4 GW of power from western Kansas to Missouri, Illinois, Indiana, and neighboring states, tapping into one of the strongest renewable energy resources in the United States [8]. In Texas, to connect areas with abundant wind resources to more highly populated load centers, the Competitive Renewable Energy Zones (CREZ) project was completed in 2013 to include 3,500 miles of transmission lines capable of carrying 18,500 MW of electricity [9].

The rapid evolution and expansion of the electric power industry introduced new challenges in

power system operations. California is the first state in the United States to debut its deregulated electricity market in the year 1996 [10]. Under a deregulated environment, power system operators maintain overall system reliability while managing the competitive electricity market [11]. Six years after its debut, the new environment of system operations was profoundly disturbed by the shortage of generating capacity, wholesale generator market power, and flaws in the market design, where several participants have seized the opportunity to manipulate the market price [10, 12]. An energy crisis that lasted a few years was experienced with high wholesale electricity prices, intermittent power shortages, and deterioration of the investor-owned utilities' financial stability in the California electric power industry [13].

On August 14, 2003, the North American electric power network experienced its largest blackout ever. This incident started with a few unplanned outages in Ohio, then escalated into a cascading blackout under territories of several entities and ultimately affected a total of over 50 million people and more than 70,000 MW of load in both United States and Canada [14]. According to the final report of investigation, human errors in real-time operations and ineffectiveness of the control room energy management system (EMS) usage are some of the leading causes of the blackout [15]. The failure to understand the inadequacies of the system, and the lack of situational awareness about the deteriorating condition of the system contributed to the escalation of some relatively local events to a wide-area system-level blackout.

The emerging advancements in the electric power industry are redefining the scope of power system planning and operations, where engineering decisions made are reliant on the most up-to-date knowledge of the grid [16]. As we are learning lessons from previous incidents, and gaining more experience from today's power industry, the fast-evolving electric grid will continue challenging power system engineers and operators in the future. Managing and operating a conventional electric grid is a task in which engineers and operators have a century of collected experience [17]. However, the understanding we obtained from previous experiences might not be sufficient enough to plan and operate the future grid in a secure and reliable manner [18]. Because of this, the electric power industry needs a more systematic and robust way of gaining knowledge.

Over the past 30 years, electric power systems are taking drastic advances in implementing information and communication technologies. Numerous measurement devices such as advanced metering infrastructure, distributed energy resources monitoring system, and high frequency wide-area situational awareness system are generating immense volume of energy data. With the sensors and information technology increasingly permeating modern power systems, the ubiquity of electricity market and system operation data can potentially provide enormous opportunity to further optimize the power systems. Valuable information and insights can be discovered from the massively collected data as aiding knowledge for engineering decisions for electric power systems [19, 20, 21]. In addition, various visualization techniques have been developed by the power industry to visualize the data, and improve the situational awareness of wide-area systems over the past decades. [22, 23, 24]

Building on the industry's current practice of power system planning and operation, this work proposes a systematic approach to create remedial action schemes using data-driven machine learning techniques and power system sensitivity analysis. The current practice of remedial action scheme development is passive, and potentially not accurate in real-time operation. The creation or update of one remedial action scheme is usually proposed by its owner, which usually are generator owners, transmission owners and distribution providers. The remedial action scheme is then approved, monitored, and implemented by system operators [25]. The need for a remedial action scheme development is not actively recognized in the control room. Also, the design of remedial action scheme is a time-consuming process. Even though remedial action schemes are used in real-time operations to take corrective actions to make sure the system is operated in a reliable and effective manner, they are often created using off-line studies on a limited set of test cases, which may not represent the real-time operating conditions accurately.

In this work, a planning tool called Auto-RAS is developed to systematically identify the need of remedial action schemes under critical system conditions, and to generate and test the conditions and corrective actions of remedial action schemes in a automated manner. The creation of remedial action schemes will utilize large volume of power system operation data and machine

learning techniques, which can identify the system stress and violations pattern that is not obvious to human eyes, so that the most up-to-date insights can provide more informed, effective, and reliable decision making in the power system control room. The creation, testing and applications of Auto-RAS are demonstrated using synthetic power systems, while the approach is general enough to be applied to any power systems, regardless of being synthetic or real.

"It is always wise to look ahead, but difficult to look further than you can see." The quote of Winston Churchill is appropriate here to guide the research work that hopes to develop robust techniques that will be useful for the operations and plannings of power system in the future. It is insufficient to think that the future is just some extrapolated views of the present, and impossible to exhaustively list all the possibilities of the future trajectory [18]. When the electric power industry is advancing at an unprecedented rate, there exist uncertainties in the applicability of previous engineering experiences. The development of Auto-RAS can provide assistance to the adaptivity and flexibility of system planning and operation, where the need of new remedial action schemes can be identified, and RAS models and their settings can be determined in a systematic and automated manner. Leveraging data-driven techniques, responses and control practices of the current system can be contextualized as statistical characteristics and mathematical expressions.

2. BACKGROUND

2.1 Power System Planning and Operation Overview

Power system planning is the development of plans for designing and construction of the system and its elements, which will satisfy assumed future needs, starting from the current state [26]. The planning of bulk power system contains two parts, the generation planning, and transmission planning. Power generation planning is performed in the context of improvements and modifications to the existing system. This process begins with electricity load forecasting, followed by reliability evaluation to determine if and when additional generation is needed. Finally, optimal capacity expansions are selected based on economic considerations [27].

Transmission planning ensures that the transmission facilities can deliver power from the generators to the demands, and that all the equipment will remain within its operating limits in both normal operation and during system contingencies and events when unexpected failure of system elements occurs. During contingencies and events, power systems transition to new operating conditions. Studying these transitions and ensuring that a stable operating points can be reached after any contingency is an essential part of transmission system planning [28].

The power system operation's main objectives are safety, reliability and efficiency. System operation has always been considered as a critical function of modern society. It affects people's safety, impacts system reliability and influences the operational cost associated with the deployment of transmission and generation resources [16].

The U.S. Department of Energy has summarized the following four rules for the reliability of power system operation [29]:

- Power generation and transmission capacity must be sufficient for peak demand
- Power system must have enough flexibility to address variable and uncertainties in demand and generation resources
- Power system must be able to maintain steady frequency

- Power system must be able to maintain steady voltage at various points on the grid

Those operation rules ensure the power system is operated to supply the aggregated demand of customers at all times, taking into account scheduled and reasonable expected unscheduled outages of system elements, on the other hand to withstand sudden disturbances such as electric short circuits, unanticipated loss of system components [17, 16]. In today's power industry, North American Electric Reliability Corporation (NERC) and Federal Energy Regulatory Commission (FERC) establish and oversee standard reflecting aspects such as resource and demand balancing, transmission operations, and interchange scheduling and coordination to ensure reliability of power system operations [30].

In the book Practical Power System Operation, the role of system operator is analogize to drivers operating a vehicle, where the operator needs to make control room actions while examine the driving condition by watching the surroundings of the car, and processing the information provided by the car dashboard [17]. Similar to features like lane departure warning, cruise control, and airbags in a vehicle, in the control room setting, there are programs and techniques developed to assist operator's decision making, and also to automatically control the system when prompt and complicated actions are needed under some conditions.

2.2 Remedial Action Schemes

According to NERC, remedial action scheme is designed to detect predetermined system conditions and automatically take corrective actions that may include, but are not limited to, adjusting or tripping generation (MW and Mvar), tripping load, or reconfiguring a system [31]. It is in place to provide assistance to the system operators to ensure that the system is reliable and stable after any credible contingencies and disturbances.

RAS accomplish objectives such as:

- Meet requirements identified in the NERC Reliability Standards
- Maintain bulk electric system (BES) stability
- Maintain acceptable BES voltages

- Maintain acceptable BES power flows
- Limit the impact of cascading or extreme events

The need to modify existing RAS or design a new RAS is usually identified during system planning studies, which can range from utility to interconnection scales [25]. From the utility scale, RAS entities can submit proposals to their reliability coordinator (RC) for RAS introduction, modification, and retirement. The RC is responsible for approving or providing revision comments after reviewing the proposal. On an interconnection scale, RCs are required to review the existing RAS within their territories every five years. All the deficiencies identified during this review process should be addressed by the corresponding RAS owners within six months [32].

The remedial action scheme design guide established by Western Electricity Coordinating Council (WECC) summarized four typical RAS features, including arming criteria, initiating conditions, action taken, and the time requirements [25]. The arming criteria are critical system conditions for which a step-wise RAS should be ready to take action when required. In many cases this can be the path flows being close to their operational limits [33]. The initiating conditions are the critical contingencies that have been known to cause violations or instabilities from the previous studies, which will initiate RAS action if the scheme is armed. The imitating conditions can be event-based, parameter-based, response-based, or the combination of the above. Event-based schemes directly detect outages and/or fault events and initiate actions to fully or partially mitigate the event consequence. Parameter-based schemes measure variables for which a significant change confirms the occurrence of a critical event. Response-based schemes monitor system response during events and disturbances and incorporate a closed-loop process to react to actual system conditions [34].

For a critical contingency, various potential remedial actions are usually available to mitigate the violations and improve system performance. The RAS actions should at least result in an acceptable system operating condition after the occurrence of a critical contingency or a disturbance. The corrective action should not introduce unintentional risks or interfere with the performance of other RAS in the system. The corrective actions may include but are not limited to [25]:

- Islanding or other line opening
- Generator tripping
- Generator runback
- Load shedding
- Braking resistors
- Static VAR control units
- Capacitor and/or reactor switching

As mentioned in the previous section, for current practice, RAS parameters are usually determined ahead of real-time using a set of pre-determined test cases that extrapolates the system operation conditions. The design decisions such as what planning scenarios to use, and what corrective actions to take are heavily dependent on engineering judgements from industry experts. With the rapid advancements in the electric power industry, power system operations and RAS implementation are compelled to keep up with the rate of progression. With the increasing penetration of renewable energy resources such as solar photovoltaic (PV) systems, wind farms and battery storage systems, the evolving electricity market policies and resulting market participant behaviors, and more frequent occurrence of unpredictable grid disturbances caused by natural disasters and malicious attacks, the understanding we obtained from previous operational experiences might not be enough to operate the future grid in a secure and reliable manner, and a traditional manual and holistic RAS design process may not suffice [35, 36, 37].

2.3 Synthetic Power Systems

With concerns over the security of power grids against malicious physical or cyber attacks, much of the data associated with the United States electric grid is considered to be Critical Energy Infrastructure Information (CEII). The access to CEII is very restricted, many times it is only available to regulators, utilities, and some researches under non-disclosure agreements (NDAs) [38]. To enable the free access of electric grid data of large scale and realistic complexity, many research efforts and progress have made recently in the creation of synthetic power systems.

Synthetic power systems are public grid models and data set that are created to have similar size, structure, and features as the actual electric grid [39, 40, 41, 42, 43, 44, 45]. Many of the synthetic power systems are also located on real geographic footprint, which can be tied to existing public energy data [42, 43]. However, the synthetic networks do not represent, or contain any confidential information of the actual grid. Synthetic power systems are also validated against a wide range of statistical and functional characteristics found in the actual grids, so that they are trustworthy and effective to be used for power system simulations and analysis [46, 47].

This work utilizes synthetic power systems developed at Texas A&M University. All the cases are public and can be accessed from [48]. Those synthetic power systems are used as the base cases in the proposed work, based on which scenarios representing different system operating conditions under various uncertainties are developed.

Figure 2.1 and 2.2 show the one-line diagram and the geographic footprint of the ACTIVSg200 synthetic system. This synthetic grid is located in Central Illinois, United States. It contains 200 transmission buses, 49 generators, four switched shunts, and 180 transmission lines with voltage levels ranging from 115 kV to 230 kV.The 230 kV network is colored with purple, and the 115 kV network is green in Figure 2.1.

The ACTIVSg2000 synthetic system is a fictitious power network on the geographic footprint of Texas, United States. It includes detailed modeling of generators, loads, transmission lines, and other power system elements. Figure 2.3 shows the one-line diagram of the ACTIVSg2000 synthetic transmission system. It has 2000 buses and 3206 transmission branches. The total electric load is 67 GW and the total generation capacity is 100 GW. The 500-kV, 230-kV and 115-kV network in this test case are represented as the orange, purple and green lines in the one-line diagram, respectively. Similar to the actual grid in Texas, the wind generation resource is designed to be abundant in the far west and panhandle area. Load centers are in the east and central side of the system, where Dallas, Houston, and Austin metropolitan areas are located. Higher voltage transmission infrastructures are constructed to connect the generation facilities to the load clusters.

Those synthetic test cases have initial ac power flow solutions, and they are N-1 secure. They

also contain several common control schemes and models such as switched shunts and load tap changers for reactive power control, generators that can remotely regulate voltages, phase shift transformers, automatic generation control (AGC), etc. On the dynamics side, they model the key generator controls such as the exciter, governor, and stabilizer, and even protection devices such as under-frequency load relays. Thus, while there are several controls in place to aid system security and stability in case of contingencies, there are no particular RAS developed for this test case. By developing RAS from scratch for a synthetic, realistic system, we propose a method to design new RAS in real grids.

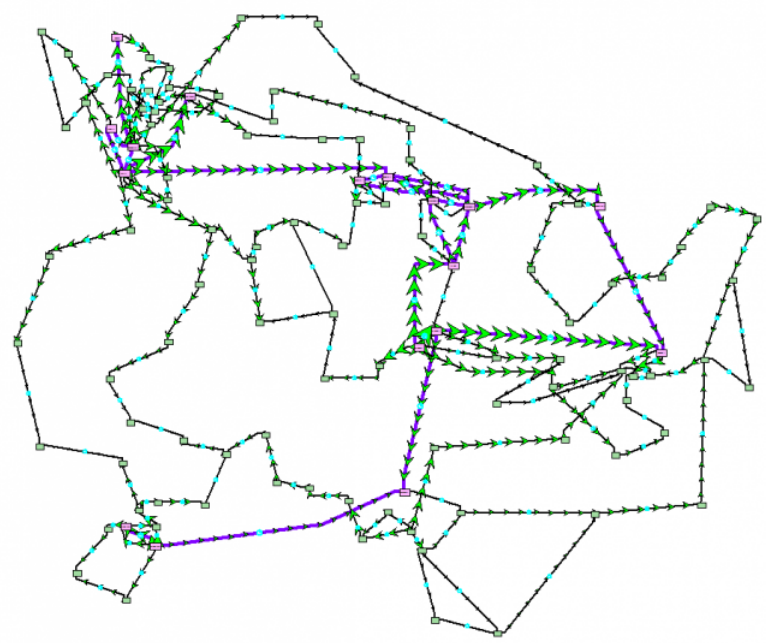

Figure 2.1: One-line diagram for ACTIVSg200 synthetic system

Figure 2.2: ACTIVSg200 synthetic system geographic footprint

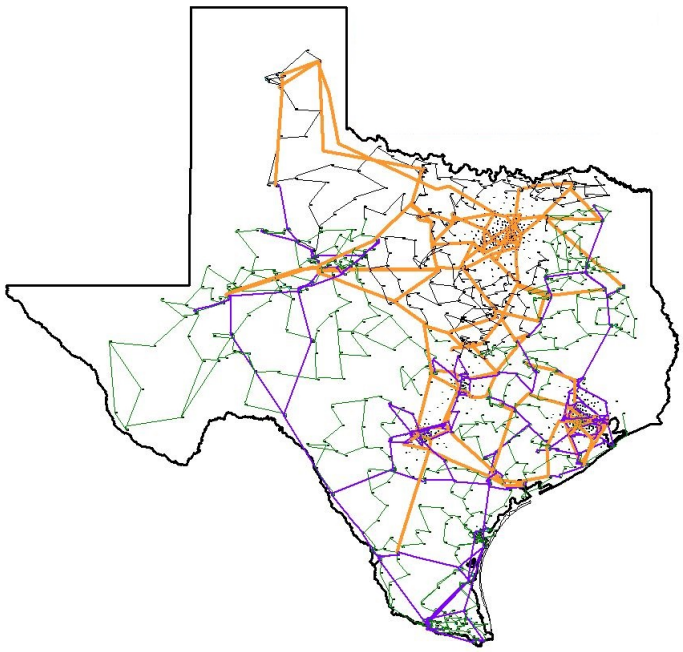

Figure 2.3: One-line diagram for ACTIVSg2000 synthetic system

2.4 Machine Learning Techniques

Machine learning techniques are algorithms that use statistics to find patterns in large amount of data in order to make decisions and predictions based on new data provided. It is a branch of artificial intelligence (AI) focused on building applications which learn from data, and improve their accuracy over time [49, 50]. In today's world, examples of machine learning applications are all around us. Smart digital assistants can initiate google search, play media and control small electronic devices in response to voice commands. Robots vacuum cleaners can learn the floor plan without having an embedded camera, and optimize the vacuuming route to reduce the cleaning time. Websites and social media can recommend us products and services based on our demographic group, or browsing history. Medical image processing and analysis systems can help doctors to spot abnormalities they might have missed. As the data being collected from digital devices getting bigger and more diverse, and computing becomes more powerful and affordable, machine learning is becoming more common and driving greater efficiencies in many aspects of the societies [51, 52, 53, 54].

There are three typical categories of machine learning methods: supervised learning, unsupervised learning, and semi-supervised learning. Supervised machine learning trains on a labeled data set. It refers to a class of algorithms that determine a predictive model using data points labeled with known outcomes. A supervised learning model typically works through some type of optimization routine to minimize a loss or error function. In other words, supervised learning is the process of training a model by providing it input data as well as desired output data. This input/output pair is usually referred to as "labeled data". For example, a computer vision model designed for driverless vehicles can be trained to identify stop signs based on a data set of various labeled road signs images. Supervised machine learning usually requires less training data than other machine learning methods because the output of the model can be directly compared to expected labeled results.

Unsupervised machine learning ingests large volumes of unlabeled data and utilizes algorithms to extract meaningful features, and discover patterns and relationships in data set that is not obvious

to human eyes. Its ability to identify similarities and differences in information make it an ideal solution for exploratory data analysis, cross-selling strategies, customer segmentation, and image recognition. Take spam email detection for example, the amount of emails people send out and receive everyday is enormous, which it is impossible for data scientist to analyze and label all of them in a short period of time. An unsupervised learning algorithm can analyze huge volumes of emails and uncover the features and patterns that indicate spam without human intervention.

As the name suggests, semi-supervised learning is a medium between supervised and unsupervised learning. During training, only a small portion of the data set is labeled to guide the classification and feature extraction from a larger, unlabeled data set. Semi-supervised learning can solve the problem of having not enough labeled data to train a model using a supervised learning algorithm.

2.4.1 Support Vector Machine

Support vector machine is a supervised learning technique. The objective of the SVM algorithm is to find a hyperplane in a multi-dimensional space that is able to directly separate data points of different classes. Here the number of dimensions equals the number of features representing each data point. If the number of features is two, then the hyperplane is a straight line. If the features are three-dimensional, then a two-dimensional plane can be founded to separate the data.

To separate the two classes of data points, there are many different hyperplanes that can be chosen, and one of the most common one is the plane that has the maximum margin between data points of both classes. This optimal separating hyperplane not only provides a unique solution to the classification problem, by maximizing the margin distance between the two classes on the training data, a better classification performance on the testing data can be achieved [55, 50].

Define a hyperplane L by

$$\{\mathbf{x} : \mathbf{f}(\mathbf{x}) = \mathbf{x}^{\mathbf{T}}\beta + \beta_{\mathbf{0}} = \mathbf{0}\}, \tag{2.1}$$

it is the linear algebra property that $\beta^* = \beta/\|\beta\|$ is the vector normal to the surface of L, and for

any point x_0 on the hyperplane L, $\beta^T x_0 = -\beta_0$. The signed distance of any point x to L is given by

$$\beta^{*\mathbf{T}}(\mathbf{x} - \mathbf{x_0}) = (\beta^{\mathbf{T}}\mathbf{x} + \beta_\mathbf{0})/\|\beta\| = \mathbf{f}(\mathbf{x})/\|\beta\|. \tag{2.2}$$

Hence $f(x)$ is proportional to the signed distance from x to the hyperplane defined by $f(x) = 0$.

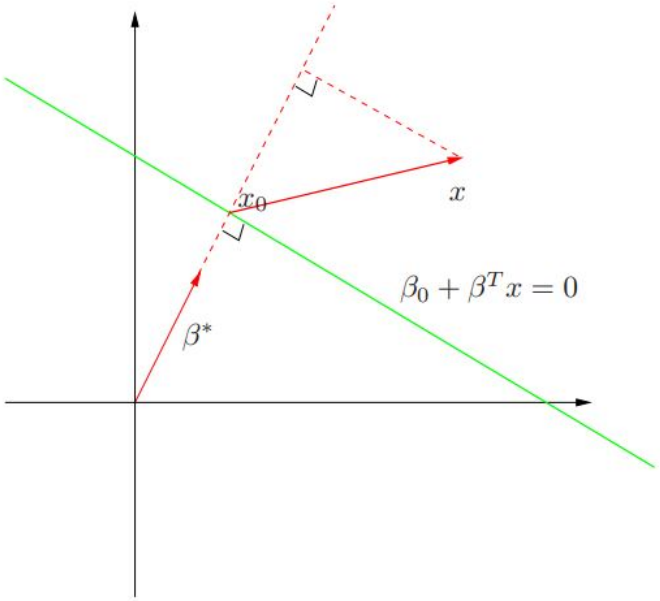

Figure 2.4: The linear algebra of a hyperplane (affine set)

Consider training data consists of N pairs $(x_1, y_1), (x_2, y_2), ..., (x_N, y_N)$, with $x_i \in \mathbb{R}^p$ and $y_i \in \{-1, 1\}$. If the classes are completely linearly separable, there exist a hyperplane $f(x) = x^T\beta + \beta_0$ with $y_i f(x_i) > 0\ \forall i$. With $1/\|\beta\|$ set to be the smallest distance between data points and the hyperplane, the optimization problem formulation that creates the largest margin between the training class 1 and -1 is

$$\begin{aligned} &\min_{\beta,\beta_0} \|\beta\| \\ &\text{subject to} \quad \mathbf{y_i}(\mathbf{x_i^T}\beta + \beta_\mathbf{0}) \geq \mathbf{1}, \quad \mathbf{i = 1, ..., N}. \end{aligned} \tag{2.3}$$

If for certain problems the classes overlap in the feature space, we can still find a hyperplane

that maximizes the margin, but allow of some points to be on the wrong side. This is done by defining the slack variable $\xi = (\xi_1, \xi_2, ...\xi_N)$, and modifying the constraints in 7.1 as:

$$\begin{aligned} &\mathbf{y_i}(\mathbf{x_i^T}\beta + \beta_\mathbf{0}) \geq \mathbf{1} - \xi_\mathbf{i}, \\ &\xi_\mathbf{i} \geq \mathbf{0}, \quad \mathbf{i} = \mathbf{1}, ..., \mathbf{N} \end{aligned} \tag{2.4}$$

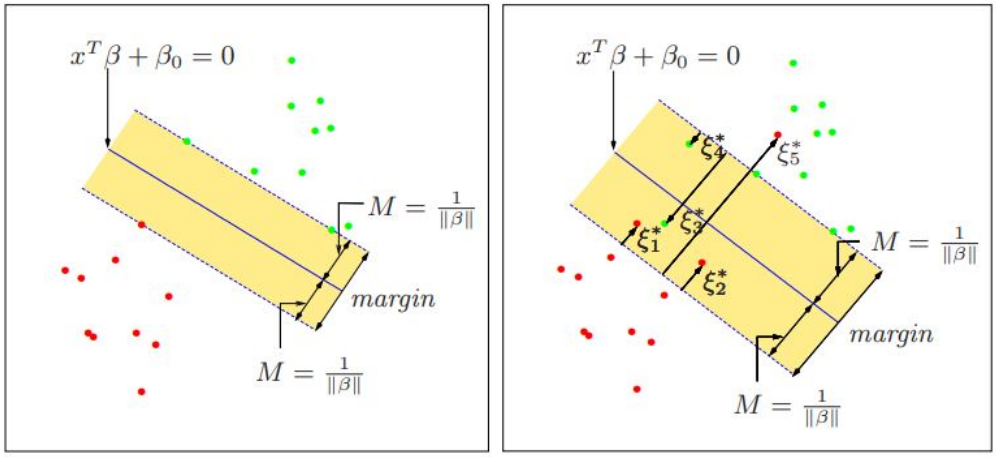

Figure 2.5: Support vector machine classifiers. The left panel shows a separable case. The right panel shows a non-separable (overlap) case.

If the classes of input data is not linearly separable in the original, feature space, kernel transformations can be applied to the data, and map the data from the original feature space into a higher dimensional space. The goal is that after the kernel transformation, the classes of data points are linearly separable in the higher dimensional feature space. We can then fit a decision boundary to separate the classes and make predictions. The decision boundary will be a hyperplane in this higher dimensional space.

2.4.2 Hierarchical Clustering

Clustering is an unsupervised learning technique. It aims to group and segment a collection of objects into subset of clusters, so that the data points within each cluster are more closely related to

one another than objects assigned to other groups. K-means clustering is one of the most common clustering algorithms used. To group the training data into pre-dertermined number of clusters, K-means clustering starts with a first group of randomly selected centroids, with the number of centroids equals to the number of desired clusters. Given the initial set of cluster centers, the K-means algorithm alternates between the following two steps until convergence:

- for each centroid we identify the subset of training data that is closer to this cluster center than any other cluster center;
- the mean value of each feature for the data points in each cluster are computed, and this mean vector updates the new center for that cluster.

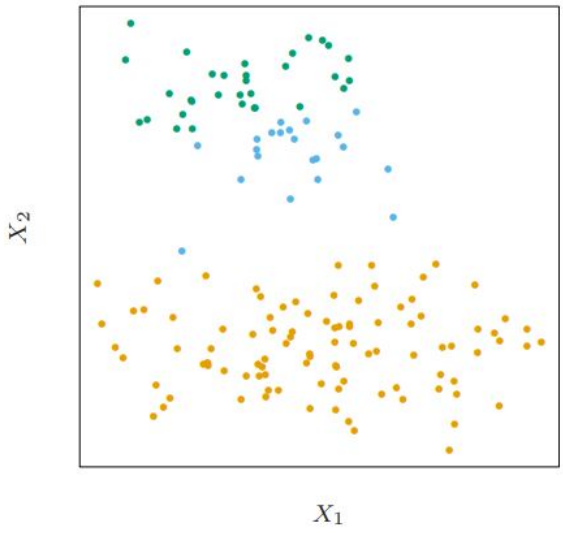

Figure 2.6: Two-dimensional data, clustered into three classes (represented by orange, blue and green) by the K-means clustering algorithm

Sometimes the data can be represented in terms of similarities between pairs of objects. This type of data can be represented by an $N \times N$ matrix D, where N is the number of data points, and each element d_{ij} quantifies the proximity between the ith and jth objects. This matrix is often provided as an input to hierarchical clustering algorithms.

While the results of applying k-means clustering algorithms is dependent on the choice for the number of clusters to be searched and the set of initial centroids, hierarchical clustering methods do not require such specifications. Using the proximity matrix, hierarchical clustering algorithms produce hierarchical decisions in which the clusters at each hierarchy level are created by merging clusters at the next lower level. At the lowest level, each cluster contains a single data point. At the highest level there is only one cluster containing all of the data points.

Strategies for hierarchical clustering can be divided into two basic categories: agglomerative (bottom-up) and divisive (top-down). Agglomerative hierarchical clustering starts at the bottom and at each level iteratively merge a selected pair of clusters into a single cluster. This creates a grouping at the next higher level with one less cluster. The pair chosen for merging consist of the two groups with the smallest intergroup dissimilarity. Divisive hierarchical clustering methods start at the top and at each level iteratively split one of the existing clusters at that level into two new clusters. The split is chosen to produce two new groups with the largest between-group dissimilarity. With both categories there are $N - 1$ levels in the hierarchy.

2.5 Power System Sensitivity Analysis and Sparse Vector Methods

Sensitivities are linearized relationships, which help determine the impact of small changes in a variable on the system. TLR sensitivities gauge the sensitivity of a line flow to multiple power transfers, whereas LODF depicts how the outage of one more more lines, taken one at a time, affects the flows on the other lines. Both these sensitivities can be derived from the Injection Shift Factor (ISF) matrix represented by $\mathbf{\Psi}$ [2, 56, 57], which is constructed using the lossless dc power flow assumptions. The change in real power line flows is given as,

$$\mathbf{\Delta f} = \mathbf{\Psi \Delta p} \tag{2.5}$$

where $\mathbf{\Delta p}$ is the vector of changes in the real power injection at each bus, and the element ψ_l^n in row $\boldsymbol{l}$ and column $\boldsymbol{n}$ of $\mathbf{\Psi}$ is the ISF of line $\boldsymbol{l}$ with respect to the injection at node $\boldsymbol{n}$. For a system

with L lines and N nodes (buses), Ψ is defined as,

$$\Psi = \tilde{\mathrm{B}}\mathrm{A}[\mathrm{B}']^{-1} \tag{2.6}$$

where $\tilde{\mathrm{B}}$ -*diag*$\{b_1, b_2, ..., b_L\}$ (b_l is the series susceptance of line l), $\mathbf{A}$ is an $\mathbf{L} \times \mathbf{N}$ incidence matrix where the element a_{ij} is non-zero only when line l_i is coincident with node j, and $\mathrm{B}' = \mathrm{A}^{\mathrm{T}}\tilde{\mathrm{B}}\mathrm{A}$.

For the sensitivity calculations, all the elements of Ψ are not needed as only certain lines and nodes are considered at a time, as will be shown below. Hence the full matrix as described in (2) does not need to be computed. Rather, the sensitivities of the elements of interest can be found using sparse vector methods [58]. For example, $\mathbf{\Delta p}$ is sparse in this case since the injection(s) contributed by the corrective action(s) are limited in number due to a judicious selection of the candidate nodes and the design criteria used. Because of this, fast forward substitution can be used. Similarly, monitoring a reasonable number of lines can help make $\mathbf{\Delta f}$ sparse, and enable the use of fast backward substitution. This is able to substantially reduce the computational complexity. Once the matrix is factored (an $\mathbf{O(n^{1.4})}$) operation, doing a fast forward and fast backward substitution involves very few (much less than $\mathbf{O(n)}$) operations [59, 58].

3. DISSERTATION OVERVIEW AND LITERATURE REVIEW

Many research efforts, from both industry and academia, have gone into improving the flexibility and dynamics of existing RAS implementation. In [60], an event-based method was proposed to enhance RAS that are created to address specific frequency and voltage instability issues. Using transient energy analysis, the conventional RAS implementation can be adjusted with flexible triggering thresholds [61], and also adaptive corrective actions [62]. To mitigate the risk of voltage instability and voltage collapse, BC Hydro developed a methodology to determine the magnitude of load shedding based on real-time measurement data [63]. Recent work of [64] proposed an approach to adaptively set the arming parameters of existing RAS based on realistic and near real-time operation conditions. All the existing research on remedial action scheme are mainly focused on improving only one of the RAS elements, such as the the enhancement of corrective action design, and increasing the adaptivity of conditions logics on the current remedial action scheme models. To the author's best knowledge, a research work that systematically develops a remedial action scheme from scratch does not exist.

This work introduces the creation, testing, and implementation of a power system planning tool, Auto-RAS. Auto-RAS is designed to systematically identify the need of remedial action schemes under critical system conditions, and to create and test the remedial action schemes conditions logics and corrective actions in an automated manner. To obtain a comprehensive understanding of the type of RAS implemented within the United States electric power system, to recognize the gap between the current RAS design process and the need from power industry planning and operations, and to identify typical characteristics of industrial RAS guiding the design of Auto-RAS, this work firstly reviewed over 150 models of RAS implemented on the WECC footprint in Chapter 4. These RAS models are designed, implemented and managed by 21 utilities in the United States Western Interconnection.

To maintain a similar level of size and modeling complexity as the real power system while still be share the research result publicly and freely, synthetic electric network models are used as

test cases in this work for the development and testing of Auto-RAS. Since the current practice of synthetic grid creation for a specific geographic footprint only contains one base case representing a snapshot of the peak load operating condition, this work developed a methodology to generate and validate unique bus-level hourly time series based on the geographic location of substations and buses for a synthetic system, so that the understating of how a synthetic system "move" over time can be obtained. This synthetic time series creation methodology is explained in detail in Chapter 5.

With the temporal information of a synthetic grid, large volume of operating scenario can be created as input for data-driven applications, so that the patterns of power system operations that are not obvious to human eyes can be identified, and informed decision regarding the design of Auto-RAS can be made. This work developed a chronological power system simulation framework in Chapter 6 that considers the time-varying load and generation level, scheduled outage, area interchange schedules, unit commitment, and economic dispatch, with permutations being implemented at each hour of the simulation year. This simulation method provides large amount of scenarios representing a wide spectrum of system operating conditions for the development and testing of Auto-RAS.

There are two main components in the Auto-RAS framework, the determination of condition logic, and the creation of corrective actions. Condition logic is one of the most important features to distinguish remedial action schemes from other operation control techniques. In urgent situations where the operating conditions are deteriorating at fast rates, or the operational mitigation plan is complicated or not obvious, system operators potentially do not have enough time to react to system events, and conduct control actions in a timely manner. With arming and triggering conditions being defined for RAS ahead of time, and the measurement data from the energy management system in the control room, the condition logic are constantly being monitored. If certain anomalous system condition is detected, the corresponding corrective action can be initiated instantly to protect the system from severe damage or collapse.

The determination of Auto-RAS condition logic is covered in Chapter 7. It takes a data-driven

approach and is reliant on the simulation of diverse operating scenarios. With operational scenario analysis, a list of severe violation system elements is identified, where remedial action schemes are needed. Those severe violation elements are then grouped together, and remedial action scheme is developed for each cluster. For each RAS cluster, a two-stage linear SVM algorithm is implemented to selected features that can best represent the operational scenarios, and learns a hyperplane that can optimally divide the scenarios with and without risks of severe violations. The selected scenario features, along with the learned hyperplane, and list of violation-causing contingencies are then leveraged as Auto-RAS condition logic.

In chapter 8, this work develops a sensitivity-based methodology to create the corrective actions that can be deployed adaptively for each RAS cluster. Leveraging network connectivity analysis, a subset of power system elements that can be controlled as part of the corrective action scheme is selected. Sensitivity analysis such as line outage distribution factor and transmission loading relief are used to quickly determine the most effective controllable elements and the corresponding corrective actions to address specific operational violations in one RAS cluster.

The testing and example results of Auto-RAS developed for the the 200-bus (ACTIVSg200) and the 2000-bus (ACTIVSg2000) synthetic systems are presented in chapter 9. A summary of this research work, as well as insights regarding the future direction are provided in Chapter 10.

4. REVIEW OF THE CURRENT RAS MODELS

4.1 Objective of the Review

There are three main objectives to review the current RAS models, which are fundamental for the development of new techniques and the improvement of RAS design procedures:

- Obtain a comprehensive understanding of the types of RAS implemented within United States electric power systems.
- Recognize the gap between the current RAS design process and the need from power system planning and operations.
- Identify typical characteristics to guide the design of Auto-RAS.

4.2 Scope of the Review

This work reviewed over 150 models of RAS implemented on the WECC footprint. These RAS models are designed, implemented and managed by 21 utilities in the United States Western Interconnection. For confidentiality reasons, the name of the utilities, and some technical details regarding the RAS models are not discussed. The key information summarized from the industrial RAS includes the following items:

- Model Format: The type of software and data format the RAS model is developed in and saved as.
- Impact Magnitude: If the RAS is a Local Area Protection Scheme (LAPS), Wide Area Protection Scheme (WAPS), or Safety Net (SN).
- Objective: The type of operational problem the RAS is designed to protect against.
- Arming Criteria: Critical system conditions for which a step-wise RAS should be ready to take action when required.

- Triggering Condition: The monitored topology or system conditions that trigger the RAS to operate.
- Remedial Actions: The mitigation actions taken.
- Other Technical Details: Parameters and values regarding the arming criteria, triggering condition, and the corrective actions.

4.3 Review Results

It is observed that PSS/E, PSLF and PowerWorld are the three most commonly used software in which RAS is designed and modeled. For PSLF and PowerWorld, *.wcrf and *.aux files are used to store both the remedial action scheme logic and actions respectively. For PSS/E, RAS definitions are defined by the user as Python functions that inform PSS/E what actions to take if a RAS is triggered under certain conditions.

According to NERC, remedial action schemes can be categorized into three types based on the magnitude of impact: Local Area Protection Scheme (LAPS), Wide Area Protection Scheme (WAPS), or Safety Net (SN). WAPS is a RAS whose failure to operate would result in the violations of planning criteria for system performance, or load loss more than 300 MW, or generation loss more than 1000 MW. If a RAS failure does not result in violations of planning criteria for system performance, or significant load loss or generation loss, it is considered a LAPS. Safety Net refers to extreme events with low probability of occurrence but high magnitude of impact. In the WECC area, 62% of the installed RAS are LAPS, 31% of the installed RAS are WAPS, and 7% are SN.

Figure 4.1 shows different types of RAS objectives observed from the industrial RAS models, and their number counts. Thermal overloading, voltage violations and instability, transient instability and cascading outages are five main operational problems RAS are designed to address. Nearly 65% of the RAS are designed to mitigate thermal overloads on transmission lines, transformers, and transmission interfaces and paths. Preventing voltage instability is the second most common single objective of the industrial RAS. System instability and cascading outages, relatively less common, are two other single objectives observed from the industrial RAS models. Additionally,

a notable amount of RAS have multiple objectives. This means that some critical contingencies can potentially cause more than one type of operational problems.

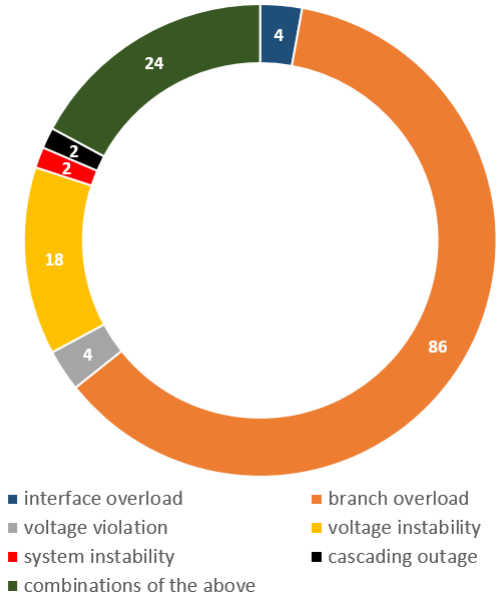

Figure 4.1: Industrial RAS model objectives by type

The arming criteria for almost all RAS definitions depend on system conditions such as injection group generation MW output, transmission line and corridor path flow, rather than changes to system topologies or contingency events. During power system real-time operation, RAS can be armed either manually or automatically. Many times for manually-armed RAS, separate files or programs are developed outside of the RAS definition for monitoring purposes. For automatically-armed RAS, the arming system conditions or topologies changes are determined and included in the contingency definition within the RAS models.

It is also observed that most remedial action schemes trigger upon changes in system topologies, instead of changes in system conditions. Out of all the 152 RAS models reviewed in this work, about one fifth have a single logic triggering condition. In other words, those remedial action schemes each monitors only one variable associated with a particular system element, such as

the status of a transmission line, as part of its triggering condition. The rest of the industrial remedial action schemes implement more complexed logics, with at least two variables of the same system element, or at least two different system components being involved.

Figure 4.2 shows the summary of corrective actions included in the industrial RAS. Transmission branch, substation bus, user-defined injection group, transmission interface, line shunt, electric load, and shunt capacitors are power system elements used in corrective actions of RAS models. There are 25 different types of corrective actions in the industrial RAS models, where generator opening, interface opening, branch opening, generator MW adjusting, and load opening are the five most popular ones. Some of the more specific industrial RAS characteristics used directly to guide the design of Auto-RAS are further discussed in Chapter 8.

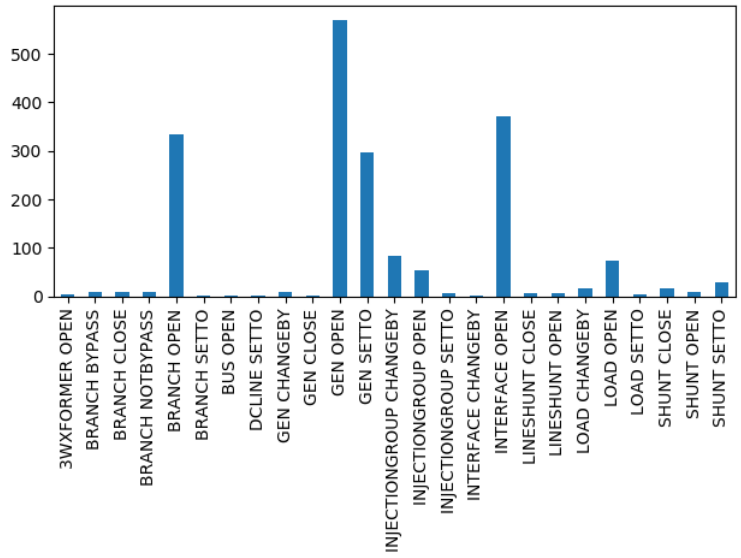

Figure 4.2: Industrial RAS model corrective actions

4.4 Key Findings and Interpretations

With the review of industrial RAS models, typical characteristics, such as the common logics used for arming conditions and triggering conditions, as well as the voltage levels, network properties and element sensitivities of specific RAS cases are summarized and used as Auto-RAS

design references. In the 152 remedial action schemes, inconsistency in modeling convention, decision preference, and design philosophy have been observed among different managing utilities. Only a small number of RAS models come with explanations of the intention and the expected performance of the design. It is also observed that need of a RAS is identified after the system has experienced certain operational issues, and the solutions are developed after the fact based on the EMS data and other measurements of the specific event. This serves as a major motivation for the development of Auto-RAS. Building on the industry's current practise of RAS design, in Auto-RAS, the need of new remedial action schemes are identified, and RAS models and their settings are determined in a systematic and automated manner.

5. CREATION OF LOAD TIME SERIES FOR SYNTHETIC ELECTRIC GRIDS

To create a large volume of power system operation data, where Auto-RAS logic condition patterns can be analyzed, and corrective actions can be tested, varying the load level is the first step. Electric load time series reflect electricity consumption patterns and provides insight on the absolute level and changing rate of load at different times. In real power systems, this information can be provided by historic load values, or load forecasting values. For synthetic power systems, the base case is usually developed to reflect only a one-time snap shot of the system operating condition, synthetic load time series need to be created to represent the load behaviors over time.

This chapter introduces the approach taken to create load time series on bus-level granularity. It is the first and also necessary step for the scenario development for synthetic power systems, which provides a wide spectrum of system operating conditions. The load time series developed for synthetic power systems have hourly resolution with the duration of a whole year, and are created on the bus level so that every bus in the synthetic grid model has a unique profile. Each bus-level load time series is created using an iterative aggregation approach, where prototypical building load profiles are aggregated based on the size and composition of load buses.

5.1 Location and Size of Bus-Level Electric Load

The location and size of the electric loads are determined during the creation of the synthetic base case discussed in [43, 47]. The load buses are located based on the clustering of geographic coordinates associated with postal codes that are obtained from the public U.S. census database. The size of each load bus is then scaled according to the population of the corresponding postal code and the per-capita MW consumption. A fixed power factor is assigned to each load as an

assumption.

The base case is used as a reference to create load time series. The size of load buses in the base case are considered to be the peak value of each bus-level time series, and the geographic coordinates assigned to each load bus are then used to determine the unique location-dependent load features such as load composition ratio and prototypical building load time series.

5.2 Load Composition Ratio

The assignment of a composition ratio of residential, commercial and industrial load on each bus is helpful to realistically represent the uniqueness of load. It establishes the geographic and demographic dependence of electric load similar to reality.

U.S. utility companies' service territories as well as their residential, commercial and industrial megawatt-hours sales values from the Annual Electric Power Industry Report are used to determine the bus load composition ratios [65]. Each load bus is assigned to one utility company based off its geographic coordinates, and the company's sales ratio of the three load types is used as the average bus load composition ratio.

5.3 Prototypical Building- and Facility-Level Load Time Series

To bridge between the bus load composition ratio, and a unique hourly profile, prototypical end user level load time series under residential, commercial and industrial load types are synthesized from public data. Building- and facility-level time series gives the desired bus load a good base to incorporate both individual user load patterns and the aggregation effect. Different categories of buildings and facilities and their prototypical load time series are realistic approximations to represent the most common and important load features.

5.3.1 Prototypical residential and commercial building load

Open source data of simulated hourly building energy consumption are used is used to create prototypical residential and commercial building load [66].

The residential data contains buildings' hourly electricity usage value from space heating/cooling, High Voltage Alternating Current (HVAC) fan, interior/exterior lighting, as well as appliances and

miscellaneous loads. Each data file covers one typical meteorological year 3 (TMY3) location in the United States, which represents geographic locations with different meteorology[67]. For commercial load, under each TMY3 location, 16 building electric load profiles are simulated using the Department of Energy (DOE) commercial reference building models[68], and the contents in each time series data are like that of residential data set.

Under the United States footprint, 1020 residential and 16,320 commercial building time series are calculated as a summation of all electricity consumption categories under each building type. They are created to exhibit unique profiles of residential and commercial buildings, and electricity consuming variations over time and geographic location.

Figure 5.1, for example, shows prototypical residential building load time series in a winter week and a summer week. The load shapes in two seasons are distinguishable, where winter profiles tend to have two peaks in one day due to winter heating, while summer profiles only have one peak per day. The magnitude of load can also be very different in each season, depending on geographic locations. In winter, regions with colder climate such as Helena, Montana, would have higher average load. While in summer, load within hot and arid climate zones, such as Phoenix,Arizona, will have much more electricity consumption.

Similarly, figure 5.2 shows prototypical commercial load time series for the large office building type. The load shape of a specific building type is generally consistent regardless of the location, and the load level is slightly higher in summer season compared to that in winter. In figure 5.3, weekly load profiles of three commercial building types (full-service restaurant, small office, and strip mall) in Los Angeles, California are shown. The load shape and size under each building type is unique. Small offices have steady load during weekdays and low load during weekends. For full-service restaurants and trip malls, load levels are constant through out the week, while full-service restaurants observe two peaks near lunch and dinner time, strip malls only peak once every day.

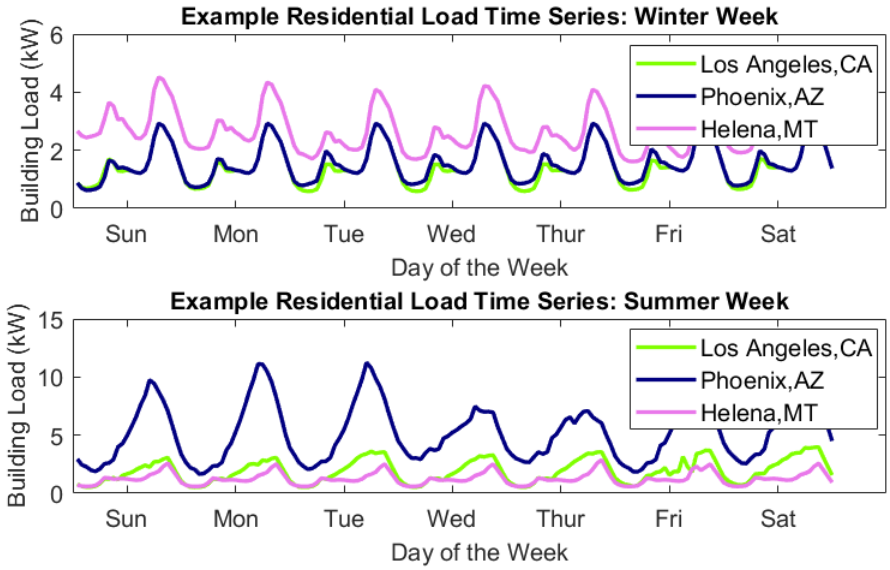

Figure 5.1: Prototypical residential building load time series examples by location

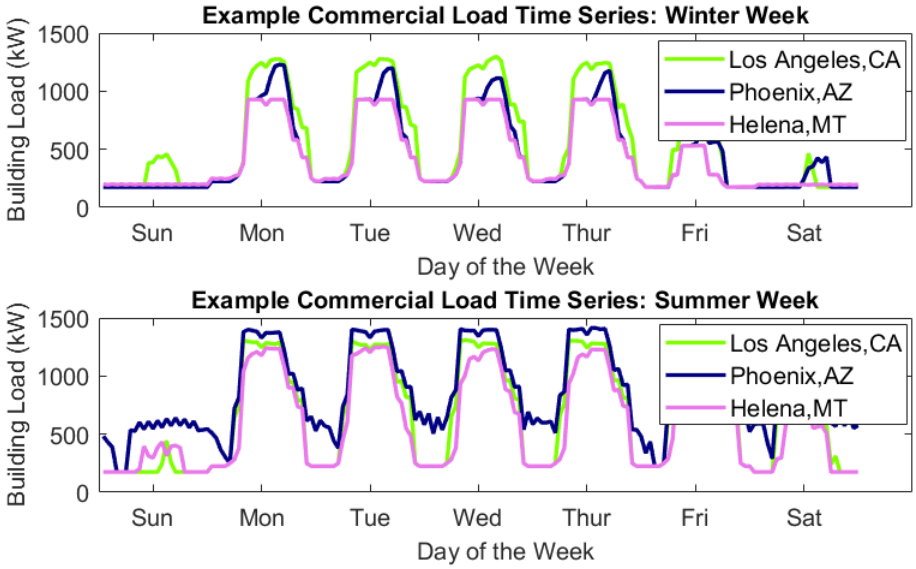

Figure 5.2: Prototypical commercial building load time series examples by location

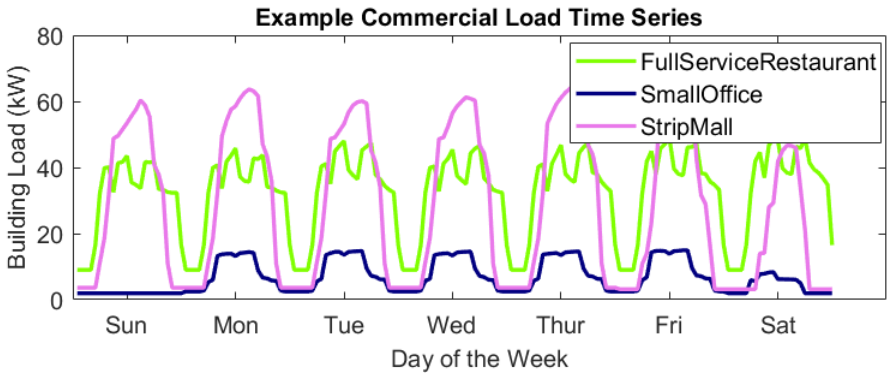

Figure 5.3: Prototypical commercial building load time series examples by type

5.3.2 Prototypical industrial facility load

Prototypical industrial facility load time series are created based on publicly available per-unit industrial load curves from Oak Ridge National Laboratory (ORNL) [69] and the industrial assessment data in Industrial Assessment Centers (IAC) Database [70].

The ORNL per-unit curves provide daily profiles of different industrial sectors, presented by different Standard Industrial Classification (SIC) codes with their unique load factor. The IAC Database contains information on the industry SIC code, total electricity usage and yearly operating hours of over 14,000 facilities in the United States, which are used to modify the ORNL curves into facility-specific load time series for a year.

For each industrial facility, the yearly operating hour is first used to determine the total number of operating days. The ORNL daily curves of the same SIC code is then expanded to a yearly load curve, with small white noise imposed and a random selection of starting day of the year. The synthesized yearly curve is then scaled so that the integral value of the curve matches the total electricity usage.

Figure 5.4 presents prototypical industrial load time series for four facilities from food, petroleum and refining, primary metal, as well as electronic and electrical equipment industries. As those load curves are adopted from the ORNL per-unit daily curves, they have similar daily variations and weekly shapes, with different load levels and load factors.

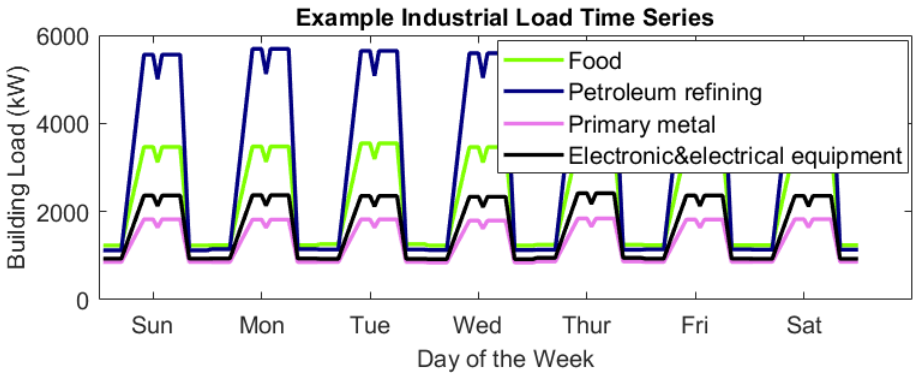

Figure 5.4: Prototypical industrial facility load time series examples by type

5.4 Aggregation of Load

The bus-level load time series is created by iteratively aggregating prototypical building and facility load time series of each load type. This aggregation process has three main aspects: integrating realistic amounts of end users under each load type, selecting representative prototypical time series, and mimicking the effect of load aggregation described in [71].

A flow chart of this aggregation process is shown in figure 5.5. The reference peak values of residential, commercial and industrial bus load type of each bus are first determined by the multiplication of bus load size and the load composition ratio. This is used to integrate realistic amount of end users under each load type, where the peak component values are the stopping criterion for the iterative aggregation process.

A pool of representative prototypical building load time series are then selected for each load bus. For residential and commercial load, the selection used the top five shortest distances between the load bus geographic coordinates and TMY3 locations. All industrial facility load time series that have smaller maximum value than the calculated peak industrial load component are included in this pool since industrial loads are less correlated with geographic locations.

Under each load type, within one iteration, only one building or facility load time series is picked based on a predetermined probabilistic distribution. In the work of [44], a combined transmission and distribution synthetic data set is created using commercially obtained parcel data as

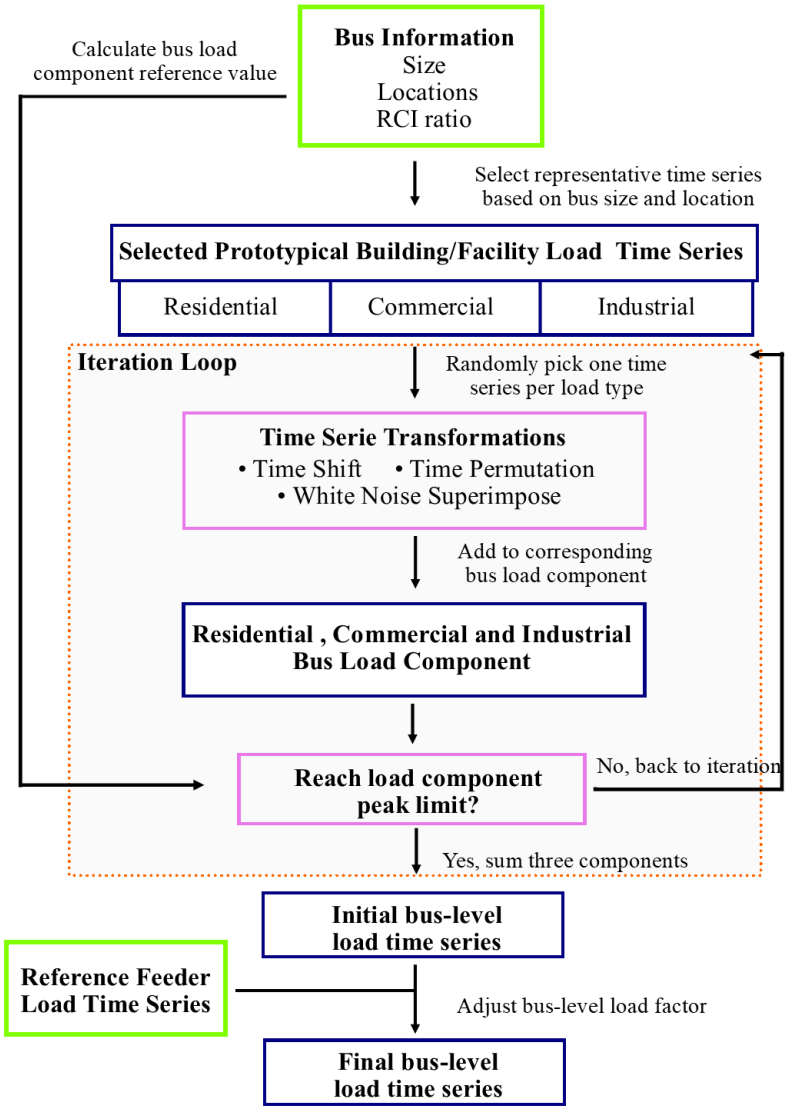

Figure 5.5: Flow chart of heuristic load aggregation approach

one of the inputs, where different building types within the geographic footprint of the test case are extrapolated based on the parcel usage categories. This publicly available data set [72] is utilized to summarize the typical percentage of building types within the service territory of a transmission substation or bus. The region of a transmission substation is defined as the polygon boundary of all the distribution feeders that are serviced by this substation. This selected building- or facility-level

load time series is then processed through three types of transformations: time shift, time permutation, and noise insertion. Those transformations diversify the load profiles of end users, so that the smoothing effect for load aggregation can be produced.

The original prototypical load time series can be shifted both forward and backward up to 12 hours following pre-defined probability mass functions, where the time to be shifted is a discrete integer variable. For residential load class, the distribution of time shift is summarized from publicly available household metering data [73]. The Pecan Street electricity consumption data has 15-minute resolution, and is collected from residential homes mostly located in the state of Texas, California and New York. To be consistent with the prototypical load time series, the metered load time series is downsampled to hourly granularity. Detrended cross-correlation analysis is then conducted for the non-stationary residential load time series [74], where the time lag yields to the peak cross-correlation is considered to be the hour shifted in between two time series. The distribution of shifting hours is then summarized among all the time series pairs as probability mass functions. Due to the lack of reference data in commercial and industrial load classes, heuristically, we assume the probabilities of shifting the time series of those two load classes to be 30% and 50% lower than the residential load class.

To imitate the random surges or drops of load for individual customers, certain hours (100, 100 and 50 hour pairs for residential, commercial and industrial respectively) are randomly chosen within the year to be permutated. As the prototypical load time series used as the input to this process is simulated data, it reflects an expected level of electricity consumption every hour, but does not account for the stochastic behavior of electricity users. For example, figure 5.6 shows a comparison between the prototypical and actual residential load time series from the same geographic region. The upper plot in figure 5.6 is the simulated data used to create bus-level load time series in Austin, TX. The lower plot is the one residential electricity consumption measurement from the same city [73]. It is observed that while the two time series are on the same load level, and share similar daily trend, the actual load time series exhibits more jitters than the simulated data.

To introduce the stochastic behaviour back into the simulated data, and avoid bus load time

series being overly conforming due to the use of similar prototypical building or facility time series, a small noise is also imposed. Since the prototypical building and facility load time series already included the seasonal, weekly and daily variations, a Gaussian noise is added to the base loading to not introduce seasonality [75]. This transformed time series is then added to the corresponding type of load component, and the iteration would stop once after the load component maximum value calculated in previous step has been reached.

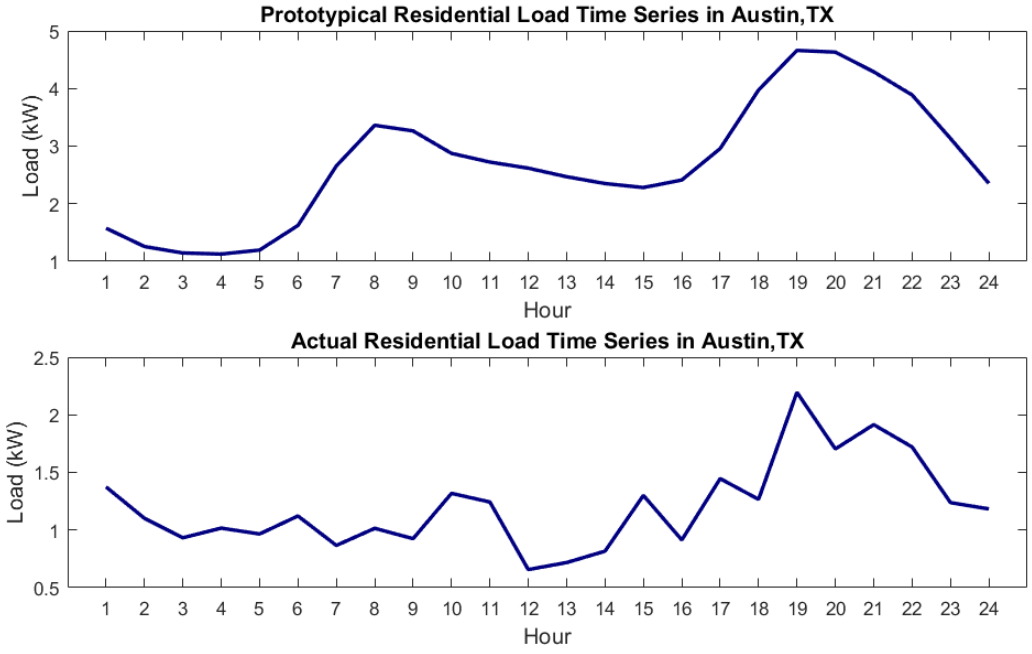

Figure 5.6: Simulated and Actual Residential Load Time Series

To mimic the effect of increasing load factor as load aggregates to a higher level [71], public feeder load time series managed by National Renewable Energy Laboratory (NREL) [76] is used to adjust the bus-level load factors to a realistic range. This distribution feeder load time series is populated from the taxonomy distribution feeders of different geographic regions.

The geographic coordinates of each load bus in the synthetic system are used to randomly select a subset of taxonomy feeders from the same geographic region, so that the summation of feeder

load time series is on the same scale as the bus load. The load factor of the aggregated feeder load time series is calculated to be the reference value. A constant component is added to the created bus-level load time series to adjust its load factor to a realistic value according to equation (5.1).

$$\frac{\boldsymbol{Constant + Average\ Load}}{\boldsymbol{Constant + Max\ Load}} = \boldsymbol{Reference\ Load\ Factor} \tag{5.1}$$

5.5 ACTIVSg200 and ACTIVSg2000 Load Time Series

On the bus-level, each load time series in the ACTIVSg200 and ACTIVSg2000 synthetic system is unique based off the location and load composition ratio of the load bus. Average bus-level load time series of different dominant load types are shown in figure 5.7. Residential-dominated bus load time series exhibit noticeable seasonal differences, where the electricity consumption in summer and winter seasons tend to have higher average values as well as higher variations. Commercial-dominated bus load time series have distinct daily patterns, while the electricity consumption base line stays relatively constant throughout the year. Industrial-dominated bus loads usually have the lowest variation and highest load factor. The average size of industrially-dominated buses are larger than the other two types.

The bus-level ACTIVSg200 synthetic load time series are shown in figure 5.8. It shows that each load bus in the synthetic system are assigned with an unique load profile, with different load sizes, as well as time-varying patterns. There are in total 160 load buses in the ACTIVSg200 synthetic system. On average, the largest load bus has a little over 45 MW of mean load value, and the smallest load bus has 0.11 MW.

The system-level ACTIVSg2000 synthetic load time series and the actual system load from their footprint regions are shown in figure 5.9. Although duplicating system-level load time series is not the desired outcome, synthetic load time series on the system level should exhibit similar general load shapes and trends compared to the actual system. The synthetic load shares similar size with the load of the actual system in the corresponding service territory. ACTIVSg2000 synthetic system has 71.1 GW of peak load, and 48.7 GW of average load, and the actual load of

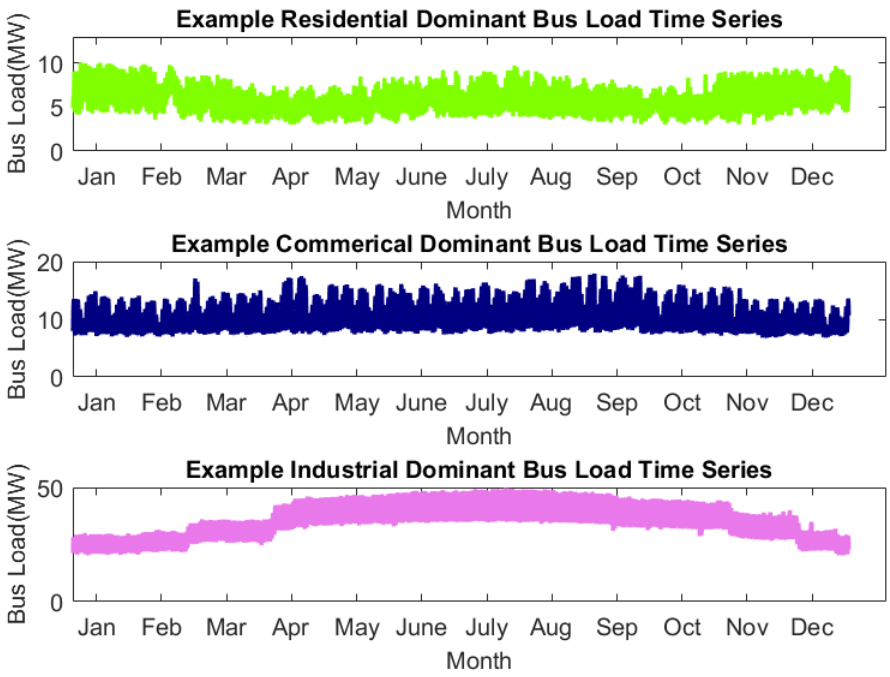

Figure 5.7: Bus-level synthetic load time series of different dominant load type

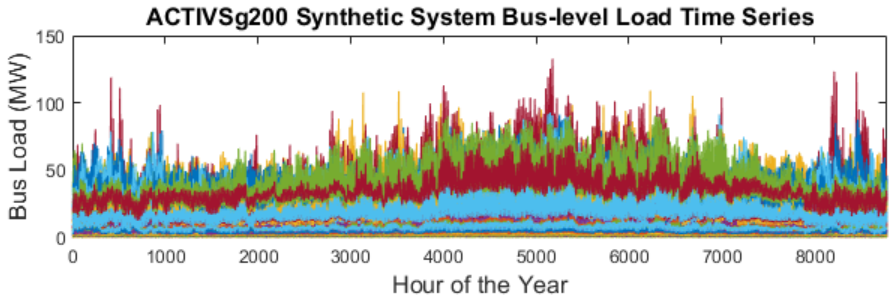

Figure 5.8: Bus-level load time series for the ACTIVSg200 synthetic system

ERCOT system has 71.2 GW of maximum load and 41.0 GW of average load. Daily and weekly patterns can be seen from the ACTIVSg2000 synthetic load time series. It is also observed that the synthetic system has similar seasonal trends compared to the actual system. ACTIVSg2000 and ERCOT systems both experience peak load in summer, and also have some high load weeks in winter.

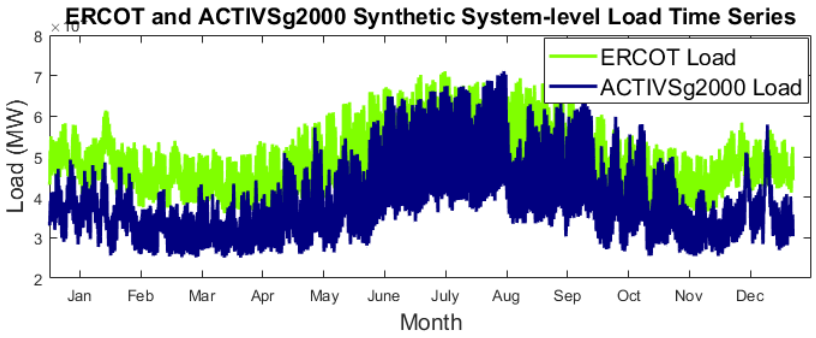

Figure 5.9: System-level load time series for the ERCOT and ACTIVSg2000 synthetic system

Since synthetic time series are fictitious, the validation of the created data against the actual data is critical to determine the quality and realism of the time series. A statistical based validation approach is implemented in this paper, as synthetic time series aim to realistically represent behaviors of load over time, instead of being an exact duplicate or forecast of the actual system time series. A comprehensive set of validation metrics enables researchers to use synthetic time series with ease, but at the same time to be aware of the underlying assumptions.

Validation metrics that are generic and independent from geographic locations are summarized using statistical characteristics found in public load data of 37 European countries [77] and 66 United States Balancing Authorities [78], so that synthetic load time series without a geographic footprint or have no availability of actual load time series can also be validated.

It is important to note that the validation of synthetic load time series is only conducted on the aggregated system level due to the lack of available real data on bus-level load time series. For an unbiased validation, aggregated reference data used during the construction process and the real data used for the validation process are independent and kept separate.

5.5.1 Load factors validation

Load factor is defined as the ratio of average and peak value of a load time series. It is one effective metric to quantitatively validate the overall shape of the synthetic load profile. For profiles with relatively constant load level, such as regions with a high industrial composition, load factors are usually higher; while heavily residential areas tend to have lower load factors due to light

occupation during the day [79].

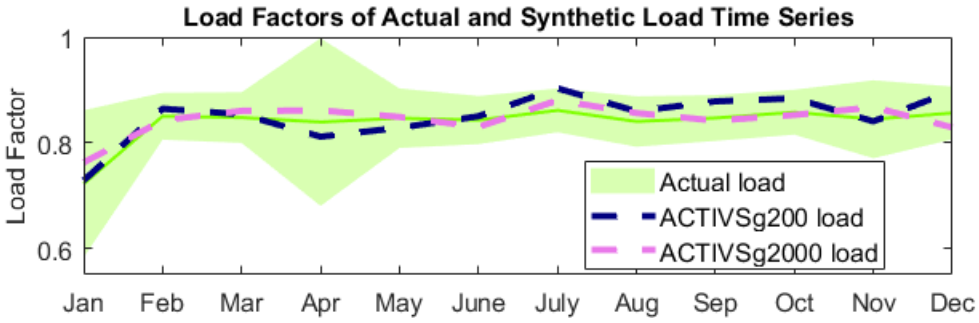

Figure 5.10: Monthly load factors of actual and synthetic load time series

The range of load factors of each month is summarized from public load data mentioned above and shown in figure 5.10 as the green shaded region. It is observed that as a general trend, the value of load factors are slightly higher in summer months, due to the increase of base electricity consumption from spacing cooling. There is also a consistent difference between the lowest and highest load factors of actual load every month, where systems with smaller size and less industrial load often have lower load factors. The load factors of ACTIVSg200 and ACTIVSg2000 load time series lie inside the range observed from actual load time series, and also follow the same monthly trend.

5.5.2 Load distribution curves validation

Load distribution curves show the percentage of time that load is at different levels relative to its mean value. The load time series is normalized based off its mean value, where load levels exceeding yearly average would have per unit values larger than one. The vertical axis of a load distribution curve is the percentage of time points.

The green shaded band in Figure 5.11 shows the range of load distribution curves found in real load time series, where load levels are scattered in between 0.5 and 2.0 per unit, with a denser distribution in the range from 0.8 to 1.2. The distribution of ACTIVSg200 and ACTIVSg2000

load time series follow the same general trend, load at most of the time points are within 0.8 to 1.2 times its yearly average.

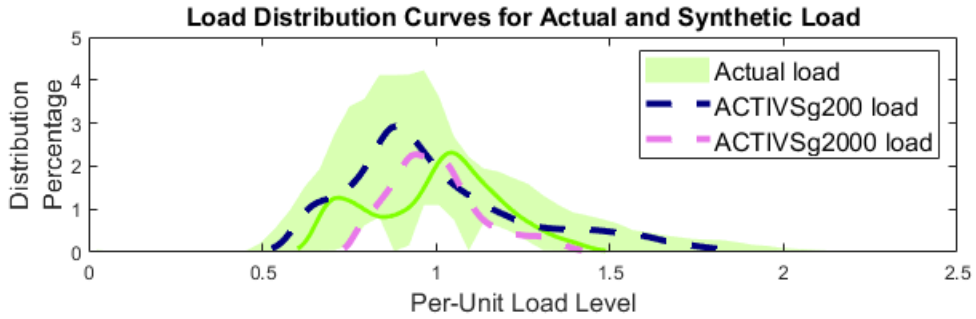

Figure 5.11: Load distribution curve validation

5.5.3 Autocorrelations validation

Autocorrelation exhibits the relationship between time points of load time series that are certain time lags apart. It provides a validation perspective in time sequence order, instead of observing time series values as if they are independent recordings. Since the load time series data is not stationary, where the average value is not constant,and the variance grows with the level of the time series, the log and differencing transformation are used to stabilize and remove the mean trend from the original time series.

Figure 5.12 shows the autocorrelations of actual and ACTIVSg synthetic load time series for time lags up to 48 steps. According to the real load time series data, the autocorrelation plot appears to be periodic with a 24-hour cycle, with its magnitude slightly decreasing every cycle. All the load time series autocorrelation exhibit a similar trend, within each cycle, the autocorrelation drops from 1 to below 0 and then increases from negative correlation back to almost unity correlation by the end of the cycle. It is also interesting to note that during the middle of each cycle, around 12 hour time lag, the autocorrelation of the differenced, logged load time series has a local maximum. The plots of synthetic load's autocorrelation lie within the upper and lower bound established by the autocorrelation of actual load time series.

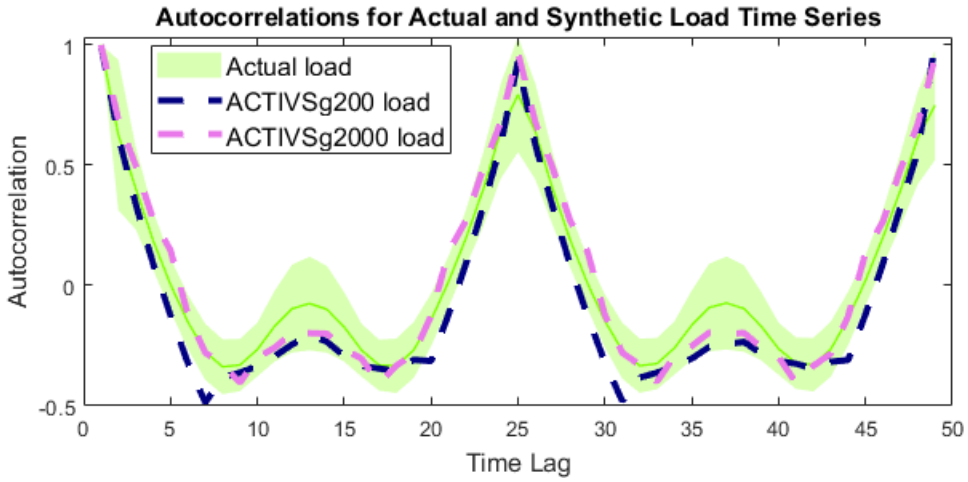

Figure 5.12: Load autocorrelation validation

5.5.4 Power spectral density validation

Power spectral density measures the the distribution of power content versus frequency of a time series. It is a technique that enables us to discover underlying periodic behaviors. The spectral density can be estimated using periodogram, which establishes the squared correlation between the targeted time series and sinusoidal waves at different frequencies spanned by the time series. Similar to the autocorrelation analysis, the log and differencing transformation are used to stabilize and remove the mean trend from the original time series.

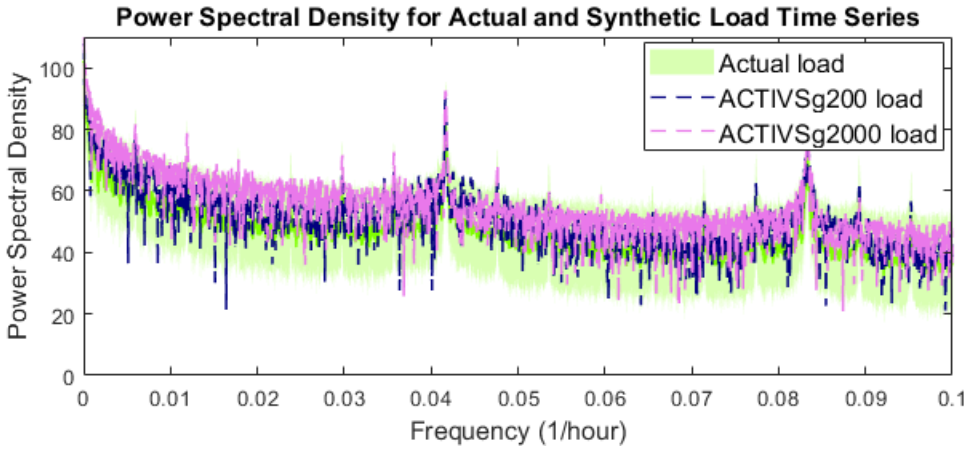

Figure 5.13: Load power spectral density validation

Figure 5.13 shows the power spectral density of the actual system-level and synthetic system-level load time series. The horizontal axis of this figure is the frequency, which presents periodic behaviors with longest period of every year, and shortest period of every 10 hours. The reference range of power spectral density of at each frequency is summarized from the public system load data, shown in the green shaded region in Figure 5.13. It is observed that the power spectral density of ACTIVSg200 and ACTIVSg2000 synthetic load generally lie in the defined upper and lower bounds. The power spectral density of the synthetic loads also exhibit distinctive sharp peaks around the same frequencies as those of the actual load time series, where the highest spike occurs at the daily 24-hour period with frequency equals to 1/(24 hour) = 0.0417/hour. Power spectral density spikes are also observed around periods such as half-day (12 hours), half-week (84 hours),one-week (168 hours), three-months (2160 - 2190 hours) and six-months (4320 - 4380 hours) .

6. OPERATION SCENARIO SIMULATION AND DATA PREPARATION

As mentioned in earlier chapters, the current RAS design process includes repeated off-line simulations on a set of predefined scenarios to identify the problem, and to ensure that the candidate corrective action is sufficient, and does not introduce unintentional risks to the system. The design decisions of which planning scenarios to use are heavily dependent on engineering judgements from industry experts. The planning scenarios used in the RAS creation process are usually hand-picked, and the total number of scenarios used to determine RAS parameters and corrective actions is often relatively small [64].

This chapter introduces the method used to generate large volume of realistic power system scenarios, and to prepare the data for machine learning applications, so that the system stress and violations pattern that is not obvious to human eyes can be identified, and the most up-to-date insights can be provided for more informed, effective, and reliable decision making in the power system control room. This method is demonstrated using synthetic transmission networks and the corresponding time series developed in chapter 5, and it is general enough to be easily applied to real power systems. To construct a specific scenario, the overall procedure used to determine the load and generation level, scheduled outages, area interchange schedules, as well as unit commitment and dispatch is firstly discussed. The techniques needed for the scenarios to converge to initial power flow solutions are then introduced. To create large amount of scenarios representing a wide spectrum of system operating conditions, a chronological power system simulation is performed, with permutations being implemented at each hour of the simulation year. Operation mode analysis is then conducted to identify if the operational scenarios created can cause severe violations where RAS is needed. These power system operation scenarios and analysis results are obtained as the input data to learn the Auto-RAS operational patterns through machine learning algorithms.

Part of this chapter is reprinted with permission from "Steady-State Scenario Development for Synthetic Transmission Systems" by Hanyue Li, Ju Hee Yeo, Jesscia L Wert, and Thomas J

Overbye, 2020. 2020 IEEE Texas Power and Energy Conference, Page 1-6, Copyright 2020 by IEEE.

6.1 Scenario Creation: Load and Generation Level

To create a wide spectrum of scenarios as input data of machine learning techniques, where the operating condition patterns can be learned, both business-as-usual and event conditions are considered. Under normal system operating conditions, electric loads have constant variations by nature as electricity consumers react to the change of time, weather, and other day-to-day events. In the case of a special or an extreme event, load in a power system can be found at an unusual level. For example, in a system with a large integration of behind-the-meter solar, the load level at specific locations may experience a sharp drop as a result of power line outages following extreme weather conditions. For the business-as-usual scenarios that reflect typical conditions, the synthetic bus load time series can be used directly to set the load level at each bus. To develop an event scenario, the original bus load time series can be used as a benchmark, where changes can be made depending on the details of situations.

The generation level of renewable energy resources can change constantly due to the varying nature of factors such as solar radiant, cloud coverage, wind speed and wind directions. The capacities of renewable energy units in the synthetic system are determined by the MW output of actual solar and wind units provided by the National Renewable Energy Laboratory (NREL) [80]-[81]. The geographic coordinates of wind and solar units in the synthetic system are mapped to the closest renewable sites in the actual power system. The 5-minute resolution data of real power output is then up-sampled to hourly resolution to be used in the synthetic scenarios. These values are considered as the maximum capacity of renewable generation units at given hours.

The generation levels of other fuel types are less variant, and usually only change on seasonal basis. For generators with coal, hydro, gas and nuclear fuel types, their seasonal capacity are synthesized from public utility data such as ERCOT's Seasonal Assessment of Resource Adequacy (SARA) and the Capacity, Demand and Reserves (CDR) Report [82, 83], and PJM's Capacity Market Manual [84]. The typical capacity factors of different fuel types in each season are shown

in 6.1. Noting that the actual value would vary depending on the specific cases and the geographic locations.

Table 6.1: Maximum Generation Capacity Factor by Fuel Type

Generator by Fuel Type	Spring	Summer	Fall	Winter
Coal	93.7	100	100	96.4
Hydro	83.5	83.2	70.8	82.3
Gas	87.5	100	92.7	80.6
Nuclear	100	100	100	100

6.2 Scenario Creation: Scheduled Outages

At any point in the year, system elements (i.e. generators, transmission lines) may be scheduled for maintenance outages. These are necessary to allow regular maintenance on system components to be performed in a safe manner (i.e. while the components are de-energized). In practice, these scheduling requests are made by utilities and, following system studies, approved by the regional balancing authority [85], [86]. The coordination of these outages is governed by NERC TOP-003-1 [86], a standard designed to ensure that system reliability is maintained even with scheduled generator and transmission outages.

The scheduled generation outages applied in these scenarios are based on the maintenance outages reported from ERCOT and PJM [82, 84], which provides a total of the generation outages for the season of study. The generation maintenance outages in these scenarios are based on the 2019 reports and are summarized in Table 6.2 [87, 88, 89, 90, 91, 92]. The values provided, however, are representative of the scheduled maintenance outages over the course of an entire season. In order to quantify outages in a scenario's snapshot depiction, a simplifying assumption that the average duration of a scheduled maintenance outage is 4 days, meaning that the seasonal outage capacity can be reframed as a seasonal outage energy. Capacity outage in the scenario can be represented per the following equations, where $\boldsymbol{P_{out,low}}$ and $\boldsymbol{P_{out,high}}$ are the scheduled maintenance outage in the scenario, $\boldsymbol{t_{avg}}$ is the average outage duration in hours, $\boldsymbol{P_{m,seas}}$ is the

Table 6.2: Typical Generator Maintenance Outages

Season	Maintenance Outages Capacity Percentage
Winter	2% - 3.8%
Spring	4.2% - 5.7%
Summer	0.1% - 3.6%
Fall	5.6% - 9.5%

Table 6.3: Maintenance Frequency of Generators by Fuel Type

Generator by Fuel Type	Period between Maintenance (months)
Coal	12
Hydro	7-12
Natural Gas	7-24
Nuclear	18-24
Solar	6-12
Wind	13

anticipated seasonal maintenance outage, $\boldsymbol{t_{seas}}$ is the number of hours in the season, and $\boldsymbol{P_{L,high}}$ is the seasonal peak load and $\boldsymbol{P_{L,low}}$ is the seasonal minimum load.

$$\boldsymbol{P_{out,low}} = \frac{\boldsymbol{t_{avg}} * \boldsymbol{P_{m,seas}}}{\boldsymbol{t_{seas}}} \frac{\boldsymbol{P_{L,high}}}{\boldsymbol{P_{L,low}}} \tag{6.1}$$

$$\boldsymbol{P_{out,high}} = \frac{\boldsymbol{t_{avg}} * \boldsymbol{P_{m,seas}}}{\boldsymbol{t_{seas}}} \frac{\boldsymbol{P_{L,low}}}{\boldsymbol{P_{L,high}}} \tag{6.2}$$

Specific generators are scheduled for maintenance outages during each season's scenarios according to the relative frequency of outages by fuel type informed by the data in Table 6.3 and industry practice, i.e. nuclear outages are also informed by data published by the U.S. Energy Information Administration [93].

Transmission outages are scheduled to allow for maintenance or construction that require specific branches to be de-energized. The transmission outages are applied to the scenarios such that connectivity of the system is maintained considering N-1 security. In order to do this, the system is considered as a graph. Candidate transmission lines for outages are screened using the bus

admittance matrix as a proxy for a connectivity matrix. If removing the candidate line in combination with previous accepted scheduled transmission outages creates a bridge in the system, the candidate is rejected and thus must remain in the case to ensure connectivity under N-1 conditions. Outages of lower capacity lines are prioritized by having a relatively high probability of selection when randomly identifying candidate lines. This process is repeated until the desired number of outages was reached for each scenario.

6.3 Scenario Creation: Area Interchange Schedules

In an interconnected electric power system, there usually exist multiple balancing authorities (BA) which reside on different geographic footprints, and have control rights over different set of resources and interchange meters. Traditional BAs have dispatchable generation, load, and interchange. The main goal of BA operation is to manage the balance of load, generation and net interchange schedule at all times to maintain a stable voltage and frequency profile within it's responsible boundary. For the creation of power system operation scenarios, with the same loading level, changes in area interchange schedules can greatly impact the dispatch of generations due to this net load and generation balance criteria, and resulting in different operating conditions.

6.4 Scenario Creation: Unit Commitment and Dispatch

The commitment and dispatch of generators schedule the specific amount of power that each unit should generate in scenarios. Unit commitment determines the on/off states of units that are not in scheduled outages from the previous section. This work formulates unit commitment as a priority list optimization problem [94]. At each hour, with the cost function of generators sorted, the subset of generators that can supply the system load with the lowest operation cost while satisfying the generator min on/off and ramping constraints is set to be online [94]. Economic dispatch calculates the specific generation amount of each unit knowing the on/off states from the unit commitment [95].

6.5 Scenario Creation: Power Flow Solutions

Convergence to a power flow solution is an indicator that the developed scenario has a feasible solution. Having an initial power flow solution is a good starting point to initialize other studies such as contingency analysis, optimal power flow, unit commitment, and transient stability simulations, et cetera. If the created scenario represents an event that spans over time, the power flow solution of the first time point is also used as the initial guess for power flow solution of the following time step.

Since the system operating condition of a customized scenario might be very different compared to the base case, its power flow solution can be far from the base case's solution. Sometimes special techniques are needed to aid the power flow algorithm to converge to an acceptable solution, especially for large scale power system cases.

6.5.1 Incremental Steps

The selection of initial guess is a critical key for the power flow convergence [96]. In the case of a customized synthetic scenario, the initial guess usually includes the voltage magnitude and angle that are obtained from the base case AC power flow solution. However, the power flow may take more iterations to converge or sometimes divergence can happen due to a big discrepancy of the operating conditions between the base case and scenario.

To address this issue, incremental steps can be established to gradually move the solution from the base case to the designed scenario case, where the solution from the previous step is used as the initial guess of the current step. For example, the spring minimum load scenario for the ACTIVSg2000 synthetic system experienced a non-convergence issue when using the base case solution directly as the initial guess. When 100 intermediate steps are added to gradually bring down the system load of 67 GW from the base case to 23 GW from the scenario case, Newton-Raphson power flow algorithm is successfully able to converge to a solution.

6.5.2 Alternative Voltage Solutions

As power flow problems are inherently non-linear, multiple solutions usually exist. While power systems are normally operated at the solution with highest voltage, sometimes the power flow algorithm might converge to an alternative voltage solution, resulting in being an inaccurate reflection of the real system values [97].

For each bus in the system, the self sensitivity of the voltage magnitude and reactive power injection can be used to confirm if an alternative voltage solution has been reached [98]. If $\boldsymbol{dV/dQ}$ is negative at one bus, it indicates the occurrence of an alternative voltage solution, where locally increasing reactive power would not provide voltage support. Starting with a higher-valued initial guess on those buses, and temporarily disabling controls are common techniques to help the returning of a high voltage solution [99]. If alternative bus voltage occurs but the sensitivity of voltage with respect to reactive power injection stays positive, the result is still at a high voltage solution, where reactive power devices can be implemented to help with the voltage profile.

6.5.3 Reactive Power Devices

Reactive power devices such as shunt capacitors, reactors, and load tap-changing (LTC) transformers are commonly used for bus voltage regulation that improves the voltage profiles of a solution [100]. The reactive power planning for the base case of synthetic transmission system recognizes that the placement and settings of reactive power devices are designed to optimize the operating condition of the base case, and leaves a margin to allow extra shunt devices to be added as needed for special-case situations and future development [43]. To accommodate the varieties of scenario development for the ACTIVSg2000 synthetic system, 24 additional shunts are added to the system to address the voltage violations.

6.6 Operation Scenario Simulation

The framework of chronological power system operation data simulation aims to create large amount of operating scenarios representing a wide spectrum of system operating conditions, and generates input data for the creation and testing of Auto-RAS. There are three inputs to this frame-

work: hourly bus-level load time series,hourly renewable generation unit MW capacity time series, and topologies of the electric grid of interest. To simulate the power system operation scenarios, the hourly time series are applied to the system sequentially for an initial power flow solution. Here the solution of the previous hour is used as the initial guess for the current hour for the Newton Raphson algorithm.

For every hour of this chronological simulation, the schedule generation and transmission outage amount is estimated based on the season of the particular time point. Additionally, multiple interchange schedule combinations are considered. The targeted MW flow between two areas are set to 11 different levels, with values ranging from 50% to 150% of MW flow amount in the initial power flow solution. If the system failed to converge at a certain interchange schedule combination, it is considered as an infeasible operating condition and discarded. With the load, online generation, online transmission, and interchange schedule being determined, This work formulates unit commitment as a priority list optimization problem. At each hour, with the cost function of generators sorted, the subset of generators that can supply the system load with the lowest operation cost while satisfying the generator min on/off and ramping constraints is set to be online. With the unit commitment solution, an economic dispatch is performed to calculate the specific generation amount of each unit knowing the on/off states of each generation unit.

6.7 ACTIVSg200 and ACTIVSg2000 Operation Scenarios Overview

This section provides an overview of the ACTIVSg200 and ACTIVSg2000 scenarios created as inputs to the Auto-RAS work. To visualize the information contained in the large volume of scenarios, techniques such as contouring and pseudo-geographic mosaic displays (PGMD) are used [101, 102]. PGMD is a display technique that fully utilizes the monitor screen space to represent large amount of detailed information. In the PGMD of power system operation scenarios, each mosaic tile represents a system element such as a bus, or a transmission branch. The size, and color of each tile is determined by the field associated with one bus or one transmission line object. The location of the mosaic tile is reflecting the object's relative latitude and longitude with only an approximate geographic precision [102]. For example, mosaic rectangles being placed at the top

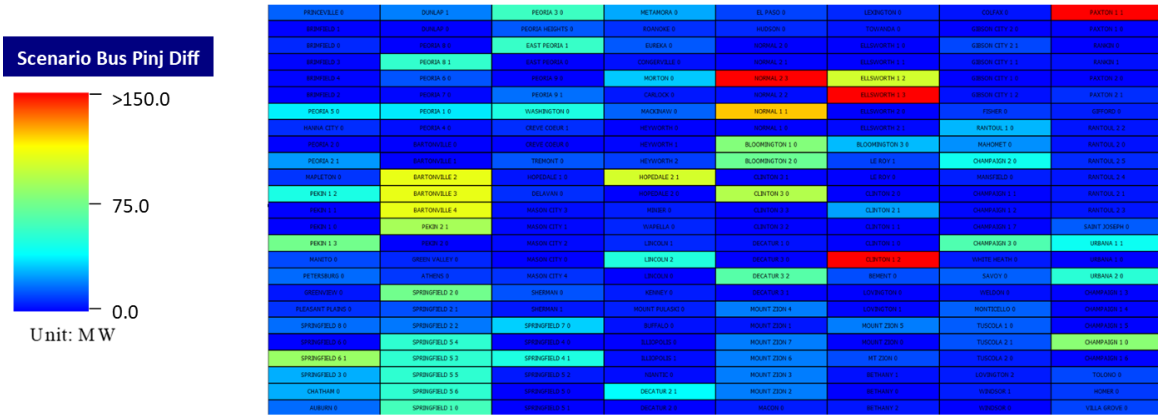

Figure 6.1: Differences in bus real-power injection for ACTIVSg200 operation scenarios

of the display usually represents objects located in the north.

A total of 8784 scenarios are initially created for the ACTIVSg 200 synthetic system. This presents a snap shot of the system operating condition for every single hour of a leap year. Since this system is relatively small in size and only have one control area, there are no permutations on area interchange schedules for each operating hour. Figure 6.1, 6.2, and 6.3 are some visualizations showcasing the differences among these hourly operating scenarios. Figure 6.1 is the PGMD showing the difference between the maximum and minimum real power injection at each bus. With the time varying load time series, and the implementation of generation dispatch and unit commitment, the power injection at each transmission node vary. As a result, different system operating conditions occur. As the slack bus, CLINTON 1 2 has the largest power injection variations among operational scenarios, because the generators attached to this bus have to constantly adjusting their MW output to balance the system load and generation. This change can also be observed in Figure 6.2 , where the transmission lines closer to the slack bus have higher MVA flow variation values. Figure 6.2 shows the contour of bus voltage per unit magnitude changes among all 8784 operational scenarios. Usually the buses with higher voltage magnitude variations have electric load attached, and are relatively further away from generator voltage regulators and reactive power control devices such as switched shunts.

For the ACTIVSg2000 synthetic system, over 35,000 operational scenarios are simulated. With

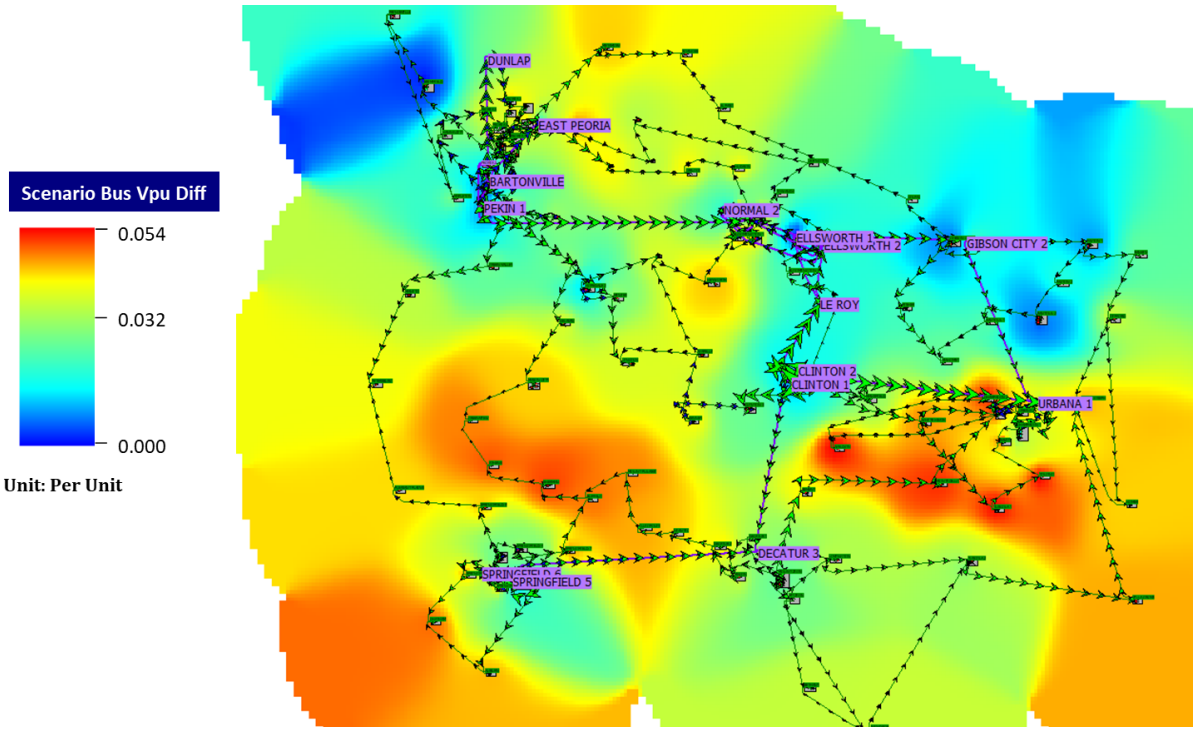

Figure 6.2: Differences in bus voltage magnitude for ACTIVSg200 operation scenarios

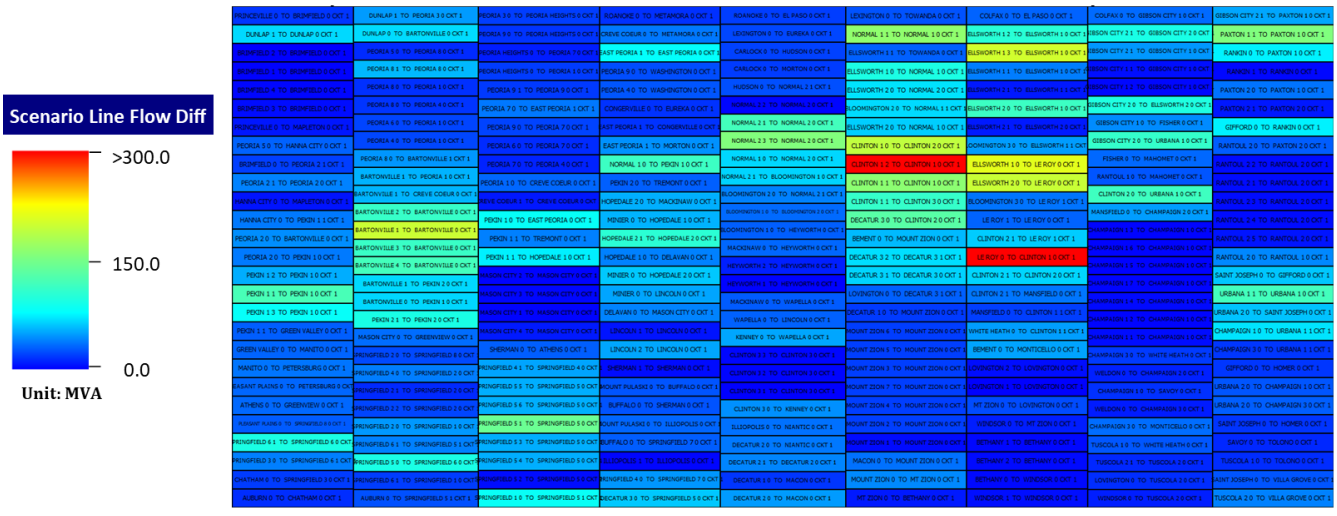

Figure 6.3: Differences in transmission line MVA flow for ACTIVSg200 operation scenarios

eight control areas being defined in this synthetic grid, on average every hour has four area interchange schedule combinations. The interchange schedule combination considers four major interfaces in the synthetic 2000-bus system: west to north central, coast to south central, north to north central, and east to north central. Since north central and south central are the two most populated areas on the footprint of Texas, those four interfaces provide main paths for remote generations to supply the load. Eight scenarios of the ACTIVSg2000 system, representing the maximum and minimum load operating conditions for each season are shown here as an overview. The descriptions of those ACTIVSg2000 representative scenarios can be found in table 6.4.

The voltage contours of the eight representative scenarios are depicted in Figure 6.4. It is interesting to note that the voltage profile of this synthetic system varies with the pattern of load curves. Since Texas has both summer and winter electric load peak due to cooling and heating, the overall voltage magnitudes are lower in those two seasons. Additionally, since the usage of electric heater in winter is more common during the night, the load values at more populated areas are higher during the low load hours. As a result, a lower voltage pocket can also be observed in the winter low load scenario.

Figure 6.5 provides a snapshot representation of the line loading in the eight representative ACTIVSg2000 scenarios. The PGMD snapshots contain 3206 individual tiles, each representing one transmission line object in the system. The color of each tile represents the line loading in the scenario, ranging from blue representing 0% line loading to red representing 100% line loading relative to the line limits. It is observed that the system loading is overall higher in the high load scenarios, compared to the low load scenarios throughout the four seasons. The transmission lines being loaded more than 75% are fairly evenly distributed everywhere in the system for the four high load scenarios. This changes for the low load scenarios, where the branches being loaded more than 75% are concentrated in the western side of the grid during spring and winter seasons.

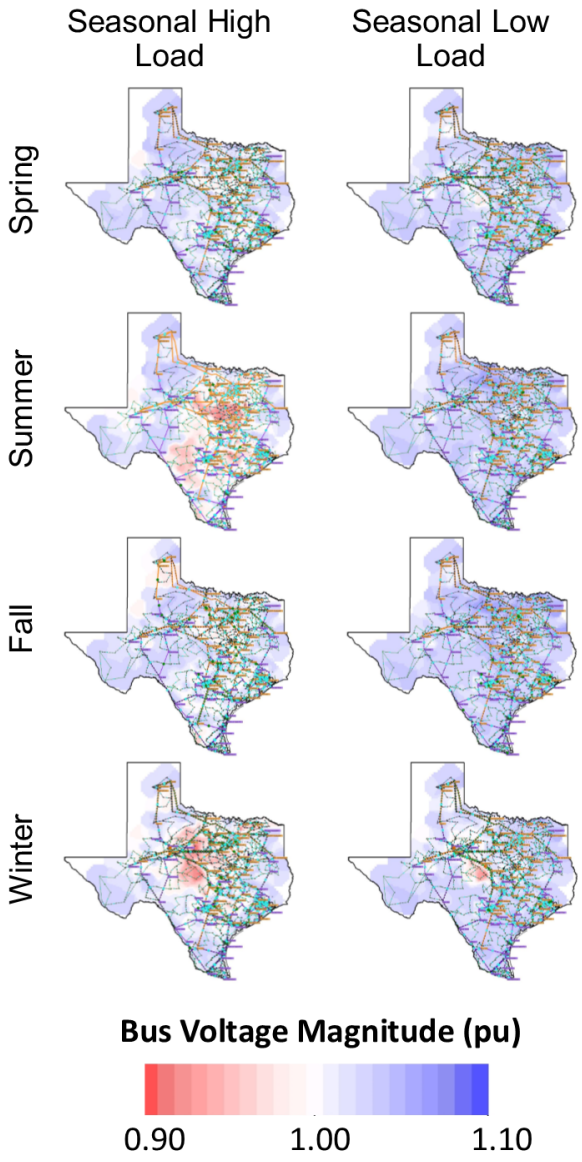

Figure 6.4: Voltage contours of ACTIVSg2000 operation scenarios

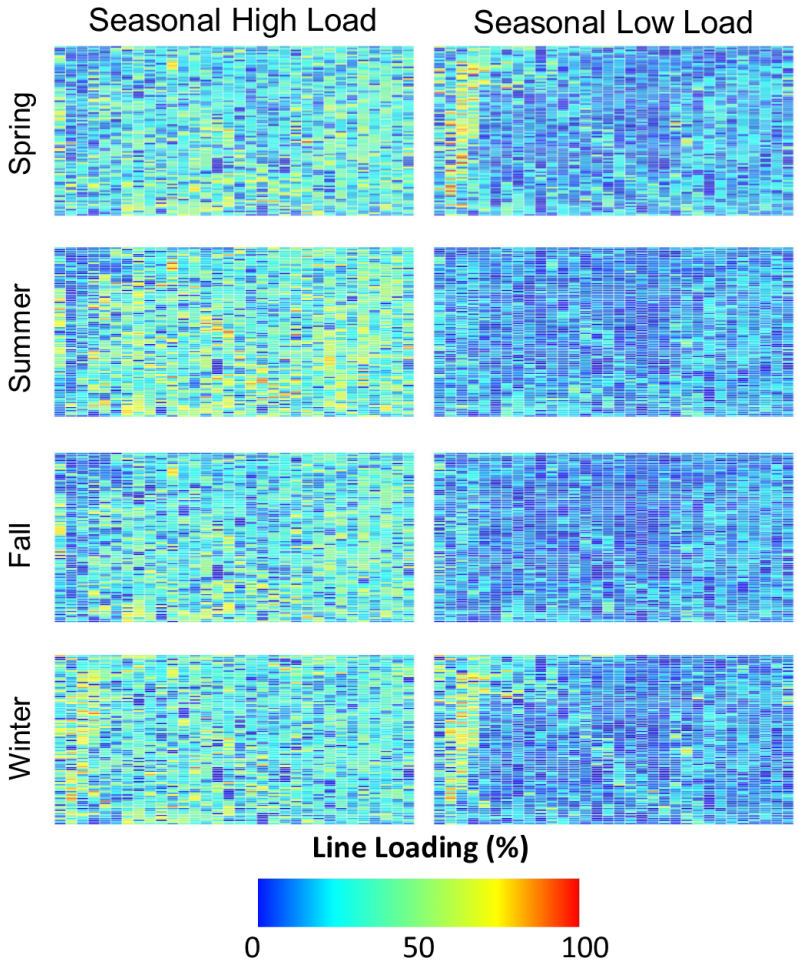

Figure 6.5: Transmission line loadings of ACTIVSg2000 operation scenarios

Table 6.4: ACTIVSg2000 Operation Scenarios Description

Case Name	**Case Description**	**Load (MW)**	**Wind Generation (MW)**	**Solar Generation (MW)**	**Generation Outage Capacity (MW)**	**Transmission Outage Capacity (MVA)**
SPR1	Maximum expected load in Spring	53329.7	5443.4	175.4	114.0	31590.4
SPR2	Minimum expected load in Spring	23072.3	7788.7	0.0	610.7	37998.0
SUM1	Maximum expected load in Summer	66275.7	1482.6	198.3	6.4	15942.6
SUM2	Minimum expected load in Summer	25350.1	3849.1	0.0	44.2	19324.1
FAL1	Maximum expected load in Fall	55848.5	2376.3	110.7	188.0	51627.5
FAL2	Minimum expected load in Fall	23379.4	8625.7	0.0	1215.0	45258.9
WIN1	Maximum expected load in Winter	53964.3	11514.2	3.9	75.0	23518.0
WIN2	Minimum expected load in Winter	23295.2	10152.9	0.0	407.7	21387.0

7. DATA-DRIVEN AUTO-RAS CONDITION ANALYSIS

One of the most important features to distinguish remedial action schemes from other operation control techniques is the RAS condition. In urgent situations where the operating conditions are deteriorating at fast rates, or the operational mitigation plan is complicated or not obvious, system operators potentially do not have enough time to react to system events, and conduct control actions in a timely manner. With arming and triggering conditions being defined for RAS ahead of time, and the measurement data from the energy management system in the control room, the condition logic are constantly being monitored. If certain anomalous system condition is detected, the corresponding corrective action can be initiated instantly to protect the system from severe damage or collapse.

Traditionally, the design of RAS conditions is based on repeated off-line studies on a limited number of predefined scenarios. The selection and conditioning of the scenarios, and the determination of RAS condition logic are heavily dependent on engineering judgements and familiarity with the specific power system. The operation scenario simulation method introduced in the previous chapter offers a significant advantage where large volumes of scenarios can be created in an automated way, representing wide-spectrum of system operating conditions. This opens doors to data-driven applications, where machine learning techniques can be applied to obtain more complete understanding of system behaviors and system elements' response to events in a systematic way.

This chapter explains the determination of Auto-RAS condition logic in details. This process includes operational scenario contingency analysis, connectivity-based clustering, and support vector machine learning. The operation scenario analysis step provides a thorough understanding on how the system respond to contingency events under different operating conditions. This helps to identify system elements subject to severe violations, where remedial action schemes are needed. Those severe violation elements are then grouped together based on network connectivities and their violation-causing contingencies. Elements belonging to the same cluster are protected and

addressed by the same remedial action scheme. For each identified RAS cluster, scenarios are labeled as a binary value, where a True label means this scenario can potentially lead to severe violation of specific system elements from this RAS cluster, and a False tag means this scenario is safe. The operational values associated with those scenarios are considered as scenario features. The scenario features as well as scenario labels are used as input data to a support vector machine algorithm. This two-stage SVM algorithm reduces the dimension of the scenario features, and learns a hyperplane that can optimally divide the scenarios of different labels. The selected scenario features, along with the learned hyperplane, and list of violation-causing contingencies are then leveraged as RAS condition logic. The overall process of the Auto-RAS condition determination is demonstrated in Figure 7.1, and explained in details in the rest of this chapter.

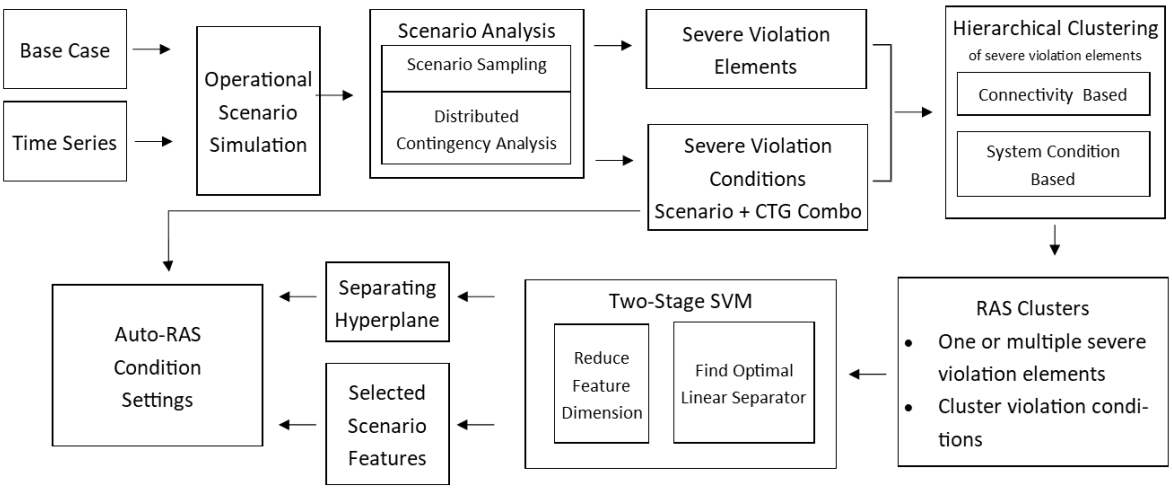

Figure 7.1: Auto-RAS condition setting diagram

7.1 Operation Scenario Analysis

Since it's impossible to run analysis on every single operation scenarios created, a subset of the scenarios are sampled, and analyzed in details. The result of the sampled scenario analysis is utilized to understand if certain operating conditions are more subject to severe violations, and

identify the specific severe violation system elements. Since the simulation of operational scenarios is chronological, the sampling process is based on the time tag associated each scenario to make sure that number of data points for each season, and time of the day is on the same order of magnitude. If a specific hour is selected during the scenario sampling process, all the area interchange variations for this hour are considered. This work is focused on severe branch violations. For the synthetic networks, the emergency MVA ratings of transmission branches are firstly defined. A contingency list is then constructed, followed by contingency analysis on all the sampled scenarios, with the help of distributed computing.

The violation of a transmission line normal thermal limit is usually not an urgent problem. In control rooms, established procedures exist to inform the operators on the actions needed to bring the transmission line flows down to acceptable thresholds [17]. Generator dispatch programs in real-time operation, such as security-constrained economic dispatch and security-constrained optimal power flow, aim to alleviate line violations by changing generator outputs as well [2]. However, in case of a severe event, transmission line loading may escalate quickly to violate reliability standards by exceeding its short-term emergency rating [103], or causing cascading tripping of lines [104, 105]. Under such extreme scenarios where operators and embedded control programs do not have enough time to react, automatic corrective actions are needed to respond to the event immediately to prevent unacceptable system operating conditions.

Based on typical current ratings of conductors at each nominal voltage level, normal MVA limits of transmission lines are calculated and included during the creation process of the 200-bus and 2000-bus synthetic systems. In this work, for different voltage levels, the ratio distribution of emergency MVA rating to normal MVA rating is summarized based on real transmission networks, shown in Figure 7.2, 7.3, and 7.4. It is observed that most of the emergency ratings of the transmission branches are no more than 1.25 times more than their normal MVA ratings. The range of ratio between branch emergency and normal MVA rating is higher for 115-kV and 230-kV networks, where the maximum ratio can reach up to 3.5 on rare occasions. For the 500-kV network, half of the branches have their emergency ratings less than 1.1 times of their normal ratings, and almost

30% of them have a ratio between 1.3 and 1.4. The maximum ratio observed for a 500-kV network is 2.4. This ratio distribution is then used to the estimate the emergency rating of each transmission line in the 200-bus and 2000-bus synthetic networks.

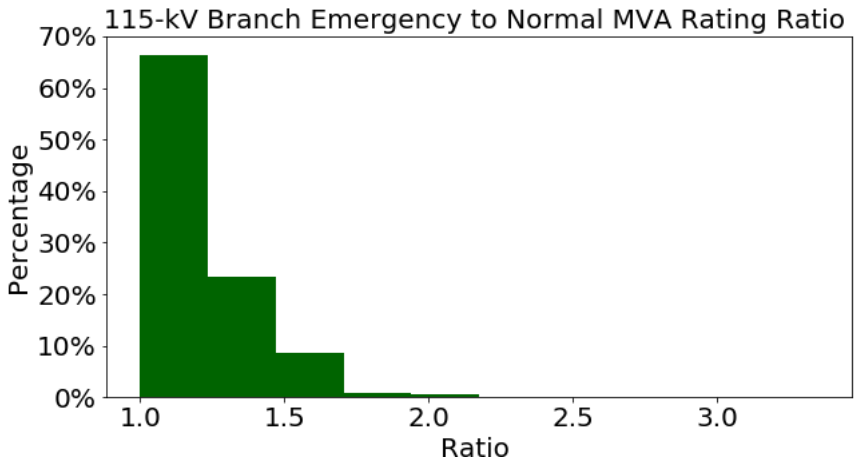

Figure 7.2: Branch emergency to normal MVA ratio statistics for real 115-kV networks

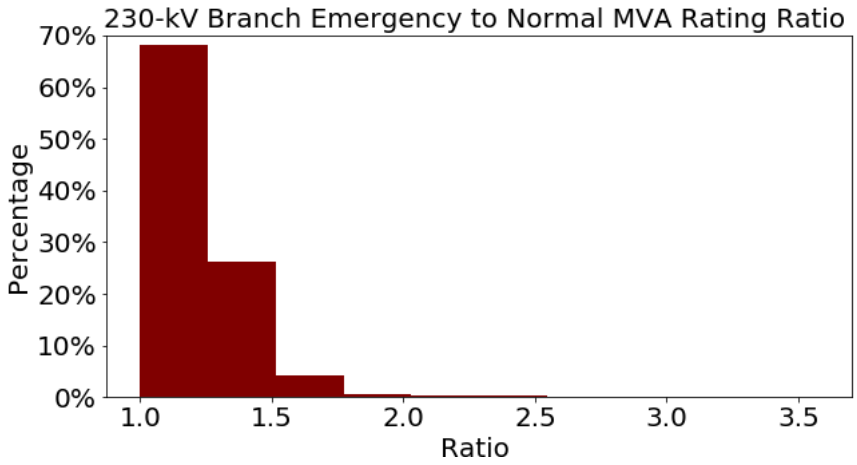

Figure 7.3: Branch emergency to normal MVA ratio statistics for real 230-kV networks

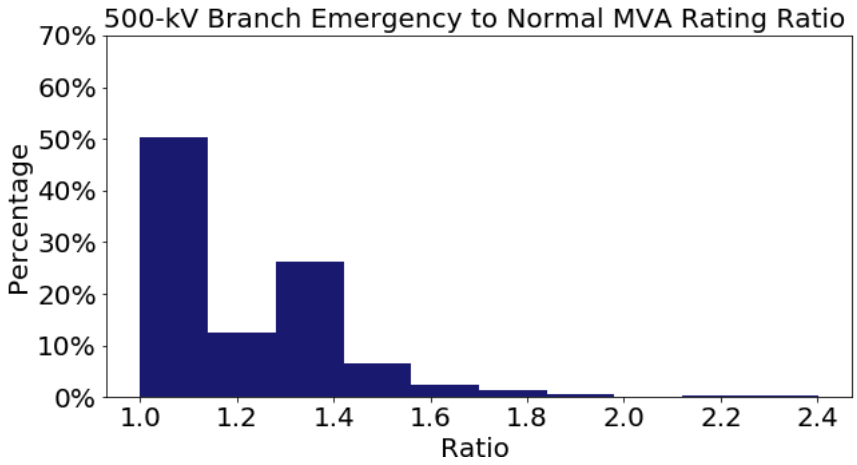

Figure 7.4: Branch emergency to normal MVA ratio statistics for real 500-kV networks

To identify severe overloading conditions where RAS are needed in the two synthetic systems, contingency lists modeling the loss of bulk electric system elements are created. Those lists first include two N-2 contingency categories defined in North American Electric Reliability Corporation transmission planning performance requirements [28], where the multiple contingencies from a common structure (P7), and two overlapping single contingencies (P6) are considered. Additional events up to N-4 contingencies are also included, where the contingency elements are within the same substation, or the first neighbor substation where the source of the disturbance is located at.

As there are over 3,000 branches in the 2000-bus synthetic system, it is unrealistic to use a complete contingency list that includes every possible combination. To effectively identify the set of contingencies with severe consequences, the list is reduced to only consider double line outage of branches 161 kV and above for P7 events. For P6 events, an LODF screening tool is used to identify pairs of contingencies that are significant without solving all the contingencies [106]. Only those contingency pairs that create line limit monitoring violations are included in the contingency list.

These contingencies are then modeled in both steady state contingency analysis and transient stability simulation. The steady state contingency analysis will evaluate the system voltages and

branch flows when the post-contingency power flow analysis converges. If the power flow fails to converge, transient stability study is conducted with simulation duration set to be 20 seconds, with a half-cycle time step. For both ACTIVSg200 and ACTIVSg2000 synthetic systems, the MVA flows of transmission lines 115 kV and above are monitored for emergency rating violations. Leveraging distributed computing techniques, contingency analysis is conducted for every sampled scenarios using the constructed contingency list. The list of severe violation element, and the combination of scenario name and contingency label that causes these violations are recorded.

For the ACTIVSg200 synthetic system, 1744 unique operational scenarios are sampled. In total, 596 contingencies are considered in the scenario contingency analysis. From the scenario analysis, two severe violation elements, the transmission line between bus number 29 and 30, and the transformer between bus 128 and 129 are identified. The approximate location of those two elements, and close-up looks of their one-line diagrams are shown in Figure 7.5. The severe violation elements are marked in red in the close-up diagrams. It is also worth noting that the locations of the buses and substations in the close-up diagrams are adjusted to show the connectivity of the network, and does not reflect the geographic coordinates. The detailed information of those two severe overloading elements can be found in Table 7.1. Under the majority of operating conditions, there are no severe violations according to the contingency analysis.

Table 7.1: ACTIVSg200 Severe Violation Elements Information

Element Name	Element Type	Voltage Level (kV)	Emergency Rating (MVA)	# of CTGs Causing Violation	# of Scenarios Causing Violation
Branch '29' '30 '1'	Line	115	100	596	2
Branch '129' '128' '1'	Transformer	230/115	250	2	2

For the ACTIVSg2000 synthetic system, 5432 scenarios are sampled for studies from the 35,000 simulated operational scenarios. In total, 4215 contingencies are considered in the scenario contingency analysis. From the scenario analysis, 10 severe violation elements are identified, shown in Figure 7.6 in red circles. The detailed information is shown in Table 7.2. It is observed that most of the severe violation elements are identified from the 161 kV network, besides one

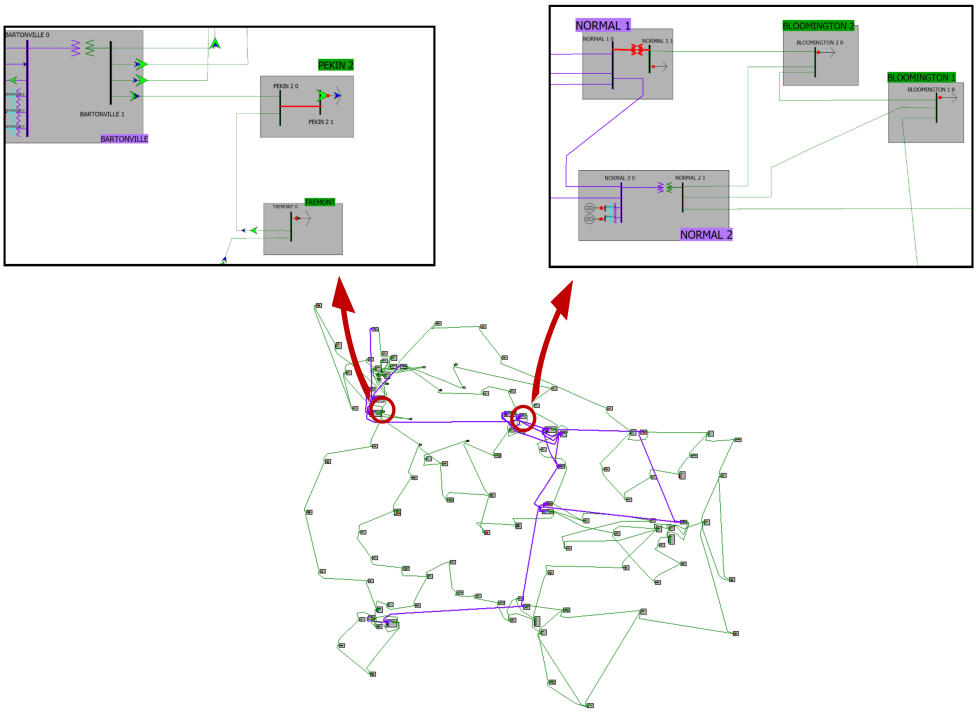

Figure 7.5: ACTIVSg200 synthetic grid severe violation elements

transmission line being identified from the 115 kV network, and one 230/115 kV transformer. Branch '2081' '2129' '1',Branch '2038' '2003' '1' and Branch '2081' '2038' '1' are located near the 500 kV network shown in orange just north of Dallas, Fort Worth area. Branch '5450' '5331' '1' belongs to a different area as the three above mentioned transmission lines, but it sits closely to them at in the southwestern direction. Branch '2093' '2012' '1', Branch '2018' '2093' '1', Branch '2058' '2026' '1', and Branch '2128' '2026' '1' are four 161 kV transmission lines in the Texas panhandle area. Branch '3102' '3091' '1' is a 115 kV transmission line just south of the Texas panhandle area. Branch '1033' '1032' '1' is a 230/115 kV transformer located in far west Texas.

7.2 Severe Violation Elements Hierarchical Clustering

With the identification of severe violation elements from the operational scenario analysis, those transmission branches are then grouped together as RAS clusters based on network connectivities and the list of contingencies that cause each targeted branch to overload. One RAS cluster

Table 7.2: ACTIVSg2000 Severe Violation Elements Information

Element Name	Element Type	Voltage Level (kV)	Emergency Rating (MVA)	# of CTGs Causing Violation	# of Scenarios Causing Violation
Branch ’2081’ ’2129 ’1’	Line	161	219.4	1	52
Branch ’2038’ ’2003’ ’1’	Line	161	251.0	1	6
Branch 2081’ ’2038’ ’1’	Line	161	251.0	1	6
Branch ’5450’ ’5331’ ’1’	Line	161	495.2	3	22
Branch ’2093’ ’2012’ ’1’	Line	161	241.8	2	1888
Branch ’2018’ ’2093’ ’1’	Line	161	239.2	2	1888
Branch ’2058’ ’2026’ ’1’	Line	161	199.9	4	1911
Branch ’2128’ ’2026’ ’1’	Line	161	187.5	4	1740
Branch ’3102’ ’3091’ ’1’	Line	115	200.0	2	1460
Branch ’1033’ ’1032’ ’1’	Transformer	230/115	170.0	3	811

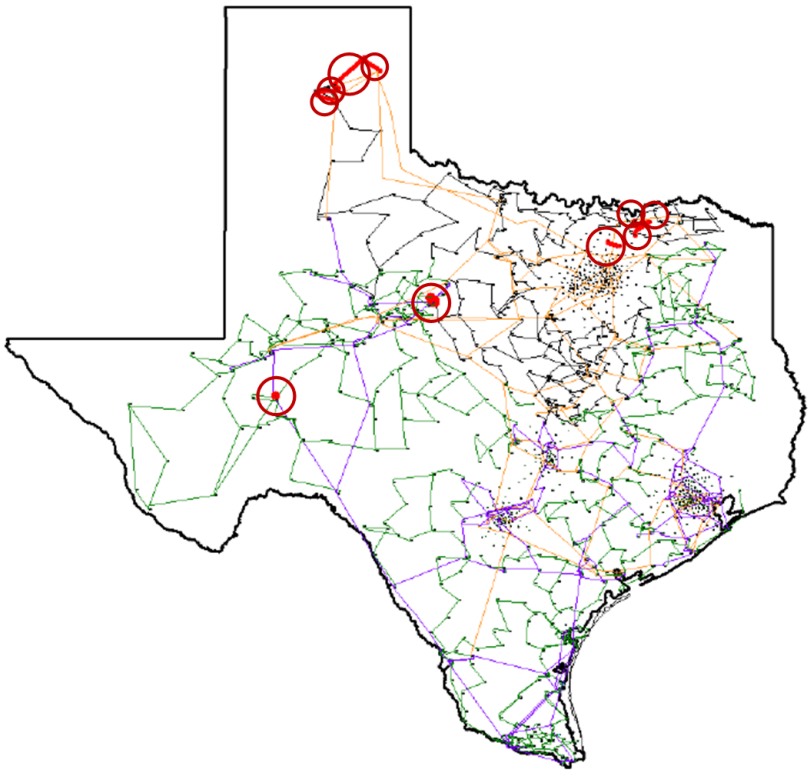

Figure 7.6: ACTIVSg2000 synthetic grid severe violation elements

can contain one or multiple severe overloading elements. Branches belong to the same group should react the system events in similar manners, and electrically close to each other. They are protected under the same remedial action scheme.

A connectivity-based hierarchical clustering is first implemented for the initial grouping of se-

vere violation elements. The distance between each pair of severe violation elements is considered as the number of vertices on the shortest path between the two nodes, calculated using the Dijkstra's algorithm. This sequential single-source method is given in Algorithm 1, where the sources are defined as the buses associated with each severe violation elements. It maintains $\boldsymbol{T}$ as the set of vertices for which the shortest paths have not been found, and $\boldsymbol{d_i}$ as the shortest known path from source node $\boldsymbol{v_s}$ to vertex $\boldsymbol{v_i}$. Initially, $\boldsymbol{T} = \boldsymbol{V}$, all the vertices in the graph, and all $\boldsymbol{d_i} = \boldsymbol{\infty}$. At each step of the iteration, the vertex $\boldsymbol{v_m} \in \boldsymbol{T}$ with the smallest distance value is removed from the set $\boldsymbol{T}$. Each neighbor of $\boldsymbol{v_m} \in \boldsymbol{T}$ is examined to see if a path through $\boldsymbol{v_m}$ is shorter than the currently known shortest path. In this case, the iteration is stopped when all the nodes associated with severe violation elements have been removed from set $\boldsymbol{T}$.

Algorithm 1 Dijkstra's shortest-path algorithm

1: $\boldsymbol{d_s} = \boldsymbol{0}$
2: $\boldsymbol{d_i} = \boldsymbol{\infty}$, for $\boldsymbol{i} \neq \boldsymbol{s}$
3: $\boldsymbol{T} = \boldsymbol{V}$
4: **for** $\boldsymbol{i} = \boldsymbol{0}$ to $\boldsymbol{N} - \boldsymbol{1}$ **do**
5: find $\boldsymbol{v_m} \in \boldsymbol{T}$ with minimum $\boldsymbol{d_m}$
6: **for** each edge $(\boldsymbol{v_m}, \boldsymbol{v_t})$ with $\boldsymbol{v_t} \in \boldsymbol{T}$ **do**
7: **if** $\boldsymbol{d_t} > \boldsymbol{d_m} + \boldsymbol{distance}((\boldsymbol{v_m}, \boldsymbol{v_t}))$ **then**
8: $\boldsymbol{d_t} = \boldsymbol{d_m} + \boldsymbol{distance}((\boldsymbol{v_m}, \boldsymbol{v_t}))$
9: **end if**
10: **end for**
11: $\boldsymbol{T} = \boldsymbol{T} - \boldsymbol{v_m}$
12: **end for**

The shortest nodal distance identified between severe violation element pairs forms the proximity matrix as the input for the hierarchical clustering. This work utilizes the bottom-up hierarchical

clustering approach, given in Algorithm 2. Initially, every severe overloading branch is a single-point cluster. The two closest severe overloading branches are joined as one cluster. The next closest clusters are then grouped together and this process continues until there is only one cluster containing the entire data set. The proximity matrix is also updated every iteration. In this case, complete linkage is considered, so that the distance between two clusters is defined as the longest distance between two severe violation branches in each cluster.

Algorithm 2 Agglomerative hierarchical clustering algorithm

Given:

A set of X of objects $\{x_1, ...x_n\}$

Proximity matrix where $D(c_i, c_j)$ is the distance between cluster i and cluster j

for $i = 0$ to n **do**

    $c_i = \{x_i\}$

end for

$C = \{c_1, ..., c_n\}$

$l = n + 1$

while $C.size > 1$ **do**

    (c_{min1}, c_{min2}) = minimum $D(c_i, c_j)$ for all c_i, c_j in C

    remove c_{min1} and c_{min2} from C

    add $\{c_{min1}, c_{min2}\}$ to C

    $l = l + 1$

end while

In this work, the cutoff distance for the hierarchical clustering is set to be three nodes. For the ACTIVSg200 synthetic system, the two identified severe violation elements are more than 3 nodes away, so that they belong to two different hierarchical clusters. For the ACTIVSg2000 synthetic system, the dendrogram is shown in Figure 7.7. Dendrogram is a diagram showing the

hierarchical relationships between objects. Initially, all the objects are a cluster of their own, with the distance value set to to zero. Every iteration, the two clusters with the shortest distance are grouped together. In the dendrogram, those two clusters are joined. The height of the joined clusters indicates the proximity value. From figure 7.7 we observe that with the cut off distance being defined as three electric nodes, there exist 5 clusters. Branch ‘2093’ ‘2012’ ‘1’, Branch ‘2018’ ‘2093’ ‘1’, Branch ‘2058’ ‘2026’ ‘1’, and Branch ‘2128’ ‘2026’ ‘1’ are grouped together, shown in color green. Branch ‘2081’ ‘2129’ ‘1’,Branch ‘2038’ ‘2003’ ‘1’ and Branch ‘2081’ ‘2038’ ‘1’ are also grouped together, shown in color red. The rest of severe violation elements are not grouped with other branches, and forms an individual cluster containing only one object.

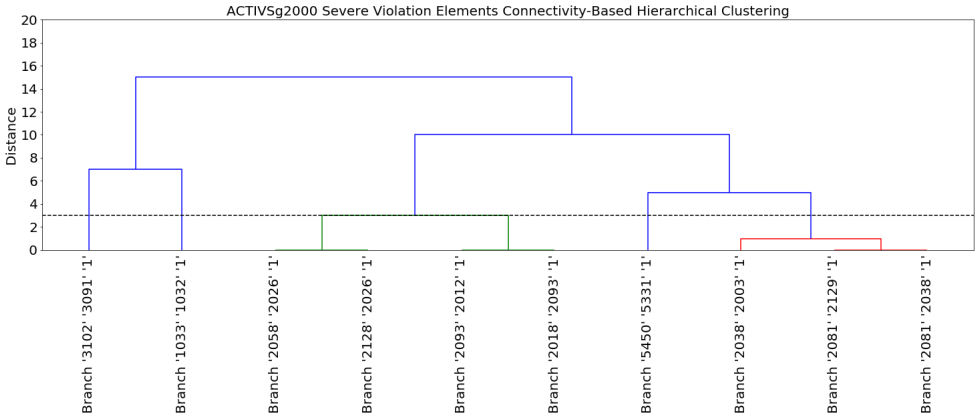

Figure 7.7: ACTIVSg2000 severe violation elements hierarchical clustering dendrogram

For the two clusters with multiple severe violation elements, the contingencies that can trigger the overloading violation of each transmission branch is further reviewed to determine if the objects belong to the same hierarchical cluster can be protected by the same remedial action scheme condition. Branch ‘2081’ ‘2129’ ‘1’,Branch ‘2038’ ‘2003’ ‘1’ and Branch ‘2081’ ‘2038’ ‘1’ show violation under the same contingency, L_002021PARIS20-002002SAVOY0C1L_002002SAVOY0-002101SHERMAN10C1, so that those three severe violation elements are grouped as one single RAS cluster. The four branches belonging to the same hierarchical cluster show violations under

two different sets of contingencies. Branch '2093' '2012' '1' and Branch '2018' '2093' '1' show severe violation when contingencies L_002127MIAMI0-002011PANHANDLE20C1L_002012PANHANDLE21-002052PANHANDLE60C1 and L_002127MIAMI0-002011PANHANDLE20C1L_002128MIAMI1-002026WHEELER0C1 are simulated. Branch '2058' '2026' '1' and Branch '2128' '2026' '1', however, show violations under four contingencies: L_002127MIAMI0-002011PANHANDLE20C1L_002093PANHANDLE50-002012PANHANDLE21C1, L_002127MIAMI0-002011PANHANDLE20C1L_002017PANHANDLE40-002127MIAMI0C1, L_002012PANHANDLE21-002041PANHANDLE30C1L_002012PANHANDLE21-002123PANHANDLE10C1, and L_002127MIAMI0-002011PANHANDLE20C1L_002012PANHANDLE21-002052PANHANDLE60C1. Thus, this hierarchical cluster is split into half. One contains Branch '2093' '2012' '1' and Branch '2018' '2093' '1', one includes Branch '2058' '2026' '1' and Branch '2128' '2026' '1'. Although those two clusters are electrically close to one another, they are addressed by two different remedial action schemes. All the finalized RAS clusters for the ACTIVSg2000 synthetic system are shown in Figure 7.8.

7.3 RAS Cluster Scenario Classification

7.3.1 Input Data Preparation

With the identification of RAS clusters, and the understanding of which scenarios can lead to certain violations based on the sampled scenario contingency analysis result, a two-stage support vector machine based learning approach is adopted to find the patterns in system operating conditions that are stressed or are at higher risk for violations. Those conditions are represented as operating features of sampled scenarios. Initially, operating features of a scenario include the MVA flows on all transmission branches, voltage per unit magnitude and angle at all buses, as well as the real and reactive power injection value at all buses. For each identified RAS cluster, sampled scenarios are labeled as binary values, where a True label means this scenario can potentially lead to severe violation of specific system elements in one RAS cluster, and a False tag means this sce-

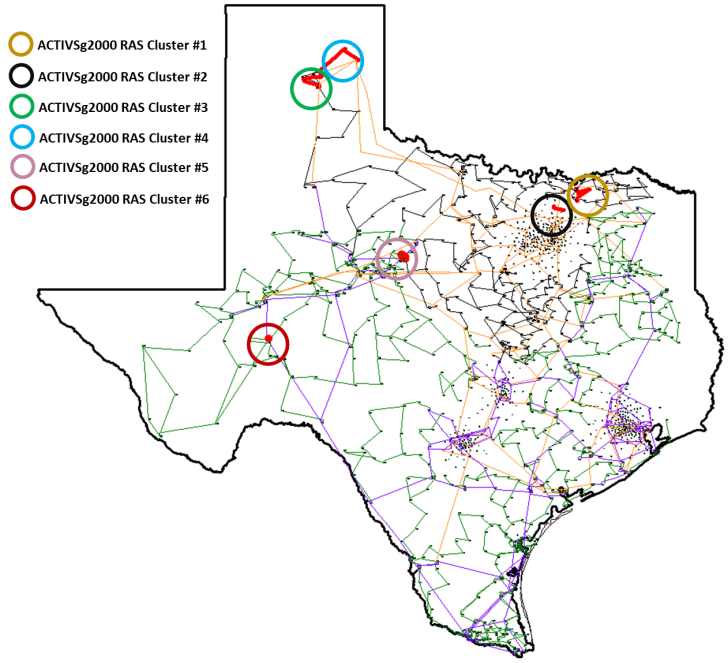

Figure 7.8: ACTIVSg2000 synthetic grid RAS clusters

nario is safe. With the scenario features being extracted and scenario data points being labeled, the support vector machine based learning method will reduce the number of features representing the operation scenarios, and find a hyperplane in the feature space that can best divide the scenarios with True and False labels.

It is worth noting that for some severe violation elements, there is a significant discrepancy between the number of True labeled scenarios and False labeled scenarios. Since remedial action schemes are usually created for urgent situations with relatively rare occurrence, the number of sampled scenarios having severe violations in the contingency analysis can be sparse. To effectively address the data label imbalance issue, artificial True labeled data can be created. This is achieved by creating permutations of features based on existing True tagged scenarios identified from the contingency analysis. Given one True labeled scenario, one feature is randomly selected and adjusted by a small amount between 80% to 120% to its original value. For example, one load of 100 MW in an existing scenario can be adjusted to a value between 80 to 120 MW as one

permutation to the original True scenario. Then the same contingency that have caused violations in the original scenario is studied on the mutated scenario, with the identified severe violation element being monitored. If the targeted element shows severe violation again in the permutation scenario, this scenario is added as an additional True labeled input data.

7.3.2 Two-Stage SVM for Scenario Classification

A two-stage support vector machine based model is developed in this work to classify scenarios for each RAS cluster. SVM algorithms evolve from theory to applications and results, whereas other classification techniques such as neural networks follow relatively more heuristic path from implementations to experiments. Consider training data consists of N pairs $(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)$, with $x_i \in \mathbb{R}^p$ and $y_i \in \{-1, 1\}$. Here p is the number of features representing a scenario, x_i are the feature values of each scenario, and $-1, 1$ represents the False and True labels of each scenario. If the classes are completely linearly separable, there exist a hyperplane $f(x) = x^T\beta + \beta_0$ with $y_i f(x_i) > 0 \; \forall i$. With $1/\|\beta\|$ set to be the smallest distance between data points and the hyperplane, the optimization problem formulation that creates the largest margin between the training class 1 and -1 is

$$\begin{aligned} &\min_{\beta,\beta_0} \; \|\beta\| \\ &\text{subject to} \quad y_i(x_i^T\beta + \beta_0) \geq 1, \quad i = 1, \ldots, N. \end{aligned} \tag{7.1}$$

If for certain problems the classes overlap in the feature space, we can still find a hyperplane that maximizes the margin, but allow of some points to be on the wrong side. This is done by defining the slack variable $\xi = (\xi_1, \xi_2, \ldots \xi_N)$, and modifying the constraints in 7.1 as:

$$\begin{aligned} &y_i(x_i^T\beta + \beta_0) \geq 1 - \xi_i, \\ &\xi_i \geq 0, \quad i = 1, \ldots, N \end{aligned} \tag{7.2}$$

SVM based classification models offer the advantage of non-abstract parameters, which can have strong physical implications. If the number of features is two, then the hyperplane is a straight

line. If the features are three-dimensional, then a two-dimensional plane can be founded to separate the data. The mathematical expression of the dividing hyperplane $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{\beta} + \boldsymbol{\beta_0}$ is a linear combination of selected features, and it provides the boundary condition where severe violations of targeting element can happen. In addition, since sensitivity based approach will be implemented for the development of corrective actions for Auto-RAS, a linear SVM can potentially provide additional insights on corrective actions that can move the system operating conditions to the side of the hyperplane that is considered to be safe.

In this case, a two-stage linear SVM model is developed to identify a hyperplane that can best divide the scenarios with and without the violation of specific system elements in one RAS cluster. The training process of this two-stage SVM model is illustrated in Figure 7.9. For each RAS cluster, the unique scenario labels, as well as all the representing features are used as input to the first stage linear SVM. The first stage SVM aims to find a subset of features as candidates to classify the scenarios of different classes. During the first stage linear SVM, if the number of True label is significant lower than that of False labeled data, artificial True data will be generated and added to the training data set. Penalty factors are also introduced during this stage to account for the imbalance of data tags. To optimize the first stage SVM parameters, grid search method is implemented to give a range of penalty factors, divide the grid under a given step size, traverse each data point in the grid by cross validation and select the penalty parameter with the highest accuracy rate. With the determination of penalty factors for the first stage SVM, the coefficient of each scenario feature is utilized to determine five candidate features for future steps. This coefficient vector with the same size as the number of total features is normalized to unity, the absolute size of one coefficient entry relative to other ones can give an indication of how important this specific feature is for the separation. Thus, the features with the top five highest absolute weight values are chosen as candidate features and used as inputs to the second stage SVM.

In the second stage SVM, different combinations of the five candidate features selected during the first stage are considered to reduce the dimensions of scenario feature space. With all the combinations considered, the scenarios can be represented in 31 different ways. For each feature

combination, a second stage linear SVM is trained to separate the scenarios of contrasting classes in a different feature space. There are two metrics used to quantify the ability of each second stage SVM model in separating the scenarios: accuracy rate and F1 score. Accuracy rate is the percentage of correct predictions on the testing data set, defined as follows, where TP stands for True Positive, TN is True Negative, FP is False Positive, and FN is False Negative.

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} = \frac{\boldsymbol{TP + TN}}{\boldsymbol{TP + TN + FP + FN}} \tag{7.3}$$

Since there exist some label imbalance in the data set, F1 score is also introduced to provide a closer look into the prediction errors. Since in this application, having false negative, or predicting True labeled scenarios as False, is less tolerable, the F1 score gives more weights to false negatives and false positives while not letting large numbers of true negatives influence to accuracy metrics. The formulation of F1 score is:

$$\begin{aligned} \text{F1} &= \boldsymbol{2} \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \\ \text{Precision} &= \frac{\boldsymbol{TP}}{\boldsymbol{TP + FP}} \\ \text{Recall} &= \frac{\boldsymbol{TP}}{\boldsymbol{TP + FN}} \end{aligned} \tag{7.4}$$

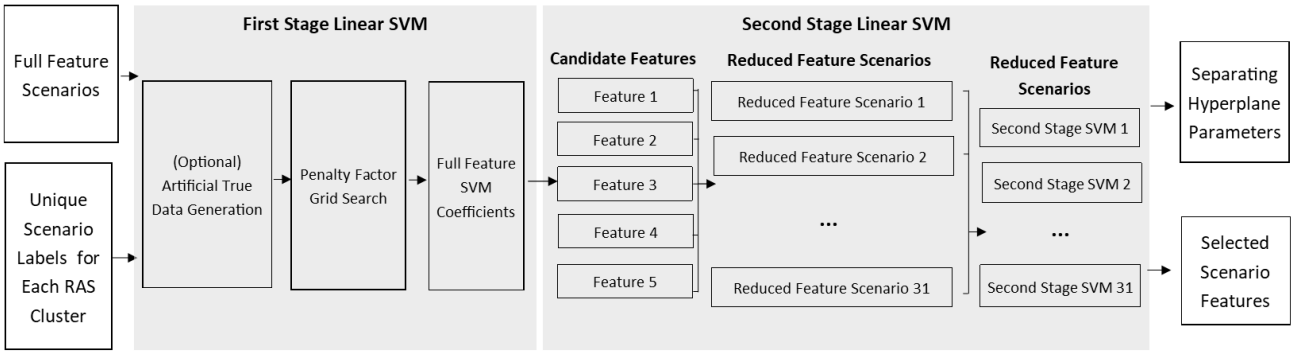

Figure 7.9: Two-stage linear SVM training process

Considering the accuracy rate and F1 score of all second stage linear SVM models, the one with the best performance is chosen. The parameters associated with the separating hyperplane, and the selected scenario features in this SVM model are used as the Auto-RAS condition settings. For the majority of the time, the severe violation elements will not experience overloading situation under the True labeled scenarios unless the occurrence of certain contingencies is present. In those cases, the separating hyperplane is modeled as the arming condition, while the critical contingencies are modeled as triggering logic of a remedial action scheme. In other situations, base case severe violation can exist under certain operating scenarios. In this case, the hyperplane is modeled as the triggering condition, and this remedial action scheme is always armed.

7.3.3 ACTIVSg200 and ACTIVSg2000 RAS Cluster Scenario Separating Hyperplanes

This section provides a summary of the ACTIVSg200 and ACTIVSg2000 RAS cluster scenario initial separating hyperplanes. Those initial hyperplanes are either directly translated into RAS conditions, or indirectly used as a reference for RAS condition settings. The detailed information of the developed RAS is introduced in Chapter 9. For the ACTIVSg200 synthetic system, each single element cluster finds an optimal separating hyperplane with one feature. Since the scenarios with different classes are linearly separable, 100% accuracy rate and a F1 score of 1 can be achieved. Figure 7.10 shows the visualizations of the one-dimensional hyperplanes for both RAS clusters. Table 7.3 provides a summary of those hyperplanes.

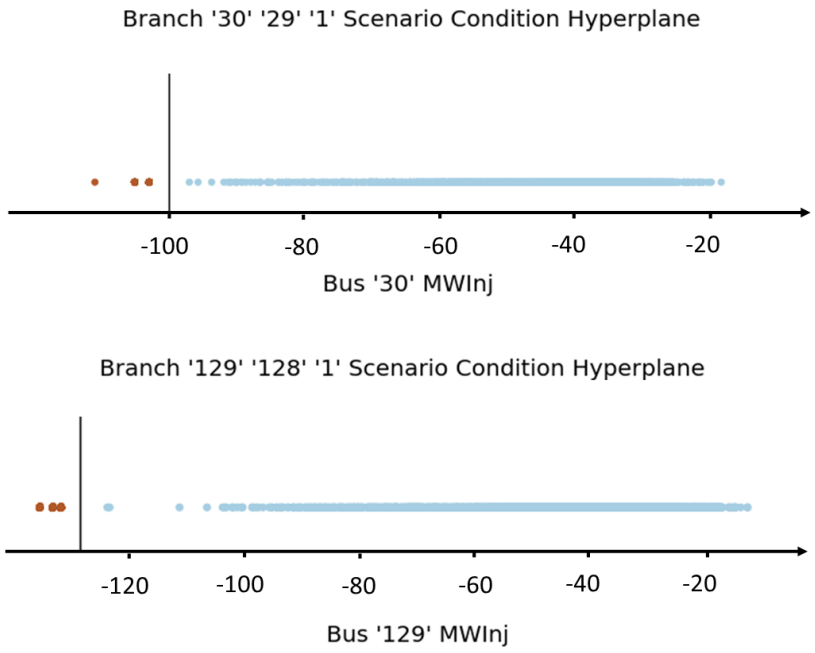

Figure 7.10: ACTIVSg200 scenario separating hyperplanes

Table 7.3: ACTIVSg200 Synthetic System Scenario Separating Hyperplane Summary

Cluster Number	Severe Violation Elements	SVM Feature Number	SVM Feature Type	SVM Accuracy	SVM F1 Score
1	Branch '29' '30 '1'	1	Bus MW Inj	100%	1
2	Branch '129' '128' '1'	1	Bus MW Inj	100%	1

Table 7.4: ACTIVSg2000 Synthetic System Scenario Separating Hyperplane Summary

Cluster Number	Severe Violation Elements	SVM Feature Number	SVM Feature Type	SVM Accuracy	SVM F1 Score
1	Branch '2081' '2129' '1' Branch '2038' '2003' '1' Branch '2081' '2038' '1'	3	Branch MVA Flow Bus MW Inj	100%	1
2	Branch '5450' '5331' '1'	1	Bus MW Inj	100%	1
3	Branch '2093' '2012' '1' Branch '2018' '2093' '1'	2	Bus MW Inj	96.15%	0.94
4	Branch '2018' '2093' '1' Branch '2128' '2026' '1'	1	Bus MW Inj	92.34%	0.91
5	Branch '3102' '3091' '1'	1	Branch MVA Flow	96.77%	0.94
6	Branch '1033' '1032' '1'	1	Bus MW Inj	98.70%	0.96

The initial hyperplanes achieving the best performance for the ACTIVSg2000 RAS clusters are shown in Figure 7.11, and a summary of those hyperplanes are provided in Table 7.4. All the Auto-RAS hyperplanes for the ACTIVSg2000 synthetic system are linear, with dimensions ranging from one to three. Two common types of SVM feature are the MVA flow on transmission lines, and the MW power injection on electrical buses. Additionally, two out of six RAS clusters are completely linearly separable, both SVM models can achieve an accuracy rate of 100% and a F1 score of 1. For the four RAS clusters where the scenarios are not completely linearly separable, the accuracy rate ranges from 92.34% to 98.70%, and the F1 score ranges from 0.91 to 0.96.

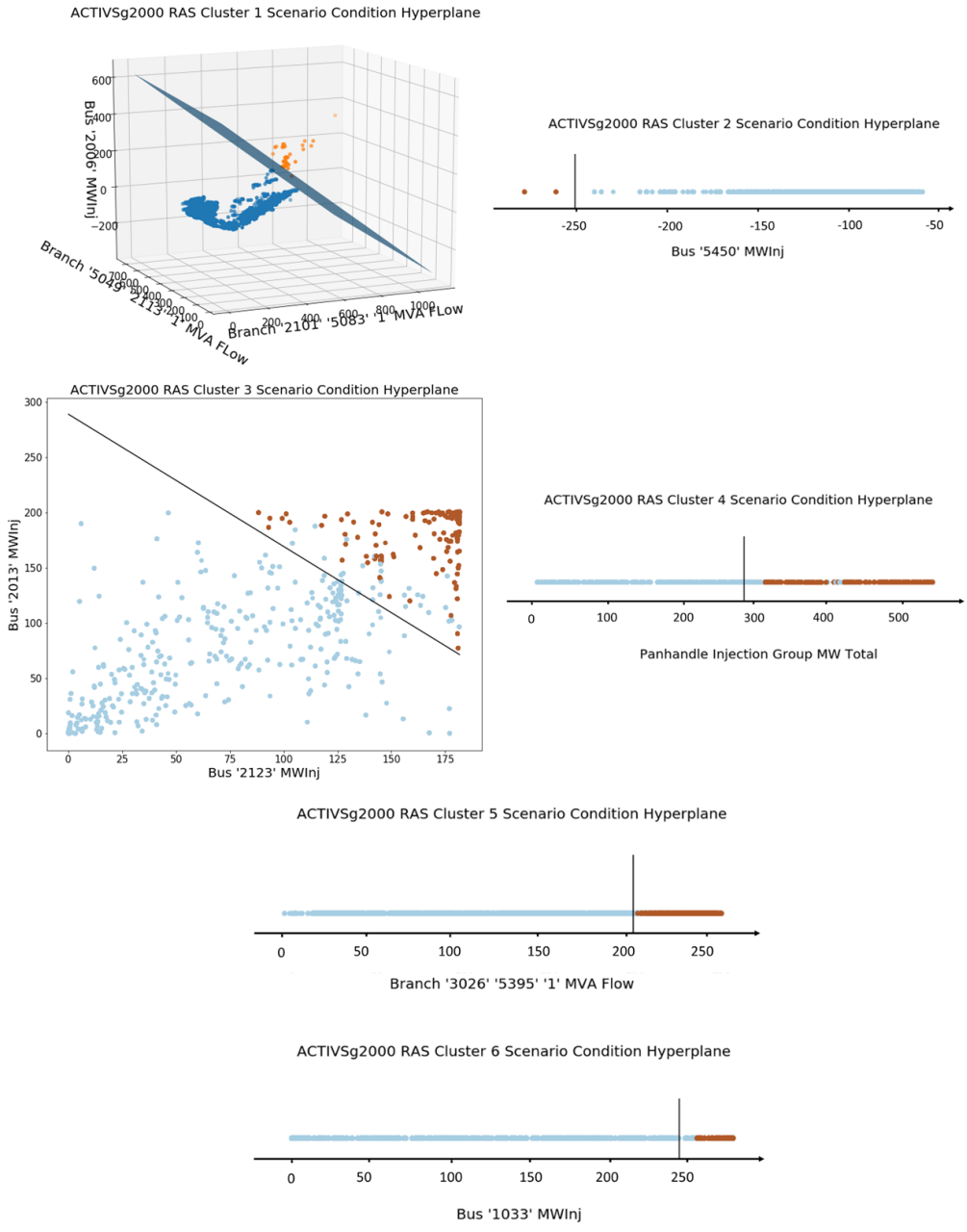

Figure 7.11: ACTIVSg2000 scenario separating hyperplanes

8. DETERMINATION OF AUTO-RAS CORRECTIVE ACTIONS

The previous chapter covers a data-driven approach that systematically identifies power system operating conditions under which auto-RAS should be armed and triggered. Utilizing chronological simulation of scenarios and the operational condition analysis, the system's need of a new RAS, and its corresponding logic conditions are identified in the planning time frame. This chapter introduces a sensitivity-based methodology to determine the corrective actions that can be deployed adaptively to prevent or mitigate operational violations during real-time operation time frame. Using the topology and parameters of the current system operating state, this corrective action design procedure automatically generates action plans for every targeting element, or elements in the same RAS cluster, with a focus on transmission branch overloading mitigation. Leveraging network connectivity analysis, a subset of power system elements that can be controlled as part of the corrective action scheme is selected based on their electrical properties and proximity to the targeting violation elements. Sensitivity analysis such as line outage distribution factor and transmission loading relief are used to quickly determine the most effective controllable elements and the corresponding corrective actions to address specific operational violations or potential violation risks, eliminating the need for repeated numerical simulations to determine the corrective actions. The parameters associated with auto-RAS corrective actions are selected based on statistical and functional characteristics derived from actual power system RAS implementations.

Sensitivity analysis is a common technique implemented in power system operations and controls. This technique approximates relationships between different power system components, and can provide useful insights in an efficient manner even for large-scale power networks. The work of [107] and [108], for example, used the sensitivity between bus voltages and nodal real and reactive power values to determine a group of coherent buses, and establish the voltage security margins. In [109], the analytical eigenvalue's sensitivity to generation schedule is calculated to enhance the damping of critical oscillation modes. A sensitivity-based under frequency load shedding is proposed in [110], utilizing linearized power flow analysis and generation droop characteristics.

Many sensitivity-based applications are also focused on overload alleviation, such as the work of [111], [112], and [113]. In this dissertation, two common sensitivity calculations, the transmission line relief, and line outage distribution factor are leveraged to identify the most effective course of action to mitigate the branch overload violation in real-time operation time frame. While most of the existing work only consider one control scheme, such as reactive power devices, generation re-scheduling, or branch switching, this work looks into three types controllable elements. In addition, the sensitivity analysis is embedded in an adaptive Auto-RAS framework, where the created corrective action might vary depending on different real time operation situations. This method to determine auto-RAS corrective actions is illustrated using the ACTIVSg200 and ACTIVSg2000 synthetic systems. The detailed corrective actions and their design considerations for those two systems are given in chapter 9.

Part of this chapter is reprinted with permission from "Towards the Automation of Remedial Action Schemes Design" by Hanyue Li, Komal Shetye, Thomas J Overbye, Katherine Davis, and Shamina Hossain-Mckenzie 2021. Proceedings of the 54th Hawaii International Conference on System Sciences, Page 3378, Copyright 2021 by HICSS.

8.1 Reference Statistics on RAS Corrective Actions

To guide the design of corrective action for each targeting transmission element or branch groups identified in the previous steps, statistical analysis was conducted on RAS implemented in real power systems to summarize their key features and range of parameter values. All of the industrial RAS used here are designed to mainly alleviate one or multiple transmission line thermal violations. Figure 8.1 shows the four categories of corrective actions adopted by these industrial RAS. Generator adjusting, which includes the opening of generator units and changing generator output set points, is the most commonly used action to mitigate line overloads. Branch switching, and load shedding are two other popular schemes adopted by the industry. Some thermal violations also require hybrid actions with a mix of generator adjusting, branch opening and load shedding.

Table 8.1 describes the characteristics and range of key parameters for each corrective action type. For branch switching schemes, the total number and MVA capcity of lines opened, the

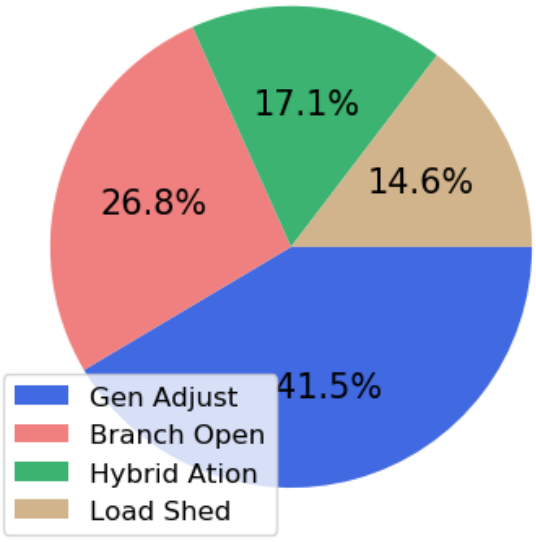

Figure 8.1: Corrective action types for overload mitigation

number of nodes away from the overloading line, and their nominal voltage levels are recorded. For generator adjusting, the number and MW value of generators adjusted, their distance from the overloading line, and their fuel types are used as reference for the corrective action design. For load shedding, the number and MW value of load opened, and the number of nodes away from the overloading line are also noted.

8.2 Auto-RAS Corrective Action Design

Given the network and electrical properties of the overloading branch or branch group, auto-RAS determines the corresponding corrective actions in a systematic and effective manner. This is achieved by network connectivity analysis that selects only a subset of system elements to control, and power system sensitivity studies that look into the linear relationship between the targeting overloading branch, and the controllable elements. This process is illustrated in Figure 8.2 and explained in details in this section.

Element Type	Characteristics	Range	Average
Branch	Number of branch switched	1-2	1
	Capacity of branch	30- 500 MVA	223 MVA
	Distance from overloading line	1-3 node(s)	2 nodes
	Nominal voltage	69-76 kV	230 kV
Generator	Number of units adjusted	1-10	3
	Size of gen adjusted	42- 350 MW	120 MW
	Distance from overloading line	1-12 node(s)	4 nodes
	Fuel type	wind, solar, natural gas	_
Load	Number of load opened	1-8	5
	Size of load adjusted	12- 75 MW	21 MW
	Distance from overloading line	1-10 node(s)	6 nodes

Table 8.1: Industrial RAS Corrective Action Statistics

8.2.1 Selection of Controllable Elements

To reduce the computational complexity of auto-RAS corrective action design, depending on the properties of the overloading branch or branch group, only a subset of system elements, such as generators, loads, and transmission lines are chosen as candidate controllable elements that have the potential to be included as part of the corrective action scheme. This selection process is based on the network connectivity analysis, where the shortest path distance between the targeting overloading branch/branch group, and the rest of the system nodes are calculated.

In the controllable element selection process, the electric grid network is considered as an unweighted, undirected graph. The source vertices are buses associated with the overloading branch, or overloading branch groups identified in the auto-RAS scenario analysis step. Given the network adjacency matrix and the list of source vertices, the shortest paths from the source to all vertices in the given graph are found using the Dijkstra's algorithm mentioned in Chapter 7. In this case, the process is repeated M times, where M is the total number of buses associated with an overloading branch group. The shortest path from the source to each graph vertex is then updated as the minimum value among all iterations. Figure 8.3 shows an illustration of this process. The vertex marked in color pink represents all the buses associated with one overloading branch group, and

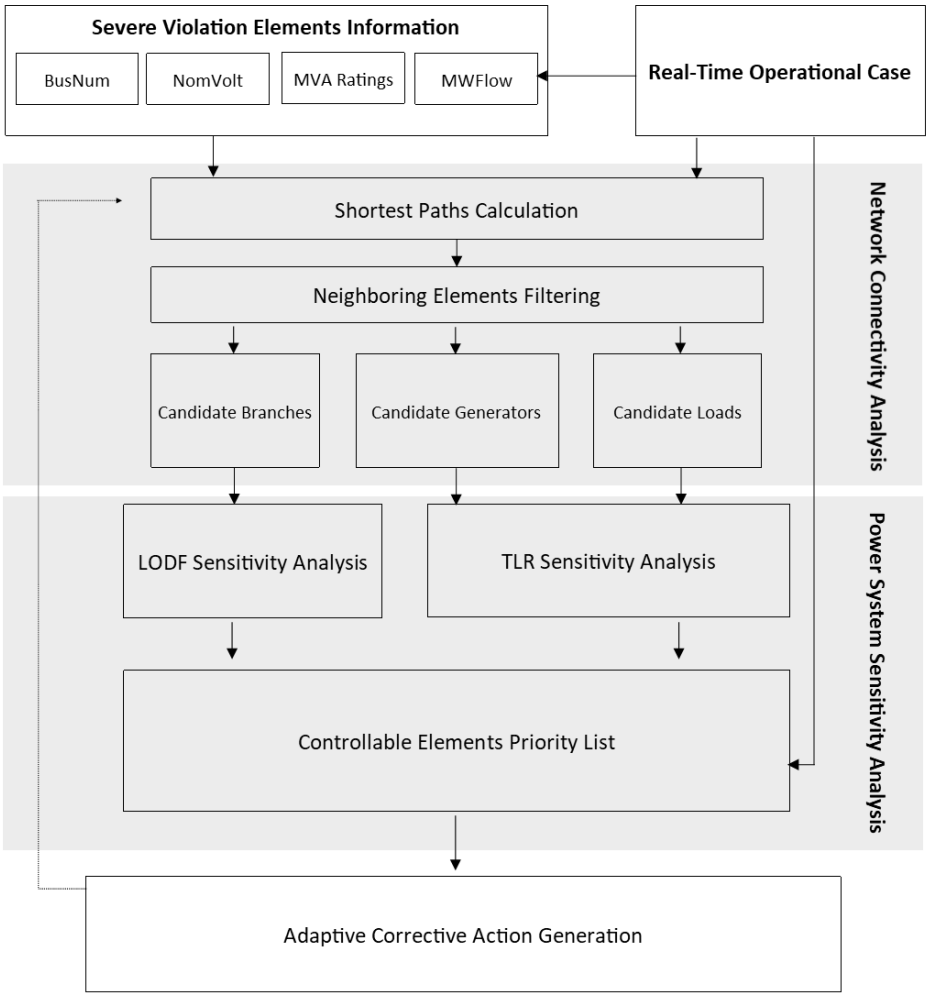

Figure 8.2: Auto-RAS corrective action generation flow chart

the green dots presents the rest of the nodes in the electric grid network. The unweighted shortest path distances are determined between the overload branch group and the rest of the system, marked as numbers in each node.

Using the statistics summarized from the RAS models implemented in the real electric grid, a subset of system elements attached close to each branch overloading group is selected. Transmission branches above 69-kV and also within three nodes away are considered to be controllable equipment. Besides the distance from thermal violation location, generators are filtered out based on their fuel types, where only wind, solar and natural gas units will be used to develop cor-

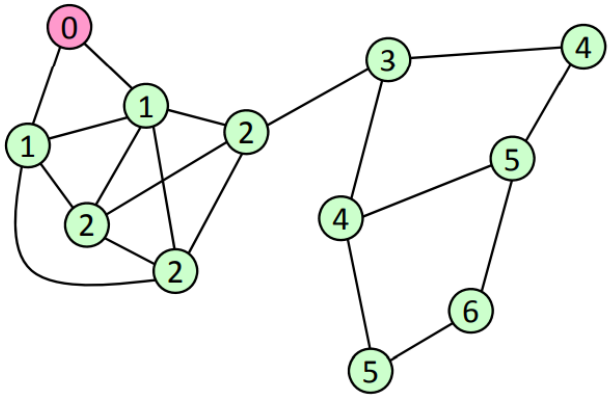

Figure 8.3: Determine shortest path distances to overloading branches/ branch groups

rective actions. Loads that can be reduced to mitigate the overloads are selected only based on their locations relative to the targeting overloading branches. This step limits the list of candidate controllable elements to a small number, so that solution space of the effective remedial action is reduced to optimize the computation time. Figure 8.4 and Figure 8.5 show visualizations of candidate controllable elements for each RAS cluster. The substations associated with the severe violation elements in one RAS cluster are shown in red rectangles, and the substations associated with candidate controllable elements are shown in dark green.

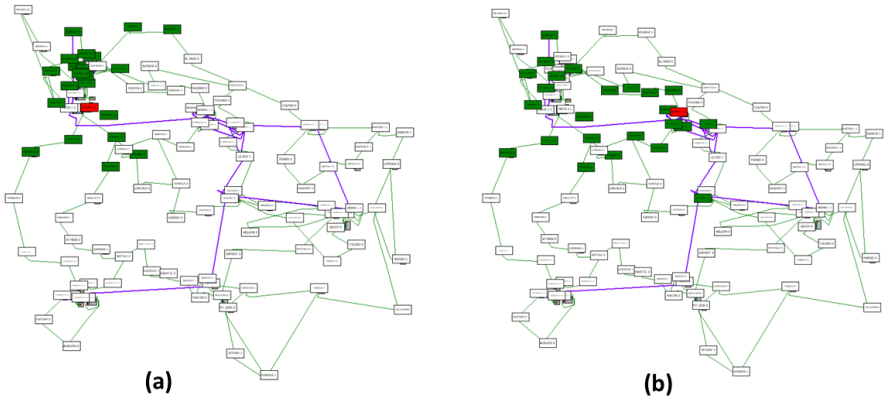

Figure 8.4: Visualization of candidate controllable elements for ACTIVSg200 (a) RAS Cluster #1 (b) RAS Cluster #2

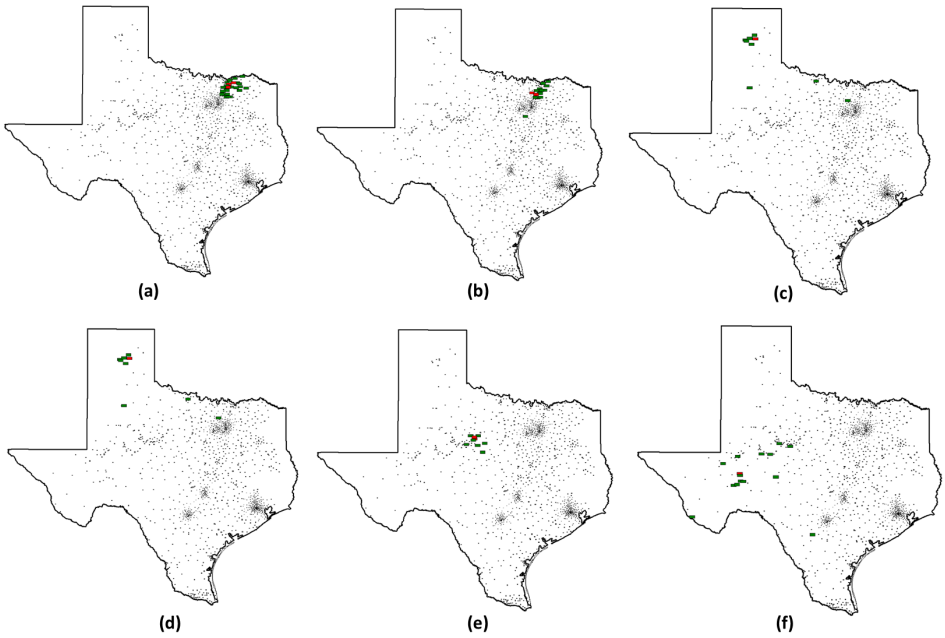

Figure 8.5: Visualization of candidate controllable elements for ACTIVSg2000 (a) RAS Cluster #1 (b) RAS Cluster #2 (c) RAS Cluster #3 (d) RAS Cluster #4 (e) RAS Cluster #5 (f) RAS Cluster #6

8.2.2 Sensitivity Relation Between Controllable Elements and Auto-RAS Branch Groups

To ensure that corrective actions can be automatically created effectively, where action plan parameters can be adjusted according to specific system conditions during system operation, sensitivity-based techniques are leveraged to study the relationships between the controllable elements and the overloading lines. Depending on the network topology and its associated parameters of an operational scenario, injection shift factor $\mathbf{\Psi}$ is constructed using the lossless dc power flow assumptions. The change in real power line flows is given as,

$$\mathbf{\Delta f} = \mathbf{\Psi \Delta p} \tag{8.1}$$

where $\mathbf{\Delta p}$ is the vector of changes in the real power injection at each bus, and the element ψ_l^n in row l and column n of $\mathbf{\Psi}$ is the ISF of line l with respect to the injection at node n. For a system with L lines and N nodes (buses), $\mathbf{\Psi}$ is defined as,

$$\mathbf{\Psi} = \tilde{\mathbf{B}}\mathbf{A}[\mathbf{B'}]^{-1} \tag{8.2}$$

where $\tilde{\mathbf{B}}$ -*diag*$\{b_1, b_2, ..., b_L\}$ (b_l is the series susceptance of line l), $\mathbf{A}$ is an $\mathbf{L} \times \mathbf{N}$ incidence matrix where the element a_{ij} is non-zero only when line l_i is coincident with node j, and $\mathbf{B'} = \mathbf{A^T}\tilde{\mathbf{B}}\mathbf{A}$.

Auto-RAS utilizes two sensitivity techniques to rank the controllable elements. To study the effectiveness of a branch switching action, the LODF matrix represented by $\mathbf{d}$ is derived from the $\mathbf{\Psi}$ matrix. An element d_l^k of this matrix $\mathbf{d}$ can be written as,

$$d_l^k = \frac{\psi_l^m - \psi_l^n}{1 - (\psi_k^m - \psi_k^n)} \tag{8.3}$$

where m and n are the two nodes associated with the switched line k. The real power flow change on line l due to the switching of line k is given as,

$$\Delta f_l = df_k \tag{8.4}$$

where f_k is the pre-switching real power flow on line k.

The impact of generator adjusting and load shedding are determined using TLR sensitivities. The TLR sensitivity matrix is also derived from Ψ. It estimates the impact of multiple transfers on one transmission line. An element ϕ_l^w of this TLR sensitivity matrix Φ can be written as,

$$\phi_l^w = \psi_l^m - \psi_l^n \tag{8.5}$$

where l is the transmission line of interest, and two nodes m and n define one transaction w. The real power flow change on line l due to the transaction w is given as,

$$\Delta f_l = \phi_l^w \Delta t \tag{8.6}$$

where Δt is the amount of MW transfer associated with transaction w. In this work, the adjustments made on generation and load values and status are modeled as transactions, where generations are usually modeled as sellers, and loads usually as buyers.

8.2.3 Determination of Adaptive Corrective Actions Plan

With the network connectivity analysis and power system sensitivity analysis, a list of candidate controllable elements and the sensitivities between them and the severe violation elements are obtained. This information is utilized to construct an initial corrective action priority list, illustrated in Table 8.2. The MW value of each controllable element is obtained from the real-time operational case, and the sensitivity factors are also calculated based on the real-time system topology. For each controllable element, the MW reduction capacity, defined as the the MW value times the sensitivity factor, is calculated. Those controllable elements are firstly sorted by the device type, where generators is the most preferable option, while load has the lowest priority. This priority list is then sorted by the controllable elements' MW reduction capacities from the greatest

to the smallest, where the element with the largest value will have the priority to be used as part of the corrective action scheme. The accumulative value of the MW reduction capacities is also calculated along with the sorted list. It serves as a reference point for the adaptive auto-RAS corrective action generation procedure. Depending on the MW value that needs to be reduced from the severe violation element, the number of controllable elements associated with the corresponding corrective action may vary.

Table 8.2: Corrective Action Priority List Example

Element ID	Element Type	MW Value	Sensitivity	MW Reduction Capacity	Accumulative MW Reduction Capacity
Element 1	Gen	X1	Y1	X1*Y1	X1*Y1
Element 2	Branch	X2	Y2	X2*Y2	X1*Y1 + X2*Y2
...	...	...	...	...	...
Element n	Load	Xn	Yn	Xn*Yn	X1*Y1 + + Xn*Yn

In cases where branches are selected before the accumulative MW reduction capacity cut off line in the priority list, additional iterations of network connectivity and sensitivity analysis are needed, because the electric node distance and sensitivity factors will vary with the change in system topologies. This process is illustrated in Figure 8.6. For each iteration, if the total number of controllable elements included in the proposed adaptive corrective action plan is $\boldsymbol{m}$, and there exist one branch element, the first branch controllable element is in position $\boldsymbol{n}$, all the actions for the generator and load elements $\boldsymbol{n-1}$ are applied to the network model, where the system topology, and the desired MW reduction value for the severe violation elements are updated. The updated network model is then used as input for the network connectivity and sensitivity analysis, where a new corrective action priority list will be generated.

Since the construction of corrective action priority list is dependant on the topology and parameters of a specific system operating state, the corrective actions in the remedial action scheme is subject to change to adapt to different operating conditions. Examples of the corrective action priority list, and the adaptive corrective action plan for the remedial action schemes created for each synthetic system are provided and explained in details in Chapter 9.

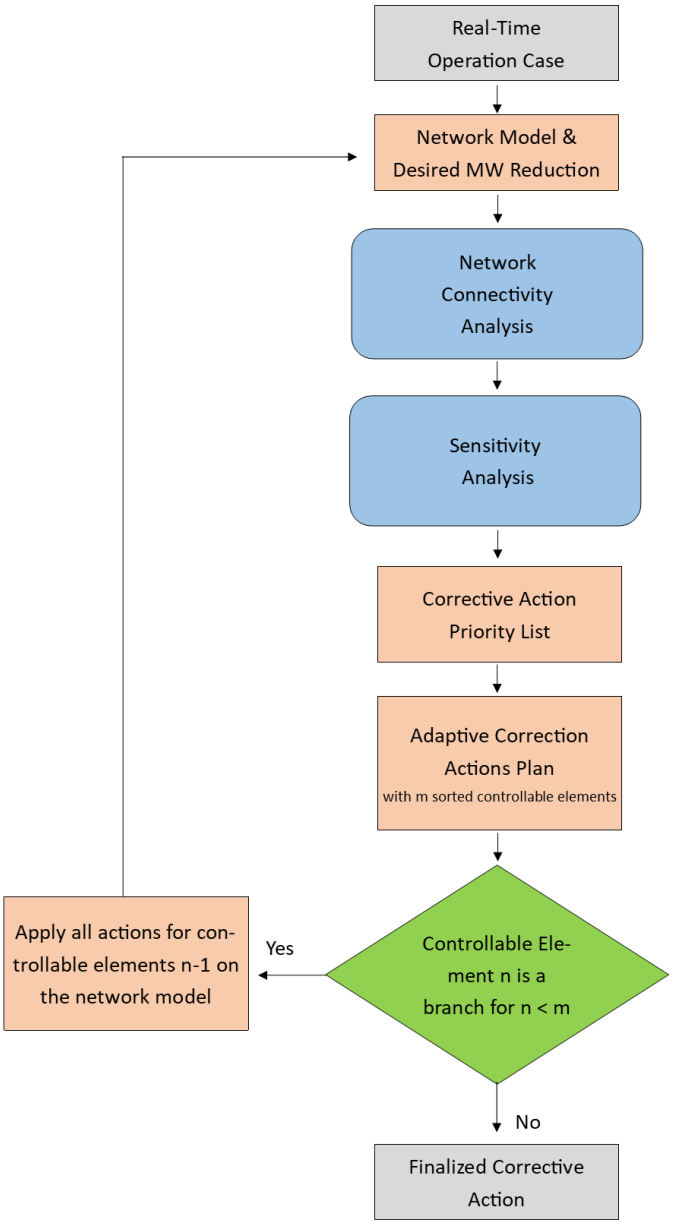

Figure 8.6: Determination of adaptive corrective actions plan

9. AUTO-RAS TESTING AND RESULT

The testing of the Auto-RAS follows the NERC reliability standards PRC 12-2 and PRC 17-1 [30], as well as the RAS review process established by WECC [114] to ensure that the designed remedial action schemes can operate to perform their intended functionalities, and do not introduce unintentional or unacceptable reliability risks to the bulk electric power system. Utilizing the large volume of operational scenarios created from the chronological simulation framework, a subset of testing cases are sampled from a pool of operational scenarios that are not selected during the Auto-RAS creation process. This step confirms that the Auto-RAS arming and triggering conditions are effective and appropriate to its system performance objectives, and the adaptive corrective action determination framework is generic enough, and can adjust to different operational system conditions.

During the Auto-RAS development process, a contingency list is constructed to include events up to N-4 contingencies. This contingency list was used for the operational scenario analysis process, where severe violation elements are identified. For the testing of Auto-RAS, the contingencies that can cause the overload of severe violation elements in at least one operational scenario are chosen. Using the shortened contingency list, each steady-state contingency is executed two times on the testing scenarios, the first time without the created Auto-RAS on the specific system to establish a performance benchmark, and the second time with the inclusion of all Auto-RAS models. By comparing the result of first round and second round contingency analysis, the checklist shown in Figure 9.1 is reviewed to evaluate the performance of Auto-RAS models.

The testing of ACTIVSg200 Auto-RAS studies 300 operational scenarios and 2 different contingencies. For ACTIVSg2000 synthetic system, 500 scenarios are sampled for the testing, and 14 contingencies are studied. All the remedial action schemes developed using the Auto-RAS framework passed the testing checklist, and are explained in details for the rest of this chapter.

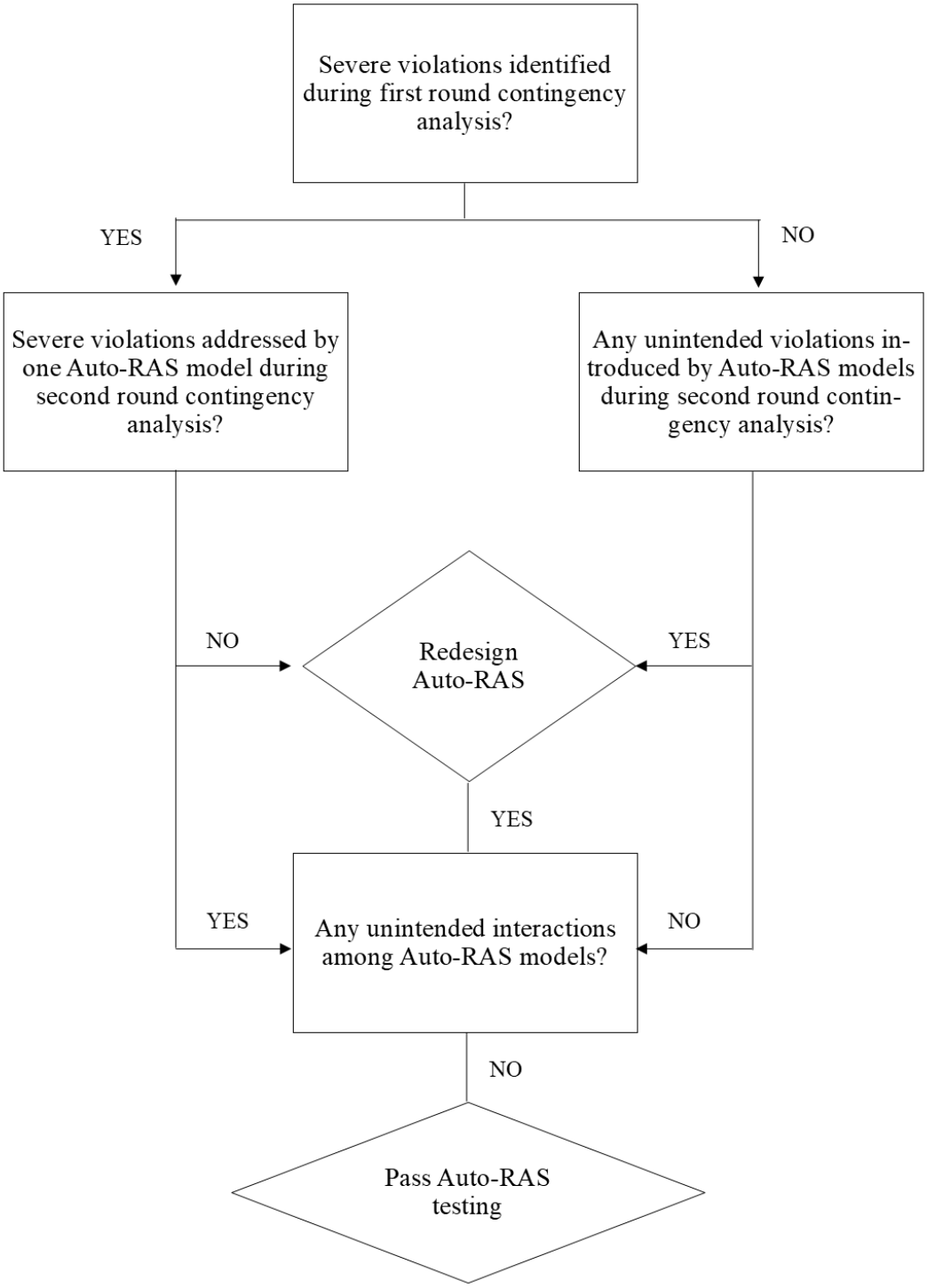

Figure 9.1: Auto-RAS testing checklist

9.1 ACTIVSg200 RAS Result

9.1.1 ACTIVSg200 RAS #1

The first RAS cluster for the ACTIVSg200 synthetic system contains only one severe violation element, the 115 kV transmission circuit between Bus PEKIN 2 0 and Bus PEKIN 2 1, shown in Figure 9.2. The bus number associated with those two buses are 29 and 30, respectively. This transmission branch a radial path with load serving purposes. Out of the 1744 operational scenarios, only two of them exhibit violation risks and are labeled as True data. To address the data label imbalance issue, 500 additional artificial scenarios are created as permutations of the original two True-labeled operational scenarios. A linear SVM model is trained with 1520 operational scenarios, initially each scenario is represented by 1045 features. The SVM model sweeps through different number of selected features, ranging from 1 to 10, and finds that solely using the feature "Bus '30' MWInj", the real power injection at Bus 30, can achieve the best result with an accuracy rate of 100% and F1 score of 1. Figure 9.3 shows the hyperplane that divides the training and testing scenarios of different classes. Since only one feature is selected in this case, the separating hyperplane classifies all the scenarios with less than -99.82 MW real power injection at bus 30 as True. In other words, if the load at Bus 30 is more than 99.82 MW, it is considered as a dangerous situation where Branch '30' '29' '1' can be overloaded.

This learning result is validated by investigating the topology and electrical ratings of the system surrounding the severe violation branch. Since the 115 kV transmission circuit between Bus PEKIN 2 0 and Bus PEKIN 2 1 is the radial path supply load at Bus 30, the loading of this transmission branch is directly correlated with the load value over time. Since the emergency rating of this branch is 100 MVA, it is logical that if the real power load value at bus 30 is more than 99.82 MW, the branch of interest will very likely be overloaded. Additionally, since this condition causes base case violation, the violation is founded when analyzing all the steady-state events in the defined contingency list.

With the base case violation situation, the arming condition of this RAS is set to be True at all

times, so that the remedial action scheme is always armed and can be initiated when the triggering condition is true. The triggering condition monitors the MW load value at Bus 30, and will initiate the corrective action whenever the load value exceeds 99.82 MW. The elements associated with this RAS condition logic are marked in Figure 9.2. The severe violation element. Branch '30' '29' '1', is marked with an explosion icon. The load associated with the RAS triggering condition is highlighted with blue shadings.

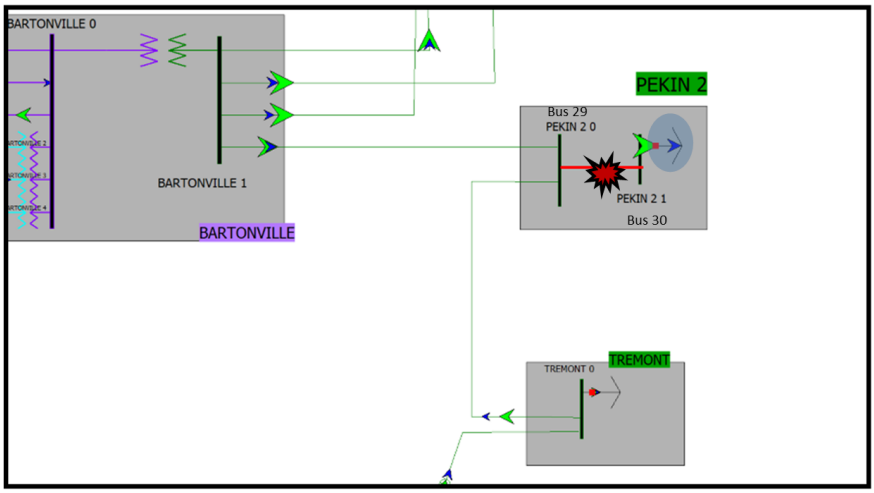

Figure 9.2: ACTIVSg200 RAS #1 condition: detailed one-line with condition elements

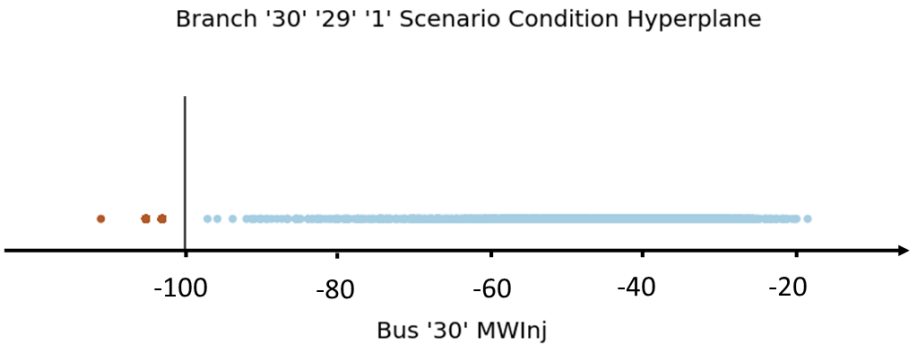

Figure 9.3: ACTIVSg200 RAS #1 condition: scenario separating hyperplane

The determination of corrective actions for this particular remedial action scheme is fairly straight forward. The initial corrective action priority list of an example testing scenario for ACTIVSg200 RAS #1 is shown in Table 9.1. Since the load value at Bus 30 the 100% correlated to the power flow on Branch '29' '30' '1', and the cause of overloading situation, load shedding at this specific bus is the only effective corrective action to alleviate the branch overloading violation. In this case, whenever the load MW value at Bus 30 exceeds 98.82 MW, it needs to be reduced to this limit value. Figure 9.4 demonstrates this corrective action plan. On a computation system with an Intel Xeon Processor and 32 GM of RAM, the corrective action for this example can be determined in 1.9 seconds.

Table 9.1: ACTIVSg200 RAS #1 Initial Corrective Action Priority List

Element ID	Element Type	MW Value	Sensitivity	MW Reduction Capacity	Accumulative MW Reduction Capacity
Load '30' '1'	Load	104.9	1	104.9	104.9

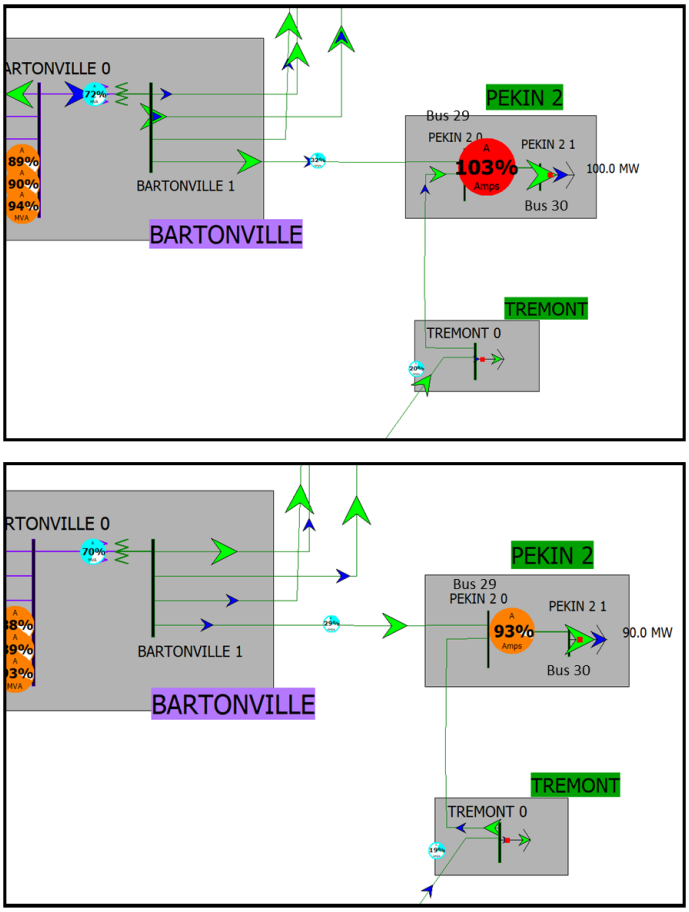

Figure 9.4: ACTIVSg200 RAS #1 corrective action demonstration

9.1.2 ACTIVSg200 RAS #2

The second RAS for the ACTIVSg200 synthetic system addresses a severe branch violation with power flow re-routing after the opening of two transmission lines from the same substation. Illustrated in Figure 9.5, NORMAL 1 and NORMAL 2 are two 230 kV substations transmitting power from a higher voltage network to the 115 kV network where electric load is aggregated to. During normal conditions, those two 230 kV substations feed the BLOOMINGTON 1 and BLOOMINGTON 2, two load substations from the 115 kV grid collectively. During one contingency named "L_000107BLOOMINGTON20-000113NORMAL21C1_000113NORMAL21-000192BLOOMINGTON10C1", two 115 kV transmission lines, Branch ‘107’ ‘113’ ‘1’, and Branch

‘113’ ‘192’ ‘1’, are out of service at the same time, cutting off the path for Substation NORMAL 2 to directly supply power to Substation BLOOMINGTON 1 and BLOOMINGTON 2. During the scenario analysis studies for the ACTIVSg200 system, 72 out of 1520 operational scenarios experienced severe overloading violations for the 115 kV transformer connecting Bus NORMAL 1 0 and Bus NORMAL 1 1 due to the power flow re-routing.

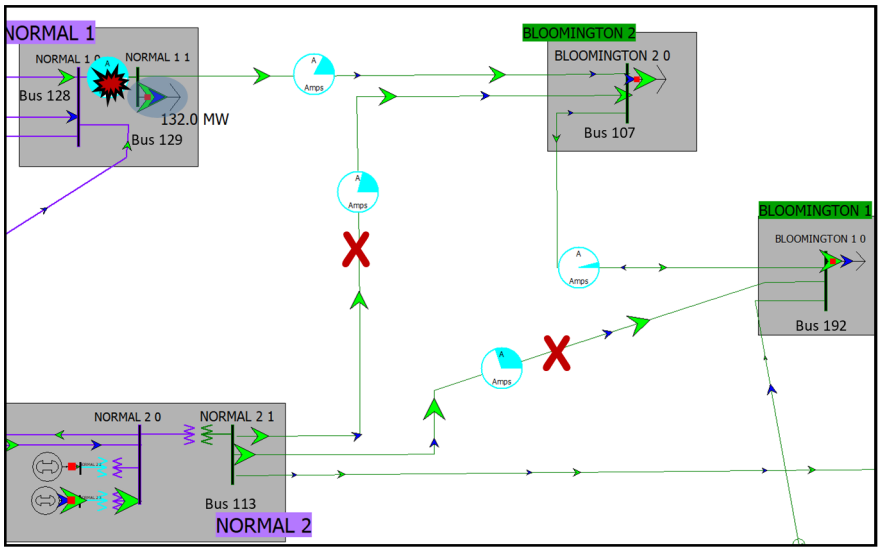

Figure 9.5: ACTIVSg200 RAS #2 condition: detailed one-line with condition elements

Using the operational scenario features as input, a linear SVM model is trained to distinguish those scenarios experiencing this severe branch overloading violation from the normal scenarios. Initially each scenario is represented by 1045 features. The 2-stage SVM model first finds 10 candidate features which are most correlated to the severe violation element, and sweeps through different features combinations, with the number of selected features ranging from 1 to 10. The SVM model training process finds that using the feature "Bus ’129’ MWInj", the MW power injection at Bus 129, can achieve the best result with an accuracy rate of 100% and F1 score of 1 on the testing data set. Figure 9.6 shows the hyperplane that divides the training and testing scenarios of different classes. Since only one feature is selected in this case, the separating hyperplane classifies the

scenarios with Bus '129' MWInj value less than -125.3 MW as True. In other words, if the load aggregation at Substation NORMAL 1 exceeds 125.2 MW, it will pre-load the transformer connecting Bus NORMAL 1 0 and Bus NORMAL 1 1 to a higher level. With the occurrence of Contingency "L_000107BLOOMINGTON20-000113NORMAL21C1_000113NORMAL21-000192BLOOMINGTON10C1", the 230 to 115 kV transformer in Substation NORMAL 1 is at risk of being severely overloaded.

ACTIVSg200 RAS #2 monitors the load MW value at bus 129, and is armed when the monitored value goes over 125.3 MW. If Branch '107' '113' '1' and Branch '113' '192' '1' are detected with change of status simultaneously, this remedial action scheme is triggered to take corrective actions immediately. An example initial corrective action priority list is shown in Table 9.2. For the determination of corrective actions, there are multiple controllable elements with high sensitivity and close to the severe violation element. The load at Bus 129 has a sensitivity of 0.84 towards the MW flow on the overloading line. The sensitivity of loads at Bus 107 and 192 are at 0.75 and 0.68 respectively. The summation of those load is 150 MW on average throughout the year, which is a significant amount for a load shedding scheme to be possible. On the other hand, the LODF of the branch between bus 107 and bus 192 towards the overloading line is -100%. In other words, if Branch '107' '192' '1' is switched off, the power flow reduction will 100% reflect on the MW flow on the targeted severe violation branch. With load shedding being at the bottom of the priority list, the branch switching scheme has the priority to be selected as the corrective action. On a computation system with an Intel Xeon Processor and 32 GM of RAM, the corrective action for this example can be determined in 2.2 seconds. Figure 9.7 shows the system condition before the contingency, and post-contingency without ACTIVSg200 RAS #2 implementation,and post-contingency with ACTIVSg200 RAS #2 implementation.

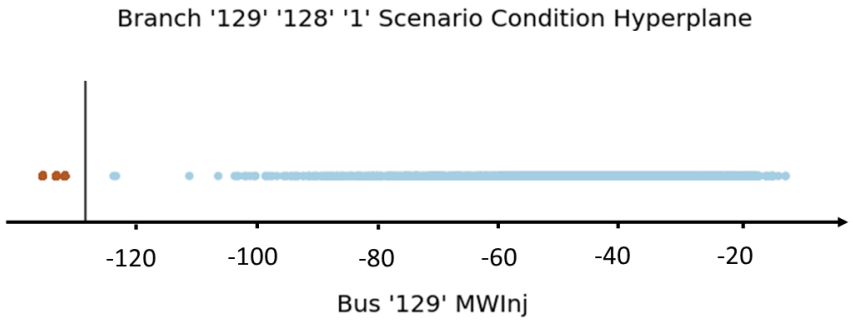

Figure 9.6: ACTIVSg200 RAS #2 condition: scenario separating hyperplane

Table 9.2: ACTIVSg200 RAS #2 Initial Corrective Action Priority List

Element ID	Element Type	MW Value	Sensitivity	MW Reduction Capacity	Accumulative MW Reduction Capacity
Branch ‘192’ ‘107’ ‘1’	Branch	43.5	1	43.5	43.5
Load ‘129’ ‘1’	Load	132.8	0.84	111.6	155.1
Load ‘107’ ‘1’	Load	73.5	0.75	55.1	210.2
Load ‘192’ ‘1’	Load	76.5	0.68	52.0	262.2

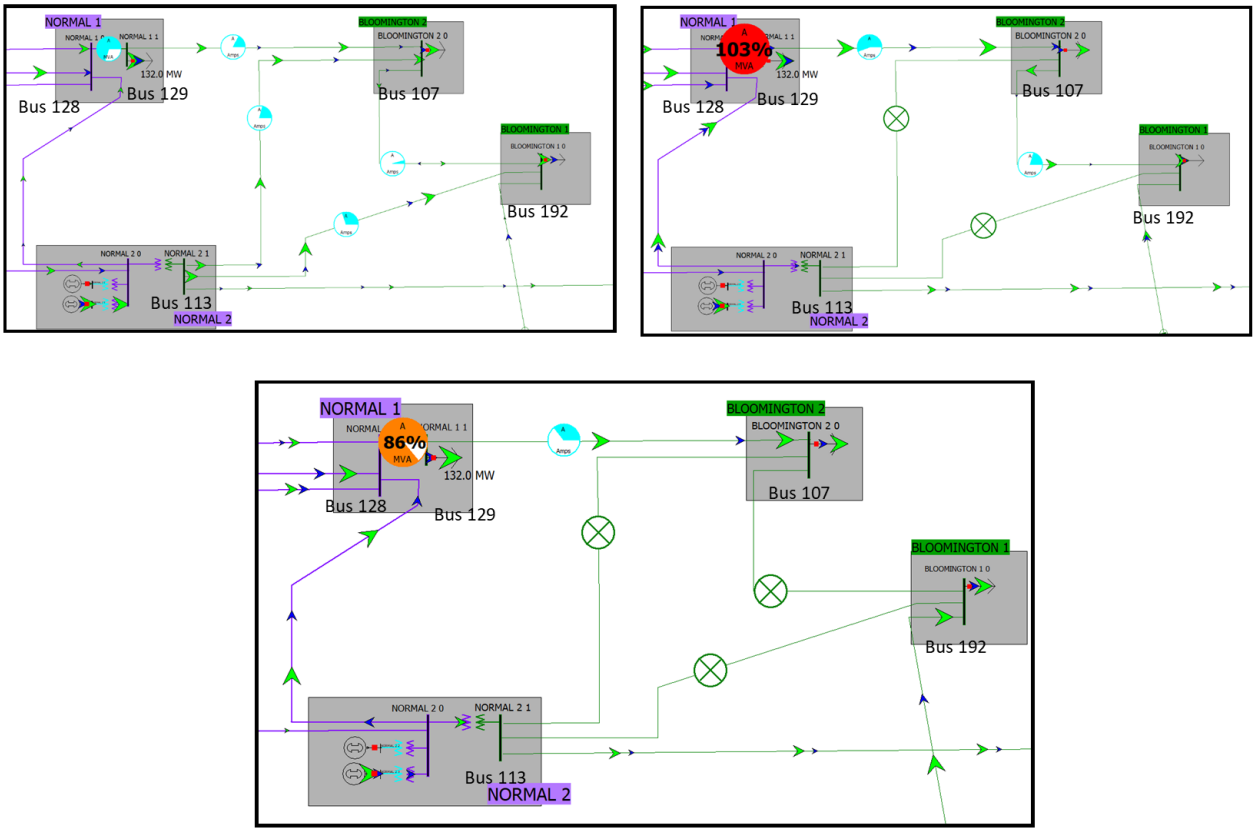

Figure 9.7: ACTIVSg200 RAS #2 corrective action demonstration

9.2 ACTIVSg2000 RAS Result

9.2.1 ACTIVSg2000 RAS #1

The first remedial action scheme for the ACTIVSg2000 synthetic network aims to protect a set of 161 kV transmission facilities close to the 500 kV network located in north Texas. The three severe violation elements, Branch ‘2038’ ‘2003’ ‘1’,Branch ‘2081’ ‘2038’ ‘1’, and Branch ‘2081’ ‘2129’ ‘1’, are three 161 kV transmission branches with emergency ratings at 251 MVA, 251 MVA, and 219 MVA respectively. They are marked with explosion signs in Figure 9.8. One of the buses associated with those targeted violation elements, Bus 2003, is located in the 500 kV substation SAVOY and connected to the higher voltage network via a 500/161 kV transformer. Bus SAVOY 0 is the 500 kV bus located in the same substation. There are three generators connecting to the 500 kV Bus SAVOY 0 with step-up transformers. All three units have the capacity of 391.8 MW, and they are directly attached to buses numbered 2004 to 2006. There are three additional neighboring 500 kV substations near substation SAVOY in north Texas, all of them have bulk-level generation attached as well. Those 500 kV substations are interconnected, and the two 500 kV branches at Substation SHERMAN 1 and PARIS 1 are two major paths to move the power from the north to central Texas, where the load centers are located in the synthetic 2000-bus system. One additional branch named "Branch ‘2113’ ‘8127’ ‘1’" feeds into Substation PAIRS from east Texas. During normal operating conditions, most of the power generated by the three units at Substation SAVOY are transmitted out by the 500 kV network either via Branch ‘2002’ ‘2101’ ‘1’ and Branch ‘2021’ ‘2002’ ‘1’. Those two transmission lines are also critical paths to move the power either from west to east (Substation SHERMAN 1 to Substation PARIS 1), or east to west (Substation PARIS 1 to Substation SHERMAN 1), so that the generation in this area can be transported to the rest of the system.

However, if both Branch ‘2002’ ‘2101’ ‘1’ and Branch ‘2021’ ‘2002’ ‘1’ are out of service at the same time, Substation SAVOY is isolated from the rest of the 500 kV network. All the power generated at Bus 2002 have to be taken out by the 161 kV network instead. Under certain operating

conditions, this can cause severe violation on Branch '2038' '2003' '1',Branch '2081' '2038' '1', and Branch '2081' '2129' '1'. The pattern of operating conditions subject to severe overloading violations after the 500 kV network double line outage is identified using the 2-stage linear SVM method. With 11206 features in total representing the branch flows, bus voltages, and nodal power injections to start with, the SVM algorithm selects three features, Branch '2101' '5082' "' MVA flow ($\boldsymbol{X_1}$), Branch '5049' '2113' '1' MVA flow ($\boldsymbol{X_2}$), and Bus '2006' MWInj ($\boldsymbol{X_3}$) to form the hyperplane $\mathbf{0.265}\boldsymbol{X_1} - \mathbf{0.004}\boldsymbol{X_2} + \mathbf{0.034}\boldsymbol{X_3} = \mathbf{18.311}$ that separates the scenarios of different classes, shown in Figure 9.9. In this case, the scenarios in the three-dimensional feature space is linearly separable. The accuracy rate of model is 100% on the testing data set, and the F1 score is 1.

The arming condition for ACTIVSg2000 RAS #1 is modeled as a mathematical expression in PowerWorld, where the value of the SVM features are monitored during the real-time operations. Whenever the value of the features' linear combination is greatly than $\mathbf{18.311}$, this remedial action is armed and ready to protect the 161 kV transmission facilities. The triggering condition of this RAS is set to be the status of Branch '2002' '2101' '1' and Branch '2021' '2002' '1'. If both lines are detected to be offline at the same time, corrective actions would be initiated.

The initial corrective action priority list for an example test case is shown in Table 9.3. For the determination of corrective actions, three generating units have the priority with significant sensitivity and capacity to adjust. Those three generators at Substation SAVOY each have a capacity of 391.8 MW. The sensitivity of the overloading line to the generator MW output adjusting is 0.61, in other words, for 1 MW generation output reduced at one of the three units, the MW flow on the overloading lines will decrease by around 0.6 MW. In the example test case, in total those three generating units can provide a 359.6 MM of line flow reduction on the targeted severe violation element. Since those three generators are attached to the same higher voltage bus via the generator step-up transformers, the generator MW output adjustment can be implemented on the three generators at the same time, where the desired MW value reduction on the severe violation branch is split three ways and shared equally. For the corrective action of ACTIVSg2000 RAS #1, the

amount of total generation MW reduction at substation SAVOY is calculated as total MW needs to be reduced from the overloading lines divided by 0.6. If additional MW reduction is needed on the severe violation branch, the two loads at Substation ECTOR and DODD CITY both have sensitivities below 0.35 towards the three overloading lines, with the size of the loads are at 2.8 MW and 1.7 MW respectively in the example test case. Figure 9.10 illustrates the implementation of the remedial action scheme. On a computation system with an Intel Xeon Processor and 32 GM of RAM, the corrective action for this example can be determined in 4.6 seconds.

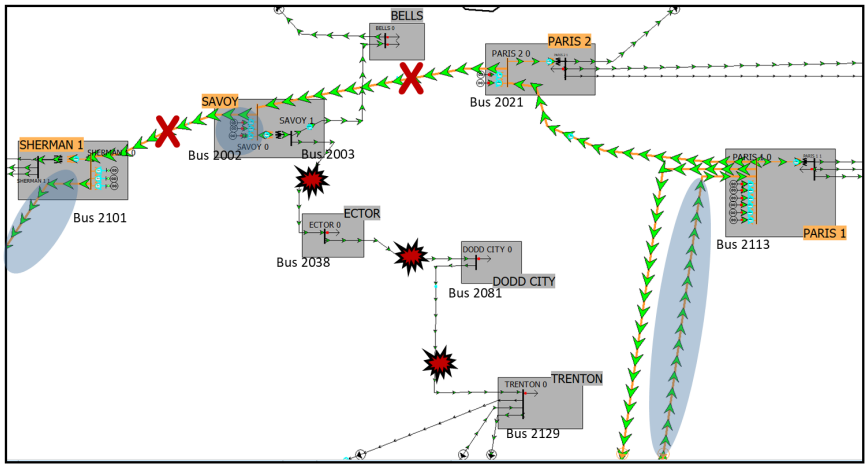

Figure 9.8: Detailed one-line for ACTIVSg2000 RAS #1

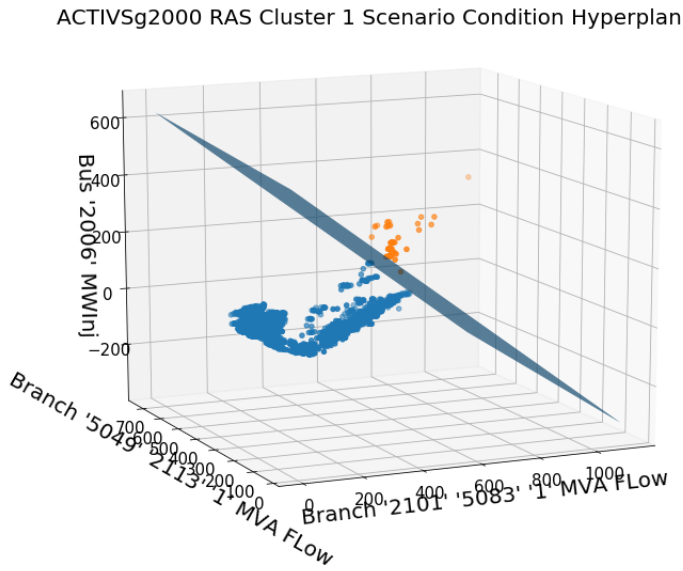

Figure 9.9: ACTIVSg2000 RAS #1 scenario separating hyperplane

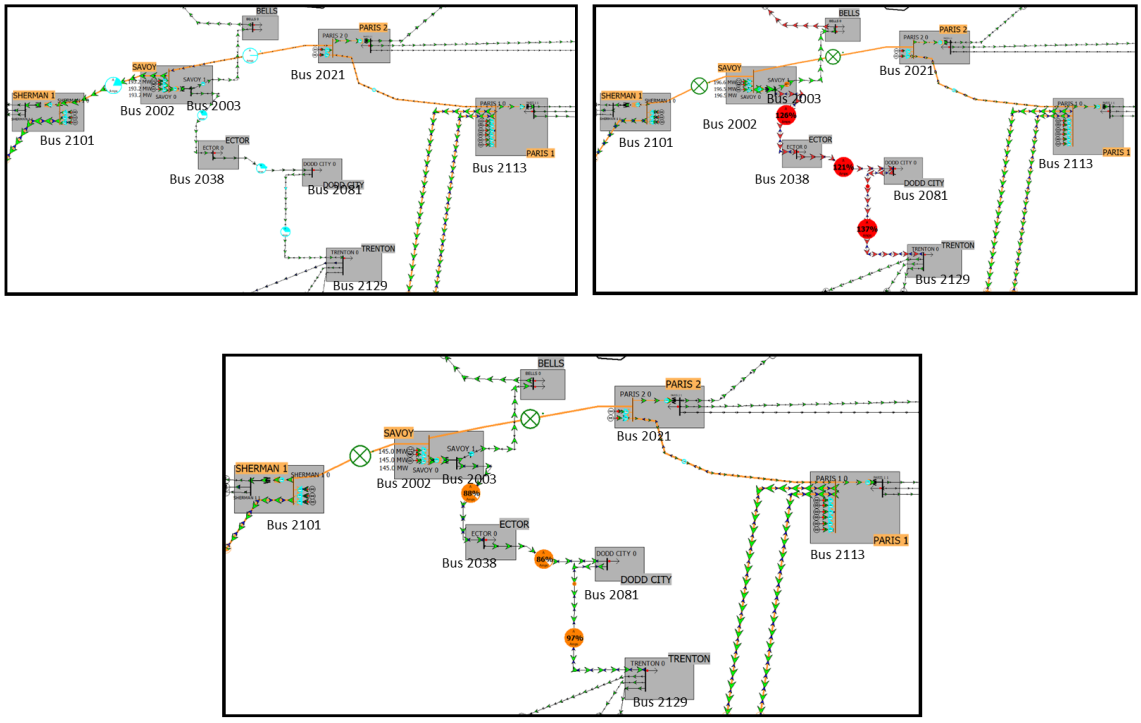

Figure 9.10: ACTIVSg2000 RAS #1 corrective action demonstration

Table 9.3: ACTIVSg2000 RAS #1 Initial Corrective Action Priority List

Element ID	Element Type	MW Value	Sensitivity	MW Reduction Capacity	Accumulative MW Reduction Capacity
Gen '2004' '1'	Gen	196.5	0.61	119.9	119.9
Gen '2005' '1'	Gen	196.5	0.61	119.9	239.7
Gen '2006' '1'	Gen	196.5	0.61	119.9	359.6
Load '2081' '1'	Load	2.8	0.24	0.7	360.6
Load '2038' '1'	Load	1.7	0.34	0.6	361.2

9.2.2 ACTIVSg2000 RAS #2

ACTIVSg2000 RAS # 2 focus on the 161 kV load serving network and protects the targeted element from severe violation of it emergency rating when certain lines from the same substation is offline at the same time. Figure 9.11 shows the detailed one-line diagram of the area of interest. This section of the 161 kV network contains four substations, Substation WHITEWRIGHT, VAN ALSTYNE, MELISSA, and CELINA. Each substation has electric load aggregation, and Substation CELINA and WHITEWRIGHT are interconnected with the rest of the system to bring powers into

this section of the grid. Without any contingency, the power flow through both Substation CELINA and WHITEWRIGHT feed the electricity demands at Substation MELISSA and VAN ALSTYNE. Under certain operating conditions, two contingencies can potentially cause severe overload on Branch '5450' '5331' '1', the transmission line between Substation CELINA and MELISSA, marked with an explosion sign in Figure 9.11. The first contingency is the opening of Branch '2124' '2106' '1'. This contingency cuts the connection between Substation WHITEWRIGHT and VAN ALSTYNE, so that the load located in Bus 2124 has to be supplied solely from Substation CELINA, via the targeted severe violation branch. The second contingency is a simultaneously line outage on both Branch '2020' '2106' '1' and Branch '2106' '2129' '1', the two branches connecting Bus 2106 in Substation WHITEWRIGHT to the rest of the system. With this contingency, the electric demands from Substation WHITEWRIGHT, VAN ALSTYNE and MELISSA have to be supplied from Substation CELINA alone, causing a great overload violation on Branch '5450' '5331' '1'.

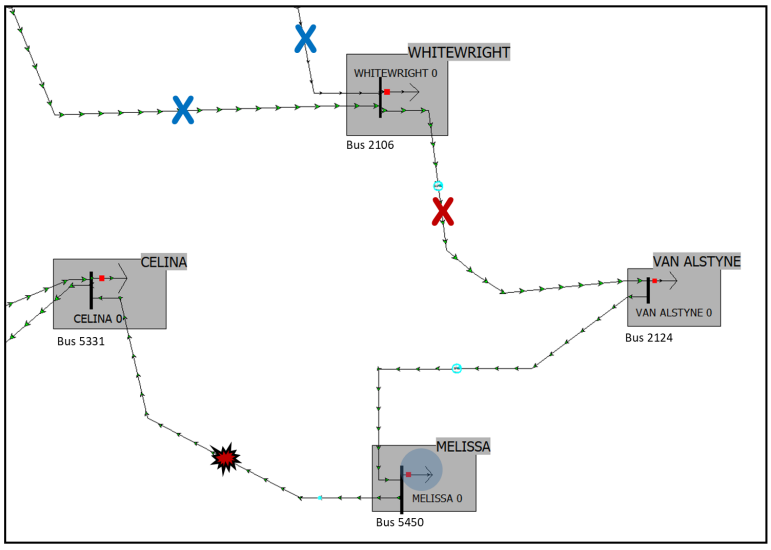

Figure 9.11: Detailed one-line for ACTIVSg2000 RAS # 2

To identify the patterns of system operation conditions, under which the severe overload can

happen with occurrence of one of the above mentioned contingencies, the 2-stage linear SVM method is utilized to select meaningful features and find the hyperplane that can best divide the scenarios of different classes. 3639 scenario samples are used in the training process, and the ratio of true to false data label is about 1 to 200. During the training process, it is observed that the scenarios can be linearly separable with only one feature, and can achieve accuracy of 100% and F1 score of 1. Shown in Figure 9.12, The single-feature SVM uses feature "Bus '5450' MWInj" to form the hyperplane.

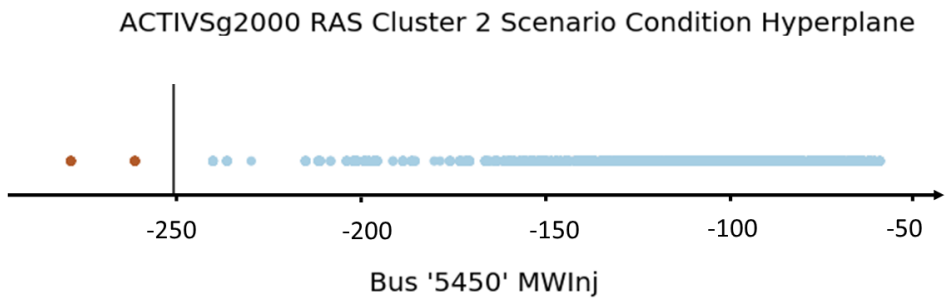

Figure 9.12: ACTIVSg2000 RAS #2 scenario separating hyperplane

If the load MW value at Bus 5450 exceeds 250 MW, ACTIVSg2000 RAS #2 considers it as a risky operating condition and is armed. With the detection of above mentioned contingencies in Substation WHITEWRIGHT, corrective actions can be automatically triggered. Figure 9.13 shows the implementation of the corrective action for ACTIVSg2000 RAS #2. Shown in Table 9.4, the load at Bus 5450 and Bus 2124 are the two controllable elements from the connectivity and sensitivity studies. Both elements have a sensitivity of 1 towards the severe violation element, Branch '5450' '5331' '1'. In this case, the load at Substation MELISSA has the priority to be chosen as part of the load shedding scheme to alleviate the overloading violation. On a computation system with an Intel Xeon Processor and 32 GM of RAM, the corrective action for this example can be determined in 3.2 seconds.

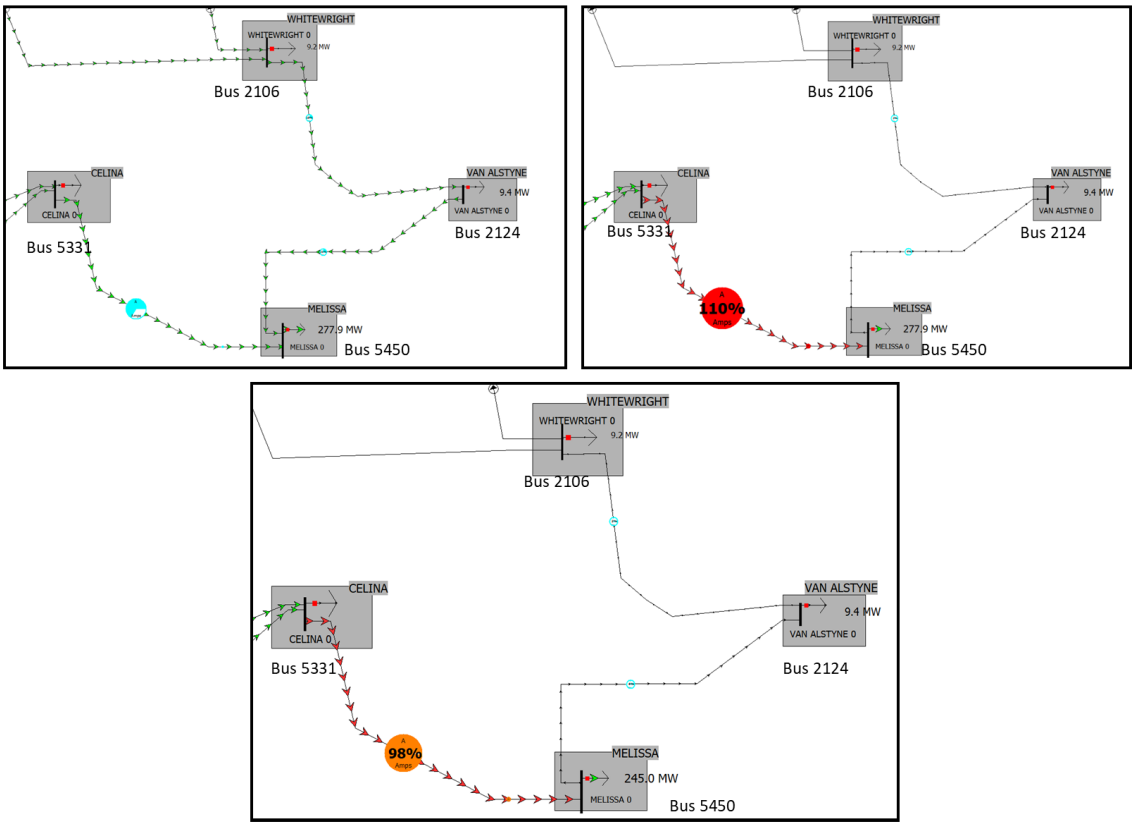

Figure 9.13: ACTIVSg2000 RAS #2 corrective action demonstration

Table 9.4: ACTIVSg2000 RAS #2 Initial Corrective Action Priority List

Element ID	Element Type	MW Value	Sensitivity	MW Reduction Capacity	Accumulative MW Reduction Capacity
Load '5450' '1'	Load	260.5	1	260.5	260.5
Load '2124' '1'	Load	11.8	1	11.8	272.3

9.2.3 ACTIVSg2000 RAS #3

Remedial action scheme #3 for the ACTIVSg2000 synthetic system targets the same section of the grid - the panhandle area on the footprint of Texas - where most of the wind generation in the system is located at. Figure 9.14 shows the topology of the network that is associated with this RAS. There are eight key substations. Substation PANHANDLE 1, PANHANDLE 3, PANHANDLE 5, WHITE DEER 0, and WHEELER are the five 161 kV substations that are interconnected

by the meshed 161 kV network. The first four 161 kV substations have bulk wind generation attached to them, with generation capacities ranging from 10 MW to more than 180 MW. Substation MIAMI, PANHANDLE 2 and PANHANDLE 4 are three 500 kV substations. Within each substation, there is one bulk-level wind generator with capacity ranging from 200 to 300 MW, and a 500/161 kV transformer connecting the 161 kV network to the higher voltage level. If this area of interest is considered as an island, there are two paths connecting to the rest of the system. The main path is consist of the two 500 kV transmission lines starting from Substation MIAMI. Most of the power generated in this area is out by those two circuits so that the wind power can be transported via high voltage transmission facilities to the rest of the system where load centers are located at. The second path is a 161 kV branch originates in Substation PANHANDLE 2, this transmission branch connects this area to the rest of the 161 kV network of the whole system, and can carry a small portion of wind generation out.

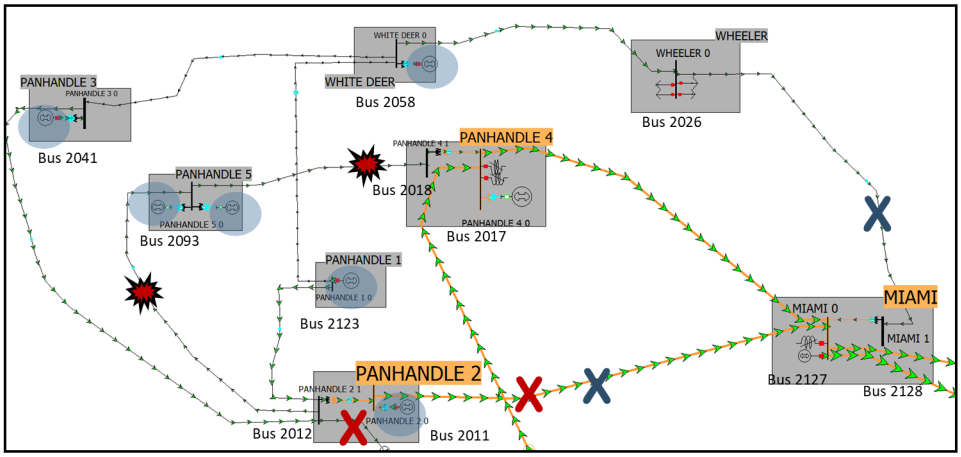

Figure 9.14: Detailed one-line for ACTIVSg2000 RAS #3

During the scenario contingency analysis, it is observed that two contingencies can cause severe violations of emergency ratings for Branch '2093' '2012' '1' and Branch '2018' '2093' '1' under certain operating conditions. The first contingency is the outage on both Branch '2012' '2052' '1' and Branch '2127' '2011' '1' simultaneously. The second contingency contains the outage on

both Branch ‘2127’ ‘2011’ ‘1’ and Branch ‘2128’ ‘2026’ ‘1’. The severe violation elements are marked as explosion signs, and the branches associated with the critical contingencies are marked as blue and red crosses in Figure 9.14. Both contingencies cut off some connections to the rest of the system in the area of focuses, so that the power flow is redirected to use Branch ‘2093’ ‘2012’ ‘1’ and Branch ‘2018’ ‘2093’ ‘1’ and cause excessive flows on those lines.

To further study the operating conditions under which the two contingencies are more likely to cause severe overload, the data driven approach is taken to identify a hyperplane that can optimally divide the scenarios with different labels. During the training process, two features, “Bus ‘2123’ MWInj”, and “Bus ‘2013’ MWIn”, are identified to achieve the best performance with 96.15% accuracy rate and 0.94 F1 score. It is worth noting that since the scenarios are not 100% linearly separable, a penalty factor is applied to the SVM model to prioritize the accuracy rate on the True data set, as it’s more important to identify the risky operating conditions correctly. The first feature, “Bus ‘2123’ MWInj”, is the MW generation in Substation PANHANDLE 1, while the second feature “Bus ‘2013’ MWInj” is the MW generation in Substation PANHANDLE 2. Since the area of interest for ACTIVSg2000 RAS #3 contains bulk wind generation and has constraint transmission facility, the summation of all the MW generation in Substation PANHANDLE 1, PANHANDLE 2, PANHANDLE 3, PANHANDLE 5, and WHITE DEER 0 is used as one single feature to train one separable SVM model. This model is able to achieve an accuracy rate of 99.39% and F1 score of 0.9913. The visualizations for the hyperplanes of both SVM models are shown in Figure 9.12.

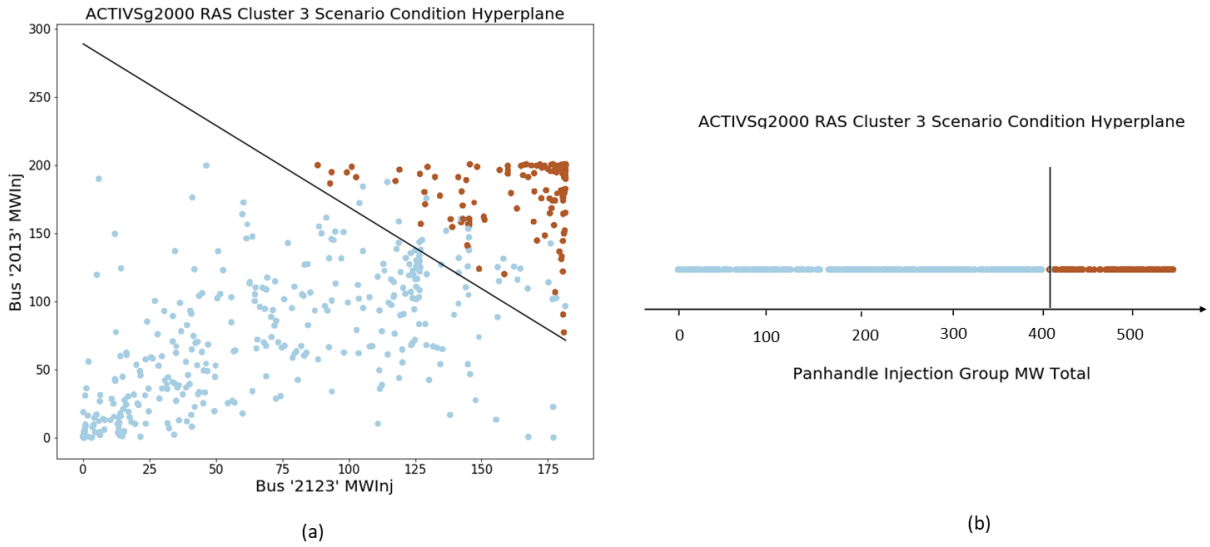

Figure 9.15: ACTIVSg2000 RAS #3 scenario separating hyperplane with (a) two Features (b) one aggregated feature

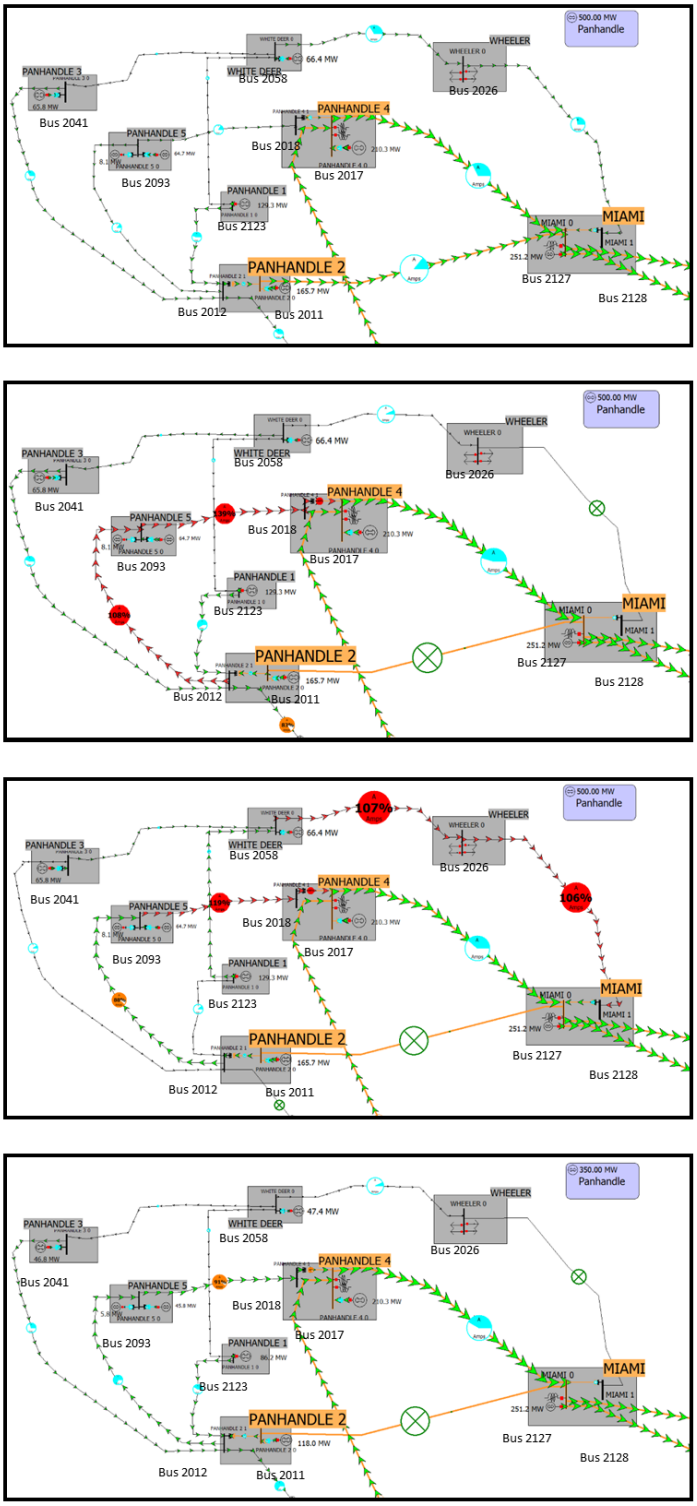

Figure 9.16: ACTIVSg2000 RAS #3 corrective action demonstration

With all things considered, panhandle injection group MW value is monitored and set as the the arming condition of ACTIVSg2000 RAS #3. If the total generation from the above mentioned substation group exceeds 410 MW, this remedial actions scheme is armed. A demonstration of ACTIVSg2000 RAS #3 is shown in Figure 9.16, the first one-line diagram shows the system operating condition before any contingency, the second and third diagram include the two contingencies that can cause the severe violation elements to overload. The last one-line diagram shows the system operating condition post-contingency with the modeling of ACTIVSg2000 RAS #3. The panhandle injection group MW value is shown in the blue box at the top right corner of each one-line diagram. The triggering condition of this remedial action scheme is the detection of any of the two contingencies. As for corrective actions, the six generators in the panhandle injection group also show up at the top of the priority list in the example test case shown in Table 9.5. Since the high generation level with an occurrence of a critical contingency is the direct cause of the overloading violation, instead of finding the sensitivity value at each of the generators in the area, the generation is reduced as a whole from the defined injection group based on a predefined participation factor. In this case, the participation factor is proportional to the maximum MW capacity of each unit. On a computation system with an Intel Xeon Processor and 32 GM of RAM, the corrective action for this example can be determined in 3.4 seconds.

Table 9.5: ACTIVSg2000 RAS #3 Initial Corrective Action Priority List

Element ID	Element Type	MW Value	Sensitivity	MW Reduction Capacity	Accumulative MW Reduction Capacity
Gen '2013' '1'	Gen	199.4	0.59	117.6	117.6
Gen '2123' '1'	Gen	159.9	0.53	84.7	202.3
Gen '2095' '1'	Gen	78.1	0.75	58.6	260.9
Gen '2042' '1'	Gen	79.2	0.54	42.8	303.7
Gen '2059' '1'	Gen	79.8	0.50	39.9	343.6
Gen '2094' '1'	Gen	9.8	0.75	7.4	351.0

9.2.4 ACTIVSg2000 RAS #4

The part of the synthetic system that is associated with ACTIVSg2000 RAS #4 is identical to that with ACTIVSg2000 RAS #3. The severe violation elements in this RAS, Branch '2058' '2026' '1', and Branch '2128' '2026' '1', originally belongs to the same RAS cluster when only the system connectivity is considered. This RAS cluster is split into two because the two sets of severe violation elements respond differently for the list of contingencies. There are four contingencies that can cause severe overloading violations on Branch '2058' '2026' '1' and Branch '2128' '2026' '1'. The first contingency is the The first contingency is the outage on both Branch '2012' '2052' '1' and Branch '2127' '2011' '1' simultaneously, marked with yellow cross in Figure 9.17. The second contingency is the outage of two 500 kV transmission lines, Branch '2127' '2011' '1' and Branch '2017' '2127' '1', marked with blue crosses. The third contingency contains Branch '2127' '2011' '1' and Branch '2093' '2012' '1', marked with green crosses. The last contingency contains Branch '2012' '2041' '1' and Branch '2012' '2123' '1', marked with red crosses in Figure 9.17. All the contingencies will cut off some connections between the wind generation substations and the rest of the system, so that the power flow is rerouted to use Branch '2058' '2026' '1', and Branch '2128' '2026' '1' more extensively and cause potential violations.

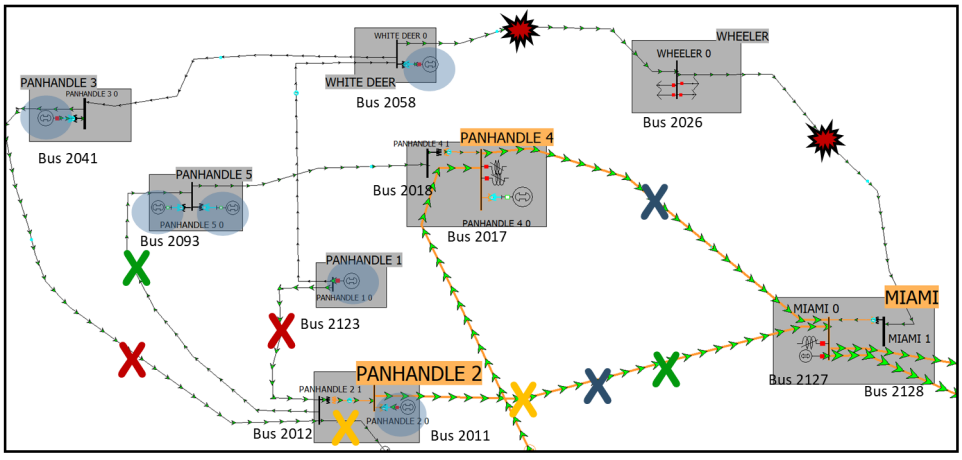

Figure 9.17: Detailed one-line for ACTIVSg2000 RAS #4

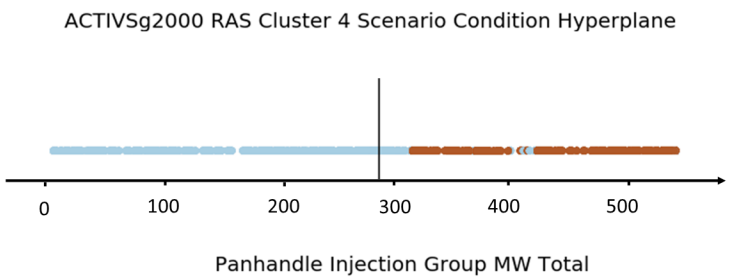

Figure 9.18: ACTIVSg2000 RAS #4 scenario separating hyperplane

Similar to ACTIVSg2000 RAS #3, an injection group containing all the generators from Substation PANHANDLE 1, PANHANDLE 2, PANHANDLE 3, PANHANDLE 5, and WHITE DEER 0 is defined. The linear SVM model is trained with one single feature, the total MW generation from the panhandle injection group. This model is able to reach an accuracy rate of 94.87%, and F1 score of 0.9315. Since the scenarios are not completely linearly separable, a penalty factor is also implemented during the training process to put more emphasize on predicting the True labeled data correctly. The cut off value learned by the SVM model is 290 MW. If the total generation from the injection group exceeds this value, ACTIVSg2000 RAS #4 is armed. A demonstration of ACTIVSg2000 RAS #4 is shown in Figure 9.19, the first one-line diagram shows the system operating condition before any contingency, the second to fifth diagram include the four contingencies that can cause the severe violation elements to overload. The last one-line diagram shows the system operating condition post-contingency with the modeling of ACTIVSg2000 RAS #4. The panhandle injection group MW value is shown in the blue box at the top right corner of each one-line diagram. The triggering condition of this remedial action scheme is the detection of any of the four contingencies. Due to its proximity in the network, ACTIVSg2000 RAS 4 has an initial corrective action priority list that is almost identical to that of ACTIVSg2000 RAS 3, shown in Table 9.6. Since the six generators in the injection group are less sensitive to the severe violation elements protected by ACTIVSg2000 RAS 4, four additional electric loads on bus 2026 are also included in this priority list. On a computation system with an Intel Xeon Processor and 32 GM of

RAM, the corrective action for this example can be determined in 3.5 seconds.

Table 9.6: ACTIVSg2000 RAS #4 Initial Corrective Action Priority List

Element ID	Element Type	MW Value	Sensitivity	MW Reduction Capacity	Accumulative MW Reduction Capacity
Gen '2013' '1'	Gen	199.4	0.32	63.8	63.8
Gen '2123' '1'	Gen	159.9	0.37	59.2	123.0
Gen '2059' '1'	Gen	79.8	0.41	32.7	155.7
Gen '2042' '1'	Gen	79.2	0.36	28.5	184.2
Gen '2095' '1'	Gen	78.1	0.23	18.0	202.2
Gen '2094' '1'	Gen	9.8	0.23	2.3	204.5
Load '2026' '1'	Load	6.2	0.76	4.7	209.2
Load '2026' '2'	Load	5.6	0.76	4.3	213.5
Load '2026' '3'	Load	2.0	0.76	1.5	215.0
Load '2026' '4'	Load	0.7	0.76	0.5	215.5

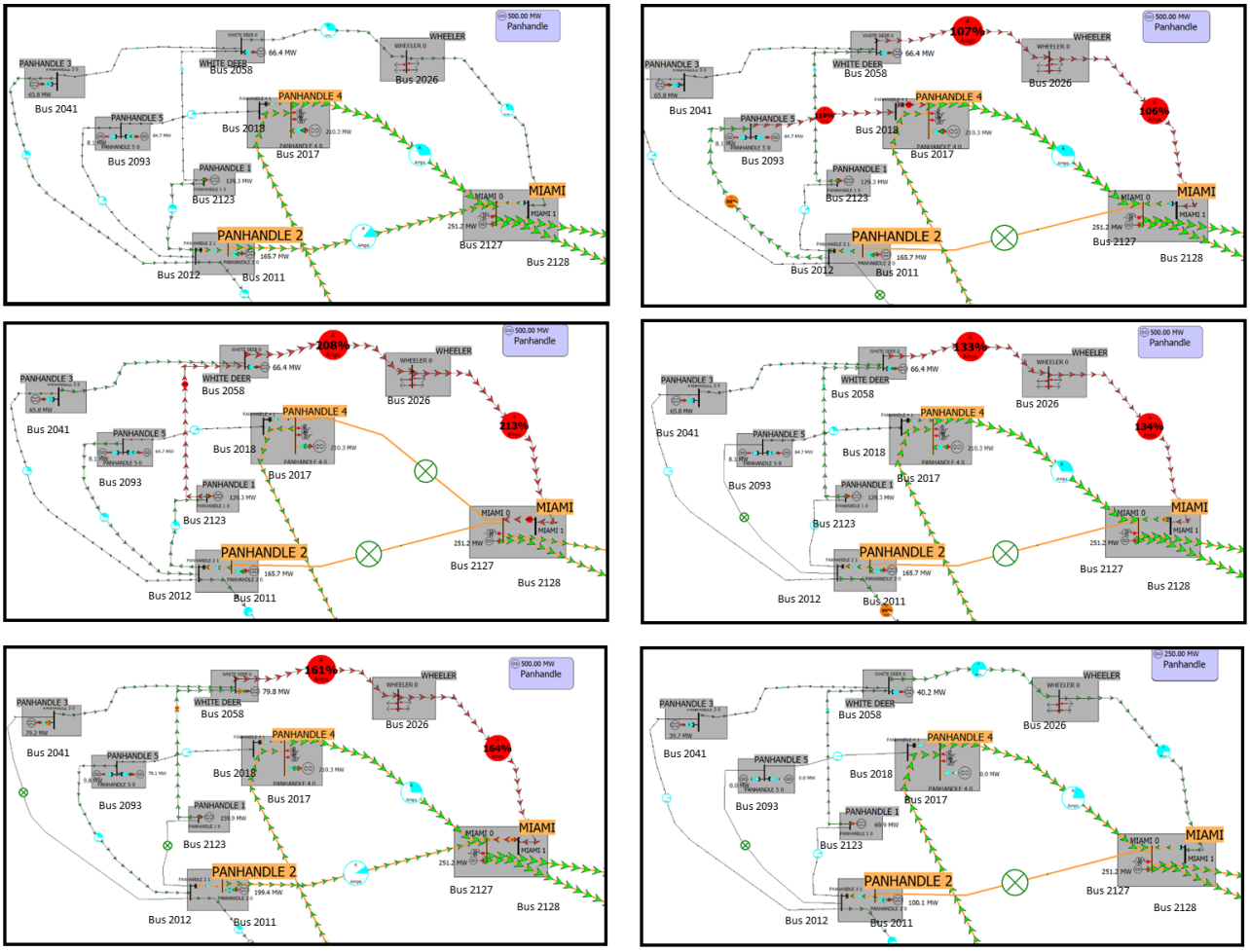

Figure 9.19: ACTIVSg2000 RAS #4 corrective action demonstration

9.2.5 ACTIVSg2000 RAS #5

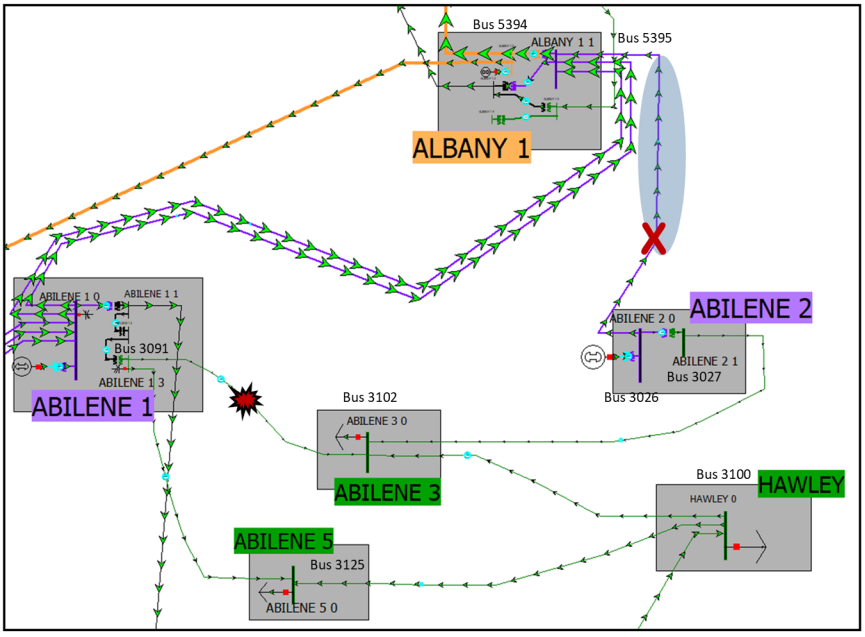

Figure 9.20: detailed one-line for ACTIVSg2000 RAS #5

There are six substations associated with ACTIVSg2000 RAS #5, shown in Figure 9.20. Substation ABILENE 3, ABILENE5, and HAWLEY are three 115 kV substations with load buses. They are interconnected with two 230 kV substations, ABILENE 1 and ABILENE 2. Substation ABILENE 2 contains one single 230/115 kV transformer, while there are three voltage levels at Substation ABILENE 1, 230 kV, 161 kV,and 115 kV. Both 230 kV substations have one generator with capacity of 200 MW. During normal operating conditions, those power generations are transmitted either through the lower voltage level at 161 kV and 115 kV, or high voltage level at 500 kV via Substation ALBANY 1, which both Substation ABILENE 1 and ABILENE 2 have connections to. If Branch '2036' '5395' '1' is out of service, the path between Substation ABILENE 2 and the 500 kV network is cut off. All the power generated from Substation ABILENE 2 now have to rely on the 115 kV network to transmit the power either into the 115 kV load serving circuits, or into the

500 kV network via Substation ABILENE 1 and ALBANY 1. Under certain operating conditions, this contingency can cause Branch ‘3102’ ‘3091’ ‘1’ to operate exceeding its emergency rating.

To identify the patterns of operational scenarios, under which the exceedance of emergency rating for Branch ‘3102’ ‘3091’ ‘1’ can happen, the 2-stage linear SVM method is applied to firstly identify a subset of scenario features important for the classification. Different combinations of those candidate features are then utilized for the training of the second stage linear SVM, where the model with highest performance metrics value is selected. For ACTIVSg2000 RAS #5, a single feature linear SVM can achieve the best performance with an accuracy rate of 96.77%, and a F1 score of 0.94. The feature associated with this SVM model is Branch ‘3026’ ‘5395’ ‘1’ MVA flow. A visualization of the hyperplane identified by this SVM model is shown in Figure 9.21.

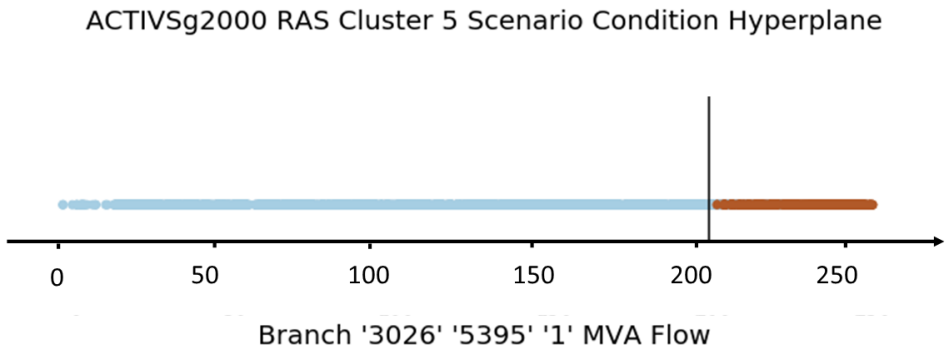

Figure 9.21: ACTIVSg2000 RAS #5 scenario separating hyperplane

The arming condition of ACTIVSg2000 RAS #5 monitors the value of the SVM feature during real-time operations. If the power flow on Branch ‘3026’ ‘5395’ ‘1’ exceeds 207 MVA, the remedial action scheme is armed and ready to protect the 115 kV transmission facilities. The triggering condition of this particular RAS is the detection of offline status for Branch ‘3026’ ‘5395’ ‘1’. The occurrence of this contingency will initiate corrective actions to alleviate the branch violation. The corrective action for ACTIVSg2000 RAS #5 considers four controllable elements from the

surround area. The first controllable element is the generator located in Substation ABILENE 2, and it has a sensitivity of 0.65 to the severe violation element. The second controllable element is Branch '3100' '3102' '1', the 115 kV transmission line between Substation ABILENE 3 and HAWLEY. This flow on this branch, along with the flow between Substation ABILENE 2 and Substation ABILENE 3, contribute to the total MVA flow on the severe violation element, and has a sensitivity factor of 1. The load on Bus 3102 and 3100 are the last two controllable elements, they have a sensitivity factor of 0.74 to 0.26 respectively. Figure 9.22 demonstrates the corrective action implementation using the first controllable element. On a computation system with an Intel Xeon Processor and 32 GM of RAM, the corrective action for this example can be determined in 4.5 seconds.

Table 9.7: ACTIVSg2000 RAS #5 Initial Corrective Action Priority List

Element ID	Element Type	MW Value	Sensitivity	MW Reduction Capacity	Accumulative MW Reduction Capacity
Gen '3028' '1'	Gen	197.8	0.65	128.6	128.6
Branch '3100' '3102' '1'	Branch	34.5	1	34.5	163.1
Load '3102' '1'	Load	26.3	0.74	19.5	182.6
Load '3100' '1'	Load	2.3	0.26	0.6	183.2

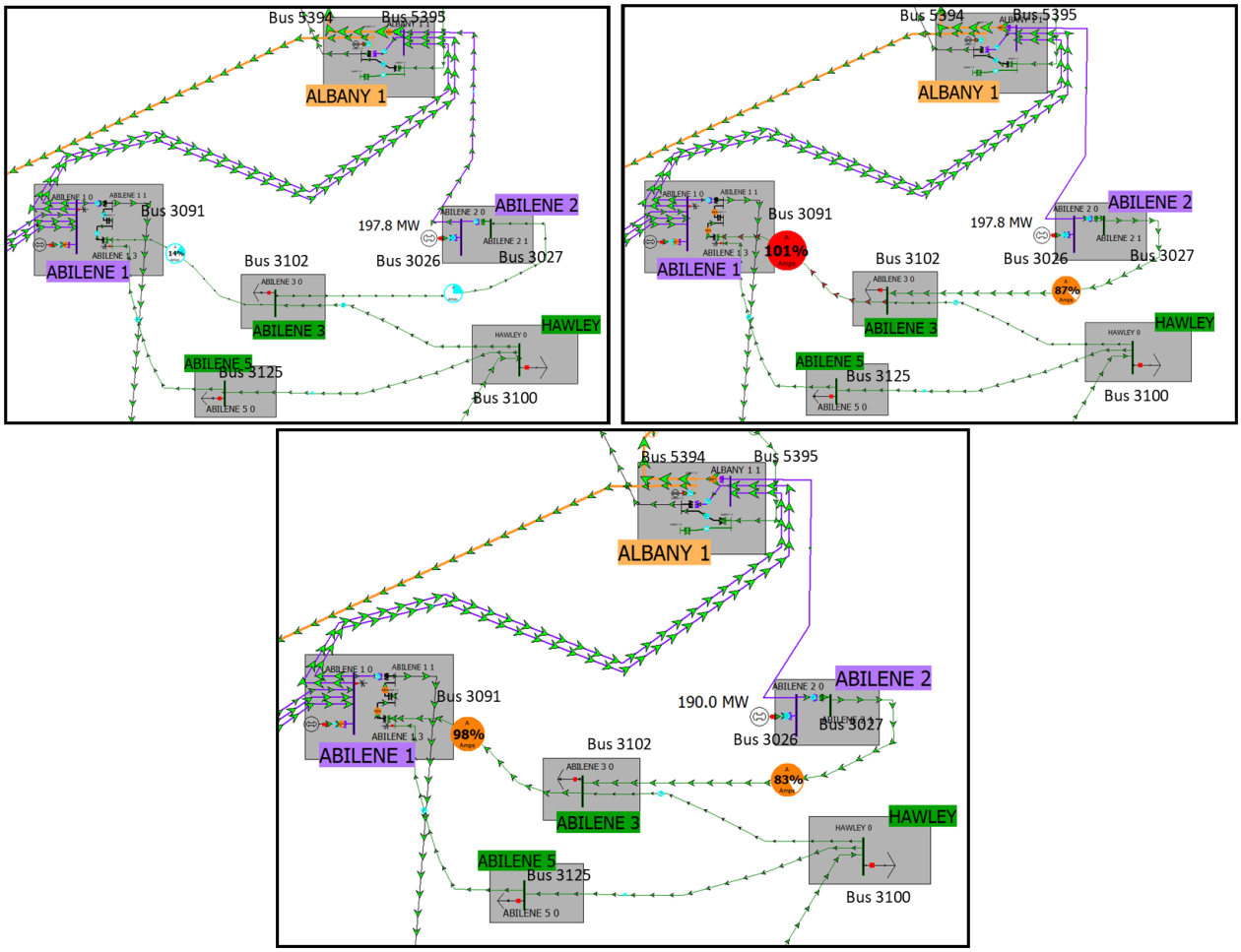

Figure 9.22: ACTIVSg2000 RAS #5 corrective action demonstration

9.2.6 ACTIVSg2000 RAS #6

ACTIVSg2000 RAS #6 protects the severe violation element Branch '1032' '1033' '1'. This remedial action scheme is geographically located in the far west area in Texas, and contains seven substations. Substation MACCAMEY 1 is the 230 kV substation that includes Bus 1032 and Bus 1033. Bus 1032 has a nominal voltage of 115 kV. It is interconnected with the rest of the 115 kV network associated with this remedial action scheme, and serves as the access point to the higher voltage grid via the severe violation element, the transformer between Bus 1032 and 1033. Bus 1033 has a nominal voltage of 230 kV, and has a 278 MW generator directly attached. It has two 230 kV transmission lines, Branch '1033' '1081' '1' and Branch '1033' '3050' '1', connecting to the rest of the 230 kV network in this synthetic system. Substation MCCAMEY 2, IRAAN 1, IRAAN 2, CRANE, FORT STOCKTON 1, and FORT STOCKTON 2 are the 115 kV substations in this remedial action scheme. All the 115 kV substations besides Substation CRANE have wind? generations, with capacities ranging from 75 to 160.5 MW. This group of generators are connected through the 115 kV network and the generation is ultimately taken out by the 230 kV transmission lines from Substation MACCAMEY 1. The one-line diagram of this remedial action scheme is in Figure 9.23.

There are two contingencies that can potentially cause the severe violation of emergency MVA rating of Branch '1032' '1033' '1'. The first contingency is the simultaneously outage of the two 230 kV transmission lines, Branch '1033' '1081' '1' and Branch '1033' '3050' '1'. Without those two higher voltage level transmission branches, all the generations in the area, including generation from Bus 1033, are reliant on the 115 kV network. Usually the flow direction on Branch '1032' '1033' '1' is from the 115 kV bus to the 230 kV bus, with the first contingency, the flow direction is reversed. Under certain operating conditions, this reverse flow will exceed the emergency MVA rating of the transformer.

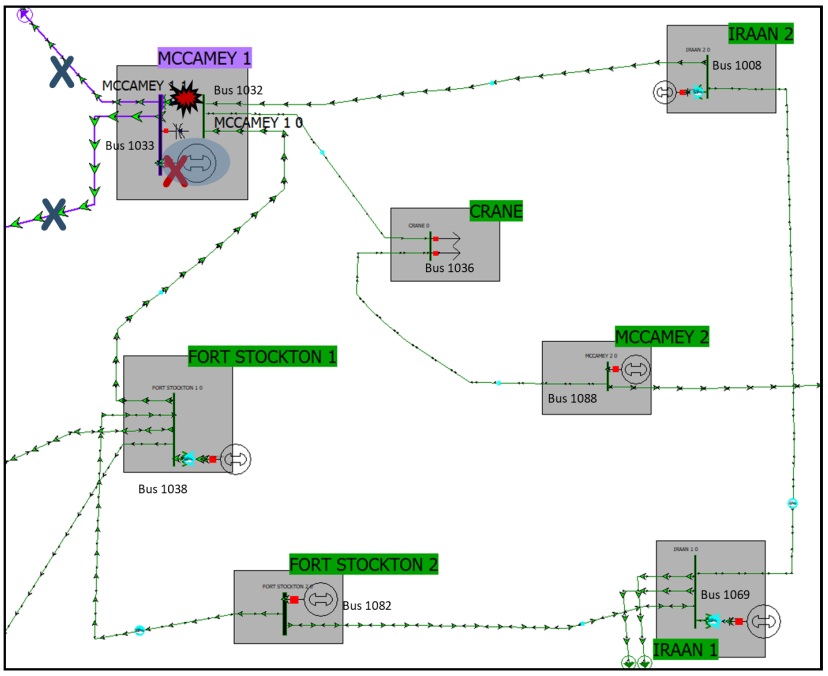

Figure 9.23: Detailed one-line for ACTIVSg2000 RAS #6

The second contingency is the outage of generator at Bus 1033. With the loss of generation in this area, other generations in the synthetic system are adjusting their MW output from the AGC response to make up for the loss and ensure the balance of electricity demand and supply. This increase of generation from the five 115 kV substations are reliant on the transformer in Substation MACCAMEY 1 for transmission, and under certain operating conditions, this can cause the overload of the transformer Branch '1032' '1033' '1'.

Since the direction of flow on the severe violation element is different for the two critical contingencies, the scenarios under which the violation occur is reviewed to determine if one or two separate remedial action scheme arming and triggering conditions are needed. It is served that the scenarios causing violations with the occurrence of the two contingencies are identical, thus only one SVM model is trained to obtain the RAS arming conditions. Figure 9.24 shows the separating hyperplane for ACTIVSg2000 RAS #6. In this SVM model, feature "Bus '1033' MWInj" is selected and the cut off value is identified to be 210 MW. This single feature SVM model can

achieve an accuracy rate of 98.70% and a F1 score of 0.96. With the identification of the separating hyperplane, the arming condition of ACTIVSg2000 RAS #6 monitors the generation level at Bus 1033. Two sets of corrective actions are initiated for two triggering conditions respectively. The first triggering condition monitors the status of Branch '1033' '1081' '1' and Branch '1033' '3050' '1'. If both 230 kV transmission lines are out of service at the same time, the generation at Bus 1033 is reduced. The second triggering condition monitors the status of generator at Bus 1033. If an outage occurs for this generator, the generation MW value for units from Substation IRAAN 1 and IRAAN 2 are reduced. On a computation system with an Intel Xeon Processor and 32 GM of RAM, the corrective action for this example can be determined in 3.7 seconds.

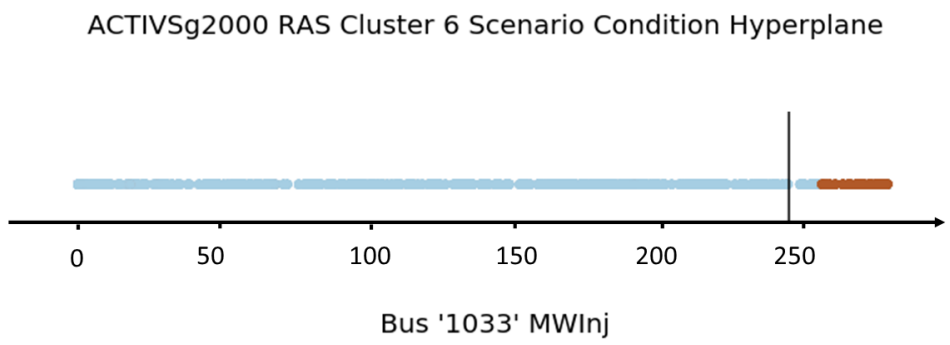

Figure 9.24: ACTIVSg2000 RAS #6 scenario separating hyperplane

Table 9.8: ACTIVSg2000 RAS #6 Triggering Condition 1 Initial Corrective Action Priority List

Element ID	Element Type	MW Value	Sensitivity	MW Reduction Capacity	Accumulative MW Reduction Capacity
Gen '1033' '1'	Gen	272.1	0.98	266.6	266.6

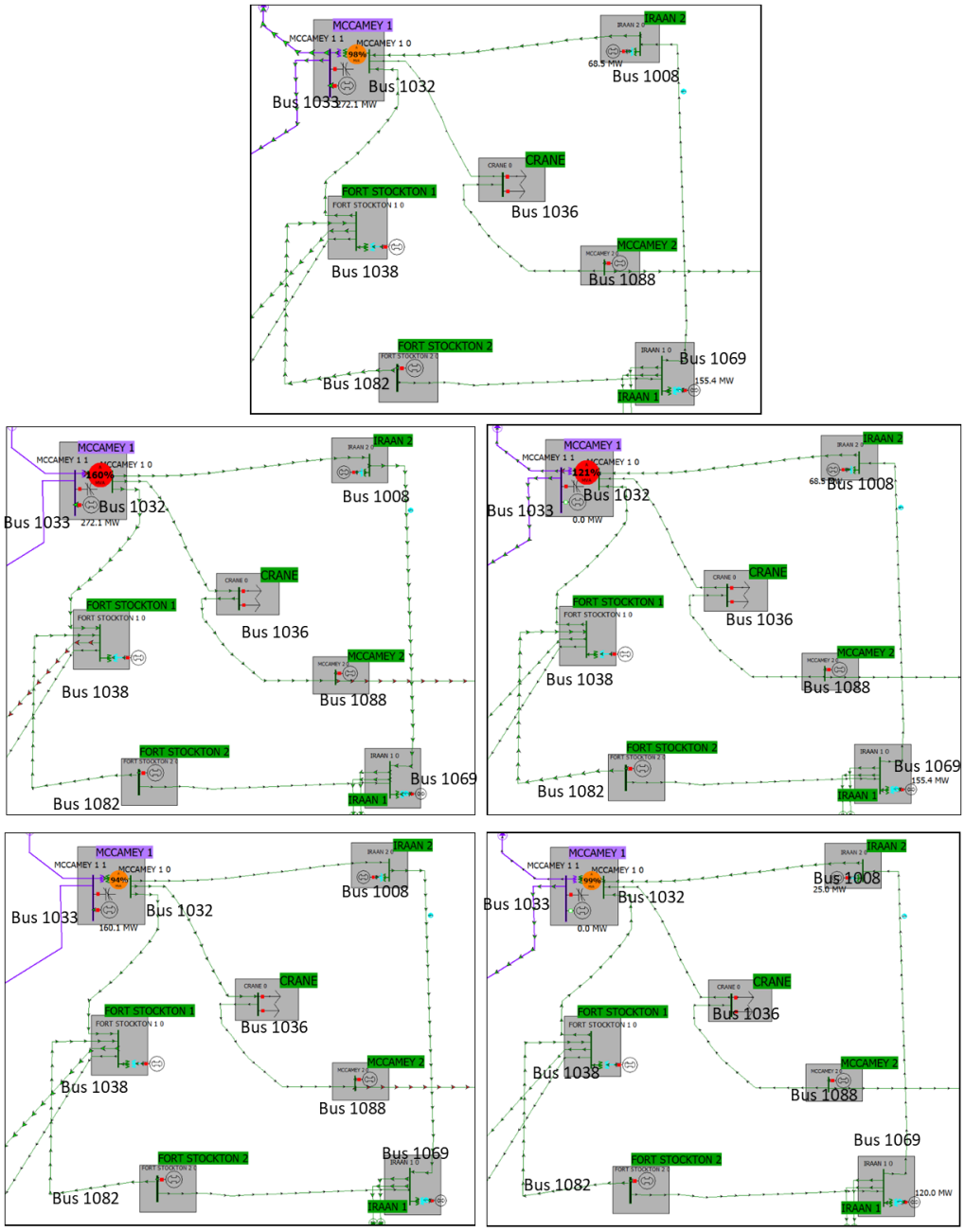

Figure 9.25: ACTIVSg2000 RAS #6 corrective action demonstration

Table 9.9: ACTIVSg2000 RAS #6 Triggering Condition 2 Initial Corrective Action Priority List

Element ID	Element Type	MW Value	Sensitivity	MW Reduction Capacity	Accumulative MW Reduction Capacity
Gen '1070' '1'	Gen	155.4	0.40	62.0	62.0
Gen '1009' '1'	Gen	68.5	0.45	30.8	92.8
Gen '1088' '1'	Gen	64.1	0.45	28.8	121.6
Gen '1039' '1'	Gen	65.9	0.43	28.3	149.9
Gen '1082' '1'	Gen	44.0	0.41	18.0	167.9

10. CONCLUSIONS AND FUTURE DIRECTIONS

10.1 Conclusions

Traditional remedial action scheme design takes a holistic approach that requires years of expertise in a specific power system. Building on the industrys current practice, this dissertation develops a power system tool called Auto-RAS that provides a systematic approach to create remedial action schemes in a robust, effective, and automated manner. Leveraging data-driven techniques, responses and control practices of the current power system are contextualized as statistical characteristics and mathematical expressions to guide the design of remedial action schemes.

There are two main components in the Auto-RAS framework, the determination of RAS condition logics, and the creation of RAS corrective actions. Utilizing chronological simulation of scenarios and the operational condition analysis, the system's need of a new RAS, and its corresponding logic conditions are identified in the power system planning time frame. In the Auto-RAS condition determination process, hierarchical clustering is implemented to group different severe violation elements into RAS clusters. For each RAS cluster, a two-stage linear SVM algorithm is implemented to selected features that can best represent the operational scenarios, and learns a hyperplane that can optimally divide the scenarios with and without risks of severe violations. The selected scenario features, along with the learned hyperplane, and list of violation-causing contingencies are then leveraged as Auto-RAS condition logic.

For the creation of Auto-RAS corrective actions, this dissertation develops a sensitivity-based methodology to identify and deploy corrective actions adaptively for each RAS cluster during real-time operation time frame. Leveraging network connectivity analysis, a subset of power system elements that can be controlled as part of the corrective action scheme is selected. Sensitivity analysis such as line outage distribution factor and transmission loading relief are used to quickly determine the most effective controllable elements and the corresponding corrective actions to address specific operational violations in one RAS cluster.

To maintain a similar level of size and modeling complexity as the real power system while still be share the research result publicly and freely, a 200 bus synthetic system, ACTIVSg200, and a 2000 bus synthetic system, ACTIVSg2000, are used as test cases in this work for the development and testing of Auto-RAS. For the ACTIVSg200 synthetic grid model, two remedial action schemes are developed to protect a 115 kV transmission line, and a 230/115 kV transformer using load shedding and branch switching corrective actions. In the ACTIVSg2000 synthetic system, six remedial action schemes are developed for three different control areas. The voltage levels of the protected elements range from 115 kV to 230 kV. Generator output adjustment, branch switching, and load shedding actions are three common corrective actions developed for RAS in ACTIVSg2000 synthetic system. All the designed remedial action scheme are tested to ensure that they can operate to perform their intended functionalities, and do not introduce unintentional or unacceptable reliability risks to the bulk electric power system.

10.2 Future Directions

The development of Auto-RAS provides an initial paths to explore how RAS design can be improved to satisfy the need of future power system planning and operations. Leveraging machine learning techniques and existing sensitivity analysis practices, a systematic way of gaining system knowledge and experience, and making more up-to-date engineering decisions in both planning and real-time operation time frame are discussed.

This dissertation primarily focus one type of operational issue, the severe violation of branch emergency limits, to prove the concept of automating the remedial action scheme design. Future work can expand the scope to address other operational problems such as voltage violations, and dynamic instabilities. In addition, future directions can also consider the development of remedial action schemes to protect power systems from volatile disturbances such as electromagnetic pulses (EMPs), extreme weather, and malicious events threatening national security. Interdisciplinary studies such as coupled-infrastructure modeling, and cyber-physical system simulations can further enhance the resilience of power system with remedial action scheme implementations.

REFERENCES

[1] W. A. Wulf, "Great achievements and grand challenges," *National Academy of Engineering, The Bridge*, vol. 30, no. 3, p. 4, 2000.

[2] A. J. Wood, B. F. Wollenberg, and G. B. Sheblé, *Power generation, operation, and control*. John Wiley & Sons, 2013.

[3] Energy Information Administration, "Natural Gas," 2021. [Online]. Available: https://www.eia.gov/naturalgas/

[4] ——, "Coal," 2021. [Online]. Available: https://www.eia.gov/coal/

[5] ——, "Electricity," 2021. [Online]. Available: https://www.eia.gov/electricity/

[6] D. Gielen, D. Saygin, and N. Wagner, "Renewable energy prospects: United States of America, remap 2030 analysis," *IRENA, Abu Dhabi, Tech. Rep*, 2015.

[7] J. H. Eto, "Building electric transmission lines: A review of recent transmission projects," 2016.

[8] "Grain Belt Express," 2021. [Online]. Available: https://grainbeltexpress.com/

[9] "CREZ Fact Sheet ," 2021. [Online]. Available: https://poweringtexas.com/wp-content/uploads/2018/12/Transmission-and-CREZ-Fact-Sheet.pdf

[10] R. D. Farmer, G. Cohen, and D. Zimmerman, *Causes and lessons of the California electricity crisis*. CBO, 2001.

[11] W. M. Warwick, "A primer on electric utilities, deregulation, and restructuring of US electricity markets," Pacific Northwest National Lab.(PNNL), Richland, WA (United States), Tech. Rep., 2002.

[12] C. Weare, *The California electricity crisis: causes and policy options*. Public Policy Instit. of CA, 2003.

[13] Energy Information Administration, "California Energy Crisis," 2019. [Online]. Available: https://www.eia.gov/electricity/policies/legislation/california/subsequentevents.html

[14] S. Abraham, H. Dhaliwal, R. J. Efford, L. J. Keen, A. McLellan, J. Manley, K. Vollman, N. J. Diaz, T. Ridge *et al.*, "Final report on the August 14, 2003 blackout in the United States and Canada: Causes and recommendations," US-Canada Power System Outage Task Force, Tech. Rep., 2004.

[15] "Technical analysis of the August 14, 2003, blackout: What happened, why, and what did we learn," *Report to the NERC Board of Trustees*, vol. 13, 2004.

[16] A. Bose and T. J. Overbye, "Electricity transmission system research and development: grid operations," *Transmission Innovation Symposium: Modernizing the U.S. Electrical Grid, U.S. Department of Energy*, 2021.

[17] E. Vaahedi, *Practical power system operation*. John Wiley & Sons, 2014.

[18] "Analytic research foundations for the next-generation electric grid," *National Academies Press*, 2016.

[19] O. Zinaman, M. Miller, A. Adil, D. Arent, J. Cochran, R. Vora, S. Aggarwal, M. Bipath, C. Linvill, A. David *et al.*, "Power systems of the future: a 21st century power partnership thought leadership report," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2015.

[20] H. Akhavan-Hejazi and H. Mohsenian-Rad, "Power systems big data analytics: An assessment of paradigm shift barriers and prospects," *Energy Reports*, vol. 4, pp. 91–100, 2018.

[21] Y. Zhang, T. Huang, and E. F. Bompard, "Big data analytics in smart grids: a review," *Energy informatics*, vol. 1, no. 1, pp. 1–24, 2018.

[22] T. J. Overbye, J. L. Wert, K. S. Shetye, F. Safdarian, and A. B. Birchfield, "The use of geographic data views to help with wide-area electric grid situational awareness," in *2021 IEEE Texas Power and Energy Conference (TPEC)*. IEEE, 2021, pp. 1–6.

[23] T. J. Overbye and J. Weber, "Smart grid wide-area transmission system visualization," *Engineering*, vol. 1, no. 4, pp. 466–474, 2015.

[24] J. L. Wert, Z. Mao, H. Li, and T. J. Overbye, "Contouring method considerations for power systems applications," in *2020 IEEE Electric Power and Energy Conference (EPEC)*. IEEE, 2020, pp. 1–5.

[25] Western Electricity Coordinating Council, "Remedial action scheme design guide." [Online]. Available: https://www.wecc.biz/Reliability/RWG%20RAS%20Design%20Guide%20_%20Final.pdf

[26] A. Demir and N. Hadžijahić, "Power system planning: Part i—basic principles," in *Advanced Technologies, Systems, and Applications II*, M. Hadžikadić and S. Avdaković, Eds. Cham: Springer International Publishing, 2018, pp. 178–188.

[27] H. M. Merrill and A. J. Wood, "Risk and uncertainty in power system planning," *International Journal of Electrical Power & Energy Systems*, vol. 13, no. 2, pp. 81–90, 1991.

[28] NERC, "Transmission System Planning Performance Requirements." [Online]. Available: https://www.nerc.com/files/TPL-001-4.pdf

[29] "Maintaining reliability in the modern power system," United States Department of Energy, Tech. Rep., 2016.

[30] North American Electric Reliability Corporation, "Reliability stadandards for the bulk electric systems of north america," 2021. [Online]. Available: https://www.nerc.com/pa/Stand/Reliability%20Standards%20Complete%20Set/RSCompleteSet.pdf

[31] ——, "Remedial action scheme definition development." [Online]. Available: https://www.nerc.com/pa/Stand/Prjct201005_2SpclPrtctnSstmPhs2/FAQ_RAS_Definition_0604_final.pdf

[32] North American Electric Reliability Corporation (NERC), "PRC-012-2 Remedial Action Schemes," 2016.

[33] Western Electricity Coordinating Council, "Major WECC Remedial Action Schemes (RAS)." [Online]. Available: https://www.wecc.org/Reliability/TableMajorRAS4-28-08.pdf

[34] P. Anderson and B. LeReverend, "Industry experience with special protection schemes," *IEEE Transactions on Power Systems*, vol. 11, no. 3, pp. 1166–1179, 1996.

[35] Q. Hou, E. Du, N. Zhang, and C. Kang, "Impact of high renewable penetration on the power system operation mode: A data-driven approach," *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 731–741, 2020.

[36] C. Glenn, D. Sterbentz, and A. Wright, "Cyber threat and vulnerability analysis of the us electric sector," Idaho National Lab.(INL), Idaho Falls, ID (United States), Tech. Rep., 2016.

[37] W. Dunn, M. Rossi, and B. Avramovic, "Impact of market restructuring on power systems operation," *IEEE Computer Applications in Power*, vol. 8, no. 1, pp. 42–47, 1995.

[38] Federal Energy Regulatory Commission , "Critical energy/electric infrastructure information (CEII)," 2021. [Online]. Available: https://www.ferc.gov/enforcement-legal/ceii

[39] Z. Wang, A. Scaglione, and R. J. Thomas, "Generating statistically correct random topologies for testing smart grid communication and control networks," *IEEE Transactions on Smart Grid*, vol. 1, no. 1, pp. 28–39, 2010.

[40] S. Soltan, A. Loh, and G. Zussman, "A learning-based method for generating synthetic power grids," *IEEE Systems Journal*, vol. 13, no. 1, pp. 625–634, 2018.

[41] J. Hu, L. Sankar, and D. J. Mir, "Cluster-and-connect: An algorithmic approach to generating synthetic electric power network graphs," in *2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2015, pp. 223–230.

[42] K. M. Gegner, A. B. Birchfield, T. Xu, K. S. Shetye, and T. J. Overbye, "A methodology for the creation of geographically realistic synthetic power flow models," in *2016 IEEE Power and Energy Conference at Illinois (PECI)*. IEEE, 2016, pp. 1–6.

[43] A. B. Birchfield, T. Xu, and T. J. Overbye, "Power flow convergence and reactive power planning in the creation of large synthetic grids," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6667–6674, November 2018.

[44] H. Li, J. L. Wert, A. B. Birchfield, T. J. Overbye, T. G. San Roman, C. M. Domingo, F. E. P. Marcos, P. D. Martinez, T. Elgindy, and B. Palmintier, "Building highly detailed synthetic electric grid data sets for combined transmission and distribution systems," *IEEE Open Access Journal of Power and Energy*, vol. 7, pp. 478–488, 2020.

[45] A. B. Birchfield, "The creation, validation, and application of synthetic power grids," Ph.D. dissertation, 2018.

[46] A. Birchfield, E. Schweitzer, M. Athari, T. Xu, T. Overbye, A. Scaglione, and Z. Wang, "A metric-based validation process to assess the realism of synthetic power grids," *Energies*, vol. 10, no. 8, p. 1233, 2017.

[47] A. B. Birchfield, T. Xu, K. M. Gegner, K. S. Shetye, and T. J. Overbye, "Grid structural characteristics as validation criteria for synthetic networks," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3258–3265, July 2017.

[48] "Electric Grid Test Case Repository." [Online]. Available: https://electricgrids.engr.tamu.edu/

[49] C. Molnar, *Interpretable machine learning*. Lulu. com, 2020.

[50] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media, 2009.

[51] R. Xu, F. Tafazzoli, L. Zhang, T. Rehfeld, G. Krehl, and A. Seal, "Holistic grid fusion based stop line estimation," in *2020 25th International Conference on Pattern Recognition (ICPR)*. Los Alamitos, CA, USA: IEEE Computer Society, jan 2021, pp. 8400–8407. [Online]. Available: https://doi.ieeecomputersociety.org/10.1109/ICPR48806.2021.9413070

[52] Z. Meng, R. Xu, and C. M. Ho, "Gia-net: Global information aware network for low-light imaging," in *Computer Vision – ECCV 2020 Workshops*, A. Bartoli and A. Fusiello, Eds. Cham: Springer International Publishing, 2020, pp. 327–342.

[53] A. Zang, R. Xu, Z. Li, and D. Doria, "Lane boundary extraction from satellite imagery," in *Proceedings of the 1st ACM SIGSPATIAL Workshop on High-Precision Maps and Intelligent Applications for Autonomous Vehicles*, ser. AutonomousGIS '17. New York, NY, USA: Association for Computing Machinery, 2017. [Online]. Available: https://doi.org/10.1145/3149092.3149093

[54] S. Zhang, Y. Zhao, D. T. Nguyen, R. Xu, S. Sen, J. Hester, and N. Alshurafa, "Necksense: A multi-sensor necklace for detecting eating activities in free-living conditions," vol. 4, no. 2, Jun. 2020. [Online]. Available: https://doi.org/10.1145/3397313

[55] L. Wang, *Support vector machines: theory and applications*. Springer Science & Business Media, 2005, vol. 177.

[56] M. Liu and G. Gross, "Effectiveness of the distribution factor approximations used in congestion modeling," in *Proceedings of the 14th Power Systems Computation Conference*, 2002, pp. 1–8.

[57] T. Guler, G. Gross, and M. Liu, "Generalized line outage distribution factors," *IEEE Transactions on Power Systems*, vol. 22, no. 2, pp. 879–881, 2007.

[58] W. F. Tinney, V. Brandwajn, and S. M. Chan, "Sparse vector methods," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-104, no. 2, pp. 295–301, 1985.

[59] F. L. Alvarado, "Computational complexity in power systems," *IEEE Transactions on Power Apparatus and Systems*, vol. 95, no. 4, pp. 1028–1037, 1976.

[60] M. D. Maram and N. Amjady, "Event-based remedial action scheme against supercomponent contingencies to avert frequency and voltage instabilities," *IET Generation, Transmission & Distribution*, vol. 8, no. 9, pp. 1591–1603, 2014.

[61] S. Wang and G. Rodriguez, "Smart RAS (Remedial Action Scheme)," in *2010 Innovative Smart Grid Technologies (ISGT)*. IEEE, 2010, pp. 1–6.

[62] Y. Zhang and K. Tomsovic, "Adaptive remedial action scheme based on transient energy analysis," in *IEEE PES Power Systems Conference and Exposition, 2004.* IEEE, 2004, pp. 925–931.

[63] H. Atighechi, P. Hu, J. Lu, G. Wang, and S. Ebrahimi, "A fast load shedding remedial action scheme using real-time data for bc hydro system," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*. IEEE, 2016, pp. 1–5.

[64] X. Fan, R. Huang, Q. Huang, X. Li, E. L. Barrett, J. G. O'Brien, Z. Hou, H. Ren, S. Kincic, and H. Zhang, "Adaptive ras/sps system setting for improving grid reliability and asset utilization through predictive simulation and controls: A use case for transformative remedial action scheme tool (trast): Jim bridger ras evaluation and analysis," Pacific Northwest National Lab.(PNNL), Richland, WA (United States), Tech. Rep., 2019.

[65] Energy Information Administration, "Annual Electric Power Industry Report." [Online]. Available: https://www.eia.gov/electricity/data/eia861/

[66] Department of Energy, "Commercial and Residential Hourly Load Profiles." [Online]. Available: https://openei.org/datasets/dataset/commercial-and-residential-hourly-load-profiles-for-all-tmy3-locations-in-the-united-states/

[67] DOE Energy Efficiency & Renewable Energy, "Building America House Simulation Protocols," 2010. [Online]. Available: https://www.nrel.gov/docs/fy11osti/49246.pdf

[68] Department of Energy, "Commercial Reference Building Models of the National Building Stock," 2011. [Online]. Available: https://www.nrel.gov/docs/fy11osti/46861.pdf

[69] Oak Ridge National Laboratory, "Assessment of Industrial Load for Demand Response Across U.S. Regions of the Western Interconnect," 2013. [Online]. Available: https://info.ornl.gov/sites/publications/files/pub45942.pdf

[70] "Industrial Assessment Centers." [Online]. Available: https://iac.university/

[71] D. Toffanin, "Generation of customer load profiles based on smart-metering time series, building-level data and aggregated measurements," Ph.D. dissertation, Zurich, 2016.

[72] "Syn-Austin-TDGrid," 2020. [Online]. Available: https://electricgrids.engr.tamu.edu/combined-td-synthetic-dataset/

[73] "Pecan Street DataPort," 2020. [Online]. Available: https://dataport.pecanstreet.org/

[74] D. Horvatic, H. E. Stanley, and B. Podobnik, "Detrended cross-correlation analysis for non-stationary time series with periodic trends," *EPL (Europhysics Letters)*, vol. 94, no. 1, p. 18007, 2011.

[75] T. Odun-Ayo and M. L. Crow, "Structure-preserved power system transient stability using stochastic energy functions," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1450–1458, 2012.

[76] National Renewable Energy Laboratory, "Taxonomy Feeders Load Time Series." [Online]. Available: https://openei.org/datasets/dataset/randomized-hourly-load-data-for-use-with-taxonomy-distribution-feeders

[77] "Open Power System Data," 2020. [Online]. Available: https://open-power-system-data.org/

[78] "Electric System Operation Data," 2020. [Online]. Available: https://www.eia.gov/realtime_grid/

[79] D. Hostick, D. B. Belzer, S. W. Hadley, T. Markel, C. Marnay, and M. Kintner-Meyer, "Renewable electricity futures study. volume 3. end-use electricity demand," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2012.

[80] National Renewable Energy Laboratory, "Wind ToolKit Data." [Online]. Available: https://developer.nrel.gov/docs/wind/wind-toolkit/wind-toolkit-extract/

[81] ——, "Solar Resource Data." [Online]. Available: https://developer.nrel.gov/docs/solar/solar-resource-v1/

[82] ERCOT, "Seasonal Assessment of Resource Adequacy for the ERCOT Region." [Online]. Available: http://www.ercot.com/gridinfo/resource

[83] ——, "Report on the Capacity, Demand and Reserves." [Online]. Available: http://www.ercot.com/content/wcm/lists/167023/CapacityDemandandReserveReport-May2019.pdf

[84] "PJM Capacity Market Manual," 2021. [Online]. Available: https://www.pjm.com/~/media/documents/manuals/m18.ashx

[85] "ERCOT Outage Coordination." [Online]. Available: http://www.ercot.com/content/wcm/training_courses/109600/TRN101_M3_Outage_Scheduling_2017.pdf

[86] "NERC Standard TOP-003-1: Planned Outage Coordination." [Online]. Available: https://www.nerc.com/files/TOP-003-1.pdf

[87] "Impact of load following on the economics of existing coal-fired power plant operations," Department of Energy, Tech. Rep. [Online]. Available: https://www.eenews.net/assets/2017/11/21/document_gw_08.pdf

[88] "GE external report on maintenance schedule checklist for hydro generators," Tech. Rep. [Online]. Available: https://www.gepowerconversion.com/sites/gepc/files/Maintenance%20Schedule%20Checklist%20for%20Hydro%20Generators.pdf

[89] "Catalog of CHP technologies," Tech. Rep. [Online]. Available: https://www.epa.gov/sites/production/files/2015-07/documents/catalog_of_chp_technologies_section_3._technology_characterization_-_combustion_turbines.pdf

[90] "Nuclear power plants undergo seasonal scheduled outages." [Online]. Available: https://www.eia.gov/todayinenergy/detail.php?id=1490

[91] "Best Practices in Photovoltaic System Operations and Maintenance." [Online]. Available: https://www.nrel.gov/docs/fy17osti/67553.pdf

[92] "GE external report on maintenance schedule checklist for wind turbine application." [Online]. Available: https://www.gepowerconversion.com/sites/gepc/files/Maintenance%20Schedule%20Checklist%20for%20Wind%20Generators.pdf

[93] "Status of U.S. Nuclear Outages." [Online]. Available: https://www.eia.gov/nuclear/outages/

[94] T. Senjyu, K. Shimabukuro, K. Uezato, and T. Funabashi, "A fast technique for unit commitment problem by extended priority list," *IEEE Transactions on Power Systems*, vol. 18, no. 2, pp. 882–888, 2003.

[95] "ERCOT Steady State Working Group Procedure Manual." [Online]. Available: http://www.ercot.com/content/wcm/key_documents_lists/73156/16._SSWG_Procedure_Manual.NewNamingConvention.doc

[96] T. J. Overbye and R. P. Klump, "Effective calculation of power system low-voltage solutions," *IEEE Transactions on Power Systems*, vol. 11, no. 1, pp. 75–82, February 1996.

[97] D. Mehta, H. D. Nguyen, and K. Turitsyn, "Numerical polynomial homotopy continuation method to locate all the power flow solutions," *IET Generation Transmission and Distribution*, vol. 10, no. 12, pp. 2972–2980, June 2016.

[98] "PowerWorld: Working with Hard-To-Solve Power Flow Cases." [Online]. Available: https://www.powerworld.com/training/online-training/security-analysis/techniques-for-conditioning-hard-to-solve-cases

[99] "PowerWorld Solution Control." [Online]. Available: https://www.powerworld.com/WebHelp/Content/MainDocumentation_HTML/RibbonGroup_Tools_PowerFlow.htm

[100] S. Noguchi, M. Shimomura, J. Paserba, and C. Taylor, "Field verification of an advanced high side voltage control at a hydro power station," *IEEE Transactions on Power Systems*, vol. 21, no. 2, pp. 693–701, May 2006.

[101] J. D. Weber and T. J. Overbye, "Voltage contours for power system visualization," *IEEE Transactions on Power Systems*, vol. 15, no. 1, pp. 404–409, 2000.

[102] T. J. Overbye, J. L. Wert, A. B. Birchfield, and J. D. Weber, "Wide-area electric grid visualization using pseudo-geographic mosaic displays," in *51st North American Power Symposium (NAPS)*, 2019.

[103] North American Electric Reliability Corporation, "System operating limit definition and exceedance clarification." [Online]. Available: https://www.caiso.com/Documents/3100.pdf

[104] W. Lachs, "Transmission-line overloads: real-time control," in *IEE Proceedings C (Generation, Transmission and Distribution)*, vol. 134, no. 5. IET, 1987, pp. 342–347.

[105] M. Vaiman, K. Bell, Y. Chen, B. Chowdhury, I. Dobson, P. Hines, M. Papic, S. Miller, and P. Zhang, "Risk assessment of cascading outages: Methodologies and challenges," *IEEE Transactions on Power Systems*, vol. 27, no. 2, p. 631, 2012.

[106] C. M. Davis and T. J. Overbye, "Multiple element contingency screening," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1294–1301, 2010.

[107] C. Aumuller and T. Saha, "Determination of power system coherent bus groups by novel sensitivity-based method for voltage stability assessment," *IEEE Transactions on Power Systems*, vol. 18, no. 3, pp. 1157–1164, 2003.

[108] F. Capitanescu and T. Van Cutsem, "Preventive control of voltage security margins: a multicontingency sensitivity-based approach," *IEEE Transactions on Power Systems*, vol. 17, no. 2, pp. 358–364, 2002.

[109] C. Li, H.-D. Chiang, and Z. Du, "Network-preserving sensitivity-based generation rescheduling for suppressing power system oscillations," *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3824–3832, 2017.

[110] C. P. Reddy, S. Chakrabarti, and S. C. Srivastava, "A sensitivity-based method for underfrequency load-shedding," *IEEE Transactions on Power Systems*, vol. 29, no. 2, pp. 984–985, 2014.

[111] J. Momoh, J. Zhu, G. Boswell, and S. Hoffman, "Power system security enhancement by opf with phase shifter," *IEEE Transactions on Power Systems*, vol. 16, no. 2, pp. 287–293, 2001.

[112] E. S. Hoji, A. Padilha-Feltrin, and J. Contreras, "Reactive control for transmission overload relief based on sensitivity analysis and cooperative game theory," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1192–1203, 2012.

[113] K. Pandiarajan and C. Babulal, "Overload alleviation in electric power system using fuzzy logic," in *2011 International Conference on Computer, Communication and Electrical Technology (ICCCET)*, 2011, pp. 417–423.

[114] Western Electricity Coordinating Council, "Western interconnection ras review," 2019. [Online]. Available: https://www.wecc.org/Reliability/Western%20Interconnection%20RAS%20Review%20Guideline.pdf