Transient Stability Prediction based on Synchronized Phasor Measurements and Controlled Islanding

Meiyan Li

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State
University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Electrical Engineering

Arun G. Phadke (Chairman)
Virgilio A. Centeno
Jaime De la Ree Lopez
Yaman Evrenosoglu
Werner E. Kohler
Mariusz K. Orlowski

May 30th , 2013
Blacksburg, Virginia

Keywords: critical locations, apparent impedance trajectory, complex synchrophasor data, decision trees, dimensional reduction, fisher's linear discriminant, transient stability, controlled islanding

Transient Stability Prediction based on Synchronized Phasor Measurements and Controlled Islanding

Meiyan Li

Abstract

Traditional methods for predicting transient stability of power systems such as the direct method, the time domain approach, and the energy function methods do not work well for online transient stability predictions problems. With the advent of Phasor Measurement Units (PMUs) in power systems, it is now possible to monitor the behavior of the system in real time and provide important information for transient stability assessment and enhancement. Techniques such as the rotor oscillation prediction method based on time series have made the prediction of system stability possible for real-time applications. However, methods of this type require more than 300 milliseconds after the start of a transient event to make reliable predictions. The dissertation provides an alternate prediction method for transient stability by taking advantage of the available PMUs data. It predicts transient stability using apparent impedance trajectories obtained from PMUs, decision trees, and FLDSD method. This method enables to find out the strategic locations for PMUs installation in the power system to rapidly predict transient stability. From the simulations performed, it is realized that system stability can be predicted in approximately 200 milliseconds (12 cycles). The main advantage of this method is its simplicity as the PMUs can record the apparent impedance trajectories in real-time without any previous calculations. Moreover, using decision trees built in CART®, transient stability prediction becomes straightforward and computationally very fast. The optimum locations for PMUs placement can also be determined using this technique.

After the transient instability prediction by the apparent impedance trajectories, a slow-coherency based intelligent controlled islanding scheme is also developed to restore the stability of system. It enables the generators in the same island to stay in synchronism and the imbalance between the generators and load demand is minimized.

This dissertation is dedicated to my parents and my husband, Baishen Li, Lihua Gao and Peihsin Liu.

For their encouragement, support and love during this period of time.

Acknowledgements

First and foremost, I would love to express my supreme respect and gratitude to my advisor and committee chair, Dr. Phadke. Dr.Phadke has provided the most precious and valuable opportunity for me to study in the power lab of Virginia Tech and granted me a great favor to instruct me during my Ph.D study. I have learned remarkably from this greatest professor in the world and benefited significantly from his genius idea, profound perspective and incisive insights of power system. He helps me very much in the past three years and I can't thank him enough.

Special thanks to Dr. Thorp. He is a very nice professor and offers me his creative thoughts and shared them with me generously. I have also benefited a great deal by discussing the academic work with him. My gratitude also goes to Dr. Evrenosoglu and Dr. Orlowski, and Dr. Kholer, for their genuine interest and support.

My appreciation also goes Dr.Centeno and Dr. Jaime De La Ree, they not only provide me the professional academic instruction but also helped me considerably during my study period. They are terrific professors and good friends of mine.

I would also thank my lab mates, Anamitra, Ryan, Rui, and Zhongyu. Thank you for their help in my study and in my life. Last but not least, I would like to thank my parents and my husband. They have always provided me the endless love and strongest support in my life.

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

The U.S. electric grid connects thousands of generators and millions of electricity customers. The interconnection of local and small power generators enhance the efficiency of power systems but may increase the possibility of wide-area blackout, which affects the U.S. economy and quality of life. The reliability of power supply influence millions of the residential, commercial and industrial users. A large contingency may potentially lead to a cascading failure or even blackouts of power system, which will dramatically harm the economy and national security. The historic blackouts such as August 10, 1996 in Western Systems Coordinating Council (WSCC) [1] and March 11, 1999 in Brazil [2] have shown that the transient instability has become one of the biggest threats to the current power systems. These events have drawn attentions from both engineers and the public. With the fast growth and large area interconnection of power system, the transient stability becomes even more complicated and hard to be predicted. Therefore, it is vital to predict the stability as soon as possible and take predetermined measures to restore the stability of power system. This demands intelligent technologies for prediction and appropriate regulation policies in the near future.

1.1 Transient Stability Definition and Prediction

When the power system is subjected to a disturbance, the initial operating condition and the severity of disturbance will affect the stability of the system. To maintain the stability of the system, different protection schemes will be put into actions with respect to different circumstances. For example, the generator and voltage regulator will operate when voltage changes; the prime mover governor will be actuated when the rotor speed varies. All these automatic controls are designed to stabilize the system at an equilibrium state. According to [3], there are mainly three types of power system stability: rotor angle stability, frequency stability and voltage stability, as shown in Figure 1.1. The rotor angle stability usually refers to the angular behavior of the generator, which reflects the relationship between the input mechanical torque and output electromagnetic torque.

Lack of sufficient synchronizing torque may lead to the acceleration of the rotor speed and cause a large deviation of the angular speed which ultimately will separate the out of step machines from the rest of the system.

Rotor angle stability includes transient stability and small disturbance angle stability. The main difference between these two types of stability is the severity of the disturbance. Small disturbance angle stability refers to when power system subjected to a small disturbance, due to the insufficiency of damping torque, the generator rotor angle will oscillate with a low frequency, between 0.1-3 Hz. The transient stability or large-disturbance rotor angle stability, is referred to with a given initial operation condition, power system is able to maintain synchronism after a severe disturbance, such as system faults, loss of generation, or circuit contingencies, etc. Due to the deficiency of the synchronizing torque, the first swing of the rotor will usually appear as an aperiodic angular separation. Therefore, instability is commonly manifested as first swing instability, which usually occurs in 3 to 5 seconds following the disturbance.

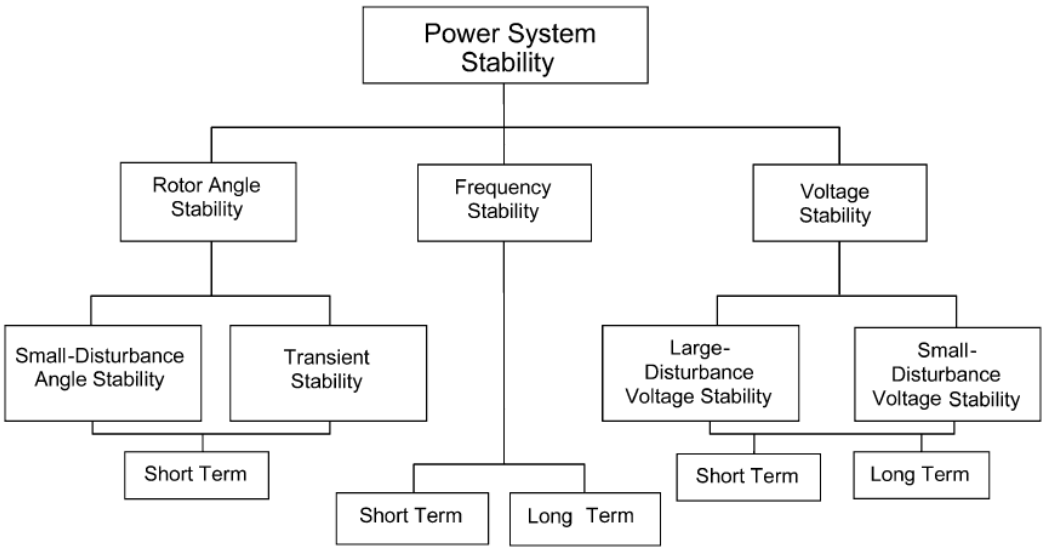


Figure 1. 1: Classification of power system stability [3]

Transient stability prediction investigates whether the synchronous machine will remain synchronized following the occurrence of the disturbance. The transient stability of the interconnected power system has become one of the largest limiting factors for

power transfer and normal operation. Fast transient stability prediction is critical for real-time controlled protection because it may prevent the cascading failure or even large area blackouts. With the help of synchronized phasor measurement, many advanced transient stability techniques, such as decision tree, support vector machine, fuzzy system, artificial neural network, etc. have been developed to assist the transient stability prediction on line.

1.2 Controlled Islanding

After suffering from a severe perturbation, the system may have a cascading failure or even uncontrolled islanding. Uncontrolled islanding means the system separate into several unplanned islands when the system has a severe disturbance [4]. In Jan.17 of 1994, five islands were formed in the WECC with lots of load shedding and generators tripping; in August 10 of 1996, four islands were formed resulted in even larger load and generators loss. During the contingencies, low voltages and power oscillations in some tie lines are observed. If controlled islanding were implemented at the early stage of the contingency, the result of the power loss and of the catastrophic contingency could be mitigated. Therefore, the controlled islanding is proposed as a remedial action to prevent the system from wide area blackouts. Controlled islanding is the last resort to alleviate the impact of catastrophic contingency in large scale power systems. Proper load shedding, generator dropping, and system splitting can maintain the system stability and much of the losses can be reduced. Customers are still able to be served at a somewhat degraded system after islanding [5].

Typical controlled islanding will be implemented in a heavy loaded system with long tie-lines and weak but stressed interconnection. In the prevailing power system, the controlled islanding boundaries are some pre-determined critical lines based on the industry practice. Considering many possible contingency scenarios, the locations and the set value of islanding devices are pre-decided by lots of offline simulations and calculations. When the disturbance happens and the threshold values are met, the circuit breakers installed in the predetermined transmission lines will trip to isolate some parts of the system. However, since the emergencies strike the system suddenly and may lead to the collapse or blackouts of the entire system in a short time, an online controlled

islanding scheme is required for real time application. In solving the online controlled islanding problem, two essential issues must be properly addressed:

1. When to implement the controlled islanding?
2. Where to implement the controlled islanding?

After the remedial scheme of controlled islanding, the generators within each island should remain in synchronism, the imbalance between generators output and the load demand should be minimized, as well as all voltage and frequency levels are stabilized at an acceptable level.

There are some common criteria for the controlled islanding [6-7].

- Islanding the system based on the coherent groups
- Two ends of line are 180° out of phase, creating the electrical center in the line.
- The oscillation of active power exceeds a threshold at the pre-specified line
- Minimizing the power mismatch between the generation and load within each island.

Also, some innovative technologies such as ordered binary decision diagram (OBDD), pattern recognition and other real time analytical techniques have been developed to determine the islanding boundaries of system.

1.3 Synchronized Phasor Measurements and Their Application in Power Systems

Synchronized phasor measurements play an important role in Wide Area Measurement (WAM). The voltage and current phasor of the system can be monitored by the Phasor Measurement Units (PMUs). PMUs using the global positioning system (GPS) was first developed in Virginia Tech and it has been widely implemented in power system and has changed the system remarkably. Thanks to the advancement of computer technology and the availability of GPS signals, this technology becomes a sophisticated and powerful tool for power system analysis. By using a common time source for synchronization, dispersed locations could be sampled and measured in real-time.

Measurement is reported at a rate of 30-60 frames per second. The classical definition of phasor $\mathbf{X}$ is:

$$\mathbf{X} = \frac{X_m}{\sqrt{2}} e^{j\phi}$$

in which, X_m is the magnitude of the input sinusoidal signal, while ϕ is the angular position with respect to an arbitrary time reference.

All the currents and voltages are sent to the analog to digital converter after filtering by the anti-aliasing filters. The cut off frequency of the filter is less than half of the sampling frequency. The PMUs measurements have been used in angle/frequency monitoring and visualization, outage prevention, state measurement, real time control and so on. The deployment of PMUs enhances the stability and reliability of power system. The architecture of PMUs is shown in Figure 1.2 [8]. The current and voltage transformers supply the currents and voltages as analog inputs. The GPS clock pulse provides the 1pps signal for the phase-locked oscillator. The microprocessor computes the phasor by the Discrete Fourier Transform (DFT). Finally, the time-stamped measurement is transferred via the modem as PMUs output.

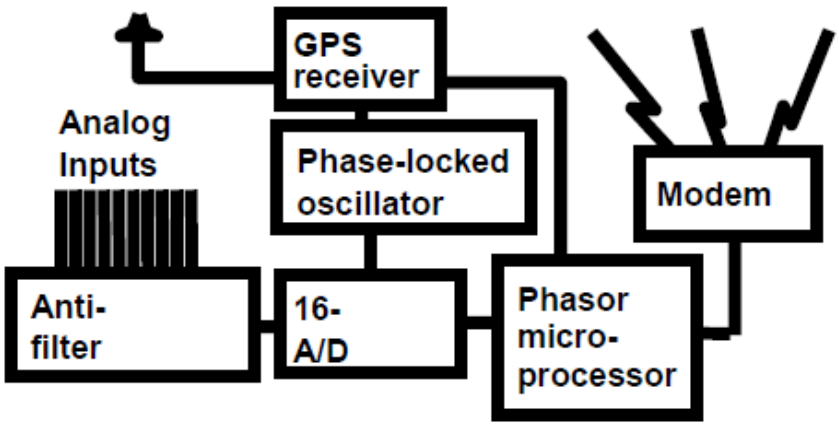


Figure 1. 2: PMUs architecture [8]

There are many applications of using synchronized phasor measurement in power system [9]:

Real time monitoring and control: PMUs provides the real time measurement data of system state instead of estimating it, which helps the operator to decrease the error margins of transmission capability and enable the transmission lines to transfer more power close to their real stability limits. Also, it facilitates the operator to monitor, detect and control system condition by the time-synchronized devices. The real time monitoring of PMUs data prevents blackouts from the events that cause voltage instability, etc.

Power system state estimation (SE): The SE is benefited by the directly measured voltages and phase angles to improve its accuracy and robustness of bad data detection, as well as linearize the SE solution. Three phase and three sequence estimation are also able to be deployed by the PMUs implement ation.

Real-time congestion management: In the past, the congestion management based on the offline calculation of the power flow only considered the thermal limits, stability limits and voltage limits. The result with large margin may be too conservative for congestion management. With the real-time and accurate data from PMUs, the congestion management becomes more economic and efficient. Operator dispatches generation to meet the economic demand without violating the transmission limits.

Benchmarking, validation and fine-tuning of system models: Accurate modeling such as the network, generator and load, etc. plays an important role in power flow calculation, stability analysis, security assessment and modal frequency response. PMUs provide the possibility for model error identification and correct value calculation. Benchmarking and tuning parameters in dynamic and oscillation modeling also becomes feasible by parameters estimation technique [10].

Post-disturbance analysis: After the contingency, the engineers can use the time synchronized data recorded by PMUs to study various dispersed places of the grid. Better fault analysis could be performed by the precise time sync which enables the time for disturbance diagnosis to be reduced to few seconds.

Power-system restoration: PMUs can avoid closing the two adjacent breakers when there is a large phase angle difference thus preventing a possible system outage. A lot of valuable time can be saved during the grid restoration.

Protection and control applications for distributed generation: As the trend of distributed generation is growing, PMUs help to monitor and island the distributed network.

Overload monitoring and dynamic rating: The application of PMUs can provide degrees of help in monitoring the load condition in real time and help to calculate the impedance of a line to estimate the average temperature of the conductor.

Adaptive protection: Synchronized phasor measurement enhances the performance of many protection devices. "Adaptive Relay" changes its characteristic according to the system circumstance, which benefits in reducing the operation margin to achieve economic and environmental advantage. Adaptive protection has been applied to Out of Step relays, line relays, etc. Accurate measurement of PMUs also helps in locating the fault, as well as diagnosing and restoring the grid after contingency.

Planned power system separation: Intelligent islanding will be applied to balance the generation and the load when power system is going through an unstable electromechanical oscillation. PMUs measurement helps the islanding by finding the coherency of generators groups which may lead to loss of stability and by finding the islanding boundaries in the prevailing system.

1.4 Motivations

Although many technologies have been studied and applied to predict the stability of the system, they are either based on offline prediction or are not fast enough for transient stability prediction. Traditional methods such as the direct method, the time domain approach, and the energy function methods do not work well for online transient stability predictions. The limitation of the direct method, mainly the equal area criterion method, is that it requires an over-simplification of the model used. This constraint makes it unsuitable for application to complex interconnected power system networks [11]. The time domain approach greatly taxes available computer resources and is slow in solving large sets of non-linear differential-algebraic equations [12]. The energy function method predicts the stability of the system based on the comparison of the transient energy and the potential energy evaluated at the unstable equilibrium point. However, it predicts the

system stability solely on the basis of the first swing stability. Unfortunately, if the second swing were to go unstable, the prediction result by this method would be incorrect [13]. Similarly, the Boundary Controlling Unstable (BCU) equilibrium point method requires that the trajectory begin inside the stability boundary, close to the controlling unstable equilibrium point [14]. It fails to predict instability if the system does not satisfy the required mathematical constraints.

With the development of Phasor Measurement Units (PMUs) in power systems, it is now possible to monitor behavior of the system in real time and provide important information for transient stability assessment and enhancement. Fuzzy-logic based methods [15,16] have been found to be successful in predicting the security of large electric power systems. Machine learning based methods have also been explored for online stability predictions. Artificial Neural Networks (ANN) were used to detect and rank the dynamic security contingency of large systems in [17,18]. Support Vector Machines (SVM) applies the kernel concept to classify the data by finding a maximum-margin hyperplane, which are able to extract features in high dimensional space for dynamic security classification as shown in [19,20]. However, in recent years, Decision Trees (DTs) have emerged as the most popular method for making transient stability predictions [21]. In comparison to the other methods, DTs are simpler to build and easier to implement. By simulating large number of contingency cases offline, a straight forward decision tree can be built for online stability assessment [22,23]. In [24], using phasor measurement and decision trees, a methodology was developed to predict loss of synchronization in real-time. But the machines used in [24] had classical dynamic models which did not accurately represent the real system. The rotor oscillation prediction method based on time series predicted the transient stability of the system in real-time [25], but it needs to initially identify the coherent groups of generators following a contingency. This requires more than 300 milliseconds after the start of the transient event to predict stability of the system. If we do not predict stability after the contingency immediately and correctly, it may lead to cascading failures eventually culminating in a widespread blackout. Consequently, it was realized that there is an imminent need for a faster and more accurate analytical real-time prediction technique.

The motivation for this dissertation is to provide an alternate prediction method efficiently and quickly for transient stability prediction by taking advantage of the available PMUs data. It predicts transient stability using apparent impedance trajectories obtained from PMUs, decision trees, and Fisher's Linear Discriminant (FLD). Examination of simulated data and field data suggest that apparent impedance might be useful in stability assessment. The emergence of dual use Line-relays/PMUs make it convenient to obtain time-tagged apparent impedances. The CART® algorithm can be used to find the locations where such measurements should be made and then communicated to the control center, where the stability assessment can be made. The trajectory of the apparent impedances recorded by the PMUs placed at strategic locations within the power system can thus be used to rapidly predict transient stability. The strategic locations are those locations where the recorded PMUs data will be most useful for system stability prediction. Using the proposed method, an optimized PMUs placement scheme can also be developed to identify critical locations from the transient stability point of view.

Based on the information about the state of the power system provided by phasor measurements, the apparent impedance trajectory is used to forecast the system stability in real-time. The first step is to record the apparent impedance trajectory of each branch for a few cycles after the clearance of the fault. This is followed by extending the dimensional reduction method to higher dimensional space, integrating the high-dimensional complex apparent impedance trajectory recorded by PMUs, to a single dimensional evaluation index for transient stability prediction. The final step involves using data mining techniques to select those critical trajectories that would characterize the system status based on the evaluation index developed in the previous step. By doing so, PMUs would only be required at the critical locations of the network which then optimizes the number of installations, shown as Figure 1.3. From the results, it is realized that very high prediction accuracy is reached by using a small decision tree with only a few nodes, i.e. with a small number of PMUs installations to observe and monitor the system, it becomes possible to perform an out of step forecasting within few cycles after the clearance of the fault. Prediction of transient stability based on the apparent impedance trajectory method was analyzed and verified on three test systems. A success

rate of more than 95% was obtained on the predictions using a 12-cycle (200 milliseconds) data window. The simulation results indicated that the algorithm developed in this dissertation is a valid method for predicting transient stability within a short time period and that it can be easily extended to larger systems.

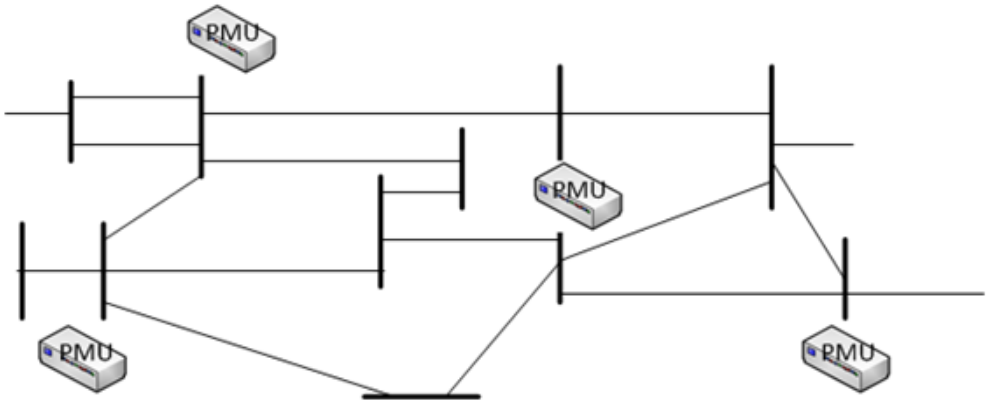


Figure 1. 3: PMUs placement in network illustration

The flow chart of the determination of critical PMUs locations by the proposed method is shown as Figure 1.4.

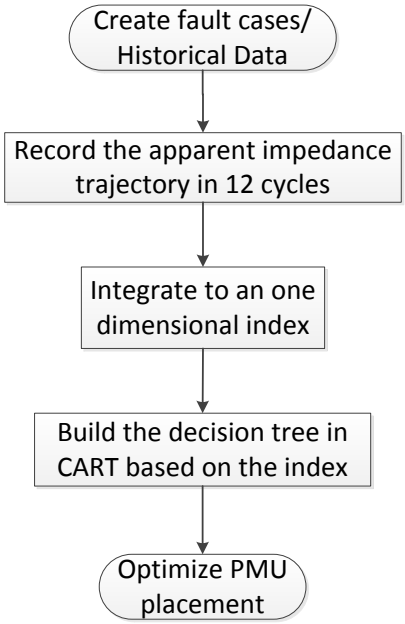


Figure 1. 4: Flow chart of building DT and PMUs placement

The optimization of the PMUs placement is done by the offline-simulation. The first step of this method is to generate the data base. The data base is built with two types

of three-phase faults that are computed at every bus and along with the N-1 operation. The first type is a three-phase short circuit fault that lasts for 80 milliseconds. The second type is a three-phase short circuit fault that lasts for 300 milliseconds. To include different operating conditions, the topology of the system is changed by tripping one of the transmission lines for every case. 12 cycles of apparent impedance trajectory are recorded at every branch after contingency. By integrating the 12-cycle data into an important index in one dimension, the critical branches characterized the system status for transient stability prediction are determined by CART®, which is a data mining tool. Finally, the buses that connect to the critical braches for PMUs installation can be selected as strategic locations by the decision tree for transient stability prediction.

Thus, when a real fault occurs in the power system, based on the trajectory detected by the PMUs previously selected, we can record the apparent impedance trajectory of the critical braches and calculate the index by extending the FLDSD method to high dimensions. Based on the decision tree that has already been built offline, the system can be predicted as stable or not. After the transient instability prediction, some pre-determined control mechanism scheme should be executed to prevent the cascading failure or even large area blackouts of system. In this dissertation, an innovative controlled islanding scheme is developed as a protection strategy to prevent the spread of contingency. The fast transient instability result generated by decision tree can assist the implementation of controlled islanding. Based on the recognition of coherency by some advanced techniques [26,27] at the early stage of contingency, the proposed controlled islanding scheme can prevent the cascading failure of the system. The generators outputs are balanced with the load demands within the island, as well as the voltage and frequency are stabilized at an acceptable level. The overall flow chart of the online transient stability prediction followed by the intelligent controlled islanding is shown as Figure 1.5:

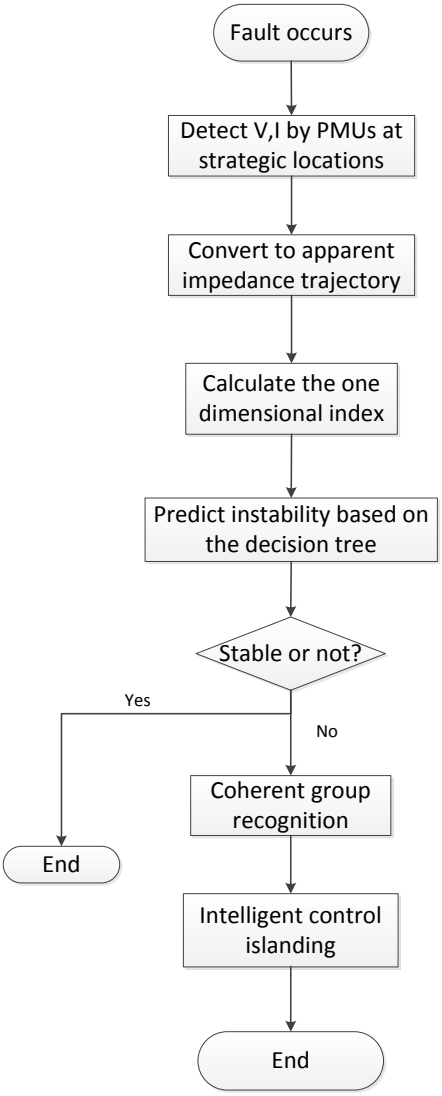


Figure 1. 5: Overall transient stability prediction followed by controlled islanding

1.5 Overview of the Dissertation

The proposed method presented in this dissertation is primarily focused on predicting the transient stability using the apparent impedance trajectory recorded by synchronized measurement. Optimal locations for PMUs placement in transient stability are also determined. Furthermore, a controlled islanding scheme is developed to restore the stability of system as a remedial protection.

The rest of this dissertation is organized as follows.

Chapter 1 introduces the transient stability prediction and the controlled islanding concept. The benefits of PMUs measurements are also discussed. The motivations and objectives of this dissertation are mentioned. Based on the decision tree built by the apparent impedance trajectories measured by PMUs located at strategic locations, the transient stability can be predicted within a short time period (12 cycles). The controlled islanding protection scheme is developed to restore the stability of system after the identification of different groups of contingency.

Chapter 2 is a literature review which includes many popular transient stability prediction technologies and the available controlled islanding schemes in the prevailing power system.

Chapter 3 illustrates the apparent impedance trajectory definition and its application in power system, such as Out of Step relay.

Chapter 4 presents a brief introduction of the data mining method and its application in transient stability prediction. CART® used as the main data mining tool in this dissertation, its function and methodology to build the decision tree are presented. The limitation of CART® in solving complex data is also included.

Chapter 5 purposes a transient stability prediction method using the apparent impedance trajectories recorded by PMUs. After raising the application of apparent impedance trajectory in this dissertation, a dimensional reduction method (FLDSD) for classifying the PMUs complex data in two dimensions developed from the Linear Discriminant Analysis and Fisher's Linear Discriminant method is explained. Extending the FLDSD method to the higher dimensional apparent impedance trajectory (12 cycles) recorded by PMUs at critical locations, the transient stability can be predicted in real time.

Chapter 6 verifies the proposed method with three test systems: 17-machine, 162-bus system, the 29-machine, 128-bus model of the Western Electricity Coordinating Council (WECC) system and the 4000 Bus WECC system. Prediction of transient stability based on the apparent impedance trajectory method was analyzed and verified on these test systems. A success rate of more than 95% was obtained on the predictions. The simulation results indicated that the algorithm developed in this dissertation is a valid

method for predicting stability within a short time period and can be easily applied to larger systems.

Chapter 7 introduces a possible pre-determined protection scheme: controlled islanding after the system is predicted as unstable and preventing the spread of the contingency in power system. The 29-machine, 128-bus WECC system is used to validate the controlled islanding scheme. After the recognition of coherent groups after instability prediction, the proposed controlled islanding method restores the synchronism of generators in each island.

Chapter 8 summarizes the contributions in this dissertation, presents conclusions drawn from the work presented, and makes several suggestions for future research related to intelligent controlled islanding.

Chapter 2 Literature Review

In this chapter, the traditional transient stability prediction and the modern transient stability prediction developed by phasor measurement unit (PMUs) are described. Controlled islanding as a final resort to prevent large area blackouts is also introduced.

2.1 Method of Transient Stability Prediction

In the modern power system planning, operation, and control, transient stability plays an important role. The objective of this section is to give an introduction to various methods for transient stability prediction, which includes the traditional methods and the modern methods. Each prediction technology has been presented with their advantages and drawbacks, as well as their developments with the deployment of synchronized phasor measurement. The traditional prediction method mainly is comprised of the time domain approach, equal area criterion, energy function method, hybrid method, etc. With the advent of phasor measurement units (PMUs) and the technology of modern communication, the prediction of transient stability can be realized in a large scale power system and in real time with high precision and great reliability. Multiple novel approaches based on PMUs have been reported, like the artificial intelligence techniques such as pattern identification, fuzzy system, artificial neural network, etc. are developed for the fast prediction of transient stability.

2.1.1 Time Domain Approach

The time domain approach is a numerical method, which gets each variable in the system as a set of values at certain intervals of time. This kind of method has been used to develop the transient stability prediction software.

In this approach, the general overall forms of system equation are step by step calculation, and could be expressed by a set of differential equations and a set of algebraic equations, as shown in (2.1) and (2.2)

$$\dot{y} = f(y, x) \tag{2.1}$$

$$0 = g(y, x) \tag{2.2}$$

The differential equation of (2.1) expresses the relationships between all machines and dynamic models, such as SVC and HVDC; the algebraic equation of (2.2) is load flow calculation in the power network. By solving (2.1), the "swing curve" of the generator can be obtained.

There are three main stages for transient stability simulation [28]:

(I) Solve the differential equations by discretization

(II) Compute the nonlinear algebraic equations iteratively

(III) Solve the linear algebraic equation.

The solution of the differential and algebraic equations could be classified as General Approach, Partitioned-Solution Approach and Simultaneous-Solution Approach:

General Approach: Solve the differential and the algebraic set simultaneously by using the implicit integration.

Partitioned-Solution Approach (also known as alternating-solution approach): Instead of solving the differential-algebraic equation at the same time, the Partitioned-Solution Approach solved the y of the differential equation by numerical integration and solved the x of the algebraic equation alternately at every time step. Explicit methods are widely used in this approach, including Explicit Euler Method, Open Multistep Formulas, Explicit Runge-Kutta Methods, Predictor-Corrector Methods, and Implicit Multistep Formulas etc. But this kind of method will create accumulated errors due to not reflecting the simultaneous nature of the solution.

Simultaneous-Solution Approach: Change the differential equation to algebraic equation by the implicit integration method. Therefore, all the variables are solved by the algebraic solution.

Since the power system is well known as high dimensional and non-linear system, the drawbacks of the time domain method are that it takes a lot of computer resources and time to solve the large set of non-linear differential-algebraic equations in the time domain method [12]. Also, the "black box" characteristic of assessments in time domain method is not able to identify and evaluate the effect of salient parameters of stability

[23]: it produces only yes or no result without giving the stability margin and its stability result based on offline simulation for a given pre-disturbance situation. Finally, the time domain prediction is only based on the three phase balanced system calculation. If the unbalanced fault happens, the transient analysis needs to be combined with the negative and zero sequence network based on symmetrical components. Therefore, this method needs to be compensated by many other stability simulations. To overcome the limitation of only yes or no answer for the conventional time domain prediction method, reference [29] developed this algorithm by deriving a stability level, which facilitated the operator to make use of the stability margin information.

2.1.2 Equal Area Criterion and Its Extensions

Traditional equal area criterion is a graphical method to find out the stability of one machine connected to infinite bus or a two-machine system. Based on the area comparison between the operation condition of power system before disturbance and the rotor angle movement after disturbance, the stability of system can be determined. The pre and post disturbance area are shown as acceleration area “abcd” and the deceleration area “edgf” in Figure 2.1respectively.

There are mainly four stages of operation points in the equal area criterion [30].

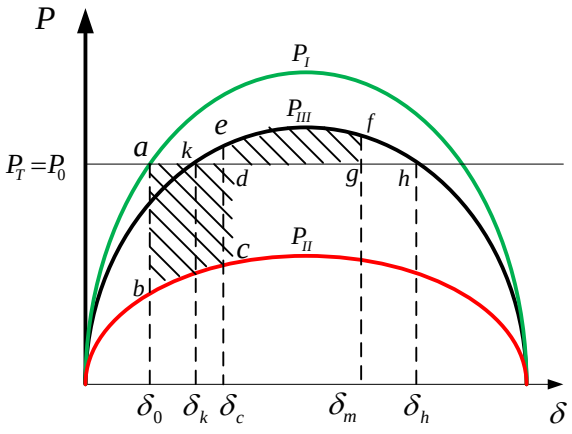


Figure 2. 1: Power–angle curve in the Equal Area Criterion

(1) Normal operation condition (Indicates by "a" in the Figure 2.1): The input mechanical power is equal to the active power of the generator transfers to the infinite system $P_T = P_0$, and the rotor angle is δ_0 during the normal operation.

(2) Fault stage (Indicates by "$b \to c$" in the Figure 2.1): After the three phase short circuit fault, the electromagnetic power decrease to P_{II} , because of the inertia of generator, the rotor angle will not change immediately ($\delta = \delta_0$,the operation point changes from point a to point b). Assuming P_T remains the same value after disturbance, therefore, the imbalance of input mechanical power and the electromagnetic power increases the velocity of rotor ($P_T - P_{II} = J\frac{d\omega}{dt}$). This results in the enlargement the rotor angle ($\delta = \delta_c$).

(3) Clear the fault on time (Indicates by "$e \to f \to k$" in the Figure 2.1): In fact, after the contingency, the relay protection will trip the fault line very fast. Assume the fault is cleared at point c , the electromagnetic power of the system will change to P_{III} . Due to the inertia of the rotor, the rotor angle remains the same and the operation point jumps from c to e . The imbalance between the output electromagnetic power and the input mechanical power brakes the rotor and slows its speed. Assuming the braking process lasts to point f when the speed of the rotor equals to synchronous speed, the rotor angle δ stops growing. However, the generator could not operate at point f since the mechanical power is less than the electromagnetic power, therefore the rotor angle will reduce. The rotor will oscillate around δ_k of the P_{III} curve until the mechanical power equals to the electromagnetic power and the oscillation of the rotor stops due to the damping torque. Finally, the generator will reach a new operation point k to continue to operate. The point k is the intersection of P_{III} and P_T .

(4) Clear the fault too late (Indicates by "$e \to h$" in the Figure 2.2): If the fault line is cut off late, the rotor speed will increase too much. Therefore, after the fault line being cleared, the rotor speed could not turn back to synchronous speed even at point h . After point h , the electromagnetic power is less than the mechanical

power, which speeds up the rotor again, and the accelerated speed becomes larger and larger, the generator and the infinite system will lose synchronism at the end.

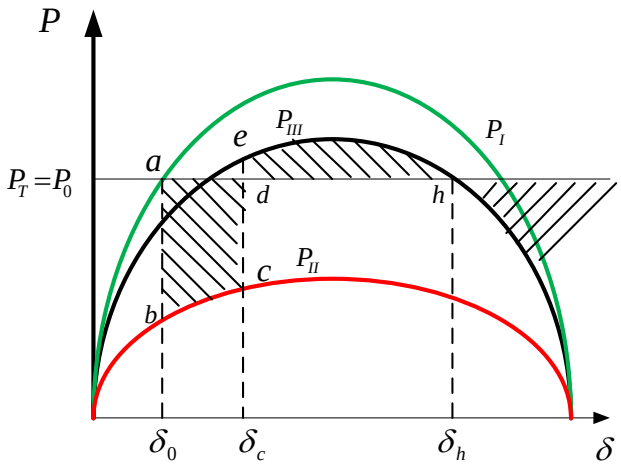


Figure 2. 2: Power–angle curve showing clear the late clearance of fault

According to Figure 2.1 and Figure 2.2, during the time between δ_0 and δ_c. The rotor will speed up because of the remaining torque,

$$\frac{T_J}{\omega_0} \times \frac{d^2\delta}{dt^2} = P_T - P_{II} \tag{2.3}$$

The area of "abcd" shown in Figure 2.1, expressed as the kinetic energy increase during the acceleration process; the area of "defg" shown in Figure 2.1 could be expressed as kinetic energy decrease during the braking process. In mathematics, if the area of the increase equals to the braking area of kinetic energy, the generator will returns to synchronism.

$$\int_{\delta_0}^{\delta_c} (P_T - P_{II}) d\delta = \int_{\delta_c}^{\delta_m} (P_{III} - P_T) d\delta \tag{2.4}$$

The equation (2.4) is called the Equal Area Criterion. The advantage of this method is its accurate results of critical clearing times of breakers besides only "yes" or "no" answer to stability of the system compared to the time domain prediction method. The disadvantage of this method is the over simplicity of the model in stability prediction. The generator uses the second order classical model and the effect of excitation system and the prime machine are neglected. The limitation of this method is its applicability to only one single

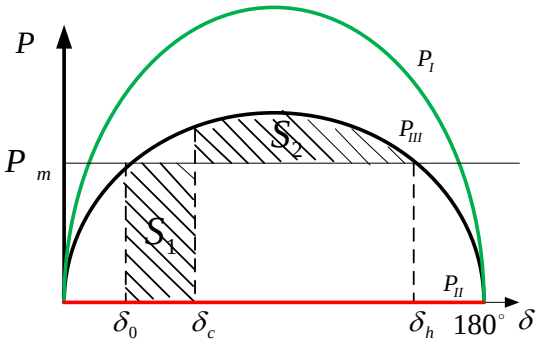
machine or one single machine equivalent system, which is not applicable to the system with complex model and multi swing transient stability analysis.

The extended equal area criterion (EEAC) was developed to improve the equal area criterion. The EEAC aggregates the coherent group of machine into two groups and reduces the system to two-machine like equivalent. Based on the EEAC methodology, the degree of transient stability can be predicted, analyzed and estimated. One significant advantage is it calculates the critical clearing time and stability margin by algebraic equations [31]. The accuracy and robustness is improved and it overcomes the limitations of the conventional equal area criterion (EAC) by analyzing only one generator to infinite bus, and extending it to multi-machine system. This is realized by clustering two subsets of generators. Therefore, the two equivalent machines are reduced into a one machine-infinite bus. A dynamic equivalent power system model based on a Taylor series expansion was developed to predict the transient stability at a fast speed [32]. Reference [33] improves the EEAC method for efficiency and accuracy by proposing sensitivity based iterative method to search the limit of multi-swing transient stability. The complicated phenomena effects such as non-monotonic variation of the stability margin and the isolated stability domain are studied.

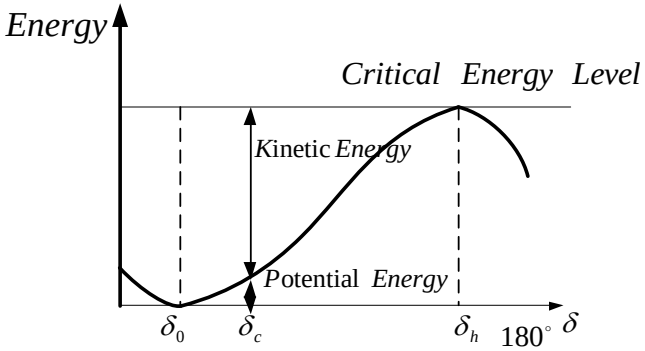
2.1.3 Energy Function

The energy function approach is a direct method, since it predicts the stability without explicitly solving the system differential equations. The time domain approach assesses the transient stability by repeatedly observing power system response after a contingency during the first few seconds, while the direct method constrains the observation within a much shorter time span, which shows its advantage of computing efficiency over the time domain method. The transient stability in energy function is well known and has been compared to the rolling ball theory [30]. If the ball remains in the bowl that means the system remain stability, if the ball rolls out of the bowl, the system falls into the region of instability. When the ball is subjected to certain kinetic energy of certain direction, the ball will roll along with the direction of initial motion. During the movement of the ball, the kinetic energy is transferred to potential energy. Therefore, the stability of system depends on the height of the rim and the initial kinetic energy of the

ball. Similarly, a power system operating at a stable equilibrium point can be seen as the ball staying at the bottom of the bowl. After a fault, the generators accelerate and the equilibrium point is disturbed. During the fault, the power system gains kinetic energy and the potential energy. After the fault clearing, only the potential energy exists, since all the kinetic energy convert to the potential energy. The Energy Function Approach is equivalent to the Equal Area Criterion as shown in Figure 2.3.



a) Power-angle relationships



b) Energy-angle relationships

Figure 2. 3: Equivalence of Equal Area Criterion and Energy function

When a system is subjected to a disturbance, the rotor of the generator will swing away from the pre-fault stable point, which is equivalent to the ball leaving its equilibrium point and moving up to the potential energy surface. At δ_0, the system is in equilibrium point as it is a pre-fault condition, δ_c is called the critical clearing angle since it is only when system clears the fault before the generator angle reaches δ_c that the system remains within the stability boundary. During the contingency process, the accelerating area of S_1 is equivalent to the kinetic energy and potential energy together in

the disturbance (before the angle δ_c). Only if the summation of the two energies is less than the critical potential energy at δ_h, can the system stay in synchronized operation. In the direct method, the degree of stability is measured by comparing the critical potential energy of the pre-fault system to the amount of energy that is injected into the system during the contingency.

There are two aspects to identify the severely perturbed machines. First is the relative kinetic energy of machine in the early stage of perturbation; second is the generator electromagnetic power variation which reflects the topology change of system. The energy function method analyzes the stability limits with analytical sensitivity approaches.

One of the advantages of the energy function method is that the system dynamic equations' integration is limited only to the fault-on period, saving computational resources, which avoids the repetitive integration of the post-fault system to decide the critical clearing time and solving the system equations continuously [34]. Like the other direct methods, the disadvantage of the energy function method is that many simplifying assumptions are needed [35]. For example, the generators of system represented as the classical models and the loads are using constant impedance models. The direct method may also converge to a wrong controlling unstable equilibrium point (UEP). Furthermore, it only predicts the first swing stability, which means if the second swing of the rotor generator goes unstable, the direct method will fail to predict it. The controlling UEP has to be calculated by an iterative procedures and may have serious convergence problems during computation. Development has been made to make the method become less conservative, for example, using the concept of potential boundary and its application, the UEP was taken into account[36,37]. By the sensitivity method, the transient energy margin could be obtained with fast deviation [38].

2.1.4 Hybrid Method

The direct method which is usually the equal area criterion method has the limitation of application in the interconnected and more complex power network. The time domain simulation method has its advantage of unlimited modeling capability. A

practical hybrid method that combines the time-domain simulation and the direct method of transient stability analysis together was first proposed by Maria et al. [39]. Several researchers in the past have incorporated the conventional time domain simulation and transient energy analysis. By the time domain simulation, the real system trajectory is computed, and by the transient energy function, a stability index of security assessment can be derived. The advantage of this method is no convergence problem will occur since it avoids calculating the controlling unstable equilibrium point (UEP). During the offline application, by reducing the times of stability cases run, the hybrid method speeds up the stability limits calculation. Multiple criteria have been applied to accelerate the online simulation of the hybrid method. The stationary phase method (SPM) extended from the hybrid method studies both the dynamic state variables and algebraic state variables at the same time. A general approach is proposed in [40] for transient stability filtering, ranking and assessment (FILTRA). This method based on making up two blocks relied on single machine equivalent (SIME) that screen contingencies in first block, rank and assess the "potentially interest" in the second one. In reference [41], a judgment of the severity index in transient stability is established. It combines the decision tree method and the integral square generator angle (ISGA) for ranking and screening the severity of the fault. Other innovative research based on hybrid methods such as the potential energy approximation [42] and transient energy function correction [43], are developed in recent years.

2.1.5 Artificial Neural Networks

Modern predictions based on phasor measurements include the fuzzy system and machine learning based method. The machine learning based method mainly consists of Artificial Neural Networks (ANN), time series prediction, Decision Trees (DTs) and Support Vector Machines (SVM), etc. They have been greatly developed in the past decades. There are mainly three types of criteria for transient stability detection [44]:

- Threshold-type criteria: rotor angle and rotor speed prediction.
- The differentiation-type criteria: energy function.
- The integration-type criteria: extended equal area criterion.

The Artificial Neural Networks (ANN) were first developed from the concept of biological neural networks. It is a mathematical model that seeks the relationship between the input parameters and the output results.

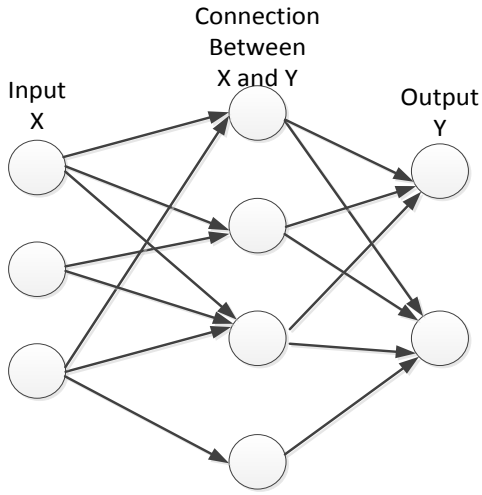


Figure 2. 4: Illustration of the relation in Artificial Neural Networks

Defines the input as X, the output as Y, by finding a class function

$$f: X \to Y$$

The elements of the class could be gained by the changing parameters, weights of connection, or the architecture, etc. The statistic modern neural network is mainly a non-linear statistical data modeling method. With the development of statistics and signal processing, the neural network is applied to larger systems combining both with adaptive and non-adaptive parameters. There are multiple layers of the system. The first layer is the input neurons, and the last layer contains the output neurons. Those layers between the first and the last layer are the connections. The ANN aims to determine the inner connections between the number of the neurons and their parameters. This is mainly done by an optimum mathematic method to minimize the cost criterion.

Neural networks have many advantages over the traditional methods in predicting the stability of large power system. In ANN application, a suitable training data set first needs to be created in the stability assessment. Usually, a large amount of different contingencies are simulated in the system so as to get various pre-disturbance system

features and the correspondence of system security index, such as the Critical Clearing Time (CCT) [45] and the system Energy Margin (EM) [46]. Feature selection is very important in choosing the features that are most useful to the discrimination ability of neural network [47]. With the advent of the application of Phasor Measurement Units (PMUs), the online stability prediction by ANN becomes available. Reference [47] applied changing rate of bus voltages and angles for six cycles after the fault clearance to train ANN with two layers to predict the system stability. Using the statistical analysis as well as the Fisher's linear discrimination, the selection of neural network training features could be determined for the dynamic security evaluation. The major drawback of the neural network is an extensive training process is required.

2.1.6 Fuzzy System

In power system, different control goals and actions should be taken according to the prevailing operation conditions. The most common approach of taking advantage of the system's qualitative knowledge is to design a fuzzy control. At the early stage, the fuzzy system is mainly based on "if-then" logic control statements. They basically follow a linguistic model, which can qualitatively illustrate two types of complex system situations, the input and output's nonlinear relationship, and the system's uncertainty [48,49].

With the development of Phasor Measurement Units, reference [16] predicts the transient stability very fast and along with high- speed control. By constructing all fault locations as well as a post-fault phasor measurement data window for prediction, it is unaffected by the power system changes.

Fuzzification means converting the variables that control a certain process to fuzzy values by using fuzzy membership function (MF) [50]. MF is a curve that indicates how each of the input values is mapped to the fuzzy value between 0 and 1. In the transient analysis, for the input fuzzy set of the power system would be the voltage magnitudes, bus frequencies, rotor angles, etc. In the doctorate dissertation [51], the voltage magnitude and bus frequency are used for load shedding control. By constructing the fuzzy rules by linguistic variables, the relationship between the inputs and outputs are described. Different combinations of the input yield different results of the output. In

reference [51], various combinations of frequencies and voltage magnitudes yield various control schemes, such as generator tripping, load shedding or no action requirement.

2.1.7 Time Series Prediction

Time series methods are mainly based on the collection of measurement along with time records the chronological behavior of power systems. The historical data is used to predict the future trend. This prediction based on the rotor angle trajectory extrapolation by curve fitting technique, approximates or estimates of the future behavior called the time series model. The most common time series models include the polynomial function model, the auto regression model and the trigonometric function model.

(1) The polynomial model is based on the prediction method. Reference [52] uses the truncated Taylor series expansion to express the generator angular speed and the rotor angle. Later, a polynomial function is proposed in reference [53] to estimate the rotor speed with the least square method. The polynomial model can be written as:

$$\theta(\mathrm{t}) = \mathrm{b}_1 + \mathrm{b}_2\mathrm{t} + \mathrm{b}_3 t^2 \ldots + \mathrm{b}_\mathrm{n} t^n \tag{2.5}$$

where $\theta(\mathrm{t})$ is the estimated value of the rotor angle at time t. $B_N = [\mathrm{b}_1, \mathrm{b}_2, \ldots \mathrm{b}_\mathrm{n}]^T$ is the parameter vector of the model. n is the order of the model. Applying the least square method, the parameter vector B_n can be solved by

$$B_N = [H^T(N)H(N)]^{-1}H^T(N) \cdot \emptyset(\mathrm{N}) \tag{2.6}$$

in which, $\emptyset(\mathrm{N})$ is the vector of generator rotor angle. $\emptyset(\mathrm{N}) = [\,\theta(0), \theta(\Delta\mathrm{t}), \ldots, \theta(\mathrm{N}\Delta\mathrm{t})]^T$. $\theta(\mathrm{N}\Delta\mathrm{t})$ is the generator rotor angle at time $\mathrm{t} = \mathrm{N}\Delta\mathrm{t}$,

$$\mathrm{H(N)} = \begin{bmatrix} 1 & 0 \cdots & 0 \\ 1 & \Delta\mathrm{t} \ldots & (\Delta\mathrm{t})^\mathrm{n} \\ \vdots & \vdots & \vdots \\ 1 & \mathrm{N}\Delta\mathrm{t} \ldots & (\mathrm{N}\Delta\mathrm{t})^\mathrm{n} \end{bmatrix} \tag{2.7}$$

The i step of the future rotor angle can be calculated through

$$\theta(\mathrm{mt}) = \mathrm{b}_1 + \mathrm{b}_2\mathrm{m}\Delta\mathrm{t} + \mathrm{b}_3(\mathrm{m}\Delta\mathrm{t})^2 \ldots + \mathrm{b}_\mathrm{n}(\mathrm{m}\Delta\mathrm{t})^\mathrm{n} \tag{2.8}$$

$$\mathrm{m} = \mathrm{N}+1, \mathrm{N}+2, \ldots, \mathrm{N}+\mathrm{i}$$

The new observation vector $\emptyset(N+1)$ can be rolling predicted with the new measurement coming.

(2) The regressive method [25] includes: Autoregressive Models (AR), Moving-Average Models (MA), Auto Regressive Moving-Average Models (ARMA).

In the AR, the p^{th} order is expressed as the polynomial function (2.9)

$$\theta_t = \emptyset_1\theta_{t-1} + \emptyset_2\theta_{t-2} + \cdots + \emptyset_p\theta_{t-p} + e_t \quad (2.9)$$

$\emptyset_i$ is the autoregressive parameter of order i, θ_{t-i} represents the past ith value of the time series, and e_t is the random error at the time period t. This equation stands for the future value θ_t that is proportional to the past p values plus a random error. If i equals one, the equation refers to an AR order one (AR(1)), which means the next time point value in the future only relates to the current time point value.

In the MA, the pth order is expressed as (2.10)

$$\theta_t = \varphi_1 e_{t-1} + \varphi_2 e_{t-2} + \cdots + \varphi_p e_{t-p} + e_t \quad (2.10)$$

φ_i is moving average parameter of order i, e_{t-i} represents the random error, past i^{th} value of the time series, and e_t is the random error during the time period t. Therefore the forecast value at time t is proportional to the random error of the past period.

In the ARMA Models, combining the AR and MA method, the prediction of the future value is related to the past values as well as the previous random errors. Take pth order AR and qth order MA combination for example, it is expressed as (2.11)

$$\theta_t = (\emptyset_1\theta_{t-1} + \emptyset_2\theta_{t-2} + \cdots + \emptyset_p\theta_{t-p}) + (\varphi_1 e_{t-1} + \varphi_2 e_{t-2} + \cdots + \varphi_p e_{t-q}) + e_t \quad (2.11)$$

All three time series models are used in forecasting the stationary series, which means if the time series data continuously changes along with time, the differentiation method needs to be implemented. Differentiation is realized by creating a new variable Z_t that equals to the difference between two successive values, shown as (2.12).

$$Z_t = \theta_t - \theta_{t-1} \quad (2.12)$$

If the order one differentiation is not enough for stationary, higher differentiation will be applied. After differentiation, the time series model type and order need to be determined by autocorrelation and partial autocorrelations [54]. After the selection of the proper estimated model, the parameters of the model need to be determined. The common method of doing so is by the least square mathematical method.

With the installation of synchronized Phasor Measurement Units (PMUs) in power systems, it is now possible to monitor the behavior of the system in real time and provide important information for transient stability assessment and enhancement. Reference [55] achieves fast stability estimation by using the reduced-order model, which is derived by detecting the post fault swing curve data and clustering them into coherent group. But this kind of method requires the topology and parameter data of the network, which are hard to satisfy in the real system. The rotor oscillation prediction method [25] based on the time series predicts the transient stability of power systems in real-time, but it also needs to initially identify the coherent groups of generators following a contingency by PMUs measurement. The second step is determining the time series model. Using the autoregressive models order two combined with differentiation process (ARI(2,1)) shown as (2.13)

$$\theta_t = \emptyset_1 z_{t-1} + \emptyset_2 z_{t-2} + e_t \tag{2.13}$$

Then the center of the coherent group is determined by (2.14)

$$\mathrm{COA} = \frac{\sum_{i=1}^{n} \mathrm{MVA}_i * \theta_i}{\mathrm{MVA}_{total}} \tag{2.14}$$

With respect to the inertial center of a power system, the rotor angle of a generator can be used to judge whether this generator loses synchronism with the system. Forecast the voltage angle 10 points in the future, COA_{t+10}, and predict the stability by comparing against the center of angles. The rule of thumb of the instability is, more than 150 degrees difference between a given voltage angle and the center of angles in two consecutive predictions. Therefore, the transient stability could be determined.

However, this method requires more than 300 milliseconds to predict the stability after the start of the transient event. If we do not predict the stability of the system after the contingency immediately and correctly, it may lead to cascading failures which will eventually culminate in a widespread blackout. Therefore, there is an imminent need for a faster and more accurate analytical real-time prediction technique.

(3) The trigonometric function model prediction [56]: the rotor trajectory prediction can also be realized with trigonometric curve fitting as below.

$$\theta(t) = \sum_{n=0}^{\infty} a_n \cos nt + b_n \sin nt \tag{2.15}$$

in which n is the model order. The parameter a_n and b_n can be estimated by the least square method.

Three steps of the time series prediction method are: First, the order of model needs to be determined. Second, identify the model parameter by the least square method. Finally, extrapolate the rotor angle trajectory. The advantage of curve fitting based on rotor trajectory prediction method is no needs for network topology, model details of power system. Therefore, the required data for the model parameters estimation are less.

2.1.8 Decision Tree

The Decision Tree Method belongs to the pattern recognition type approach, which has been widely used in transient stability prediction, about 88.6% of all the prediction methods [21]. It merits the advantage of high computing efficiency and assessing the affecting extent of different parameters in stability prediction. The method provides a straightforward analysis via mining the data and building an offline decision tree. According to this information, are based on the "if-then" criterion of certain important attributes, the transient stability of power system will be forecasted. The ultimate goal of the Decision Tree Transient Stability (DTTS) is providing the reliable information for the operator in short-term operation planning and real time monitoring. The simple controllable attributes will generate a simple tree, easy to be understood and applied to the sensitivity analysis, whereas, the complex attributes will generate a complicated but reliable decision tree [23]. By simulating many contingency cases offline,

a straight forward decision tree can be built. The decision tree is trained by a large number of cases called "training set" (LS), which is a sample set of pre-classified states and a list of attributes. The more offline cases are trained, the more real cases can be applied to the prebuilt decision tree. The decision tree mainly split the data into two subsets in each node in order to get the maximal purified partition. The tree will define the optimal attributes and their optimal value. The tree will stop growing when one of the terminal criteria is met:

- The class of a terminal node is "sufficiently" pure. The required impurity is fixed by a small threshold.
- The accuracy of the tree is not possible to improve for further partition. Therefore, the node is called a "deaden terminal node".

A testing set must be used to test the accuracy of decision tree that has been built. To get a reliable and unbiased estimation, the test data set needs to be large enough. In CART[@], all the data are divided to 10 groups randomly, nine of them are used as training data to build the decision tree, while the remaining one groups is used for testing the pre-built decision tree. The idea of decision tree is simple, illustrated as Figure 2.5

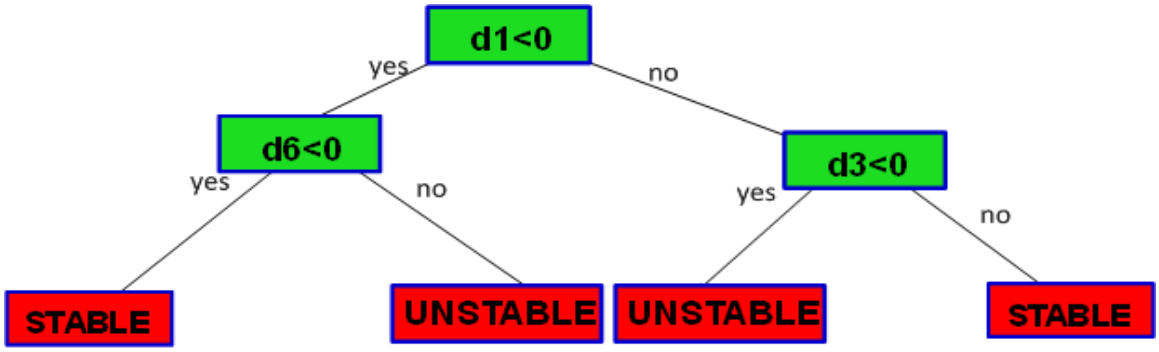


Figure 2. 5: A simple decision tree for illustration purpose

In figure 2.5, suppose there are two groups of data, "STABLE" and "UNSTABLE". To classify the two groups of data, two depths of decision tree are needed. d1, d3 and d6 are selected by the decision tree as important attributes for the data

partition. If the system satisfies the rule d1<0 and d6<0 or meet the criterion d1>0 and d3>0, then the system is classified as stable, otherwise the system is classified as unstable.

The decision tree has been widely applied to predict the power system stability. DTTS can be used to predict the stability and instability in two class classification. In reference [23], the critical clearing time (CCT) is determined by the decision tree. Reference [24] classified the transient swing as either stable or unstable based on the real time PMUs measurement. In reference [57],the decision tree not only forecasted whether the system is in a potential hidden failure or in a likelihood cascading situation, but also optimized the PMUs installation locations for monitoring the system state.

2.1.9 Support Vector Machine Method

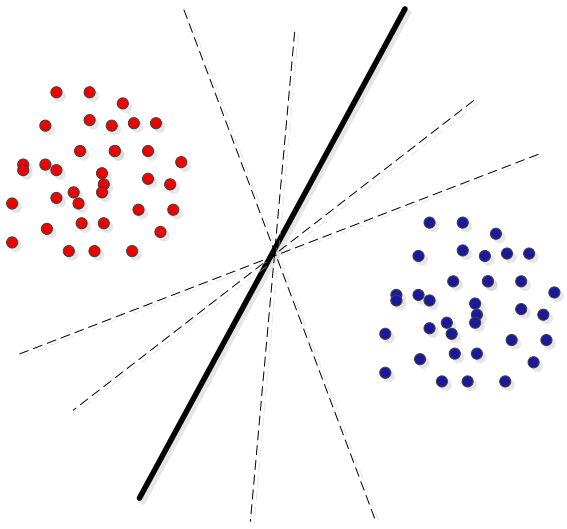


Figure 2. 6: Support vector machine classifier

The goal of the support vector machine is finding a classifier to maximally split the data [58]. Considering two groups of data, the red dots and the blue dots respectively, are in two dimensions shown in Figure 2.6. There are many possible linear classifiers that could partition the data; however, there is only one classifier that maximizes the margin, which means maximizing the distance between the nearest data point of the two groups and the classifier, shown as the solid line in the Figure 2.6.

The hyperplane that optimally split the two groups of data is expressed as:

$$|\langle w, x\rangle + b| = 0 \tag{2.16}$$

where x is the arbitrary data point in the hyperplane.

The shortest distance from a point x to the hyperplane is:

$$d = \frac{|\langle w, x\rangle + b|}{\|w\|} \tag{2.17}$$

The optimal hyperplane separating the data is the one that minimizes (2.18)

$$\varphi(w) = \frac{1}{2}\|w\|^2 \tag{2.18}$$

subject to

$$y_i(\mathrm{w} \cdot x_i - b) \geq 1 \tag{2.19}$$

The solution of the optimization problem given by the Lagrange function

$$\min_{w,b} \max_{a \geq 0} \{\varphi(w)\} = \min_{w,b} \max_{a \geq 0} \left\{ \frac{1}{2}\|w\|^2 - \sum_{i=1}^{n} a_i [y_i(\mathrm{w} \cdot x_i - b) - 1\ ] \right\} \tag{2.20}$$

where a are the Lagrange multiplier. To minimize the function with respect to w and b

$$\begin{cases} \frac{\partial \varphi}{\partial w} = 0 \\ \frac{\partial \varphi}{\partial b} = 0 \end{cases} => \begin{cases} \sum_{i=1}^{l} a_i y_i = 0 \\ w = \sum_{i=1}^{l} a_i y_i x_i \end{cases} \tag{2.21}$$

The coefficients of the hyper plane are $w = \sum_{i=1}^{l} a_i y_i x_i$; $b = \mathrm{w} \cdot x_i - y_i$. The corresponding x_i is called the supporting vectors. The supporting vectors are line on the margin which satisfies:

$$y_i(\mathrm{w} \cdot x_i - b) = 1 \tag{2.22}$$

When generalized to a high dimension, a hyperplane in the higher dimension is constructed.

The optimization equation becomes,

$$a^* = \arg\min_a \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l} a_i a_j y_i y_j K(x_i x_j) - \sum_{k=1}^{l} a_k \tag{2.23}$$

where $K(x_i x_j)$ is the kernel function which is the non-linear mapping into feature space. It is the inner-product kernel between a "support vector " x_i and the vector x from the input space, shown in Figure 2.7:

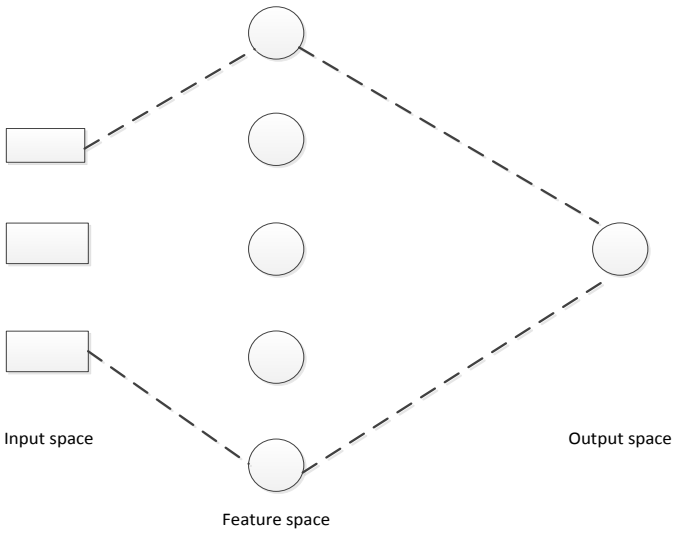


Figure 2. 7: Mapping the input space to higher dimension feature space

Some studies have shown that the SVM gain has higher prediction accuracy and less computation time compared to the conventional ANN and other machine training based methods [19,20]. Compared to the Neural Network (NN) approach in transient stability analysis, the SVM has a significant advantage extracting feature in high dimensional space. Therefore, it is more practical to be implemented. The SVM is also feasible for dynamic security classification [59].

2.2 Method of Controlled Islanding

The idea of controlled islanding based on slow coherency was first proposed in reference [7]. If the relay protection does not stabilize the system and different coherent groups are recognized, then it is necessary for the system take the controlled islanding measures to separate the system to several subsystems before the perturbation travels

from one part of the system to another. This avoids the cascading consequence of the power system. The controlled islanding is considered self-healing and can effectively constrain the disturbance within a small region.

In an n buses power system, all the possible splitting lines combination number is calculated by 2^n. This is time consuming to find the combination of tripping line to split the system into islands. Therefore, it is necessary to develop a feasible controlled islanding scheme to determine the islanding boundaries in real time. Currently, the methods developed for controlled islanding are few. One popular islanding scheme is studying the power flows on the critical lines. The controlled islanding boundaries are chosen by studying the electrical center or the active power in transmission lines in the system. These two methods require exhaustive offline simulation to select the critical buses as the potential electrical center or the critical lines with large power flow changes. Another prevalent islanding scheme is graphical based islanding, called Ordered Boundary Decision Diagram (OBDD).

2.2.1 Active Power Changed Based Controlled Islanding

One popular islanding scheme is performed by studying the active power of transmission line. In reference [60], the intersection lines for controlled islanding are determined by observing the active power on both ends of the line in the system. By simulating all possible scenarios comprehensively in the power system, and if the changing rate of the active power in the pre-specified transmission lines between two concessive sampling time excess a threshold value, then the controlled islanding is considered to be a remedial measure for the controlled protection scheme. The value of threshold is influenced by the topology change and generation/demand growth. By monitoring the active power of the pre-specified lines in real time, those lines trip to form the islands when the threshold is met.

The main disadvantages of the pre-determined boundary method are that the generators in the same island may be not coherent and it may also result in the imbalance between the load demand and generation output, which resorts to undesirable generation dropping and load shedding [61].

2.2.2 Electrical Center Based Controlled Islanding

Another feasible islanding scheme is performed by detecting the electrical center in the power system. The electrical center is formed when the difference of phasor angle between both ends of the transmission line is 180°, which is detected by the R-Rdot Out of Step relay. The electrical center may dictate the islanding boundaries [62]. This kind of method is based on some assumptions [61]:

1) The out of step scenarios that caused uncontrolled islanding are independent of fault locations, type of fault, and fault severity.

2) The locations for splitting the system into different islands only depends on the prevailing system topology and load injection, which means the Out of Step relay can only install at some critical locations to create the islands.

3) The loss of synchronism does not appear at the same time, which allows the Out of Step relay to operate sequentially.

Under large disturbances, the R-Rdot Out of Step relay installed at the pre-selected lines will trip and the chosen island will be formed. To prevent the Out of Step relay from being tripped wrongly during the oscillation of system, the ohm relay applied "blinders" to discriminate the swing locus of the out of step tripping and swing cases. The ohm relays installed at the pre-determined boundary locations are adjusted with the desired settings.

The electrical center based controlled islanding has the same disadvantage as the active power changed based controlled islanding, which cannot guarantee the load-generation balance in the island. Moreover, when the system topology changes, the electrical center will shift to other locations. Controlled islanding that is only based on the R-Rdot relay may not always split the system into separate islands.

2.2.3 Slow Coherency Based Controlled Islanding

In this approach, slow coherency based grouping is deployed to split the generators into different coherent groups. Within each island, the imbalance between the generators' output and the load demand is minimized. After the recognition of coherency,

exhaustive searching of boundaries combination is developed in reference [7]. Reference [63] further refined the searching of the islanding boundaries by a maximum flow with minimum cut, using a graph based scheme. Reference [64] implements the load shedding that takes into account the declining frequency to avoid frequency instability in the load rich system. Combining with the graphical approach, reference [65] studied the controlled islanding in the large scale system by the network simplification algorithm. Furthermore, the generators that belonged to the same coherent group are represented by a fake node, and the islands are split by resorting to the prebuilt graph partition library.

The advantage of slow coherency criterion is that it takes into account the dynamic behavior of the generators. The result of the slowly coherent grouping is independent to the disturbance, which reveals the intrinsic trends of the generators' behavior after disturbance.

2.2.4 OBDD Based Controlled Islanding

Another possible islanding scheme is determined by the graphical connection searching. Ordered Boundary Decision Diagram (OBDD) [66], based on graph theory, was proposed to investigate a proper splitting scheme to satisfy the "equality" and "inequality" constraints. "Equality" constraints mean the power balance between the generators and the loads in the static state; "inequality" constraints mean the generators or transformers load do not exceed the rated values, as well as the transmission capacity is below the thermal or static stability limits. The first step of the OBDD is applying the node-weighted graph checking algorithm so all possible islanding boundaries could be determined. The checking algorithm is a graphical search theory that builds the Boolean matrix of the branch connection and check the network connection by a set of polynomial operations. The load nodes were selected to meet the power balance, and it determines the branch needs to be tripped. The second step of the OBDD is the power flow checking with the dc power-flow method to check whether the islanding scheme satisfies the security limits. The OBDD method is able to find out all the existing islanding schemes via searching all possible combination of the islanding boundaries.

The idea of OBDD method is described below: First, it is necessary to build a real power system with a graphic model. For a system with n buses, the weighted undirected

graph can be expressed as $G(V,E,W)$, in which $V=\{v_1,v_2,...v_n\}$ is the bus set, $W=\{w_1,w_2,...w_n\}$is the weighted set of buses, and E is the edge set. In which, v_i is the i^{th} bus in system; $w_i=P_i$ when not considering the loss of the network, and P_i is the active power injection of bus i. E stands for the transmission line of the network. The elements within the edge set have the relation $e_{ij}\in E$, and if there is a transmission line between bus i and bus j, then $e_{ij}\neq 0$. The transmission line status can be dictated by the Boolean value

$$b_{ij}=\begin{cases}1 & connected\\ 0 & disconnected\end{cases} \tag{2.24}$$

Define $m=SIZE(E)$, which stands for the number of the lines. One possible islanding boundary is expressed as $\gamma=\{b_{i1j1},b_{i2j2},...,b_{imjm}\}$. All the combination of the possible Boolean values which can partition the system constitutes the strategy space.

The OBDD aims to find all the possible islanding boundaries instead of the optimal solution of islanding. Therefore, a series of constraints need to be added to sift the solution. There are three main constrains when splitting the system:

1) Separation and Synchronization Constraint (SSC): requires the generators belonged to the same coherent group to be separated to the same island.
2) Power Balance Constraint (PBC): after controlled islanding, each island remains the load-generation balanced within each island.
3) Rated value & Limit Constraints (RLC): requires that the generator or transformer load do not exceed the rated values, as well as the transmission capacity below the thermal or static stability limits

The advantage of the OBDD method is its feasibility of solving the Boolean variables based decision making problems. In power systems, the Boolean variables indicate the transmission line status, generator tripping, or load shedding. The disadvantage of the OBDD method is the large computer resources that are required for its huge searching space.

Chapter 3 Apparent Impedance Trajectory

3.1 Apparent Impedance Trajectory Definition

Apparent impedance expresses the relationship between the current and the voltage in an electrical circuit, which is the quotient of the positive sequence rms voltage between the two terminals of the transmission line by the rms value of the positive sequence current in the transmission line. The mathematical expression of it is:

$$Z = \frac{E}{I} = R + jX \tag{3.1}$$

Displayed in the R-X diagram:

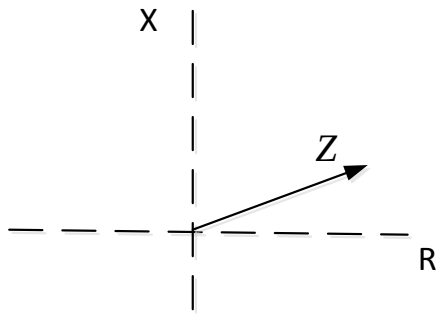


Figure 3. 1: Apparent impedance in R-X diagram

The trajectory of apparent impedance can be expressed as a data set of apparent impedances. This trajectory can be expressed with respect to time and displayed in a R-X diagram as shown in Figure 3.2.

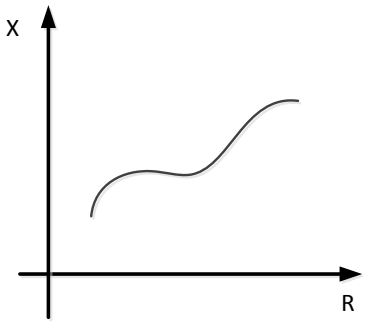


Figure 3. 2: Apparent impedance trajectory example

3.2 The Application of Apparent Impedance in Modern Power Systems

Many adaptive protection schemes have been implemented based on the apparent impedance. Two types of relay protection scheme are introduced in this section: Distance Relay and Out of Step Relay.

3.2.1 Distance Relay

The distance relay measures the apparent impedance of a line to locate the fault position. The apparent impedance is calculated by comparing the voltage drop between the fault and the device location. Here is a simple example to illustrate how the apparent impedance is calculated from relay 1.

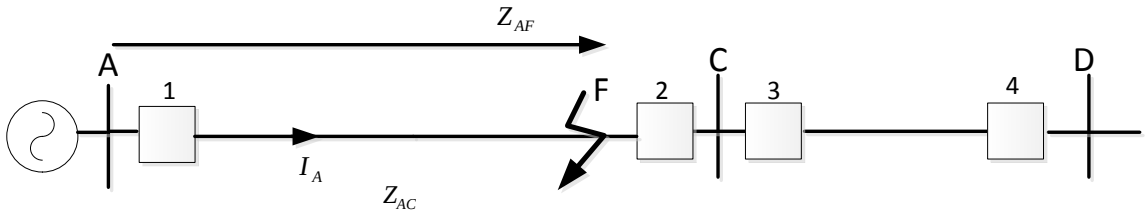


Figure 3. 3: Apparent impedance calculation example

Voltage at Terminal A is:

$$V_A = V_{AF} = I_A Z_{AF} \tag{3.2}$$

Impedance that is measured by the relay in Terminal A:

$$Z_{AF} = \frac{V_A}{I_A} = R_{AF} + jX_{AF} \tag{3.3}$$

The distance relay settings are calculated based on the apparent impedance measured at this location. The settings typically consist of three zone protection. Shown in Figure 3.4,

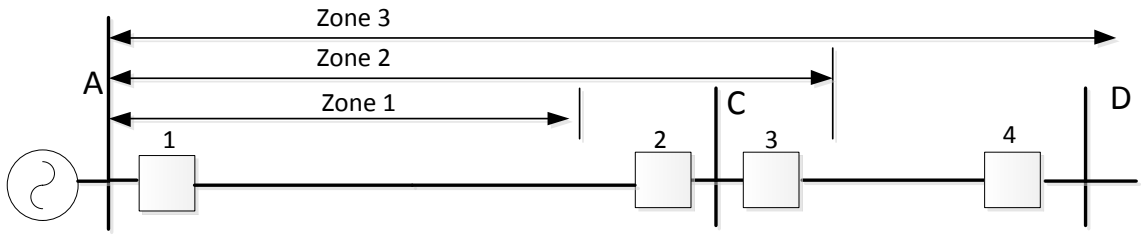


Figure 3. 4: Three zone distance relay of a transmission line

The zone 1 of relay 1 protects most of the transmission line AC, about 80% of the line's length. Zone 2 of relay 1 protects the entire line, AC, and the beginning of line CD. The protection of zone 1 will operate very fast while the protection of zone 2 operates with a time delay to make sure the zone 1 of line CD (relay 3) trips first for a fault within its protective zone. Adding zone 3 protection provides the remote backup for the adjacent circuits.

As shown in Figure 3.5, the characteristic of distance relay is a circular shape. The inner circle to the outer circle is zone1, zone2 and zone3. The fault at location C will lie in zone2 of relay protection and the fault at location D will lie in zone3 of relay protection.

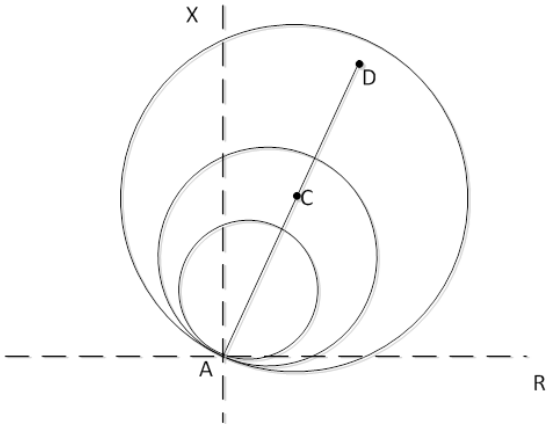


Figure 3. 5: Three zone of protection characteristic

3.2.2 Out-Of-Step Relay

The apparent impedance trajectory has been widely used in the Out of Step relay of power system protection especially for detecting loss of synchronism in Out of Step relays. Usually, Out of Step relays are located at the critical locations to detect the unstable swing of the system. The Out of Step relay is used in controlled islanding to split the system into two or more islands when instability occurs between two or more regions that may lead to loss of synchronization. Currently, the splitting places are predetermined to make a feasible balance of generators and load in each area.

The Out of Step relay needs to discriminate the system out of step condition and the system oscillations. With fiber-optic communication, the microprocessor based Out of Step relay can be set with a "blinder" to block the system oscillation condition [67]. Shown as Figure 3.6, the "blinder" refers to the angle range set by the distance relay. This angle range can be narrowed to the desired extent. The different locus of apparent impedance is drawn in the R-X plane to express various swing angles. Without the "blinders", the tripping area of the Out of Step relay range is $[90^{\circ},240^{\circ}]$. The "blinders" in the Figure narrow the range of the swing angle to $[135^{\circ},195^{\circ}]$. During the Out of Step contingencies or large swing cases, the Out of Step relay may trip if the swing trajectory passes the tripping area.

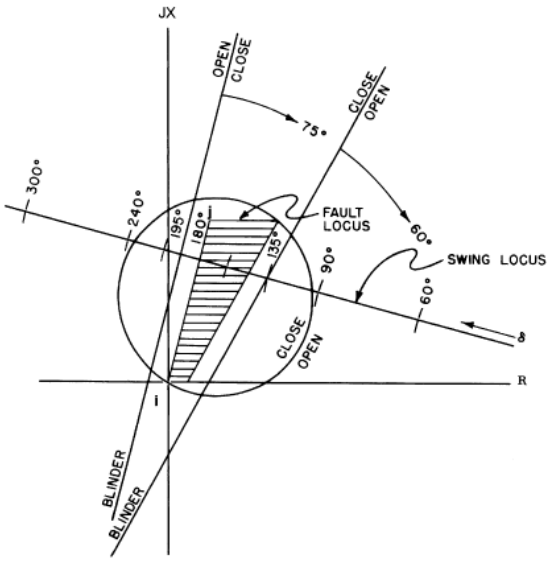


Figure 3. 6: Swing blocking relay in controlled islanding [61]

Out-of-step relays, such as the R-Rdot relays, monitor the speed of the system trajectory in the R-X subspace. If the swing between two groups of generators goes unstable, the apparent impedance trajectory will traverse the trip zones of the relay in a very short time [67]. By monitoring the movement of the apparent impedance, the Out of Step relay will judge if the system swing is stable or not. Suppose the relay detects two apparent impedance trajectories Z1 and Z2 respectively, shown as Figure 3.7 [68]. The former trajectory corresponds to a stable swing, while the latter one corresponds to an unstable swing. The stable trajectory traverses the R-X plane slowly and it is caused by the oscillation of the system, while the unstable trajectory creates a quick movement in the plane and it is caused by a fault. To further confine the Out of Step relay to trip correctly, an auxiliary relay acting as an additional parallel blinders needs to be added to the Out of Step relay. The paralleled blinders (inner blinder of the shade area in Figure 3.6) acts as an elapse time setting ΔT . The auxiliary relay will block the Out of Step relay from tripping when the time of the apparent impedance locus takes to cross the area between the outer blinder and the inner blinder is larger than the pre-set time interval ΔT . Typically, the elapsed time ΔT is about ¼ to 1 cycle.

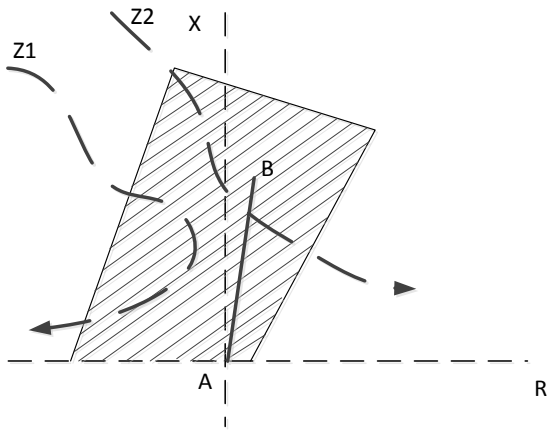


Figure 3. 7: Stability swing in the zone of Out of Step relay [68]

This adaptive relaying scheme based on a PMUs ability to monitor system stability in real-time and identify oscillations that indicate out of step conditions in the power system [69,70]. As it is mainly based on the assumption of a two-machine like system, the coherence identification must be pre-calculated to modify the relay settings. By studying the locus of the apparent impedance in the R-X plane, transient stability or instability can be detected by the Out of Step relay. However, to detect transient stability or instability, the coherencies of the two groups of generators need to be initially identified and the relay parameters adjusted as the system conditions change [71,72]. The oscillating generators may develop to many coherent groups, but the Extended Equal Area Criterion is still valid in these evolving generator groups. By studying the locus of the apparent impedance in the R-X plane, transient instability of the system could be detected by the Out of Step relay.

3.3 Summary

In this chapter, the concept of apparent impedance and its application in relay protections are presented. Distance relay operates by comparing the apparent impedance between the relay and the fault location with the set value; while the Out of Step relay detects the instability by analyzing the locus of apparent impedance trajectories. However, these kinds of relays can only detect the instability/out of step of the system instead of predicting it. In power system operation, failure to predict the instability and take immediate controlled protection in time may lead to cascading failure of the system or even a blackout. Therefore, a transient stability prediction method based on apparent impedance trajectories and data mining technology is proposed in this dissertation. It is noticed by our research that the apparent impedance trajectories can characterize the trend of the system's transient stability. Its methodology will be illustrated in more details in Chapter 5.

Chapter 4 Data Mining in Power System

4.1 Data Mining Definition and Methods

4.1.1 Data Mining Concepts

Data mining, which usually is called Knowledge Discovery in Database (KDD), has been commonly applied to many aspects in our daily life, like the financial, retail, communication and marketing areas. Data mining mainly contains three stages [73]: (1) Data preparation. (2) Data mining. (3) Result expression and explanation. It enables people to process the data through various perspectives, extract the useful knowledge, summarize their characteristic or find out the determinative factors within large amount of data. With the useful information provided by data mining, the correlations and patterns are found among the data. The two most important disciplines of data mining are statistics and machine-learning. People are able to make helpful decisions in increasing the profit of business, improving the quality of life, optimizing the patterns of operation, and etc. Different analysis levels includes: Artificial Neural Networks, Genetic Algorithms, Decision Trees, Nearest Neighbor Method, Rule Induction, Data Visualization and etc.

Two primary goals of data mining are prediction and description. Prediction means to predict the future or unknown values of interest based on some variables or parameters; description means to describe the data in an understandable way.

The main applications for data mining are:

- Classification
- Estimation
- Prediction
- Affinity grouping or association rules
- Clustering
- Description and Visualization

The examples of the above applications in daily life are introduced as below:

Classification: Select the classified training set from the data, apply the data mining classification technique on the training set and build the classification model to classify the new data.

For example:

a. Power system voltages, are classified as low, rated and high
b. Failure diagnosis. Analyze the reason for the voltage collapse and enhance the stability of voltage.

Estimation: Like classification. The difference is classification describes the output as discrete type variable while the estimation describes output as continuous type variable. The result of classification is a definite number while the result of estimation is uncertain.

For example:

a. Estimate the electricity market price based on the data mining techniques [74].
b. Estimate the predicting-aided state using the bad data mining in a power system [75].
c. Estimate the rotor angle boundary during the transients [76].

Generally speaking, estimation serves as the first step of classification. Given a certain input data, the unknown continuous value can be obtained via estimation, then the data can be classified based on the pre-determined criteria value.

Prediction: Usually, prediction takes into effect via classification and prediction. Based on the model derived from classification and estimation, the model can be used to predict the unknown variable. The goal of prediction is to forecast the unknown variable(s). This prediction should be tested from time-to-time, the prediction success rate is known only after a certain time period.

For example:

a. Predict the eigen-value and damping ratio of a power system based on the load model and nodal injection [77].
b. Predict the transient stability based on decision tree [24].

Clustering: Clustering places similar records into a group. The difference between clustering and classification is it doesn't rely on the pre-determined class, therefore clustering does not require a data set to train.

For example:

a. The extension of EAC to the multi-machine system is realized by clustering two subsets of generators.
b. The assessment for the dynamic security using the fuzzy clustering [59].

Description and Visualization: It is the expression of the data mining result.

4.1.2 Data Mining Tools

Powerful software tools need to be applied to the exploration and development of data mining. The most important tasks in data mining can be categorized in several aspects, which based on whether the output values are already known [78]:

Supervised learning: The output variable has already been known, and the analysis methods include classification, fuzzy classification and regression. The classification classifies the class as an integer output and fuzzy classification classifies the class as 0 and 1. Regression is used to predict a future real value based on the recent or past values.

Unsupervised learning: The output variable is unknown, and the analysis methods include clustering and association learning. Clustering describes groups of similar examples by fuzzy clustering, while association learning find the most typical group among frequently appearing groups.

Semi-supervised learning: The output variable is part of the known.

Depending on the frequency of tools, the above tasks can be solved.

1st frequent tools: Classifier is applied to solve the estimated density function, relevance analysis, and feature extraction, etc.

2nd frequent tools: Decision tree, clustering, regression, data cleaning, data filtering, artificial neural networks, factor analysis, etc.

3rd frequent tools: Fuzzy classification, support vector machines, k-nearest-neighboring methods, independent component analysis, Bayesian networks, etc.

Infrequent tools: Random forests, WEKA, etc.

4.2 Data Mining in Power System

The application of data mining technique in power system comes into use with the growth of the available data, which requires the cooperation of the electrical engineers and the computer engineers. Prediction and description of data mining are most

frequently used in power systems [79]. In prediction, some attributes are used to predict the unknown state or future values of other attributes of interest. For example, transient stability is predicted by decision trees, dependency stables etc.; the adjust load prediction is based on the detection of the most important attribute changes from the historical data. In description, data mining uses human readable expressions to describe the data. For example, a power system emergency state is described by equipment exceed their rated capability or voltages below their rating. The classification of power system state is also very important in data mining, for example, the state of the system is classified under normal, alert, emergency, in extremis, or restorative state. Once the state of the system is determined, proper actions can be taken to return the system back to the normal operating condition. There are mainly two types of fault classification [80]: on-line and post-fault. On-line classification requires a fast response at the moment when the fault appears, which may help the real time diagnosis and decision making by the operator during a contingency. A security adaptive protection scheme based on the classification result of data mining is developed [81] to reduce the possibility of hidden failures and potential cascading events. Post-fault classification is a sequence classification with variable time period, which can be performed offline. An overview of data mining application in power grid can be shown as Table 4.1:

Table 4. 1: Summary of tasks, techniques and applications of data mining in power systems [79]

Reference	**Technique**	**Task**	**Application Area**	**Problem**
Huisheng, 1998	Fuzzy/ neural network	Classification	Transmission lines	Faults classification.
Vasilic, 2002	Neural network	Classification	Transmission lines	Faults classification.
Kezunovic, 2002	Neural network	Detection and diagnostic	Transmission lines	Detection and diagnosis of transient and faults.
Mori, 2002	Regression tree and Neural network	Forecasting	Distribution system	Load forecasting
Tso, 2004	Statistical analysis	Detection	Transmission and Distribution systems	Detection the substations most sensitive to the disturbances
Figueiredo,2005	Decision tree	Classification	Distribution system	Electric energy consumer
Dola, 2005	Decision tree and neural network	Classification	Distribution system	Faults classification
Vasilic, 2005	Fuzzy/ neural network	Classification	Transmission lines	Faults classification.
Chia-Hun & Chia-Hao 2006	Adaptative wavelet networks	Detection and discrimination	14-bus power system	Power-quality detection for power system disturbances

Silva, 2006	Neural network	Detection and classification	Transmission lines	Faults detection and classification
Costa, 2006	Neural network	Classification	Transmission lines	Fault classification
Dash, 2007	Support Vector Machine	Classification identification	Transmission lines	Classification and identification of series compensated
Monedero, 2007	Neural network	Classification	Not defined	Classification of electrical disturbances in real time.
Hagh, 2007	Neural network	Classification	Transmission lines	Faults classification and locations
Pang & Ding 2008	Wavelet transform and self-organizing learning array	Power quality disturbances classification	Distributed power system	Power-quality detection for power system disturbances
Bhende, 2008	Neural network	Classification	Not defined	Detection and classification of Power quality disturbances
Saibal, 2008	WN	Classification	Distribution system	Classification of transients
Ramos, 2008	Decision tree	Classification	Distribution system	Characterization and classification of consumers

4.2.1 Decision Trees and CART®

As mentioned in section 2.1.8, decision trees are the most popular data mining technique currently in use in the power systems. Reference [21] states that 88.6% of all the data mining methods that have been applied to power system problems used decision trees. Decision trees are an algorithm for predicting models. Based on the characteristic of a certain object, it observes the past behavior or historical data to forecast the future value. The decision tree gains the relationship between the input and the output variables through a binary tree to visualize the problem [21]. Decision trees can provide classification and prediction in the meanwhile. There are two sets of data in building the decision trees: training set and testing set. A decision tree is trained by a large number of cases called the "training set" (TS), which are a sample of pre-classified states and a list of attributes. The decision tree splits the data into two subsets in each node in order to get the maximal purified partition. The tree will define the optimal attribute and its optimal value. The tree will stop growing when the terminal criterion is met:

- The class of a terminal node is "sufficiently" pure
- The accuracy cannot be improved by further partitioning

A set of "testing set" (TS) must be used to evaluate the accuracy of the decision tree that has been built. To get a reliable and unbiased estimation, the test data set needs to be

large. Decision trees are able to discriminate the relevant and irrelevant elements and filter out many irrelevant predictors for classification. Decision trees choose the most obvious split via calculating the node purity. The four most influential algorithms of decision trees are: Classification and Regression Tree (CART), Automatic Interaction Detection (AID), Chi Square Automatic Interaction Detection (CHAID), and C5.0 (a commercial decision-tree and rule-learning package developed by RuleQuest Research). CART uses Gini technique to build the decision trees, C5.0builds the decision tree using the concept of information entropy while CHAID based on the CHi-squared automatic interaction detection [82].

4.3 Introduction to CART®

CART is acronym for Classification and Regression Tree. Classification is used to analyze the categorical data while regression is used to analyze the continuous data. CART® is the statistical software that has been developed by Salford Systems Ltd. in 1984 [83]. The CART® technology explores the underlying mechanism based on which the data was created and uses binary recursive splitting to partition the sample space. The binary means each parent node is split into two child nodes. The splitting rules may be different at each node to get optimum separation of two groups of data. This process is repeated and each node is treated as a "parent" node with the sub-groups being the corresponding "child" nodes. The partitioning does not stop until the terminal node is reached, for which, no further split is possible as the maximum purity has been reached or the quality of the last split has fallen below the specified threshold [84]. The tree will be completed when all the nodes stop to partition. A simple example of a decision tree is given in Figure 4.1:

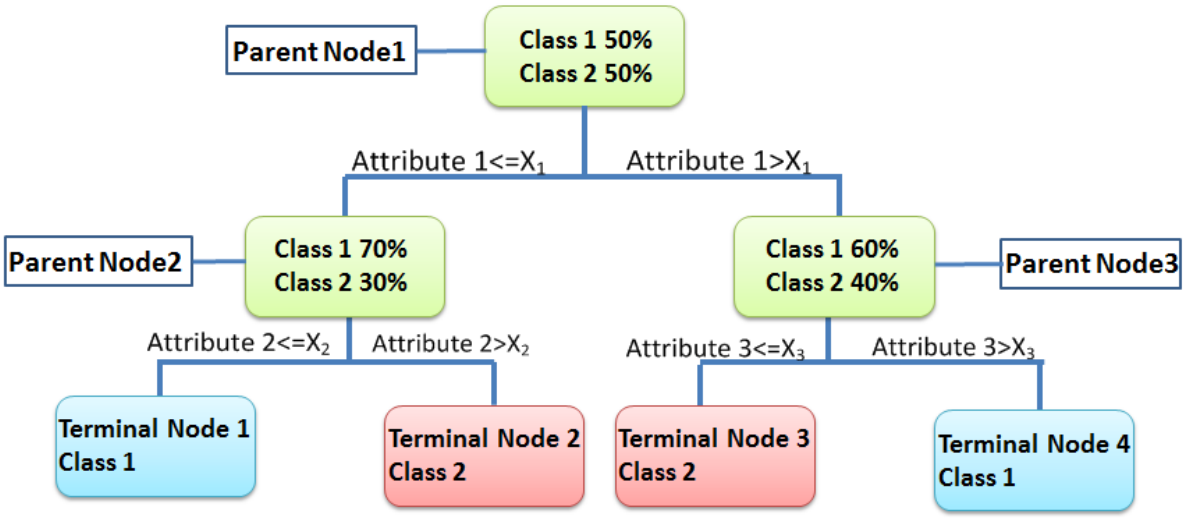


Figure 4. 1: Decision tree classification example

A decision tree has multiple branches, parent nodes, terminal nodes and attributes. All the data assembles in the Parent Node 1. Based on the first splitting rule of attribute 1, the left child node t_L (Parent Node 2) and the right child node t_R (Parent Node 3) are derived. The new parent nodes (Parent Node 2 and Parent Node 3) continue to partition their own data with the splitting rule of attribute 2 and attribute 3 respectively. This process will stop when the required purity of each class is met. In this case, the partition stops when Class 1 and Class 2 are "sufficiently" discriminated. Usually, a sufficient level is defined by the impurity of a split is less than a certain threshold, and when the threshold is met, the partition process will stop. The ending node is called terminal node which means no more splits will be implemented in that node. The terminal nodes show the classification results. The decision tree split a node based on one single variable at a time, is shown in Figure 4.1. First if the attribute 1 is less than or equal to x_1, the data will be divided to the left child node (parent node 2). Then, based on the attribute 2, if it is less than x_2, the data in parent node 1 will be classified as Class 1 (in blue) at terminal node 1, otherwise the data is classified as Class 2 (in red) when attribute 2 is larger than x_2 in the figure. Second, if the data with higher attribute1 ($Attribute1 > x_1$), additional attribute 3 is needed to divide the two classes. On the basis of attribute 3, if attribute 3 is less than or equal to x_3, the data is classified as Class 2 at the terminal node 3, the

remaining data that belongs to Class 1 are at the terminal 4. CART® first grows thoroughly to build a maximal tree until no further splits are possible. It will begin to prune the tree and move away some branches until the best tree is found, which is decided by both the accuracy rate and the applied penalty. During this process, CART® assesses the parametric model and helps in choosing the variable and detecting the inner interactions. It is worthy to point out CART® is also capable to classify more than two classes of data.

The advantages of CART® over other data mining techniques like AID, CHAID, and C5.0, are briefly summarized as below [73]:

- Optimal tree creation due to superior predictive accuracy
- Assessing goodness of fit of solution via cross-validation
- Ability to use the same variable in different parts of the tree
- Resistance to presence of outliers

4.3.1 Data preparation for CART®

The input data for CART® is in a table format, shown in Table 4.2. The rows are the cases vectors of the measurement $C = \{1,2,...n\}$; the first n columns are the attribute vectors $Attr \quad i = \{a_{i1}, a_{i2}...a_{in}\}$; the last column is the target (classification class) that the DT needs to partition: $k_i = \{k_1, k_2...k_n\}$. In a two class classification example, $k_i = \{class1, class2\}$. The value of the attribute a_{ij} can either be numerical or categorical. Take a data set $a_{ij} = \{1,2,3,4,5,6,7,8\}$ for example, if these attributes are numerical, the splitting values in Figure 4.1 can be any real value belonged to [1,8], $x_i \in [1,8]$,($x_1 = 1.5, x_2 = 2.3, x_3 = 6.6$ for example);

$$\begin{cases} t_L \le x_i \\ t_L > x_i \end{cases}, x_i \in [1,8], \tag{4.1}$$

if these attributes are categorical, the splitting values in Figure 4.1 are the attributes themselves, and the splitting rule changes to

$$\begin{cases} t_L = x_i \\ t_L \ne x_i \end{cases}, x_i \in a_{ij}. \tag{4.2}$$

CART® uses the regression method (optional least square criterion) to grow the DT when the target values are numerical, and uses the classification method (optional Gini diversity index) to grow the DT when the target values are categorical. In the DT of this dissertation, the target values are categorical, therefore it belongs to a classification problem.

Table 4. 2: Input data Table for CART®

	Attr 1	**Attr 2**	**…**	**Attr n**	**Target**
1	a_{11}	a_{12}	…	a_{1n}	Class k_1
2	a_{21}	a_{22}	…	a_{2n}	Class k_2
…	…	…	…	…	…
n	a_{n1}	a_{n2}	…	a_{nn}	Class k_n

4.3.2 Classification Method in CART®

There are two main splitting rules for classifying the decision tree in CART®: the Gini and the Twoing Criterion.

Gini: The Gini impurity criterion is given by

$$i(t) = 1 - \sum_{j=1}^{J} p(j \mid t)^2 \tag{4.3}$$

in which J is the total number of classes, $j = 1,2,...J$, t is the node number, p is the probability, and $p(j \mid t)$ indicates the probability of the *jth* Class fall into the node t.

$$p(j \mid t) = \frac{n(j)}{n(t)} \tag{4.4}$$

For a 10 cases of two classes example, $p(1 \mid t) = \frac{4}{10} = 0.4$, $p(2 \mid t) = \frac{6}{10} = 0.6$, which indicates the proportion between Class 1 and Class 2 at the node t is 4:6, 4 cases belongs to Class 1 while 6 cases belongs to the class2. The impurity is 0.48 in this case.

The split criterion goodness of a split s at the node t can be represented as:

$$\Delta i(s,t) = p \cdot i(t) - p_L \cdot i(t_L) - p_R \cdot i(t_R) \tag{4.5}$$

where $\Delta i(s,t)$ is the decrease impurity of the split s at the node t. $i(t)$ is the impurity at the node t, $p = 0.5$ means the two classes are evenly divided at the node. p_L and p_R are the probability of the number divided into the left node and the right node. $i(t_L)$ and $i(t_R)$ are the diversities for the left descendant node and the right descendant node. The optimal split at the node t will turn into a problem of minimizing the impurity $\Delta i(s,t)$.

The Twoing Criterion: The Twoing Criterion is based on the idea of class partition instead of the node heterogeneity. It only measures the difference of probability between the left descendant node and the right descendant node. The formal criterion is to maximize the value of

$$\frac{p_L p_R}{4} \sum_j (| p(j | t_L) - p(j | t_R) |^2) \tag{4.6}$$

where p_L, p_R indicate the probability of cases fall into the left descendant node and the right descendant node. $p(j | t_L)$ and $p(j | t_R)$ means the probability of the jth Class falls into the left node and right node respectively. The Twoing Criterion treats the multi-groups' classification as a two groups problem by grouping the similar classes together and spliting them into two groups. Therefore, in the multi-groups' situation, it will incline to divide the data according to the similarity among those groups. When it is a two-class classification, the calculation results of the GINI and the Twoing Criterion are the same. But it is substantially different when it comes to the multiple class classification.

The ordered Twoing Criterion developed from the Twoing Criterion. The number of the classes is "ordered". The classes to be classified are ranked according to their characteristic, which means the classes are ordinal labeled. CART® will treat the classes with farther away ordinal as the more different classes. The classes with closer number are considered to share more similarity together, therefore they are more easily to be classified as the same group and fall into the same node.

Although the Twoing Criterion has the advantage of discovering the similarity of different classes in the multiple classes' cases, it has the shortcoming that the split quality is not as good as the GINI method. Therefore, the GINI method is chosen as the splitting method in this dissertation.

4.3.3 Decision Tree Evaluation in CART®

Usually, the V-Fold Cross Validation method is used with sufficient data. With this method, all the data will be divided into a training set and a testing set. With the training set, an overly large tree is grown, while the testing set is used to reduce the misclassification rate. With the testing set, the rate of those misclassified cases is estimated. In CART®, all the data will be divided into 10 parts, 9/10 of data are randomly chosen to build a decision tree, and the remaining 1/10 of data are used to are used to test the error rate of the built decision tree. The process will continue until each 1/10 of data are being used as the test data. Furthermore, CART® can estimate the accuracy of the tree based on the frequencies of the dependent variable classes.

After the decision trees are grown, CART® decides the best tree by testing the misclassification rate or the relative cost. The relative cost is calculated based on the testing set.

Take Figure 4.1 for example again, suppose 100 cases in total, 50 cases belongs to Class 1 while the other 50 cases belongs to Class 2, after the first split of attribute 1, 60 cases fall into the left child node (parent node 2). Among it, 58 cases belong to Class 1 and 2 cases belong to Class 2. If the decision tree stops at this node, the misclassification rate/the relative cost of this node is $\max_j p(j \mid t_{parent\_node2}) = \frac{2}{58+2} = 0.03$, as shown in Table 4.3.

Table 4. 3: Misclassification rate at Parent Node 2

Parent Node 2			
Class	N cases	N Mis-Classed	Pct. Error
0	58	2	0.03

The total misclassification rate of the decision tree can be calculated by:

$$\sum_{k \in Ter\min al} (1 - \max_j p(j \mid T_k)) \cdot p(T_k) \qquad (4.7)$$

in which T_k is the k[th] terminal node.

4.3.4 Decision Tree Surrogate in CART®

CART® develops a technique to solve the problem of the missing value in the decision tree. The solution of this problem lies in the use of the surrogates to back up the primary split of the tree. The objective of a surrogate is to imitate the partition achieved by the primary split. For each split of the tree, CART® develops the alternative splits (surrogates), which can be used to classify an object when the primary splitting variable is missing. If the surrogate is a satisfactory one (indicated by high association of the surrogate split with the original split), the predictions that made by using the surrogates would be very similar to the predictions that made by using the original variable. In this sense, the concept of a surrogate is very close to the notion of the correlation in a linear regression. Not only must a good surrogate splitting the parent node into the descendant nodes is similar in size and composition to the primary descendant nodes, but, to the extent possible, the surrogate must also match the primary split on the specific cases that go to the left child nodes and the right child nodes. Therefore, in our case, to increase the robustness and the reliability in case the primary node gets disconnected, further PMUs will have to be deployed to measure the surrogate attributes.

4.3 Limitation of CART®

CART® performs the classification based on only one attribute, but the synchronized phasor measurements consist of the real and imaginary part, which cannot be considered as an integrity by CART®. Suppose we have 2 groups of complex data, the red and the blue dots as seen in the Figure 4,2. Since the current data mining tools will only split the data in 1 dimension, CART® will only split the data parallel to the x axis or the y axis. But splitting a complex number based on only its real or imaginary part may not be optimal nor be completed, as shown in the figure.

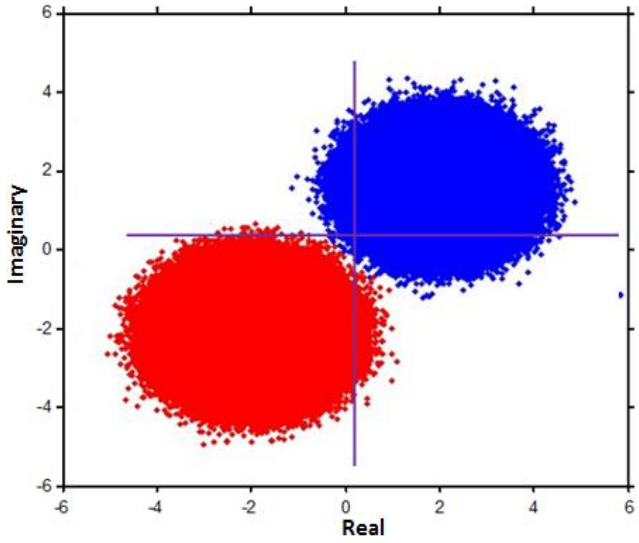


Figure 4. 2: Classification of complex values in CART®

Chapter 5 Proposed Methodology for Transient Stability Prediction

To investigate a fast and accurate transient stability forecasting method, this dissertation utilizes the apparent impedance trajectories recorded by the phasor measurement for prediction. The proposed method includes two steps: the offline simulation for the PMUs placements and the online prediction of the transient stability.

5.1 Offline Simulation for PMUs Placement Optimization

The offline simulation steps shown in Figure 5.1 are used for determining PMUs placements, which are prepared for making the online transient stability predictions in section 5.2.

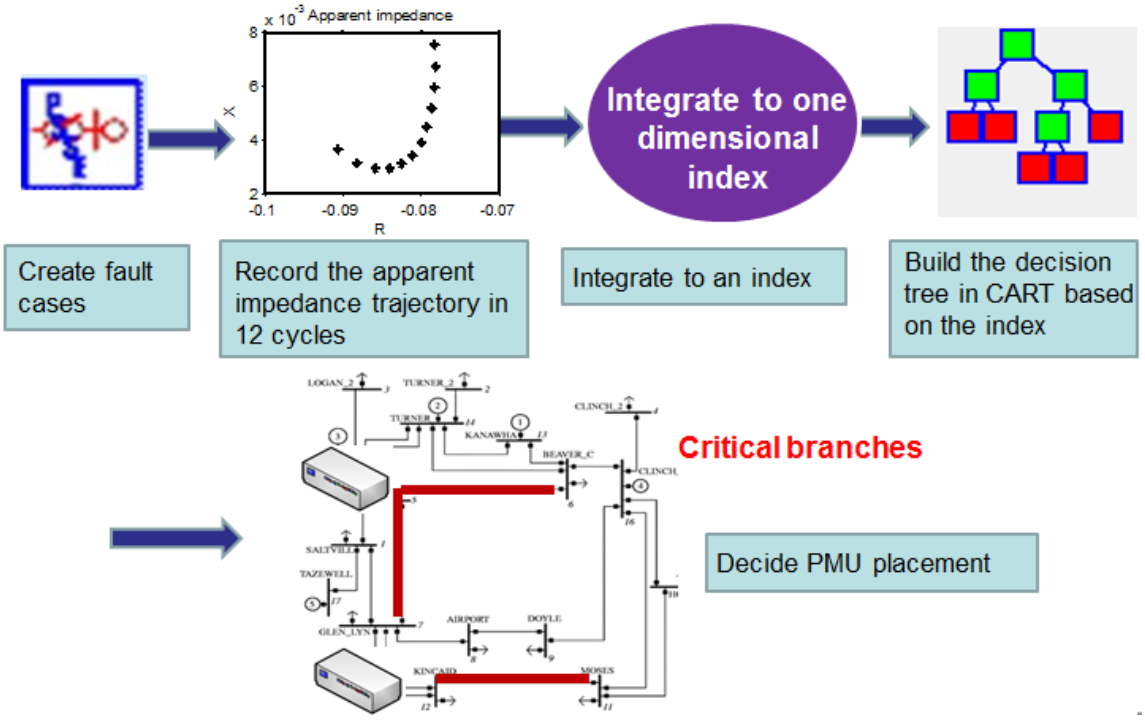


Figure 5. 1: Optimized PMUs placement scheme (pre-calculated)

The first step is creating the data base. According to reference [85], DTs are built based on the relevant pre-contingency operating states which are pre-classified by means of a transient stability program. This information can be acquired by two methods:

1) By using the historical records of the system, and taking plausible scenarios of load-generation-topology schemes into account, the corresponding operating states are generated.

2) A series of disturbances are considered for the concerned stability problem, and then considering a suitable stability method to classify each operating state with respect to each such pre-assigned contingency. These samples of pre-contingency, pre-classified operating states then make up the data base. Subsequently, it will be decomposed into the training and the testing sets.

In our case, similar to 2), we are concerned about the transient stability of the system when faults occur on all possible locations within the system along with the N-1 operation. Two types of three-phase faults are computed at every bus in the system. The first type is a three phase short circuit branch fault that lasts for 80 milliseconds. The fault is isolated by disconnecting the faulted branch. The second type is a three phase short circuit bus fault that lasts for 300 milliseconds. The fault is isolated by disconnecting all the branches that connected to the faulted bus. To include different operating conditions, the topology of the system was changed by tripping one of the transmission lines of the system for every contingency case. Thousands of fault cases with N-1 contingency were created in PSS/E®.

The second step is recording the 12 cycles (200 milliseconds) of the apparent impedance trajectory at every branch.

The third step is integrating the 24-dimensional apparent impedance trajectories into an important index in one dimension to solve the limitation of CART®. The index indicates the distance of the 24 dimensional data point to the hyperplane, which will be explained in more details in 5.4.3.

The fourth step is imputing the index of every branch to CART® for the stable/unstable cases classification. Based on the proposed methodology, a decision tree can be built. A simple example is shown in Figure 5.2. We could build the decision tree based on this one dimensional data, and the critical braches for detecting the trajectory are selected by the decision tree as important attributes. In this example, the distance of branch 1, branch 3 and branch 6 are important for the stability prediction.

Finally, the buses that connected to the critical branches are chosen as the strategic locations for the PMUs installation, as their phasor measurements are critical for transient stability prediction in the power system. For the example given in the fourth step in Figure 5.2, bus *A* is connected to branches 1 and branch 3, and bus *B* is connected to branch 6. Therefore, these two buses become optimal locations for PMUs placement to classify the stable and unstable cases. In this manner, PMUs placements for the transient stability predictions can be optimized.

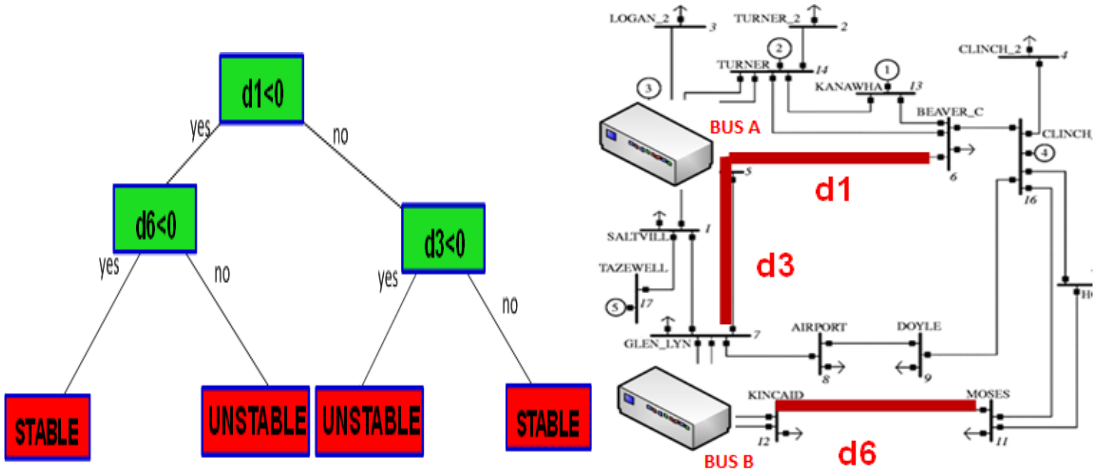


Figure 5. 2: Simple decision tree example

5.2 Online Prediction of Transient Stability

The steps mentioned above are mainly for the optimization of PMUs placement. When a real fault occurs in the power system, based on the voltages and currents measured by the PMUs that selected in the offline simulation, the 12-cycle apparent impedance trajectories of the critical braches can be obtained. Then a one dimension index needs to be calculated and input to the decision tree built from offline simulations mentioned in 5.1. Follow the binary logic of the pre-built decision tree, the transient stability of the power system can be predicted. The decision tree must arm/disarm the relay with a selected window of apparent impedance trajectory. If the system is predicted as unstable, some pre-determined control mechanisms have to be executed at once to prevent the cascading failure and restore the stability of power system, , such as the

intelligent load shedding in [86] and intelligent controlled islanding [66].The flow chart of online implementation is shown in Figure 5.3.

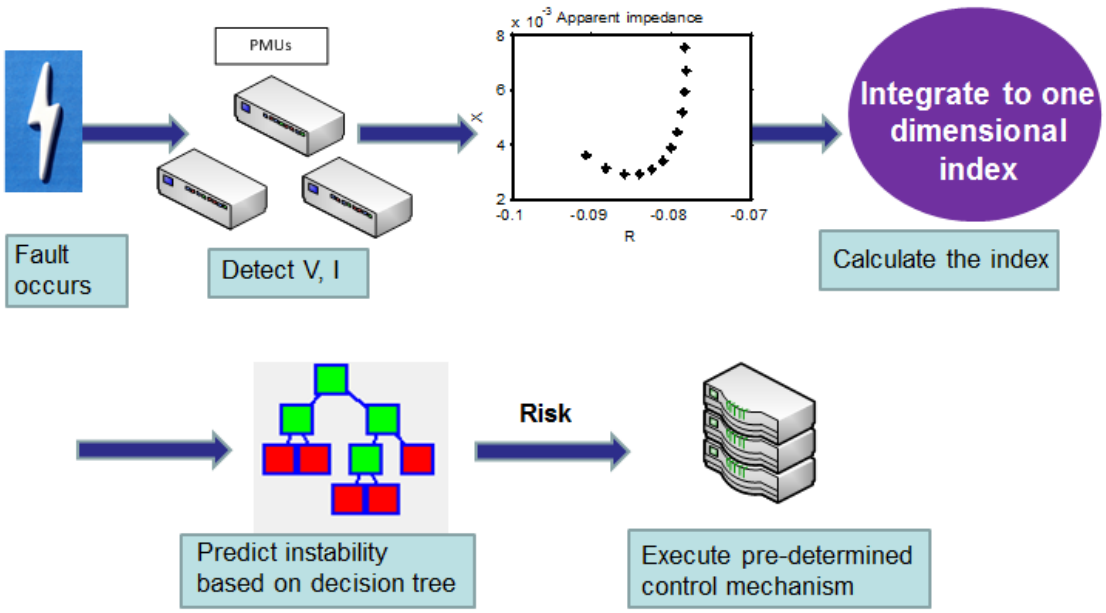


Figure 5. 3: Instability prediction process

5.3 Apparent Impedance Trajectory in the Proposed Methodology

Apparent impedance trajectories, as mentioned in Chapter 3, are mainly applied in the Out of Step relay to detect the asynchronism of the system based on the tripping zone settings. However, in this dissertation, instead of using the apparent impedance trajectory for rotor angle instability detection, the use of a 12-cycle apparent impedance trajectory to predict the rotor angle stability is proposed. The apparent impedance trajectory of every branch is computed through the offline simulations. The apparent impedance is recorded at every branch after the clearance of the three phase fault, which can be expressed as $Z = R + jX$

This measurement is recorded for a few cycles after the clearance of the fault. The time series of the apparent impedance is expressed as $T = (T_1, T_2, \ldots T_n)$ where n denotes the number of cycles recorded. Since the apparent impedance is a complex quantity, an

arbitrary trajectory Y, belonging to an $(l_1 \times l_2)$-dimensional space Z⊗T, is created, which is the tensor product of the following two spaces: Z spanned by $\{z_i\}_{i=1}^{l_1}$, $l_1 = 2$ and T spanned by $\{t_i\}_{i=1}^{l_2}$, $l_2 = n$. Therefore the projection of the data onto the space Z⊗T is $R^{l_1 \times l_2} = R^{2n}$. Thus, Y is projected onto the 2n-dimensional space $Y \in R^{2n}$ as seen in Figure 5.4. It is to be noted here that in Figure. 5.4, $n = 12$.

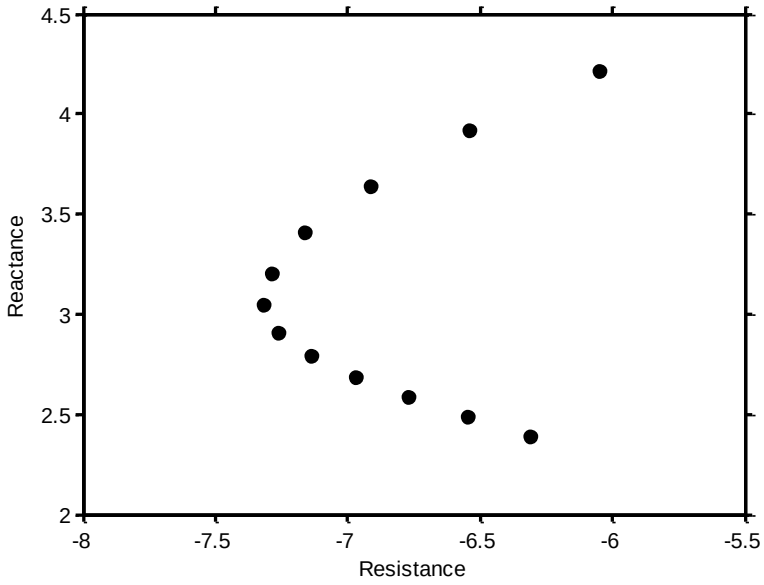


Figure 5. 4: Apparent impedance trajectory with 12-sample data (n=12)

However, the apparent impedance trajectory is in 24-dimensions. But since CART® can only handle the one-dimensional attribute, for inputting the apparent impedance trajectory of every branch into CART® and building the decision tree, a dimension reduction method developed from Linear Discriminant Analysis (LDA) and Fisher's Linear Discriminant (FLD) needs to be applied. The dimensional reduction method solves the problem of simultaneously handling complex data recorded by the PMUs.

5.4 Dimensional Reduction method

This section is arranged as follows. After a brief introduction of the LDA and the FLD method, the dimensional reduction method is presented in great detail.

5.4.1 Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) [87,88] has been extensively used in the past few decades for the feature extraction and the dimension reduction. The principle of LDA is explained below with an illustration. Suppose there are two groups of data in two dimensions, the red dots and the blue squares, as seen in Figure 5.5. Let both of the groups have the same variance ($\sigma_r = \sigma_b$) but different means ($m_r \neq m_b$), as shown in Figure 5.6. Then, the LDA aims to project the two groups of data in a certain direction, and partitions them to the maximum possible extent. Therefore, this approach seeks to maximize the distance between the mean vectors of the two groups. There are many possible directions to project the two groups of data, but there would be only one optimum direction that could maximally discriminate the two groups of data, as seen in Figure 5.5.

Assume we have two dimensional samples$\{X^1, X^2\}$, N_1 of which belong to Class r (red dot) and N_2 of which belong to Class b (blue square). The data sets can be expressed in a matrix considering their dimensions, as is given below:

$$red\ dot = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \\ \dots & \dots \\ r_{N_1 1} & r_{N_1 2} \end{bmatrix} \quad blue\ squre = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ \dots & \dots \\ b_{N_2 1} & b_{N_2 2} \end{bmatrix} \tag{5.1}$$

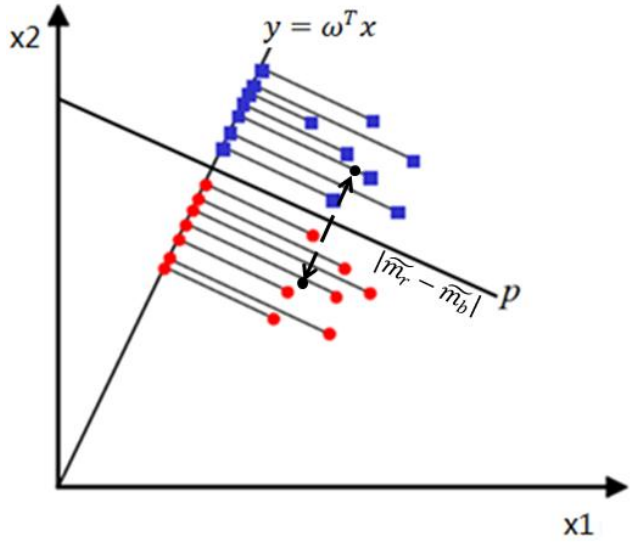


Figure 5. 5: An illustration of LDA

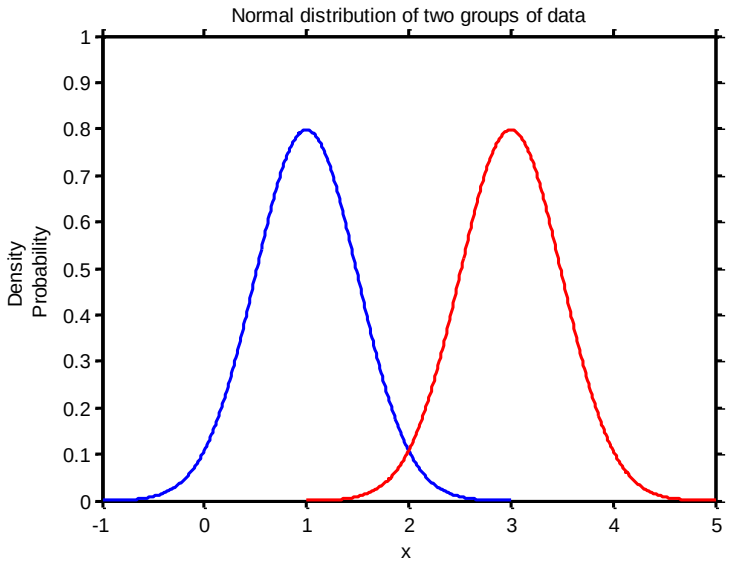


Figure 5. 6: Two groups with the same variance but different means

LDA seeks to obtain a scalar y by projecting the sample X onto a line.

$$y = \omega^T x \tag{5.2}$$

Where $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $\omega = \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix}$

In order to maximize the separation, the direction of the projected scalar needs to be found. After projection, the dimensionality of this splitting problem reduces from two dimensions (x_1, x_2) to only one dimension – scalar y. The mean vector of the i^{th} Class (c_i) in the feature space of x and y is

$$m_i = \frac{1}{N_i} \sum_{x \epsilon c_i} x \tag{5.3}$$

$$\widetilde{m_i} = \frac{1}{N_i} \sum_{y \epsilon c_i} y = \frac{1}{N_i} \sum_{x \epsilon c_i} \omega^T x = \omega^T \frac{1}{N_i} \sum_{x \epsilon c_i} x = \omega^T m_i \tag{5.4}$$

Creating an objective function that finds the distance between the projected mean values, the partition problem turns to maximizing the distance of the mean value of the two groups corresponding to the scalar y. Or, we can understand the objective function to be a maximization of the between class variances.

$$\begin{aligned} J(\omega) &= max|\widetilde{m_r} - \widetilde{m_b}| \\ &= max|\omega^T m_r - \omega^T m_b| \\ &= max|\omega^T (m_r - m_b)| \end{aligned} \tag{5.5}$$

By finding the suitable slope ω of the projected scalar y, the objective function can be formulated.

5.4.2 Fisher's Linear Discriminant

The LDA method may not always split the two groups of data maximally because it does not take the variances of the two groups into account. The Fisher's Linear Discriminant (FLD) [89,90] which is a generalization of the LDA technique to be used when the variance of the two groups of data are different, ($\sigma_r \neq \sigma_b$), shown in Figure 5.7.

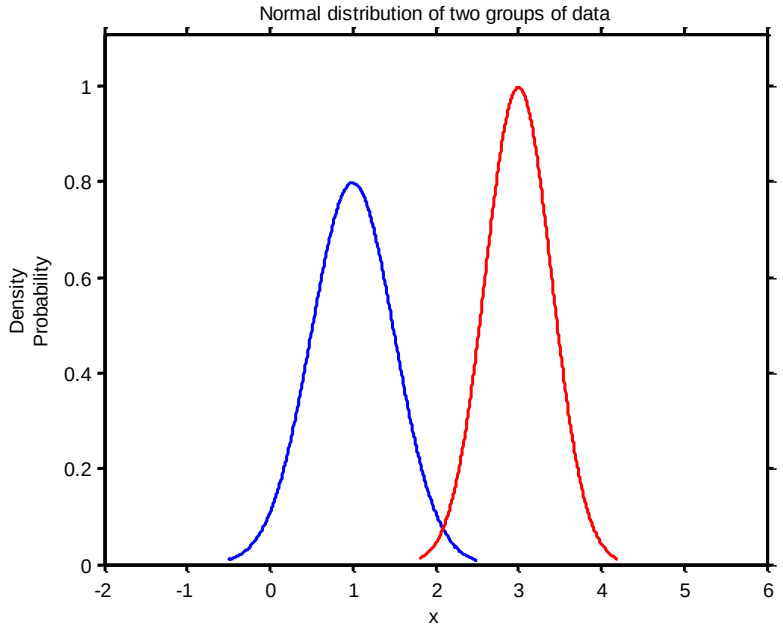


Figure 5. 7: Two groups with different variance

The main goal of the FLD method is to find the best separation by maximizing the ratio of the variance between the classes to the variance within the classes. Within-class scatter ($\widetilde{S_w}$) is the expected covariance of each of the classes, which measures the variability within class j after the projection to the y scalar. It could be expressed as

$$\widetilde{S_w} = \sum_j (y - \widetilde{m_\iota})^2 \tag{5.6}$$

Therefore, for the two classes' classification, the variability within them after projection could be expressed by $\widetilde{S_{wr}}^2 + \widetilde{S_{wb}}^2$, which is the within class classification. Combining this with the definition of between class classifications, an optimization function is defined by Fisher as shown below:

$$J(\omega) = max \frac{|\widetilde{m_r} - \widetilde{m_b}|^2}{\widetilde{S_w}} = max \frac{|\widetilde{m_r} - \widetilde{m_b}|^2}{\widetilde{S_{wr}}^2 + \widetilde{S_{wb}}^2} \tag{5.7}$$

The goal of the optimization is to find a projection vector $\vec{\omega}$ which enables the projected data within each group (${S_{wr}}^2 + {S_{wb}}^2$) as close as possible, while keeping the projected means of the two groups ($|\widetilde{m_r} - \widetilde{m_b}|^2$) as far apart from each other as possible.

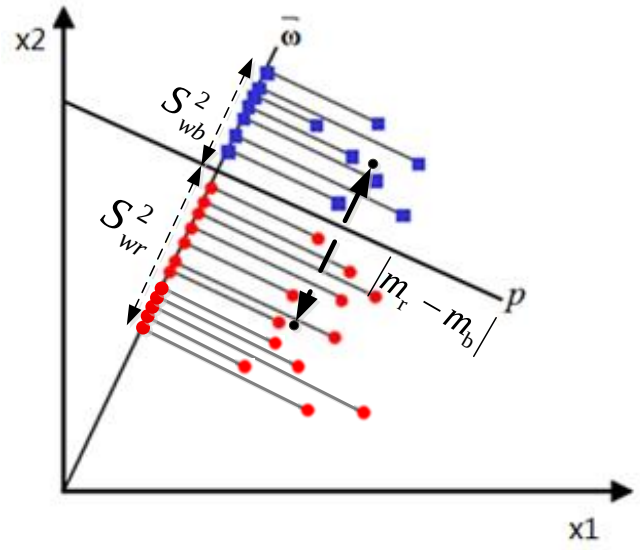


Figure 5. 8: Within class variance and the between class variance

$$\begin{aligned}\widetilde{S_{wr}}^2 + \widetilde{S_{wb}}^2 &= (y - \widetilde{m_r})^2 + (y - \widetilde{m_b})^2 \\ &= (\omega^T x_r - \omega^T m_r)^2 + (\omega^T x_b - \omega^T m_b)^2 \\ &= \omega^T (x_r - m_r)(x_r - m_r)^T \omega + \omega^T (x_b - m_b)(x_b - m_b)^T \omega \end{aligned} \quad (5.8)$$

Define a function of a scatter matrix in the feature space x

$$S_i = (x_i - m_i)(x_i - m_i)^T \quad (5.9)$$

Then, (5.8) will become

$$\begin{aligned}\widetilde{S_{wr}}^2 + \widetilde{S_{wb}}^2 &= \omega^T S_r \omega + \omega^T S_b \omega \\ &= \omega^T (S_r + S_b)\omega \\ &= \omega^T S_w \omega = \widetilde{S_w} \end{aligned} \quad (5.10)$$

Similarly, the between classes variance in the y-space can be expressed as,

$$\begin{aligned}|\widetilde{m_r} - \widetilde{m_b}|^2 &= (\omega^T m_r - \omega^T m_b)^2 \\ &= \omega^T (m_r - m_b)(m_r - m_b)^T \omega \end{aligned} \quad (5.11)$$

Defining a function of a scatter matrix in the feature space x, we get,

$$S_b = (m_r - m_b)(m_r - m_b)^T$$

Then, (5.11) will become

$$|\widetilde{m_r} - \widetilde{m_b}|^2 = \omega^T S_b \omega = \widetilde{S_b} \tag{5.12}$$

The between-class scatter ($\widetilde{S_b}$) in the projection feature space y is the expected covariance between the classes. This measures the variability between classes after the projection to the y scalar.

Using (5.10) and (5.12), the optimization equation (5.7) becomes,

$$J(\omega) = max\left(\frac{\omega^T S_b \omega}{\omega^T S_w \omega}\right) = max\left(\frac{\widetilde{S_b}}{\widetilde{S_w}}\right) \tag{5.13}$$

On doing the least squares optimization to (5.13), we have,

$$\frac{\partial J(\omega)}{\partial \omega} = 0 \tag{5.14}$$

Therefore,

$$\frac{1}{\omega^T S_w \omega} \cdot 2S_b\omega - \frac{1}{(\omega^T S_w \omega)^2} \cdot 2S_w\omega \cdot \omega^T S_b \omega = 0 \tag{5.15}$$

Substituting (5.13) in (5.15)

$$S_b\omega - J(\omega)S_w\omega = 0 \tag{5.16}$$

Then, the separation of the two groups will be optimized when

$$J(\omega) = {S_w}^{-1}(m_r - m_b) \tag{5.17}$$

The FLD has been extensively used for the feature extraction and the dimension reduction. It aims to find the optimal transformation so that the original high dimensional space is preserved in the low-dimensional space [91]. It classifies two groups of data by projecting them on an axis represented by the eigenvectors that discriminates the data to the maximum extent. It has been applied in fields like the face recognition [92], the microarray data classification [93], etc.

<u>Advantage and limitation of the LDA and the FLD:</u>

One of the reasons for popularity of the LDA and the FLD is the ability to analyze the higher dimensional data in the lower dimensional space.

Limitations of the LDA and the FLD [90]:

- Cannot classify data that do not follow a Gaussian distribution.
- Does not work when the two groups of data have the different variances but the same mean as shown in Figure 5.9.

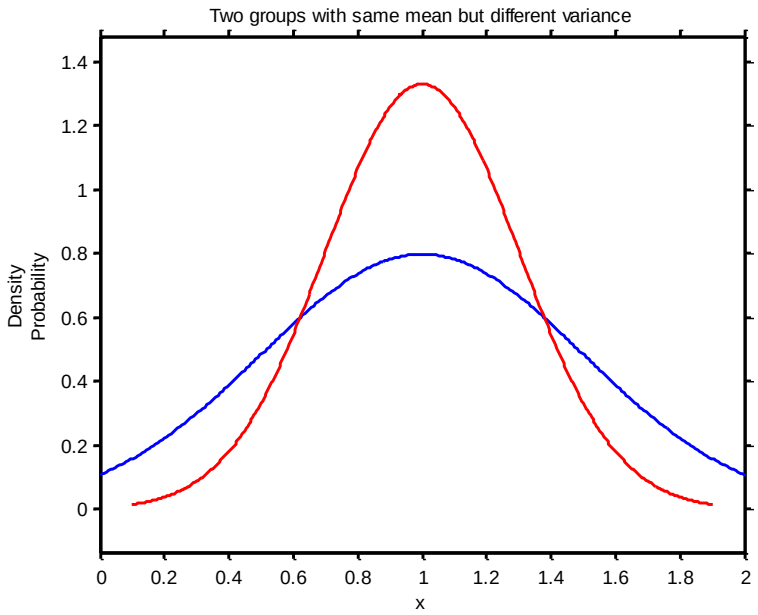


Figure 5. 9: Two groups of data having different variance but same mean

5.4.3 Dimension Reduction Method for complex PMUs data

This method, called Fisher's Linear Discriminant applied to Synchrophasor Data (FLDSD) [94], is a combination of the LDA and the more general, FLD. The concept of FLDSD is illustrated with an example as follows: The hyperplane that splits the two groups of complex data maximally should be like the line that shown in Figure 5.10

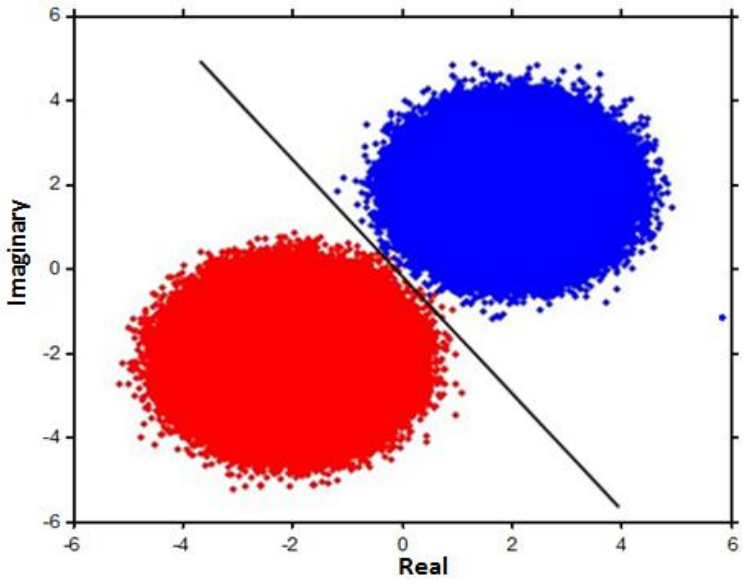


Figure 5. 10: Classification of two groups of data

Firstly, the centroids of the two groups of data are found, a plane that is perpendicular to the vector connecting the two centroids can maximally split the two groups data. One group of data is located at on one side of the perpendicular plane, while the other group is located on the other side of the plane. By calculating the distance to the plane, we can tell which group the data belongs to. Therefore the distance of the data point within each group to the perpendicular plane can be used as an attribute to build the decision tree. Thus, the two dimensional data can now be integrated to a one dimensional distance data, as shown in Figure 5.11.

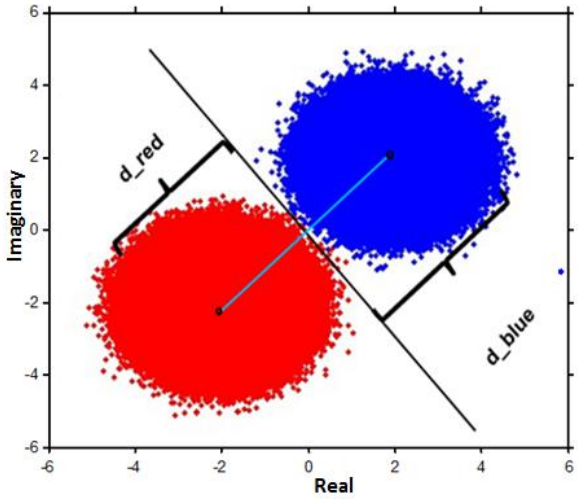


Figure 5. 11 Integrate two dimensional data to one dimension

In [94], LDA and FLD were first applied to solve the problem of splitting complex synchrophasor data using decision trees in the domain of power systems. The central notion of the dimension reduction for complex PMUs data is to come up with a single variable that is able to replace the multivariate input data. This method states that the single variable that can be used for performing the split is the distance between the complex number and an optimally selected hyperplane. Suppose there are two groups of data, having centroids $\overrightarrow{\mu_1}$ and $\overrightarrow{\mu_2}$, and co-variances Σ_1 and Σ_2; then the optimum split between the two groups of data would be obtained by a hyperplane $\vec{p}$ whose equation is given by,

$$\vec{p}: a_1 x_1 + a_2 x_2 = a_3 \tag{5.18}$$

Wherea_1,a_2, a_3 are the coefficients of the hyperplane function. The goal is to maximize the projected difference of the two groups of data by devising a hyperplane that will place the data sets on either side of it in the best possible way. This is done by using LDA which projects the two ellipsoidal shaped data along the direction represented by the eigenvectors to get the maximum separation between the two groups. As shown in Figure. 5.12, the hyperplane $\vec{p}$ that could provide the best split between the two groups of data is found by the dot product, $\vec{p} \cdot \vec{\omega} = 0$ where $\vec{\omega}$ is computed using FLD as shown below .

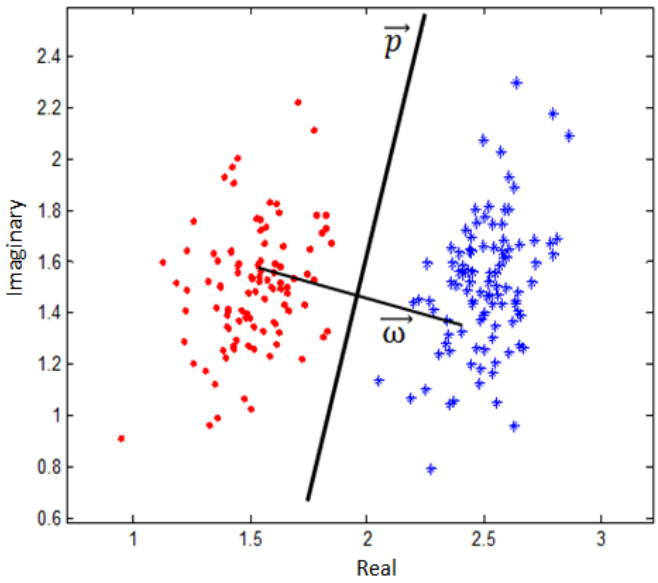


Figure 5. 12: Two groups of data partitioned using LDA

In accordance with FLD, the classification of two groups having different variances would be defined as the ratio of the between groups variance relative to the within group variance. Mathematically, this means,

$$S = \frac{\sigma^2_{\text{between}}}{\sigma^2_{\text{within}}} = \frac{(\vec{\omega}.\vec{\mu_1} - \vec{\omega}.\vec{\mu_2})^2}{\vec{\omega}^T.\Sigma_1.\vec{\omega} + \vec{\omega}^T.\Sigma_2.\vec{\omega}} \tag{5.19}$$

To achieve maximum separation between the two groups of data, the optimal split is found through the projected vector $\vec{\omega}$ given by the equation,

$$\vec{\omega} = (\Sigma_1 + \Sigma_2)^{-1}.(\vec{\mu_1} - \vec{\mu_2}) \tag{5.20}$$

Since the perpendicular plane can be found by $p \cdot \vec{\omega} = 0$

The equation of finding the perpendicular plane change from

$$(\overrightarrow{\mu_1} - \overrightarrow{\mu_2}) \cdot (x - ch) = 0 \tag{5.21}$$

To

$$(\Sigma_1 + \Sigma_2)^{-1} . (\overrightarrow{\mu_1} - \overrightarrow{\mu_2}) \cdot (x - ch) = 0 \tag{5.22}$$

Once the projected direction is identified, the bisector hyperplane $\vec{p}$can be found by $\vec{p} \cdot \vec{\omega} = 0$. Thus, by using FLD, the two groups having different distributions could be split in an optimal way. The next step is to find the shortest distance from the complex numbers (x_i, y_i) to $\vec{p}$ as shown below,

$$D(i) = \frac{a_3 - a_1 x_i - a_2 y_i}{\sqrt{a_1^2 + a_2^2}} \tag{5.23}$$

In this way, the two dimensional data has been expressed by a one dimensional variable distance of the individual points to the hyperplane. The splitting rule for the two groups of data then becomes:

$$\begin{aligned} &\mathrm{D(i)} \geq 0\mathrm{: Group\ 1} \\ &\mathrm{D(i)} < 0\mathrm{: Group\ 2} \end{aligned} \tag{5.24}$$

The process of the dimensional reduction can be illustrated using Figure 5.13 and Figure 5.14. In Figure 5.13, using the LDA method, the vector connecting the centroids of the two groups is depicted by a black line. Using the FLD improvement, by multiplying$(\Sigma_1 + \Sigma_2)^{-1}$, the vector in black rotates to a new vector $\vec{\omega}$ depicted by the green line. The plane that is perpendicular to the new vector will then maximally split the two groups of data. In other words, the previous perpendicular plane rotates by an angle to maximize the two groups' partition.

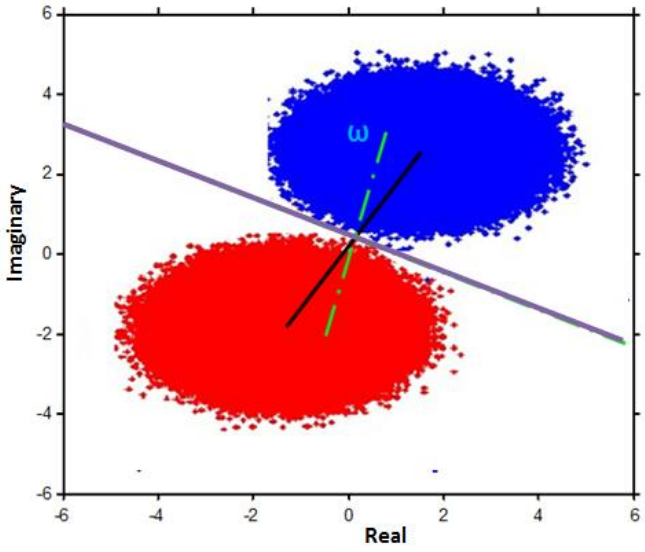


Figure 5. 13: Equivalent to a rotation of the perpendicular plane

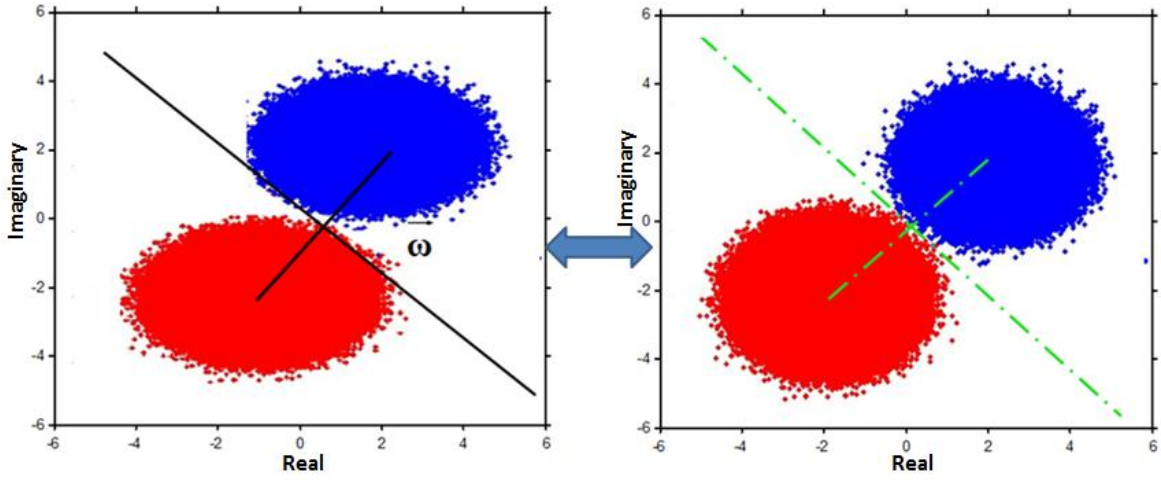


Figure 5. 14: Shape transformation equivalence

Another way to understand this process is through the perspective of shape transformation as shown in Figure 5.14. By multiplying the inverse of the sum of the variance matrix$(\Sigma_1 + \Sigma_2)^{-1}$, the FLD method changes the distribution of the two groups of data. It concentrates the data (the classes are grouped into clusters) to let the original elliptical shape change to a circular one, where $x_{new} = (\Sigma_1 + \Sigma_2)^{-1} \cdot x$. The shape transformation of the two groups of data enables the split of the two groups becoming

more feasible and obvious. By connecting the centroids of two groups of new data, the perpendicular plane can be found as done previously.

5.4.4 Test Example

Suppose we have two groups of data, both of which have normally distributed 1000 data points represented by the red dots and blue dots as shown in Figure 5.15.

$\overrightarrow{\mu_{red}} = [-0.5, 0.5]$, $\overrightarrow{\mu_{blue}} = [0.5, -1]$, $\overrightarrow{\sigma_{red}} \neq \overrightarrow{\sigma_{blue}}$

Based on the dimensional reduction method, the perpendicular line is calculated:

$$\vec{p}: -x + 1.085y = 0.6$$

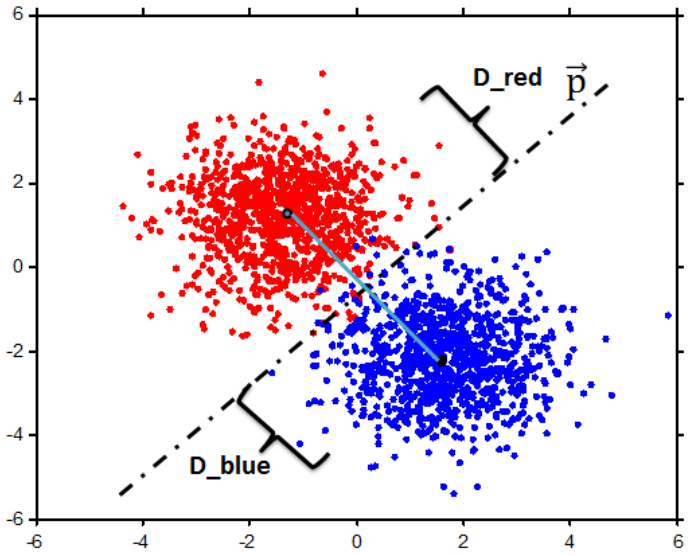


Figure 5. 15: Two group's classification example

Calculating the distance of each group with respect to the optimal hyperplane, the decision tree that is built by CART® is shown in Figure 5.16. The decision tree is able to classify the data with 99% accuracy in a single split. The splitting criterion is whether the distance between the data point and the perpendicular plane is greater than or less than 0.07. If the distance between the data and the separating hyperplane is more than 0.07 (D>0.07), the data points they belong to the red group, otherwise they belong to the blue group. CART® uses 9/10 of the sample data for building the decision tree (training set), while the remaining 1/10 of the sample data is used for testing the performance of the tree (testing set). The partition success rate of the training set is 99.1% as seen in Table 5.1. 6 out of 1000 dots of red group are misclassified to be in the blue group, while 12 out of

1000 dots of the blue group are misclassified to be in the red group. The prediction success for the testing set is given in Table 5.2, which also achieves the prediction with a high accuracy.

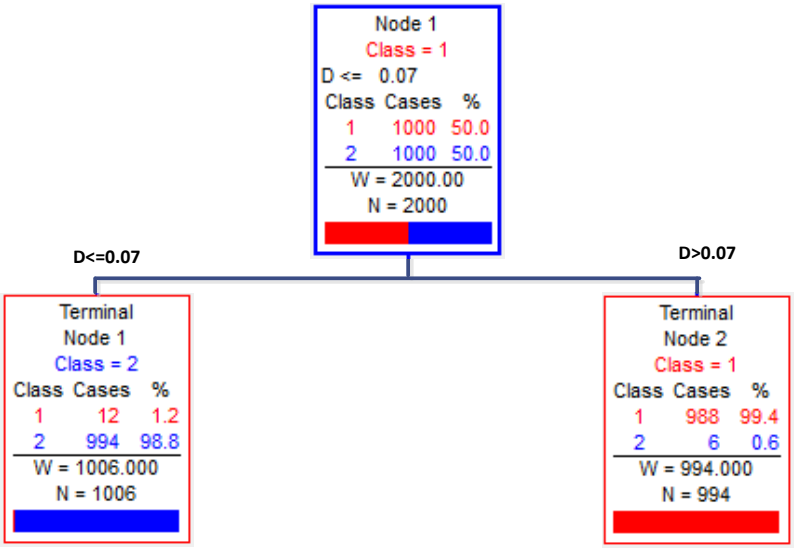


Figure 5. 16: Decision tree for two groups' classification

Table 5. 1: Classification success of the training sample (two classes)

Actual Class	Total Class	Percent Correct	Observed Class	
			1	2
1	1000	99.40%	994	6
2	1000	98.80%	12	988
	Average % Correct	99.10%		
	Overall % Correct	99.10%		

Table 5. 2: Prediction success of the testing sample (two classes)

Actual Class	Total Class	Percent Correct	Observed Class	
			1	2
1	1000	99.30%	993	7
2	1000	98.70%	13	987
	Average % Correct	99.00%		
	Overall % Correct	99.00%		

5.5 Extension to the Higher Dimensional Apparent Impedance Trajectory

By extending the method to address 24-dimensional apparent impedance trajectory, the transient stability of the system can be predicted very quickly.

Assume that stable cases, $\boldsymbol{S_1}$ and unstable cases, $\boldsymbol{S_2}$ are two subsets of the $2n$-dimensional real space $\boldsymbol{X}$ that corresponds to the n-cycle apparent impedance trajectory. Thus,

$$\begin{aligned} \boldsymbol{S_1} &\subseteq \boldsymbol{X} \\ \boldsymbol{S_2} &\subseteq \boldsymbol{X} \\ \boldsymbol{S_1} \cup \boldsymbol{S_2} &= \boldsymbol{X} \end{aligned} \tag{5.25}$$

Suppose the centroids of the two categories be $\overrightarrow{\boldsymbol{\mu_1^{2n}}}$ and $\overrightarrow{\boldsymbol{\mu_2^{2n}}}$ and the variances be $\boldsymbol{\Sigma_1^{2n}}$ and $\boldsymbol{\Sigma_2^{2n}}$. The hyperplane in the high dimensional space is then defined as the optimum plane that is perpendicular to the projection vector $\overrightarrow{\boldsymbol{\omega}}$. Extending the idea developed in part 4.1 to higher dimensions, the goal is to separate $\boldsymbol{X}$ into two half-spaces such that each open half-space contains points belonging to either $\boldsymbol{S_1}$ or $\boldsymbol{S_2}$. The hyperplane that optimally partitions the two groups could then be expressed as,

$$\vec{\boldsymbol{p}}: \langle \boldsymbol{a}, \boldsymbol{X_N} \rangle + b = 0 \tag{5.26}$$

where $\boldsymbol{a}$, $\boldsymbol{X_N}$ are $(2n \times 1\ matrix)$ each, with $\boldsymbol{X_N} = \widehat{x_1}, \widehat{x_2}, \ldots \widehat{x_{2n}}$ denoting the data points in the hyperplane and b being a scalar that shifts the hyperplane from the origin. According to FLD, the perpendicular vector to the hyperplane that maximizes separation of the two groups is given by,

$$\overrightarrow{\boldsymbol{\omega^{2n}}} = \left(\boldsymbol{\Sigma_1^{2n}} + \boldsymbol{\Sigma_2^{2n}}\right)^{-1} . \left(\overrightarrow{\boldsymbol{\mu_1^{2n}}} - \overrightarrow{\boldsymbol{\mu_2^{2n}}}\right) \tag{5.27}$$

For computing the separating hyper plane in $2n$ dimensional space, form a $2n \times 2n$ matrix $\boldsymbol{M}$ which has $(2n-1) \times (2n-1)$ identity matrix embedded in it with the column $\overrightarrow{\boldsymbol{\omega^{2n}}}$ forming the last column. Then, if $\boldsymbol{V}$ denotes the first $2n-1$ rows of $\boldsymbol{M^{-1}}$, we have,

$$\begin{aligned} \mathrm{M} &= \begin{bmatrix} \mathrm{I}_{2n-1} & \overrightarrow{\omega^{2n}} \\ 0 & \end{bmatrix} \\ \mathrm{M}^{-1} &= \begin{bmatrix} \mathrm{V}_{2n-1 \times 2n} \\ \ldots \end{bmatrix} \end{aligned} \tag{5.28}$$

It can be easily observed that the columns of $\boldsymbol{V}$ are orthogonal to $\overrightarrow{\boldsymbol{\omega^{2n}}}$. The distance between the $2n$-dimensional data and the hyperplane can now be expressed as,

$$D = \frac{1}{2}(\boldsymbol{X_N} - \boldsymbol{X})^T(\boldsymbol{X_N} - \boldsymbol{X}) \tag{5.29}$$

Now, a matrix $\boldsymbol{W}$ can be defined such that,

$$\boldsymbol{W} = \boldsymbol{I_{2n}} - \boldsymbol{V^T}(\boldsymbol{VV^T})^{-1}\boldsymbol{V} \tag{5.30}$$

where the matrix $\boldsymbol{W}$ gives the perpendicular vector to the hyperplane from a given point $\boldsymbol{x}$ in $\boldsymbol{R^{2n}}$, where $\boldsymbol{R^{2n}}$ denotes the $2n$-dimensional space. Now, since the hyperplane passes through the midpoint of the two centroids, a vector $\boldsymbol{e}$ can be defined, as shown in (5.31), to minimize the distance given by (5.29).

$$\boldsymbol{e} = \boldsymbol{W}\left(\boldsymbol{x} - \boldsymbol{0.5}\left(\overrightarrow{\boldsymbol{\mu_1^{2n}}} + \overrightarrow{\boldsymbol{\mu_2^{2n}}}\right)\right) \tag{5.31}$$

The shortest distance between the data point and the perpendicular plane can be found by,

$$|\mathrm{d}| = \sqrt{\mathrm{e^T e}} \tag{5.32}$$

Therefore, the 2n-dimensional data set can be expressed as a one-dimensional variable, distance. The stable sets and the unstable sets can be discriminated using the sign of the distance as shown in (5.33).

$$\mathrm{sign}(d) = \mathrm{sign}\left(\boldsymbol{e^T}\overrightarrow{\boldsymbol{\mu_1^{2n}}}\right) \tag{5.33}$$

The linear classifier can now be used to solve the high dimensional data efficiently. With prior knowledge of the distribution of the two categories and a pre-calculated hyperplane, a decision tree can be built in CART® to predict transient stability based on the apparent impedance trajectory. In this dissertation, $\mathrm{n} = 12$, thereby indicating that 12 data points of the apparent impedance trajectory was used for the prediction.

5.6 Extension to Multiple Groups Classfication

When it comes to multiple classes, several hyperplanes need to be created to split the data correctly. This is an extension of the technique developed in the previous section. Fig. 5.12 given below illustrates the concept.

Let us assume that we have to classify four groups of data as shown in Figure 5.17 below. Let the four classes be denoted by $\boldsymbol{X_1}, \boldsymbol{X_2}, \boldsymbol{X_3}$, and $\boldsymbol{X_4}$ such that,

$$\boldsymbol{X_1}, \boldsymbol{X_2}, \boldsymbol{X_3}, \boldsymbol{X_4} \subset \boldsymbol{R^2} \tag{5.34}$$

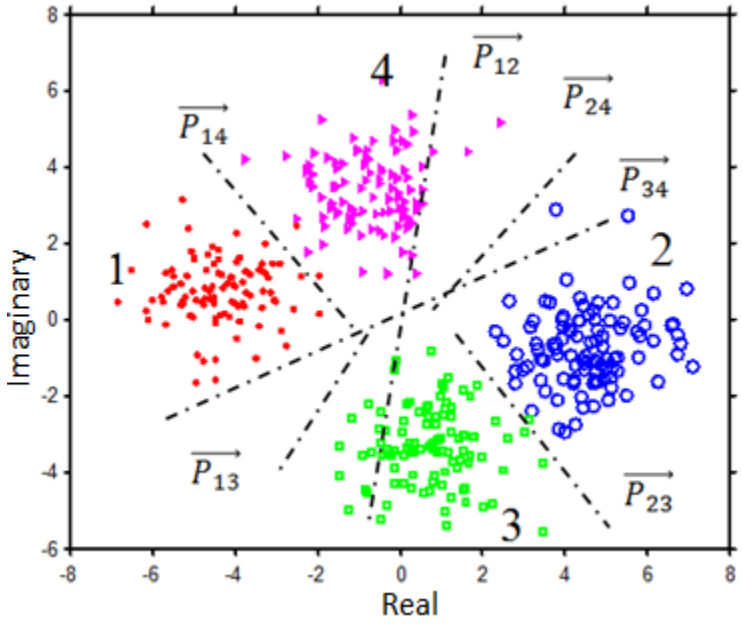


Figure 5. 17: Classification of four groups of data in two dimensions

In accordance with Figure 5.17, let Class 1 be the red dots, Class 2 be the blue circles, Class 3 be the green squares, and Class 4 be the purple triangles. The goal is then to seek the best possible split between any two groups of data. The number of hyperplanes required to do so is given by $\frac{N\cdot(N-1)}{2}$, where N denotes the number of groups to be separated. Then, by finding the distance of each group of data from every hyperplane, the decision tree can be grown as shown in Figure 5.18.

Although there are six hyperplanes to separate the four groups of data for the distance calculation, the decision tree only pick up the distance to two hyperplanes $\overrightarrow{P_{13}}, \overrightarrow{P_{14}}$ as critical attributes. After splitting twice, the decision tree succeeds to partition four groups of data. Firstly, with the help of the hyperplane $\overrightarrow{P_{13}}$, Class 2 and Class 3 are completely separated from Class 1 and Class 4. When the distance from the data to the hyperplane $\overrightarrow{P_{13}}$ is less than -0.82 ($D_{13} \leq -0.82$), the data drops to the left child node. Secondly, the Parent Node 2 continues to split using the hyperplane $\overrightarrow{P_{14}}$. If the distance

to $\overrightarrow{P_{14}}$ less is than -2.28, it belongs to Class 2 , otherwise it belongs to Class 3. The same binary splitting process is applied to classify Class 1 and Class 4.

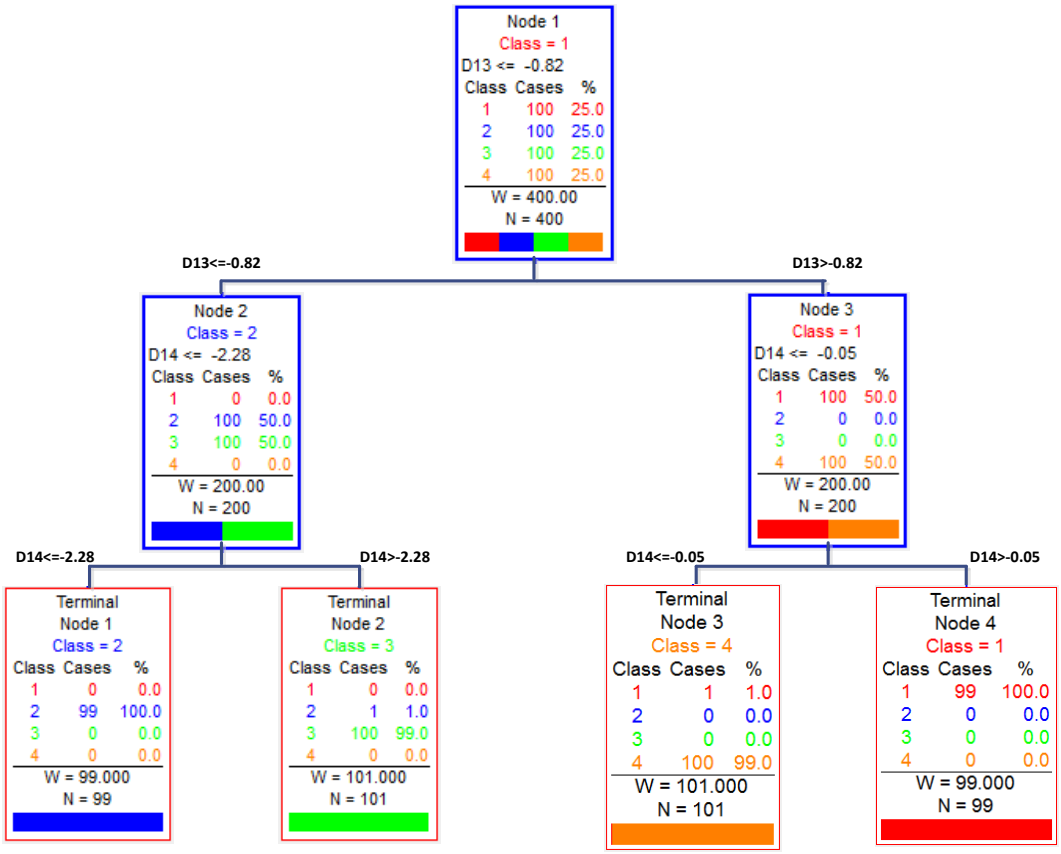


Figure 5. 18: Decision tree built for 4 groups' classification

It can be seen that the partition of the four groups are efficient and rapid with the dimensional reduction method. Two hyperplanes instead of six are needed for the classification of four classes of data, as shown in Figure 5.19. Two centroid vectors $(\overrightarrow{\mu_1} - \overrightarrow{\mu_4})$ and $(\overrightarrow{\mu_1} - \overrightarrow{\mu_3})$ are needed, as depicted by the two red lines in the figure.

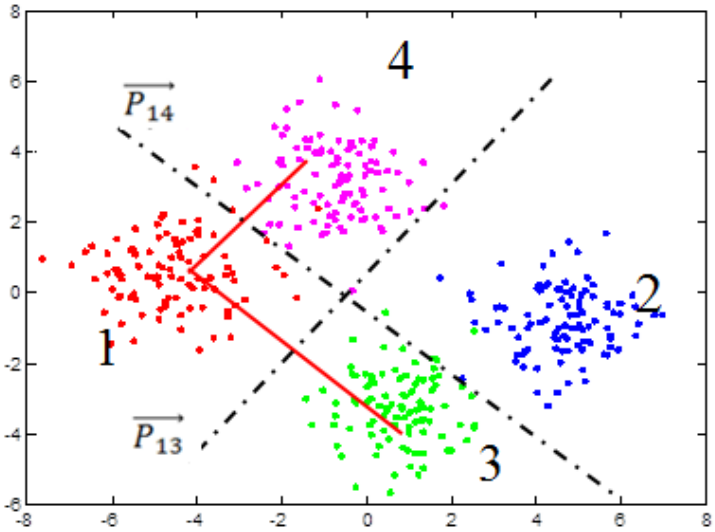


Figure 5. 19: Decision tree built for four groups' classification

From Table 5.3, it can be realized that a very high prediction success rate (99%) is obtained.

Table 5. 3: Prediction success of the testing sample (4 classes)

Actual Class	Total Class	Percent Correct	Observed Class			
			1	2	3	4
1	100	99.00%	99	0	0	1
2	100	99.00%	0	99	1	0
3	100	99.00%	0	1	99	0
4	100	99.00%	1	0	0	99
Total	400					
	Average % Correct	99.00%				
	Overall % Correct	99.00%				

Chapter 6 Simulation Result

6.1 Test System and Fault Type

Three test systems are used in this dissertation to verify the proposed transient stability prediction method. They are the 17-machine, 162-bus System, the 29-machine, 128-bus WECC system and the 4000-bus WECC system.

The transient stability definition used for the test systems is defined as follows:

1. For a small system like the 17-machine, 162-bus system and the 29-machine, 128-bus WECC system, when all the generators in the system maintain synchronized, the system is defined as stable.
2. For a large scale system, like the 4000-bus WECC System, the Out of Step effect of one small generator on the whole system is negligible. Only when a significant proportion of the total capacity or a large number of generators go out of step, then it will pose a threat to the system's stability. Keeping this in mind, in this dissertation, for the 4000-bus WECC System, when the loss of synchronized generation is more than 200 MW, the system is defined as unstable.

Two types of three-phase faults are simulated at every bus in the system, as shown in Figure 6.1 and Figure 6.2. The first type is defined as a branch fault, which is a three-phase short circuit branch fault and which lasts for 80 milliseconds. The fault is cleared by disconnecting the faulted branch. The second type is defined as a severe bus fault, which is a three-phase short circuit bus fault that lasts for 300 milliseconds. The fault is cleared by disconnecting all the branches connected to the faulted bus. Since the larger number of training cases used to build the decision tree, the confidence of the prediction is higher. To generate more cases, different operating conditions were simulated under the $N-1$ contingency criterion. The topology of the system was changed by tripping one of the transmission lines for every contingency case. The three-phase faults were uniformly applied on all the buses of the system.

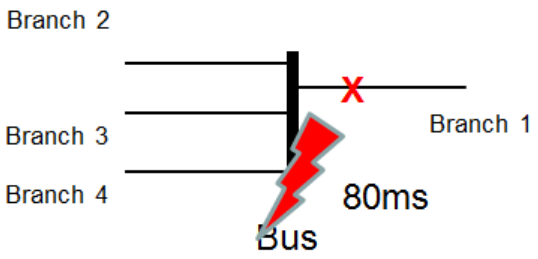


Figure 6. 1: Branch Fault

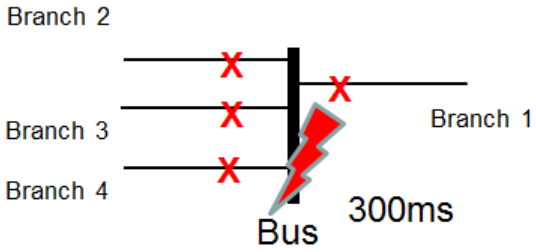


Figure 6. 2: Bus Fault

6.2 Test in the 17-Machine, 162-Bus System

Test 1: Tree Size Comparison Before and After the FLD

As mentioned in Chapter 5, the FLD is an improvement over the LDA technique when the two groups of data have different variances. In this section, the 17-machine, 162-bus system is used to validate the improvement by observing the decision tree size built by CART®. The single line topology of the 17-machine, 162-bus system is shown in Figure 6.3. By simulating three phase fault for 80 milliseconds and 300 milliseconds, training cases were created for different bus fault locations, along with $N-1$ contingency in which one of the branches were tripped at a time. 21,634 cases with different topologies were thus created. Among them, 9,442 were unstable cases and 12,192 were stable cases. Recording the apparent impedance trajectories of all the branches in the system, and integrating the 24 dimensional data into one single index (distance), the decision tree was built in CART®.

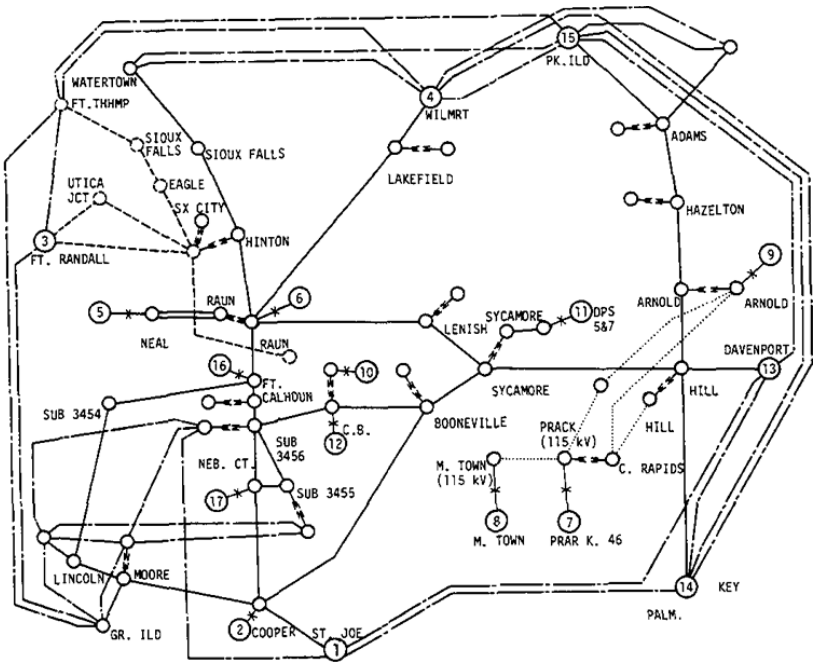


Figure 6. 3: Single line diagram of the 17 gen 162 buses System [25]

Figure 6.4 and Figure 6.5 show the size difference of the decision tree built by CART® before and after the FLD improvement. Before applying the FLD improvement, the optimal decision tree had more than 300 nodes, and the relative cost of the tree was 0.526. After applying the FLD improvement, the optimal tree had only 89 nodes, implying that the decision tree became much smaller. Furthermore, the relative cost of the decision tree reduced to 0.191 after the FLD improvement, which meant that the later tree is more efficient in separating the stable and the unstable cases.

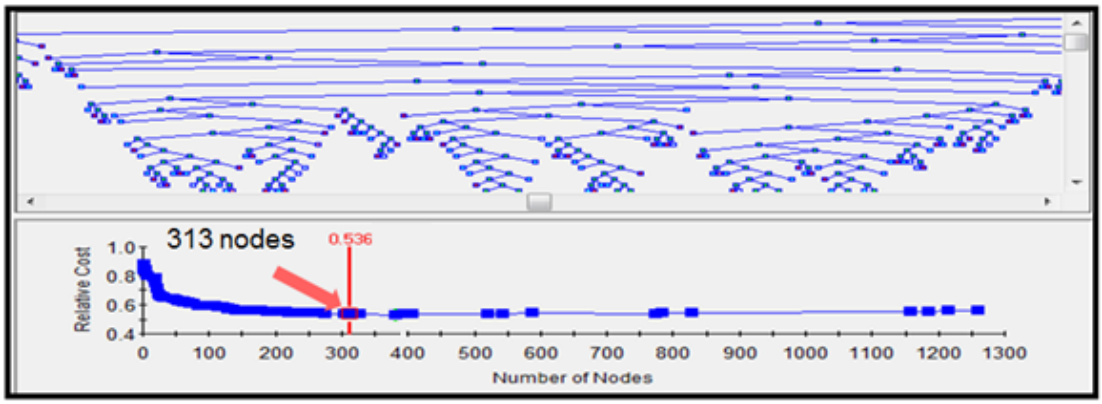


Figure 6. 4: Decision tree size in CART® before FLD

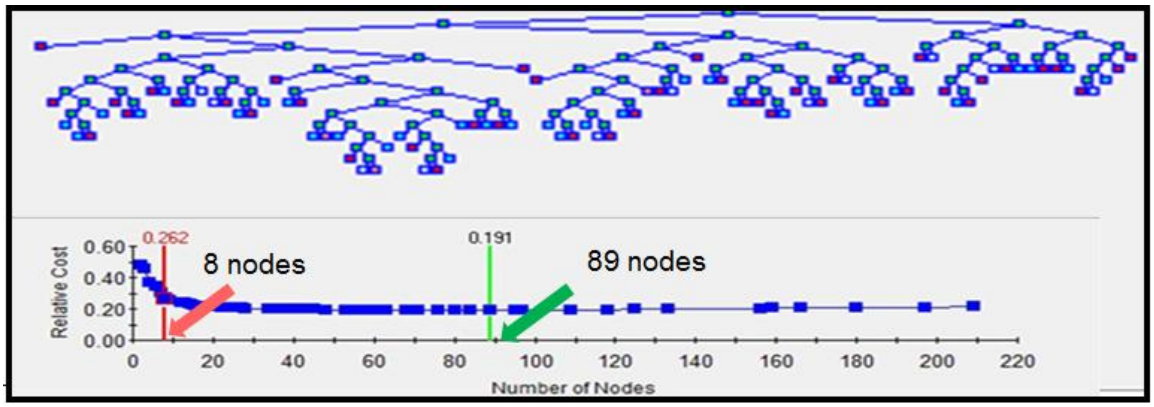


Figure 6. 5: Decision tree size in CART® after FLD

As mentioned in Chapter 4, to test the prediction success of the decision tree, ten-fold cross-validation had been employed, in accordance with which, CART® divided the entire sample into ten subsets. It took nine subsets to construct the decision tree, and used the remaining one subset to test the tree. The same process was repeated until each part of the data had been tested. The overall test error was the summation of the ten testing samples errors for each tree. The prediction success of the decision tree built in CART® using a 12-cycle data window before applying the FLD improvement is shown in Table 6.1. It needs 313 nodes to build an optimal tree with 82.94% of prediction success for the testing sample. On pruning the tree to a smaller size and maintaining an 80% prediction success rate, it still needed 85 nodes.

Table 6. 1 Prediction success of decision tree before FLD

Comparison with different nodes	313 nodes of the decision tree	85 nodes of the decision tree
Train sample success	93.96%	86.12%
Testing sample success	82.94%	80.13%

The prediction success of the testing sample after applying the FLD improvement is shown in Table 6.2. The optimal decision tree consisted of 89 nodes with a 97.25% training success rate, and a 95.14% testing success rate, which is more accurate in comparison to one obtained by the former tree (without the FLD improvement). To maintain a 90% testing success rate, only eight nodes were needed to split the stable cases

and the unstable cases. Therefore, it can be seen that the transient stability prediction based on the apparent impedance trajectory with FLD improvement works well under topology changes (the apparent impedance trajectory variances of the stable case and the unstable cases are different). Thus, with only eight nodes the proposed approach of using FLD is able to provide a prediction of higher than 90% accuracy for this system.

Table 6. 2 Prediction success of decision tree after FLD

Comparison with different nodes	89 nodes of the decision tree	8 nodes of the decision tree
Train sample success	97.25%	90.34%
Testing sample success	95.14%	90.41%

Test 2: Prediction Success with 116KV Branch Out of Service

In reality, the branches of certain voltage level are more likely to trip in different operation condition. In this system, 116 KV branch is more susceptible to the switching operation. Therefore, the topology of the system is changed by tripping one of the 116 KV transmission lines for every case. 21,546 cases are created, among them, 5,750 were unstable cases and 15,796 were stable cases. The decision tree built using a 12-cycle data window had 53 nodes with 99.12 % of overall training success, and 98.24 % of overall testing success. To maintain a 95% testing success rate, only 5 nodes were needed to build the decision tree. From the results obtained in this simulation, it is proved that the apparent impedance trajectory prediction works better when considering those branches that are most frequently out of operation.

Table 6. 3 Prediction success for the 17-machine, 162-bus system

Comparison with different nodes	53 nodes of the decision tree	5 nodes of the decision tree
Training sample success	99.12%	95.64%
Testing sample success	98.24%	95.00%

In accordance with the Ten-Fold Cross-Validation employed, CART® used 9/10 of the data to be the training sample to train the tree. The prediction success of the training sample can be seen in Table 6.4. Class 0 stands for the stable cases, while Class 1 stands for the unstable cases. 68 out of 15,796 stable cases are misclassified as unstable cases, while 122 out of 5,750 unstable cases are misclassified as the stable cases. The prediction success of unstable cases is 97.88%, while the stable cases prediction success is 99.57% in training sample.

Table 6. 4 Prediction success of training sample

Actual Class	Total Class	Prediction Accuracy	Observed Class	
			1	0
1	5,750	97.88%	5,628	122
0	15,796	99.57%	68	15,728
	Average % Correct	98.61%		
	Overall % Correct	99.12%		

Similar to what was done in the previous paragraph, CART® used 1/10 of the data randomly as the testing sample to test the decision tree built using the training sample. It can be seen from Table 6.5 that the prediction success rate of the decision tree is a slightly lower than the one obtained using the training sample, but that it was still pretty high (>95%).

Table 6. 5 Prediction success of testing sample

Actual Class	Total Class	Percent Correct	Observed Class	
			1	0
1	5,750	95.97%	5,518	232
0	15,796	99.07%	147	15,649
	Average % Correct	97.52%		
	Overall % Correct	98.24%		

6.3 Test in the 29-Machine, 128-Bus WECC System

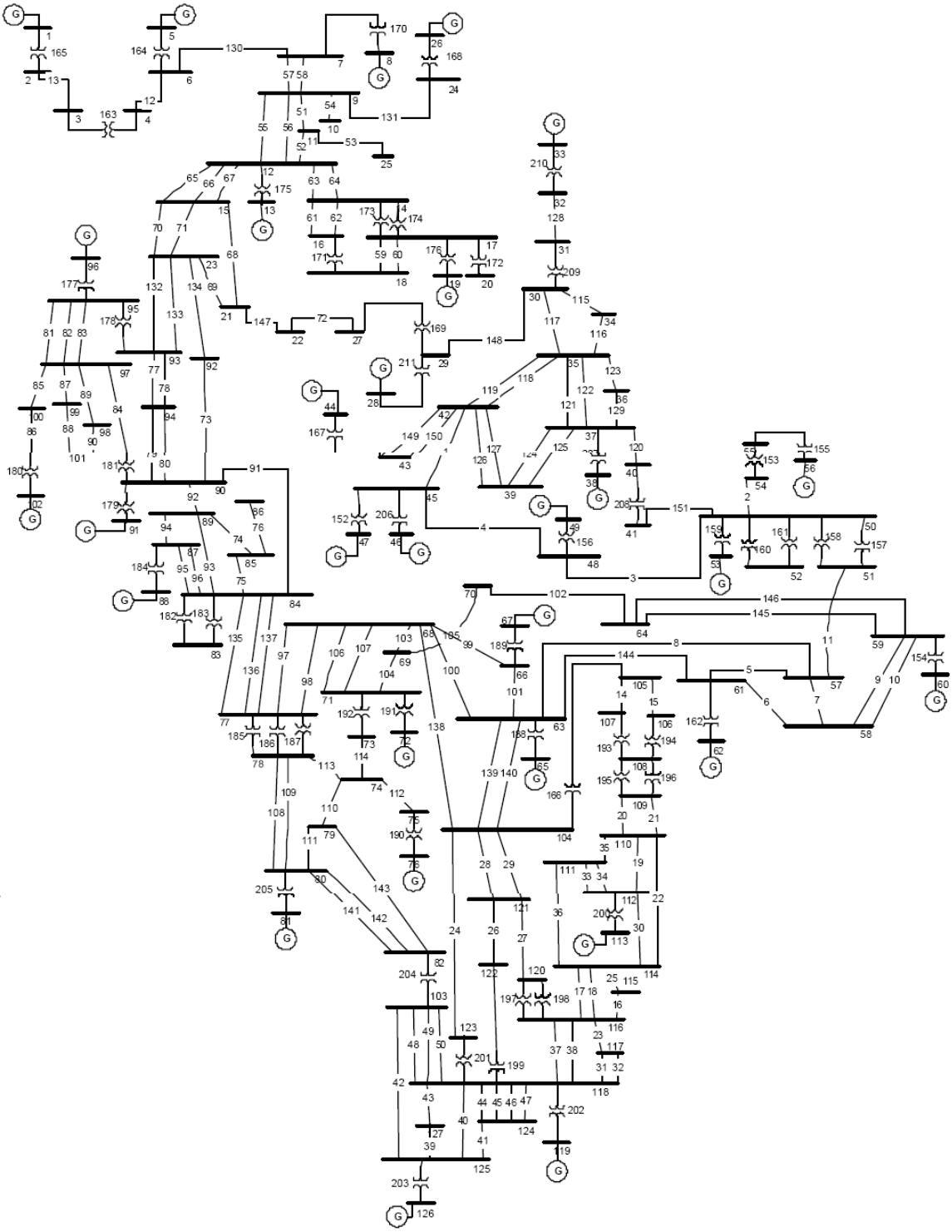


Figure 6. 6: Single Line Diagram of the 128-bus WECC system [95]

Test1: Prediction Success in the 128-Bus WECC System

In the 128-bus WECC system [95], similar $N-1$ contingencies cases were run to create different system topologies. 10,640 three phase faults were created, of which 2,333 cases resulted in a stable system and in 8,307 cases, the system lost stability. From Figure 6.7, it can be seen that the relative cost of the optimal decision tree is pretty small (0.103), indicating that the decision tree can separate the two classes efficiently. With 17 nodes of the decision tree built with a 12-cycle data window, the testing sample had 97.43% prediction success. After pruning the decision tree to 7 nodes, it still had a 95.44% prediction success rate for the testing cases.

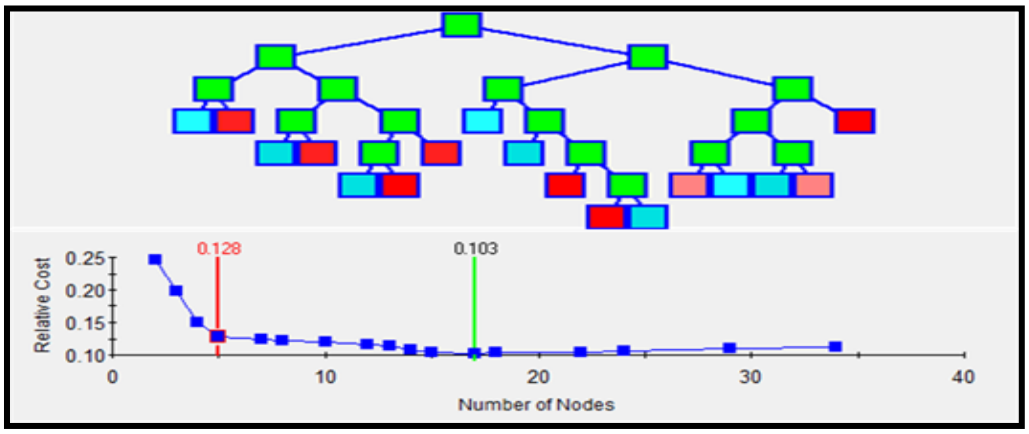


Figure 6. 7: Decision tree size in CART® of 128-bus WECC system

Table 6. 6 Prediction success for the 128-bus WECC system

Comparison with different nodes	17 nodes of the decision tree	7 nodes of the decision tree
Train sample success	98.67%	95.63%
Testing sample success	97.43%	95.44%

A. Optimizing PMUs Placement For Transient Stability Prediction

The process of building the decision tree is performed offline using simulations. In real-time the implemented decision tree is traversed using PMU measurements to classify the system as stable or unstable. As mentioned above, the 128-bus WECC system only requires 7 terminal nodes to get 95.44% of prediction accuracy. The "parent nodes" (branch nodes) of the decision tree shown in Figure 6.8 denote the optimal locations for

PMU placement. Parent Nodes A19, A62, A91, A31, A41 and A1 denote the branch apparent impedance trajectories which have been selected for performing the classification. The corresponding branches of these parent nodes can be seen in Table 6.7, it can be realized that the following branches: lines 22-27, 74-78, 104-123, 37-39, 56-128, and 2-3 should be monitored by PMUs. Therefore only six buses are required for PMUs to be installed to obtain a high-level of prediction accuracy (> 95%). The desired buses are buses 27, 78, 123, 39, 128, and 3. Of these, buses 27 and 123 were 500kV buses; bus 39 was a 345kV bus, bus 3 and 78 were 230kV buses, and bus 128 was a generator bus. The PMU placement for the transient stability in the 128-bus WECC system is shown in Figure 6.9. These locations were chosen as strategic places for installing the PMUs. In real-time the implemented decision tree is traversed using PMU measurements obtained from these buses to classify the system as stable or unstable.

Table 6. 7 Relationship among parent nodes selected by CART®, branches in the system and PMUs placement (tree 1)

Decision Tree Nodes	Parent nodes selected by CART®	Branches of the system	Buses selected for PMUs Placement
Node 1	A19	Line 22-27	27
Node 2	A62	Line 74-78	78
Node 3	A91	Line 104-123	123
Node 4	A31	Line 37-39	39
Node 5	A41	Line 56-128	128
Node 6	A1	Line 2-3	3

The decision tree built in CART® has 7 terminal nodes, and the branch nodes shown in the decision tree are the critical locations. With a few PMUs installed at these critical locations identified as a first step in the proposed method, it is possible to forecast the system stability by studying the apparent impedance trajectory of the branches monitored by PMUs in real time (about 200 milliseconds). The main advantage of this method is its simplicity as the PMUs located in these strategic places can record the apparent impedance trajectories in real time without any previous calculations. With the pre-built decision tree, it is very straightforward to predict stability, which indeed makes it a very fast prediction method.

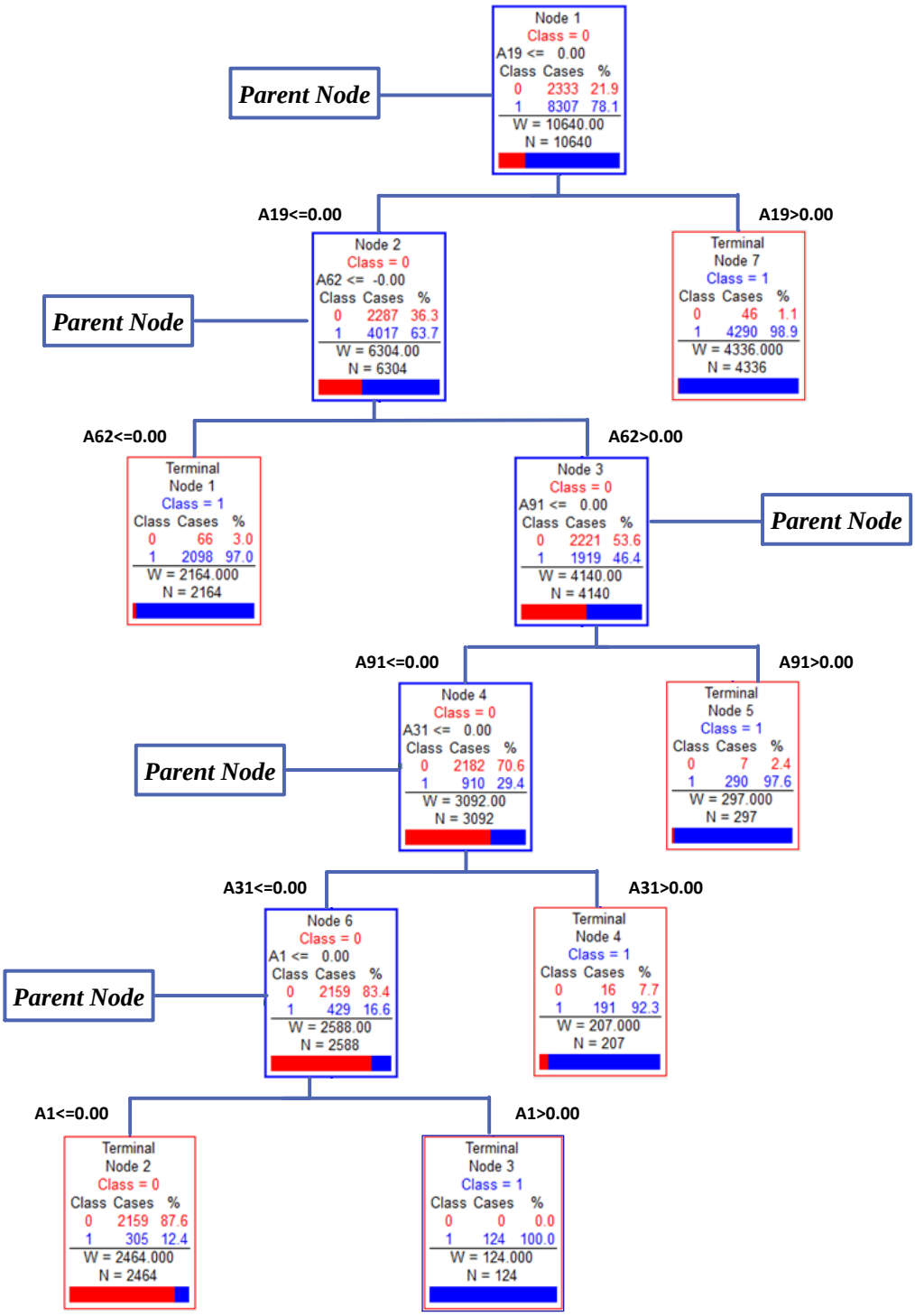


Figure 6. 8: Decision tree of the 128-bus WECC system (tree 1)

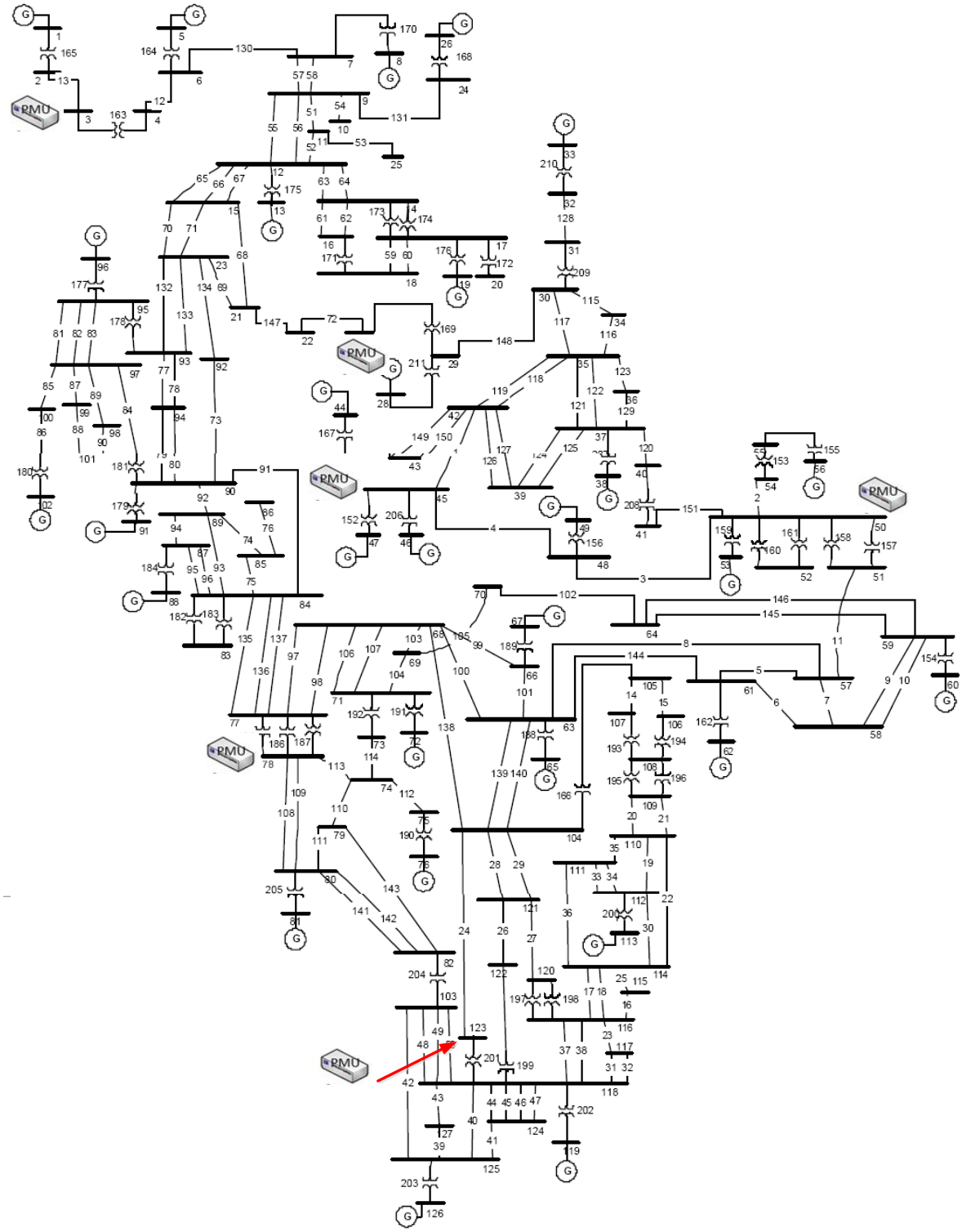


Figure 6. 9: PMUs location of the 128-bus WECC system of decision (tree 1)

B. Surrogate of the Decision Tree of 128-Bus WECC System

As mentioned in Chapter 4, CART® uses the surrogates to replace the missing attributes to build the decision tree. Suppose the primary split Attribute A19 is missing, the first surrogate of A19 is A76 with an association of 0.6. The predictive association measures how good the surrogate imitates the primary split.

Table 6. 8 Surrogate of the primary split of tree 1

Primary Node	Surrogate	Association
1	A76	0.6

In order to test the performance of the surrogate replacing, we build an alternate tree with the surrogate replacing the primary split in CART®. The optimal decision tree with the surrogate primary split A76 had 23 terminal nodes. On pruning the tree while maintaining a prediction success of more than 90% in the testing sample, only 7 terminal nodes are required.

Table 6. 9 Prediction success of the alternate decision tree

Comparison with different nodes	23 nodes of the decision tree	7 nodes of the decision tree
Training sample success	98.21%	91.16%
Testing sample success	97.85%	90.93%

The detailed decision tree with more than 90% of prediction accuracy is shown in Figure 6.10. For the alternate decision tree built by CART®, although the important attributes partitioning the stable/unstable cases are different, the tree size is more or less the same as the original tree with primary split A19 (only 6 parent nodes are needed to spit the stable/unstable cases).

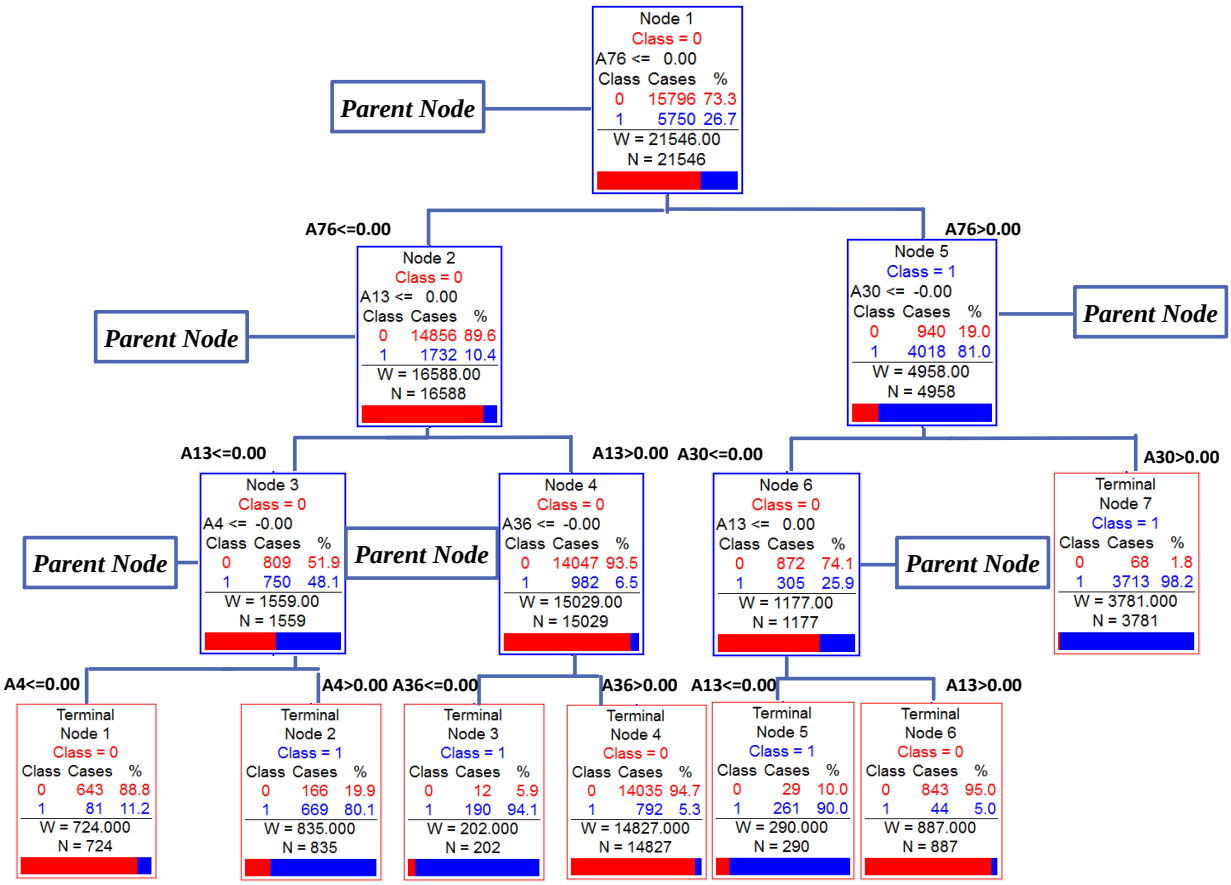


Figure 6. 10: Alternate decision tree of the 128-bus WECC system (tree 2)

The relationship among the parent nodes, the branches of the system and the PMUs placements is shown in Table 6.10. In the new tree, only 5 buses are selected as critical locations for the PMUs placements. They are buses 90, 16, 9, 45, and 37.

Table 6. 10 Relationship among parent nodes selected by CART®, branches in the system and pmu placement (tree 2)

Decision Tree Nodes	Parent nodes selected by CART®	Branches of the system	Buses selected for PMUs
Node 1	A76	Line 89-90	90
Node 2	A13	Line 14-16	16
Node 3	A4	Line 7-9	9
Node 4	A36	Line 42-45	45
Node 5	A30	Line 36-37	37
Node 6	A13	Line 14-16	16

C. Test with Different Number of the Sample Data

In the previous simulations, trajectory having 12 data points was used. In this dissertation, we try to find a suitable data window of apparent impedance trajectory to predict transient stability, the tree providing an optimal balance between size and prediction success must be chosen for execution. Different numbers of time series have also been tested and their effects on the result are compared in Table 6.11 for the 17-machine, 162-bus system. From the Table, we can see that the larger number of data points we observe during simulation, the less nodes we need to build the decision tree. To get 98% of prediction success, the 10-data point trajectory needs 56 nodes, while the 12-data point trajectory needs 53 nodes, and the 18-data point trajectory needs only 40 nodes for prediction. On pruning the tree to 5 nodes, the 10-data point can only get 89.2% of accuracy, whereas the 12 and 18-data points get about 94.5% of accuracy. From this table, two conclusions can be drawn:

- As more data from the trajectory are observed, greater accuracy is obtained.
- Balancing the decision tree size and the prediction accuracy involved, the 12-data point trajectory is accurate enough for prediction purposes.

Table 6. 11 Comparison of different sample data

	10 Data Point	12 Data Point	18 Data Point
Node no. of optimal tree	56	53	40
Train sample success	99.31%	99.45%	99.53%
Testing sample success	98.42%	98.61%	98%
Testing success at 5 Nodes	89.2%	94.5%	94.52%

6.4 Test with Multiple Classes

Multiple classes' classification was tested on the 17-machine, 162 bus system. 3,826 cases were created in total. Four classes of data were defined,

- Class 1: Branch fault with stability (1000 cases)
- Class 2: Branch fault with instability (1000 cases)
- Class 3: Bus fault with stability (1000 cases)
- Class 4: Bus fault with instability (826 cases)

The branch fault is defined as a three phase branch fault lasting for 80 milliseconds, the branch trips after the clearance of the fault. The bus fault is defined as a

three phase fault lasting for 300 milliseconds, and for which all lines connected to the faulted bus are tripped after clearance of fault. On performing the simulations, it was observed that an optimal tree with 7 nodes could split the four groups of data with more than 99% accuracy. The results are summarized in Table 6.12.

Table 6. 12 Classification of four classes in 17 gen 162 buses system

Actual Class	Total Class	Prediction Accuracy	Observed Class			
			1	2	3	4
1	1000	99.80%	998	2	0	0
2	1000	99.70%	3	997	0	0
3	1000	100%	0	0	1000	0
4	826	99.39%	0	0	5	821
	Average % Correct	99.72%				
	Overall % Correct	99.74%				

Furthermore, the trajectory of the apparent impedance could discriminate the branch fault from the bus fault completely (since there is no overlap between Classes 1 and 2 with Classes 3 and 4). The bus fault with stability (Class 3) has the highest prediction success which means its trajectory is dominant compared with the other three groups of data. The detailed decision tree built in CART® is shown in Figure 6.11. With 6 terminal nodes, the optimal decision tree can split four classes of contingency accurately. In the decision tree, 6 parent nodes are needed to split the stable/unstable cases. The relationship among parent nodes, branches of the system and PMU placement is shown in Table 6.13. The buses for PMUs placement selected by the decision tree are all 345kV. They are buses 25, 5, 126, 124, and 112.

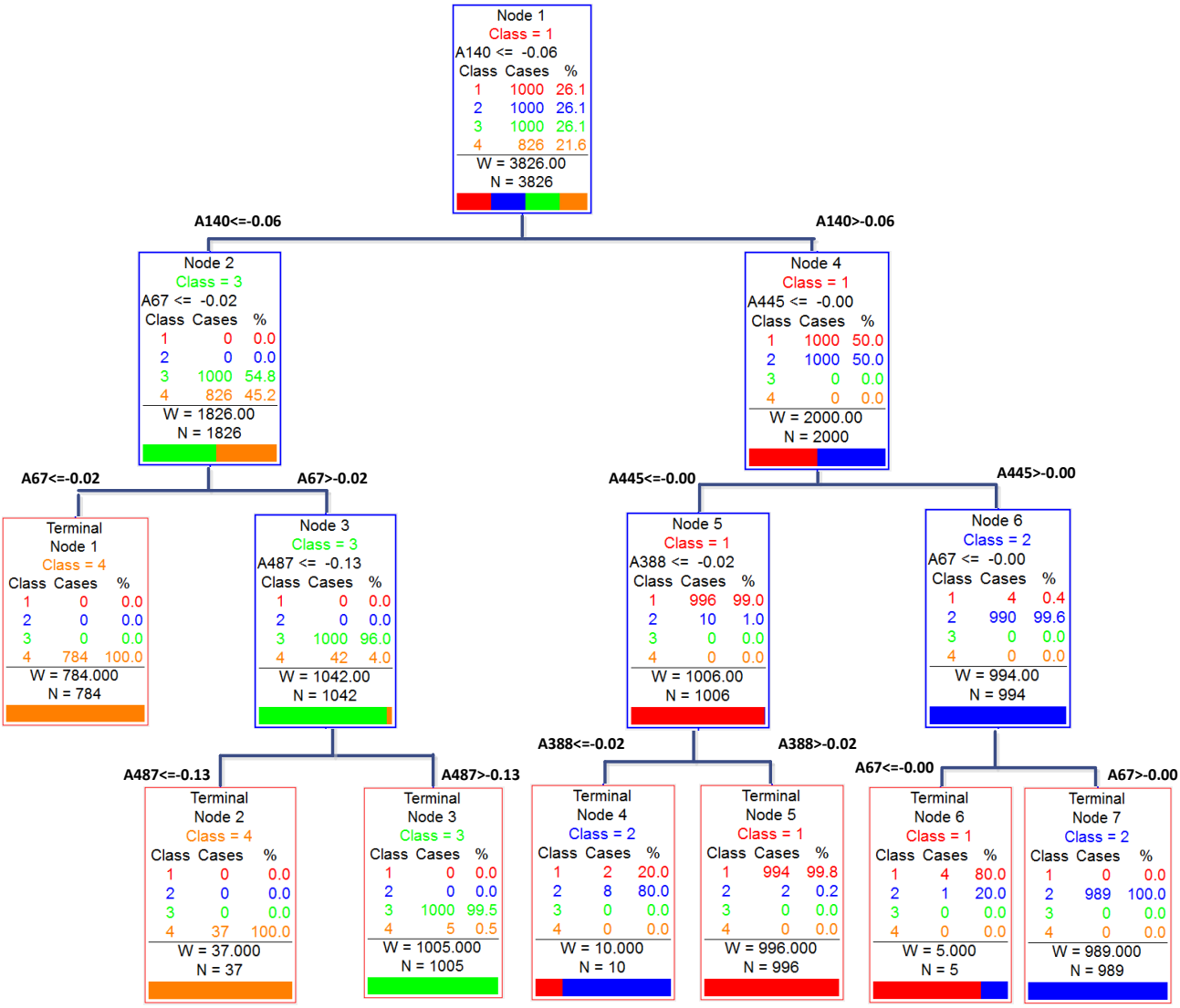


Figure 6. 11: Decision tree of the 17 gen 162 buses System (four classes)

Table 6. 13 Relationship among parent nodes selected by CART®, branches in the system and PMUs placement

Decision Tree Nodes	Parent nodes selected by CART®	Branches of the system	Buses selected for PMUs Placement
Node 1	A140	Line 25-27	25(345KV)
Node 2	A67	Line 5-131	5(345KV)
Node 3	A487	Line 126-62	126(345KV)
Node 4	A445	Line 124-3	124(345KV)
Node 5	A388	Line 112-120	112(345KV)
Node 6	A67	Line 5-131	5(345KV)

The surrogate buses for these buses are shown in Table 6.14. The high predictive associations of the surrogates indicate that, for this case, it is easy to replace the primary attributes when they are missing.

Table 6. 14 List of surrogates (four classes)

Decision Tree Nodes	Primary Split	Surrogate	Predictive Association
Node 1	A140<=-0.06	A141<=-0.014	1.00
Node 2	A67<=-0.02	A72<=-0.065	1.00
Node 3	A487<=-0.13	A488<=-1.69	1.00
Node 4	A445<=-0.00	A29<=0.0079	0.99
Node 5	A388<=-0.02	A413<=-0.083	0.6
Node 6	A67<=-0.00	A72<=-0.0053	0.8

For the final set of experiments, the same contingency cases were computed as before but with only two class classifications: stable (Class 0) and unstable (Class 1). This decision tree built in CART® also had 7 terminal nodes and is shown in Figure 6.12 with the classification results presented in Table 6.15. From the results given in Table 6.15 it can be seen that the decision tree with the same number of nodes obtains slightly lower accuracy. This is quite understandable because 6 hyperplanes would partition the data more comprehensively and efficiently in comparison to only one hyperplane. The surrogates of the two class classification are listed in Table 6.16. Compared with Table 6.14, the lower predictive association of the surrogates means that when the primary attributes are missing in the two class classification, the decision tree is more difficult/ less effective to find the substitute buses than the four classes' classification.

Table 6. 15 Classification of two classes in 17 gen 162 buses System

Actual Class	Total Class	Percent Correct	Observed Class	
			1	0
1	1826	99.18%	1811	15
0	2000	99.30%	14	1986
Total	3826		1825	2001
	Average% Correct	99.24%		
	Overall% Correct	99.24%		

Table 6. 16 List of surrogates (two classes)

Decision Tree Nodes	Primary Split	Surrogate	Predictive Association
Node 1	A46<=0.00	A44<=-0.0015	1.00
Node 2	A12<=-0.06	A59<=0.0733	0.99
Node 3	A5<=0.00	A26<=0.01509	0.68
Node 4	A33<=-0.84	A21<=-0.1654	0.95
Node 5	A23<=0.01	A80<=0.2463	0.58
Node 6	A11<=-0.09	A69<=0.1427	0.48

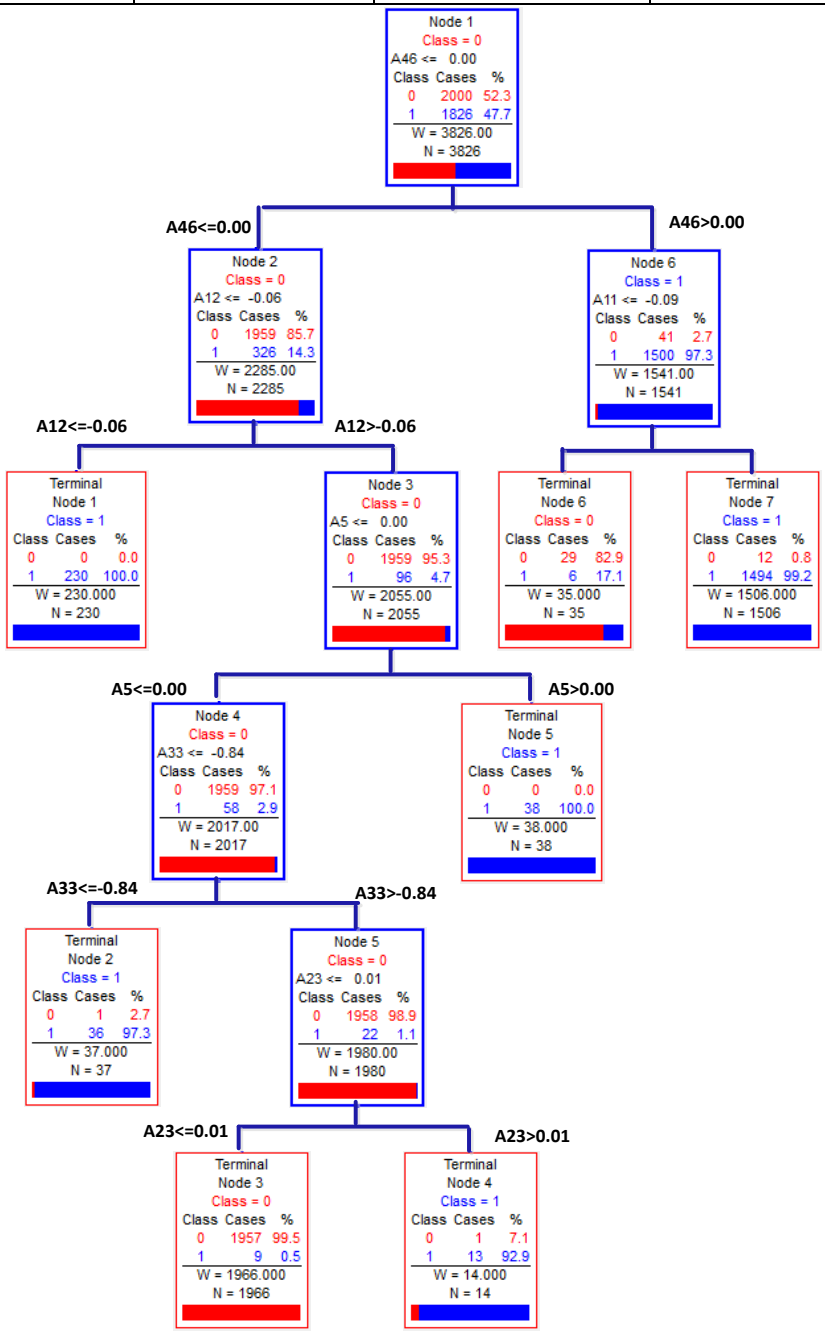


Figure 6. 12: Decision tree of the 17 gen 162 buses System (two classes)

6.5 Test in WECC 4000 Buses System

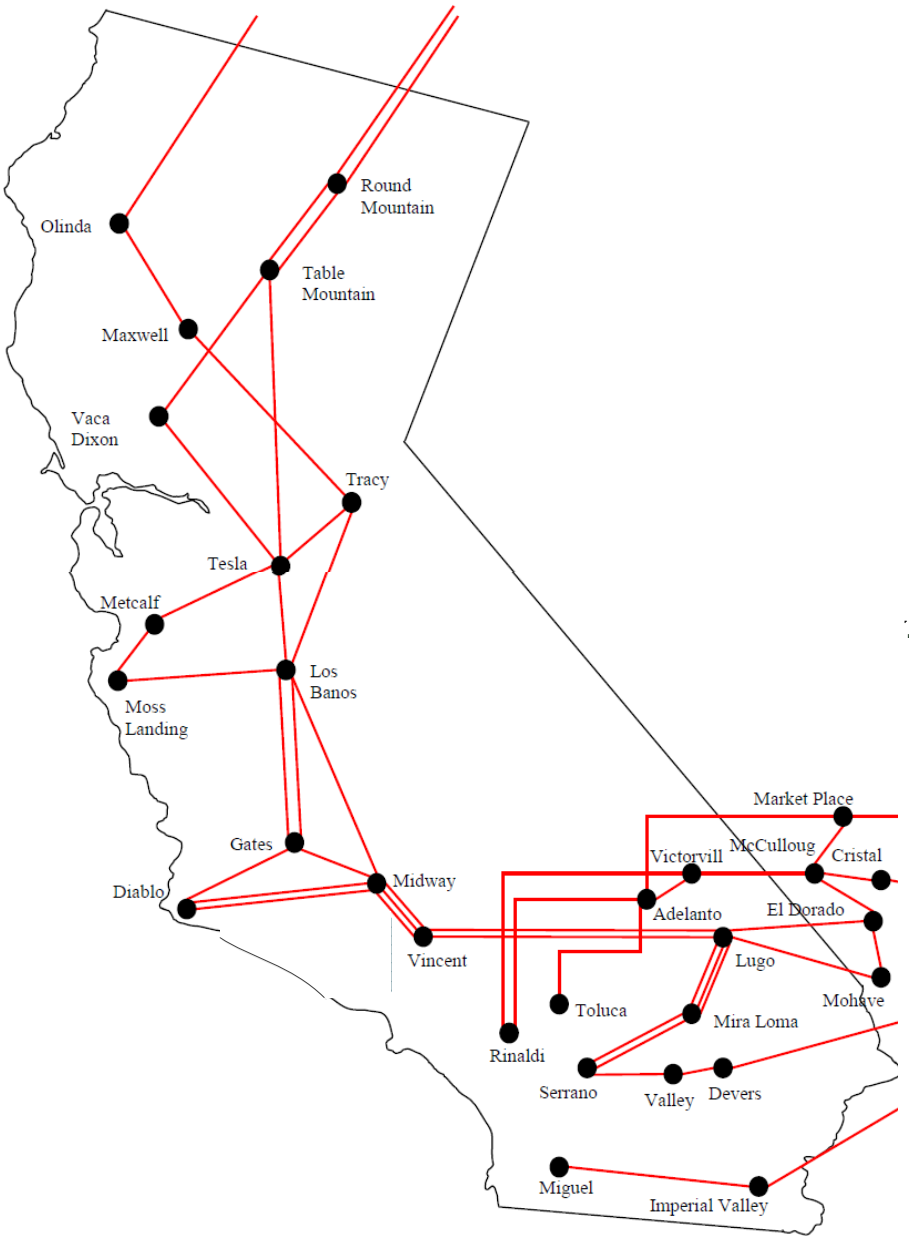


Figure 6. 13: 500 kV buses and lines in California [57]

The transient stability prediction method based on the apparent impedance trajectory recorded by PMUs has been validated in the previous sections. In this one, a

larger and more realistic 4000 bus model of the California system is used to test the proposed approach. The total actual generation capacity of this system is 2851.3 MW. The threshold of the instability set in this dissertation for this system is 200 MW, which is approximately 7% of the total generation capacity. In other words, if the total capacity of out of step generators is larger than 200 MW, the system is considered to be unstable, otherwise it is considered to be stable. By simulating three phase fault for 300 milliseconds at each bus, 9869 cases with different topologies and $N-1$ contingencies were created. 8813 cases of them were stable cases, while in the remaining 1056 cases, the system became unstable. The decision tree grown in CART® is shown as below.

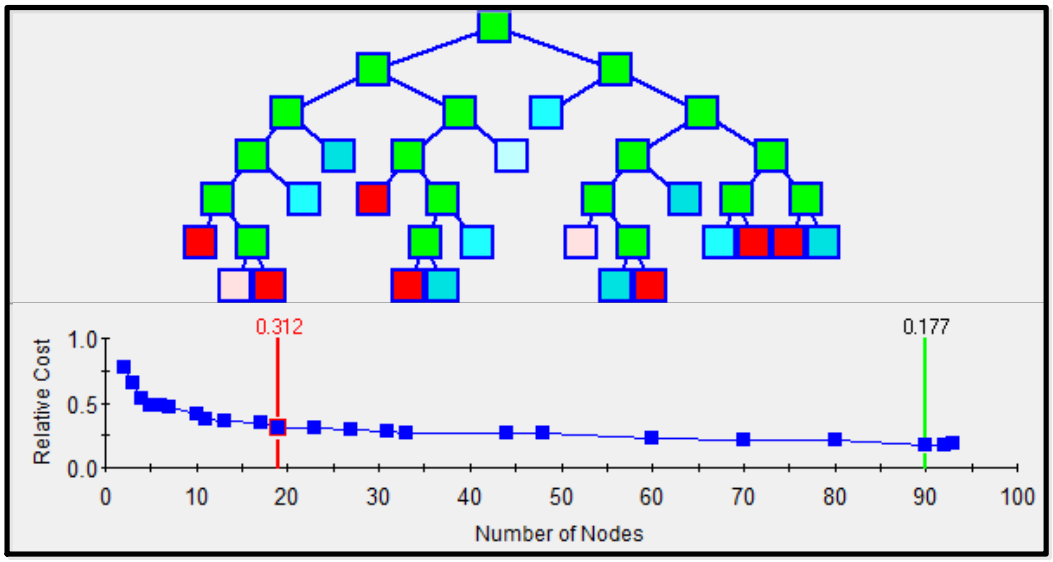


Figure 6. 14: Decision tree of 4000 buses California System built in CART®

The optimal decision tree built in CART® has 92 nodes, the prediction success of the training sample and the testing sample are shown in Table 6.17 and Table 6.18 respectively.

Table 6. 17 Prediction success of the training sample

Actual Class	Total Class	Prediction Accuracy	Observed Class	
			1	0
1	1056	97.92%	1034	22
0	8813	99.82%	16	8797
	Average% Correct	98.87%		
	Overall% Correct	99.61%		

Table 6. 18 Prediction success of the testing sample

Actual Class	Total Class	Prediction Accuracy	Observed Class	
			1	0
1	1056	90.72%	958	98
0	8813	98.99%	89	8724
	Average% Correct	94.85%		
	Overall% Correct	98.11%		

In the testing sample, 98 of the unstable cases are misclassified as stable cases while 89 of the stable cases are misclassified as unstable cases. On pruning the decision tree to 19 nodes a prediction accuracy of over 90% was still obtained as shown in Table 6.19.

Table 6. 19 Classification of two classes in WECC 4000 buses system

Comparison with different nodes	92 nodes of the decision tree	19 nodes of the decision tree
Train sample success	99.61%	92.75%
Testing sample success	98.11%	92.84%

The corresponding transmission lines selected by CART® as important attributes to build the decision tree are shown in Table 6.20. There are nearly fifty 500 KV buses in the WECC system, which are usually installed with PMUs at these buses. However, only 16 buses of the 500 KV buses needed to be installed with PMU, indicated by the red color in the table. They are NAVAJO, PALOVRDE, ADELANTO, MCCULLGH, TABLE MT, OLINDA, TESLA, LOSBANOS, MOSSLAND, DIABLO, MIDWAY, IMPRLVLY, LUGO, ELDORDO, DEVERS, and ROUND MT. With prior knowledge of the distances of the data points from a previously computed hyperplane, a decision tree can be built in CART® to predict transient stability based on the apparent impedance trajectory in real time.

Table 6. 20 Critical transmission lines selected in WECC 4000 buses system

Transmission Line selected by CART®	From bus	To bus	Buses selected for PMUs Placement
8	NAVAJO	MOENKOPI	NAVAJO
9	NAVAJO	MOENKOPI	
10	HASSYAMP	PALOVRDE	PALOVRDE
12	HASSYAMP	PALOVRDE	
13	ADELANTO	TOLUCA	ADELANTO
14	ADELANTO	VICTORVL	
17	MARKETPL	ADELANTO	
20	MARKETPL	MCCULLGH	MCCULLGH
21	MCCULLGH	VICTORVL	
24	MCCULLGH	VICTORVL	
28	ADELSVC	ADELANTO	
32	LOSBANOS	GATES	
33	ROUND MT	TABLE MT	TABLE MT
39	ROUND MT	TABLE MT	
42	TABLE MT	TESLA	
46	OLINDA	MAXWELL	OLINDA
50	VACA-DIX	TESLA	TESLA
52	TRACY	TESLA	
53	TRACY	LOSBANOS	LOSBANOS
54	TESLA	METCALF	
55	TESLA	LOSBANOS	
56	METCALF	MOSSLAND	MOSSLAND
57	MOSSLAND	LOSBANOS	
58	LOSBANOS	GATES	
60	LOSBANOS	MIDWAY	
64	GATES	DIABLO	DIABLO
65	GATES	MIDWAY	MIDWAY
67	DIABLO	MIDWAY	
68	DIABLO	MIDWAY	
69	IMPRLVLY	MIGUEL	IMPRLVLY
71	N.GILA	IMPRLVLY	
76	LUGO	MIRALOMA	LUGO
82	LUGO	MOHAVE	
86	MOHAVE	ELDORDO	ELDORDO
88	DEVERS	VALLEYSC	DEVERS
93	MALIN	ROUND MT	ROUND MT
99	CAPTJACK	OLINDA	
107	PALOVRDE	DEVERS	
122	ELDORDO	MCCULLGH	
123	LUGO	VICTORVL	

6.6 Discussion

6.6.1 Accuracy

It can be seen from the classification results that the prebuilt decision tree obtained more than 90% of prediction accuracy with few nodes. We also compare with the prediction success of the methodology proposed in this dissertation to that obtained in [96]. From the comparison it becomes clear that by using the apparent impedance trajectory directly, one can predict transient stability accurately and quickly. In [96], 90% success was achieved for the unstable cases while the success percentage for the stable cases was around 80%. In our DT, we got more than 90% success for both stable as well as unstable cases.

With the availability of the PMUs data and the historical data to replace the PSS/E simulation, which makes the proposed method be a measurement based method and the prediction of stability can be achieved by the offline prebuilt decision tree. No more simulations of the system are needed after the contingency, which make it more robust. Therefore, the application of the proposed method is model free and will be more accurate compared to the model based method. If there're not enough historical data to be traced, and the database are built on the simulations, which makes the performance of the tree depend on the accuracy of the model used for generating the training and the testing sets. In order to overcome these model-based limitations, the authors suggest letting the decision tree be put into practice in the field and allowing it to make decisions based on the measurements that it gets from the real system. However, the outputs that the tree generates are only observed, not any decisions or system modifications are made based on that output. Then, after it has run for a sufficient period of time (few months), the pattern of the outputs generated by the tree should be studied to check whether it is making the right predictions or not. Doing this will not only validate the model that had been used for making the assessment, but it will also corroborate the logic of using apparent impedance obtained directly from PMUs for making transient stability predictions in real-time.

6.6.2 Computational Burden

The technique does not present an excessive computational burden while generating the training and testing sets since all the cases can be computed in parallel. The computation process can be implemented automatically using the python programs in PSS/E®. The 10,640 cases of 128-bus WECC system took approximately 2 hours of wall clock time using 5 computers with Intel Core i3 CPU. For on line use purpose, high speed computers and the parallel computing technique will help a lot with this method. It is believed the computation process will speed up with more advanced computers.

6.6.3 Advantages of apparent impedance trajectories

The main advantages for using the apparent impedance trajectories to make transient stability prediction are:

1. Convenient: The emergence of dual use Line-relays/PMUs make it convenient to obtain time-tagged apparent impedances.
2. Fast: Based on our simulation results, only need 12 cycles (200 milliseconds) of the PMU data for transient stability prediction.
3. Accurate: According to our simulation, a more than 90% of prediction success rate can be obtained.

Chapter 7 Intelligent Controlled Islanding

The power system is operated closer to the stability limit than ever before and is vulnerable to the catastrophic failure. The prediction results of the transient stability introduced in Chapter 4 can be used as the online instability alarms and local control. When a power system encounters a series of large disturbances that may eventually lead to catastrophic blackouts, the traditional relay protection schemes to isolate the fault may not be enough to mitigate the consequences and ensure the stability of the system. Other pre-determined controlled protections, such as intelligent controlled islanding, need to be applied as well. Intelligent controlled islanding serves as the last defense of remedial action to prevent the collapse of a power system.

7.1 Methodology of Intelligent Islanding Scheme

This dissertation proposed a strategy to determine the initiated time and the boundaries of the intelligent islanding scheme. This research intends to search the load more effectively and determine the islanding boundary quickly in a large power system. The proposed method is coherency based controlled islanding which is different from reference [7,60], this approach aims to provide a general solution for determining the cut set.

For a speedy recovery of the intelligent control islanding, two main issues needs to be addressed:

- When to implement the controlled islanding.
- Where to split the system.

When: Assuming the coherency of the generators is detected at an early stage by some advanced techniques, the intelligent islanding would be considered as a possible remedial measure to restore the stability of the power system after a large disturbance when the transient instability has been predicted and the traditional contingency protections, such as relay tripping, is not likely to restore the system stability. The instability is predicted by the proposed method using 12-cycle apparent impedance trajectory that mentioned in Chapter 5.

Where: The location of intelligent islanding is determined in two steps. The first step of this method is finding the loads within the island. The load selection of the controlled islanding is based on electrical closeness instead of physical distance. The loads are selected relying on Relative Electrical Distance (RED) [97] between each generator and load bus. Ranking the magnitude of relative electrical distance from smallest to largest, we could determine which buses would be more influential on the dynamic performance of the generators, and could preliminarily select the loads within each island. The second step is determining the boundary lines for islanding by the non-zero element of the "mutual admittance matrix". The simplicity of load detection and boundaries determination makes this method very attractive.

Chapter five has already well addressed the problem of "when" to separate the system. In this chapter, we will focus on "where" to implement the controlled islanding. The proposed intelligent controlled islanding scheme is based on two assumptions:

1) The coherency of different generators is recognized by the advanced coherency identification technology after the contingency.
2) The buses with larger electrical distance must be connected by those buses with smaller electrical distance.

The intelligent controlled islanding scheme proposed in this research can be explained by the flow chart in Figure 7.1.

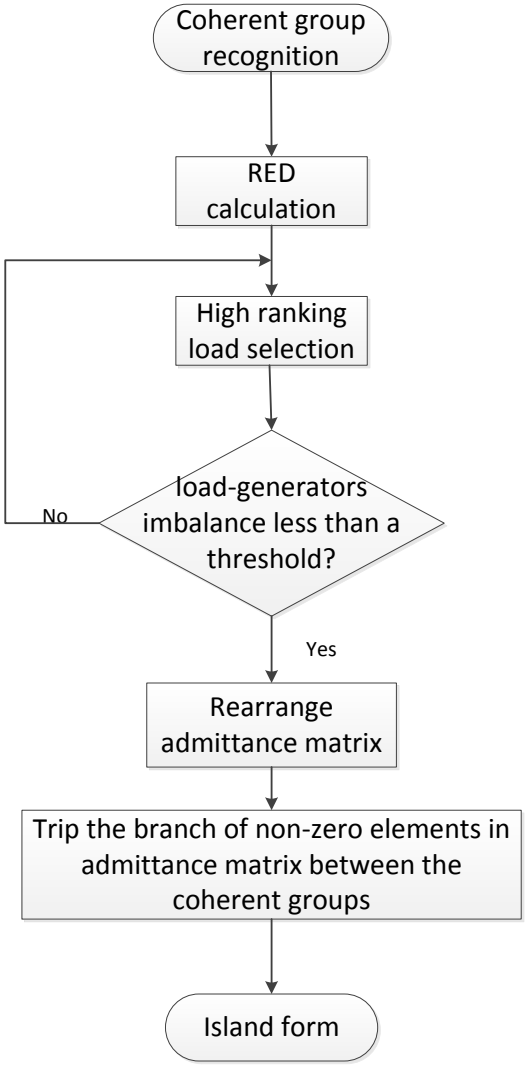


Figure 7. 1: Flow chart of intelligent islanding

A. Coherent Group Recognition

When initiating the intelligent islanding, the coherent groups of the generators need to be recognized. Generators belonging to the same group should remain in the same island. Suppose all the buses in a system are expressed as $B = B_G \cup B_L$; where B stands for all the buses in the system, B_G stands for the generator buses; and B_L stands for the other buses remaining in the system. Suppose n groups of coherent generators have been recognized, such that

$\sum_{i=1}^{n} B_{Gi} = B_G$, B_{Gi} is the i^{th}oherent group of generators. The generator buses in a different coherent group (the j^{th} coherent groups for example) belongs to the different island $B_{Gi} \cap B_{Gj} = \phi, i \neq j$

B. Load Selection by Relative Electrical Distance (RED)

The first step of the proposed controlled islanding method is to find out the loads within each island in order to minimize the generation output and load demand. Loads are selected by the relative electrical distance (RED) [97] ranking to balance the active power of the generators. The concept of the relative electrical distance was first proposed by Vishaka K. el. to estimate the relative electrical locations of the generators buses with respect to the load buses.

Ranking the RED index from smallest to largest, the loads with higher RED ranking with respect to the generator will have larger impacts on the generator rotor angle behavior compared to the lower ranking ones. Based on the topology of the network, the admittance matrix is constructed:

$$\mathbf{Y_b} = \begin{bmatrix} \mathbf{Y_{gg}} & \mathbf{Y_{gl}} \\ \mathbf{Y_{lg}} & \mathbf{Y_{ll}} \end{bmatrix} \tag{7-1}$$

in which, the subscript g stands for the generator buses and the subscript l stands for the other buses in the system. The voltages of all the buses can be obtained by (7-2)

$$\begin{bmatrix} \mathbf{E_g} \\ \mathbf{V_l} \end{bmatrix} = \begin{bmatrix} \mathbf{Y_{gg}} & \mathbf{Y_{gl}} \\ \mathbf{Y_{lg}} & \mathbf{Y_{ll}} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{I_g} \\ \mathbf{0} \end{bmatrix} \tag{7-2}$$

in which, $\mathbf{I_g}$ is the injection current of the generators. The voltage of the buses without generators can be expressed as:

$$\mathbf{V_l} = -(\mathbf{Y_{ll}})^{-1}\mathbf{Y_{lg}}\mathbf{E_g} \tag{7-3}$$

If we let $\mathbf{F}_{lg} = -(\mathbf{Y}_{ll})^{-1}\mathbf{Y}_{lg}$, we get a complex matrix that gives the relation between load bus and source bus voltages. The rows of $\mathbf{F}_{lg}$ correspond to the load buses, while the columns correspond to the generator buses.

Relative electrical distance (RED) is defined as:

$$RED = \mathbf{I} - abs(\mathbf{F}_{lg}) \tag{7-4}$$

$\mathbf{I}$ is a unity matrix.

In each coherent group of generators, the loads with shorter relative electrical distance should be selected to remain in the island.

The load searching will terminate when the following three requirements are met.

1. The next load in the ranking table is the terminal bus of a generator that belongs to other coherent groups.
2. The power imbalance between the generator-load is less than a threshold.
3. Bus/load selection of each generator via the RED ranking table must have at least one bus duplicated to guarantee the connection within the island.

According to the operation experience of the power system, the Out of Step relay trips one after another, which indicates that multiple coherent groups do not develop simultaneously. Any contingencies resulting in multiple groups of coherency are initiated by two coherent groups at the first stage. Therefore, the controlled islanding in this dissertation only focuses on two groups of contingency. In this research, only one coherent group with fewer generators will be chosen to select the load based on the RED method since it will make the speed of load searching faster. The remaining coherent group will take over the rest of the buses, which means one of the islands can guarantee the generator-load balance, but the remaining island may need proper load shedding or a generator curtailment scheme to restore the generator-load balance.

Since the basic configuration of the islands does not change dramatically within a short period of time when different lines are removed, the RED index table is prebuilt prior to the disturbance and can be applied during the contingency.

C. Islanding Boundary Determination

The second step of the proposed islanding method is to determine the boundary for controlled islanding in a short period of time using the admittance matrix. This is done by tripping the connecting lines between the two islands indicated by the non-zero elements of the "mutual-admittance matrix". The concept of the "mutual-admittance matrix" is explained later.

Take a simple 3-generator 5-bus system with two coherent groups for example,

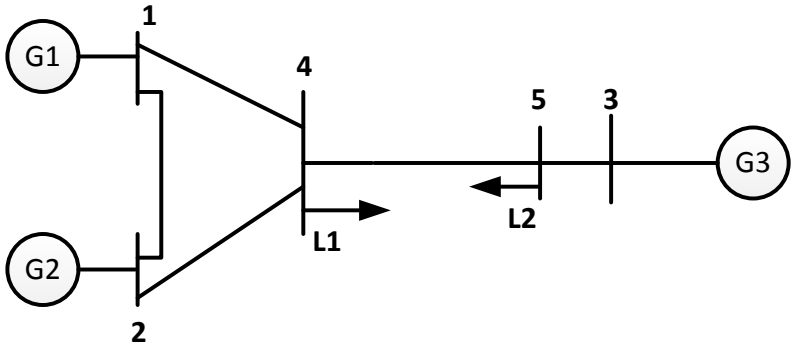


Figure 7. 2: Five buses illustration system

The admittance matrix is:

$$\mathbf{Y} = \begin{bmatrix} Y_{11} & Y_{12} & 0 & Y_{14} & 0 \\ Y_{21} & Y_{22} & 0 & Y_{24} & 0 \\ 0 & 0 & Y_{33} & 0 & Y_{35} \\ Y_{41} & Y_{42} & 0 & Y_{44} & Y_{45} \\ 0 & 0 & Y_{53} & Y_{54} & Y_{55} \end{bmatrix} \tag{7-5}$$

Only when there is connection between bus i and bus j, does $Y_{ij} \neq 0$

$$Y_{ij} = \begin{cases} c\,(c \neq 0) & connected \\ 0 & disconnected \end{cases} \tag{7-6}$$

The concept of the islanding boundaries is illustrated in Figure 7.3

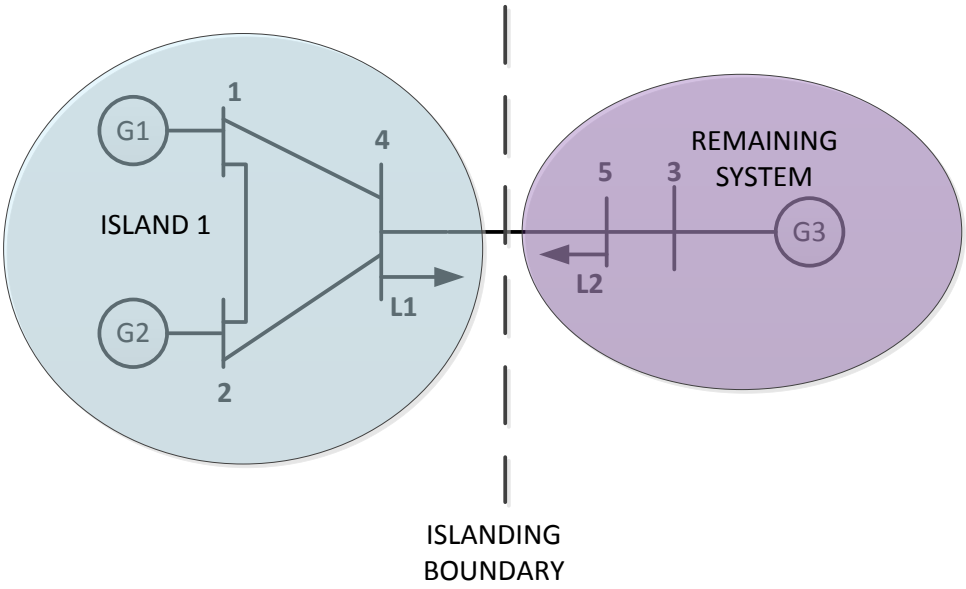


Figure 7. 3: Concept of intelligent islanding

Suppose generator G1 and G2 are identified in a coherent group, oscillating against generator G3, and load L1 is selected to balance the power output of G1 and G2 by the RED method. Therefore island1 includes G1, G2 and L1, while the remaining system, island2, which consists of G3 and L2. $\mathbf{G1, G2, L1} \in \mathbf{Y_{island1}}$, $\mathbf{L2, G3} \in \mathbf{Y_{island2}}$. The admittance matrix of each island is:

$$\mathbf{Y_{island1}} = \begin{bmatrix} Y_{11} & Y_{12} & Y_{14} \\ Y_{21} & Y_{22} & Y_{24} \\ Y_{41} & Y_{42} & Y_{44} \end{bmatrix}, \mathbf{Y_{island2}} = \begin{bmatrix} Y_{55} & Y_{53} \\ Y_{35} & Y_{33} \end{bmatrix} \tag{7-7}$$

Therefore, the structure of the admittance matrix can be expressed as

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y_{island1}} & \mathbf{Y_B} \\ \mathbf{Y_B^T} & \mathbf{Y_{island2}} \end{bmatrix} \tag{7-8}$$

Here, define $\mathbf{Y_{island1}}, \mathbf{Y_{island2}}$ as the "self-admittance matrix" of island 1 and island 2; $\mathbf{Y_B}$ is defined as the "mutual-admittance matrix" between the two islands in this dissertation. Rearranging the admittance matrix according to the sequence of islands as given in (7-7), the admittance matrix of (7-5) changes to (7-9).

$$\mathbf{Y} = \begin{bmatrix} Y_{11} & Y_{12} & Y_{14} & 0 & 0 \\ Y_{21} & Y_{22} & Y_{24} & 0 & 0 \\ Y_{41} & Y_{42} & Y_{44} & Y_{45} & 0 \\ 0 & 0 & Y_{54} & Y_{55} & Y_{53} \\ 0 & 0 & 0 & Y_{35} & Y_{33} \end{bmatrix} \tag{7-9}$$

In this case, the "mutual-admittance matrix": $\mathbf{Y_B} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ Y_{45} & 0 \end{bmatrix}$ (7-10)

The non-zero elements in the "mutual-admittance matrix" $\mathbf{Y_B}$ denote the lines that connect between island1 and island2. $\mathbf{Y_B} = \mathbf{0}$ means that the island 1 and island 2 are completely separated. Therefore, by tripping the intersection line 4-5 ($Y_{45} = 0$), island 1 and island 2 can be formed.

7.2 Simulation Results

The proposed controlled islanding scheme can be applied to both of the steady state and the transient conditions. In this section, the IEEE-39 bus system is tested in the steady state to illustrate the proposed controlled islanding scheme, while the WECC 128-bus system is tested under transient conditions to validate the feasibility of the proposed controlled islanding scheme.

Case Study 1: Steady State

During the normal operation, at any given time, it is possible to partition the system into desirable parts [61]. The locations selected to be tripped should minimize the imbalance between the load and generation within the island, which will result in the least disturbance of operation. The IEEE-39 buses system is tested for the controlled islanding method in the steady state. The topology of the system is shown in Figure 7.4

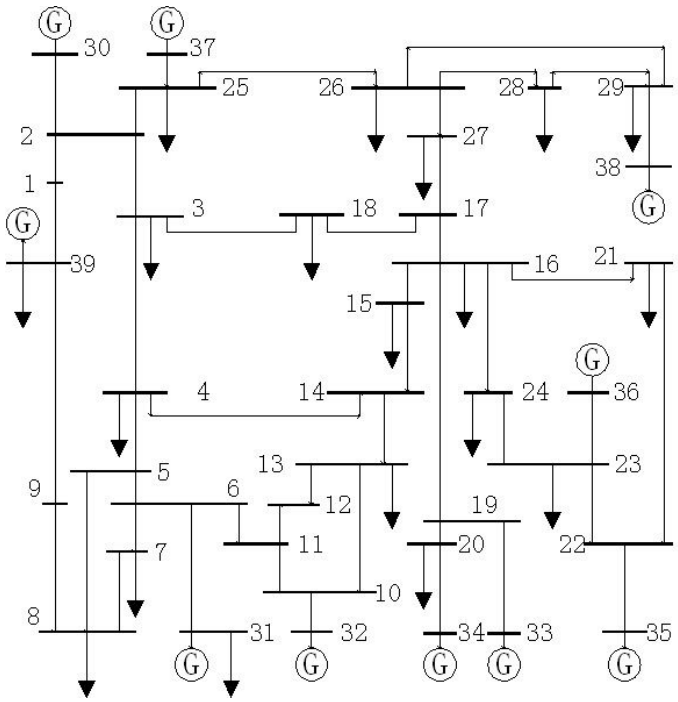


Figure 7. 4: IEEE-39 buses system

Suppose Gen 31 and Gen 32 are chosen to form a new island isolated from the other generators in the system. Following the flow chart of intelligent islanding in Figure 7.1, the steps are as follows.

Step 1: Calculate the relative electrical distance (RED) of every load bus with respect to every generator and build the load index table. The whole index table can be seen in Appendix A. Sorting the load bus by the RED from smallest to largest, the top 3 rankings with respect to GEN 31 and GEN 32 are shown in Table 7.1:

Table 7. 1 The top 3 ranking of load with respect to GEN 31 and GEN32

RANKING	GEN 31	GEN32
1	load 7	load 12
2	load 8	load 7
3	load 12	load 4

The load searching scheme is implemented according to the load sorting. For example, load 7 and load 12 are selected in the first depth of the load searching, since they are the loads with the smallest electrical distance with respect to GEN 31 and GEN

32 respectively. At the second depth of load searching, load 8 and load 7 are selected. Therefore, after two depths of load searching scheme, loads 7, 8, and 12 are chosen to match the generation output of GEN 31 and GEN 32. At t=2s, the generation output of Gen31 and Gen 32 is 1161.97MW, loads with high RED ranking should be accumulated in the island until the power imbalance between the generators and the load is minimized or less than a set threshold. In this case, only three depths of load searching are needed, shown as Table 7.2

Table 7. 2 Three depths of load searching to balance the capacity of GEN 31 and GEN32

DEPTH	Load selected	Load demand (MW)
1st	7,12	242.3
2nd	7,12,8	764.3
3rd	7,12,8,4	1264.3

From Table 7.2, the 3rd depth of load demand (1264.3 MW) is closer to generator capacity (1161.97 MW) than the 2nd depth of load demand (764.3 MW). Therefore, the load searching stops at the 3rd depth. In total, loads 7, 4, 8, and 12 are chosen in the three depths of RED load searching. The total load demand of these buses is 1264.3MW.

During the load searching, the buses with smaller RED value with respect to GEN 31 and GEN 32 than the selected load buses should also be included in the island. This is because the selected load buses must be reached by some of those buses according to the second assumption mentioned in section 7.2. Based on Table 7.3 and Table 7.4, the buses with smaller RED selected from the two generators to the load in these three depths searching are buses 6, 5, and 11 for GEN31; while buses 10, 13, 11, 14, 6, and 5 are for GEN32. Therefore, buses 5, 6, 10, 11, 13 and 14 should also be included in the island.

Table 7. 3 The RED of buses with respect to GEN 31

GEN 31		
DEPTH	BUS NO.	RED
	BUS6	0.636197
	BUS5	0.66219
1st	LOAD7	0.666766
2nd	LOAD8	0.682141
	BUS11	0.730154
3rd	LOAD12	0.754517

Table 7. 4 The RED of buses with respect to GEN 32

GEN 32		
DEPTH	BUS NO.	RED
	BUS10	0.531236
	BUS13	0.57995
1st	LOAD 12	0.585514
	BUS11	0.596021
	BUS14	0.692729
	BUS6	0.721147
	BUS5	0.729942
2nd	LOAD 7	0.740853
3rd	LOAD 4	0.749654

Step 2: Rearrange the admittance matrix with the sequence of $(\mathbf{Y}_{\mathbf{island1}}, \mathbf{Y}_{\mathbf{island2}})$

$(31,32,4,5,6,7,8,10,11,12,13,14) \in \mathbf{Y}_{\mathbf{island1}}$

$(1,2,3,9,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,33,34,35,36,37,38,39) \in \mathbf{Y}_{\mathbf{island2}}$

The non-zero elements of $\mathbf{Y}_{\mathbf{B}}$ are $Y_{3,4}$, $Y_{8,9}$ and $Y_{14,15}$. Therefore, to partition island 1 and island 2, lines 3-4, 8-9, and 14-15 need to be tripped. (The admittance matrix rearranged with the new sequence can be seen in Appendix B)

Case Study 2: Transients 1

The proposed controlled islanding scheme is also tested in the WECC 128-bus system for transient application. A three phase fault occurs at bus 50 at 0.2 seconds, and all the lines connected to the faulted bus are tripped to clear the fault after 300 milliseconds. Although load-shedding after the contingency was implemented, these three generators still lost synchronism. Without controlled islanding, the plot of all the rotor angles in the system is shown in Figure 7.5.

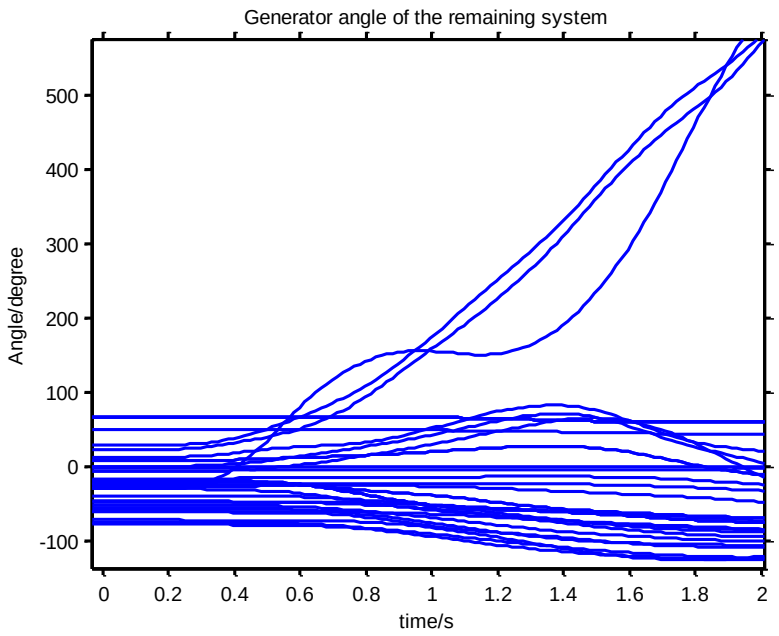


Figure 7. 5: Rotor angles without controlled islanding

At about 0.7 seconds, assuming three generators (generators CRAIG 46, HAYDEN 47 and SJUAN 49) are detected as the same coherent group oscillating against the other generators in the system by some advanced coherency recognition technology. Hence, there will be two islands in the whole system. Following the flow chart of intelligent islanding in Figure7.1:

Step 1: After recognizing that the generators CRAIG 46, HAYDEN 47 and SJUAN 49 should be isolated in an island, the relative electrical distance (RED) of every load bus with respect to the generators CRAIG 46, HAYDEN 47 and SJUAN 49 need to be calculated. At t=0.7s, the total generator capacity of the island is 3978.42MW. Built the load index table of island1, here list the top ranked of load with respect to each of these three generators:

Table 7. 5 The top ranking of load with respect to the three generators

RANKING	CRAIG 46	HAYDEN 47	SJUAN 49
1st	load 45	load 45	load 48
2nd	load 42	load 42	load 45
3rd			Load 42

It needs to be noticed that the bus 42 is the terminal bus of GEN 44 which belongs to the other coherent group. Therefore, the load buses further than bus 42 (have lower RED rankings) should not be included in the island. Therefore, in this case, only two depths of load searching are needed, as shown in Table 7.6:

Table 7. 6 One depth of load searching to balance the generation capacity of CRAIG 46, HAYDEN 47 and SJUAN 49

Depth	Load selected	Load demand (MW)
1^{st}	45,48	3284

Finally, the buses 46, 47, 49, 45, and 48 are selected to remain in the island. The total load of the island (load 45, 48) is 3284 MW. Although it is not exactly matched of the active power, it is the minimal imbalance of generator-load at that time stamp.

Step 2: Rearrange the admittance matrix according the sequence of $(\mathbf{Y}_{\mathbf{island1}}, \mathbf{Y}_{\mathbf{island2}})$.

$(46, 47, 49, 45, 48) \in \mathbf{Y}_{\mathbf{island1}}$

$(remaining\ \ buses\ \ in\ \ system) \in \mathbf{Y}_{\mathbf{island2}}$

The non-zero element of $\mathbf{Y}_{\mathbf{B}}$ is $Y_{45,42}$, therefore, to isolate the island 1 and island2, line 45-42 needs to be tripped. The islanding boundaries of CASE 1 are shown in Figure 7.9:

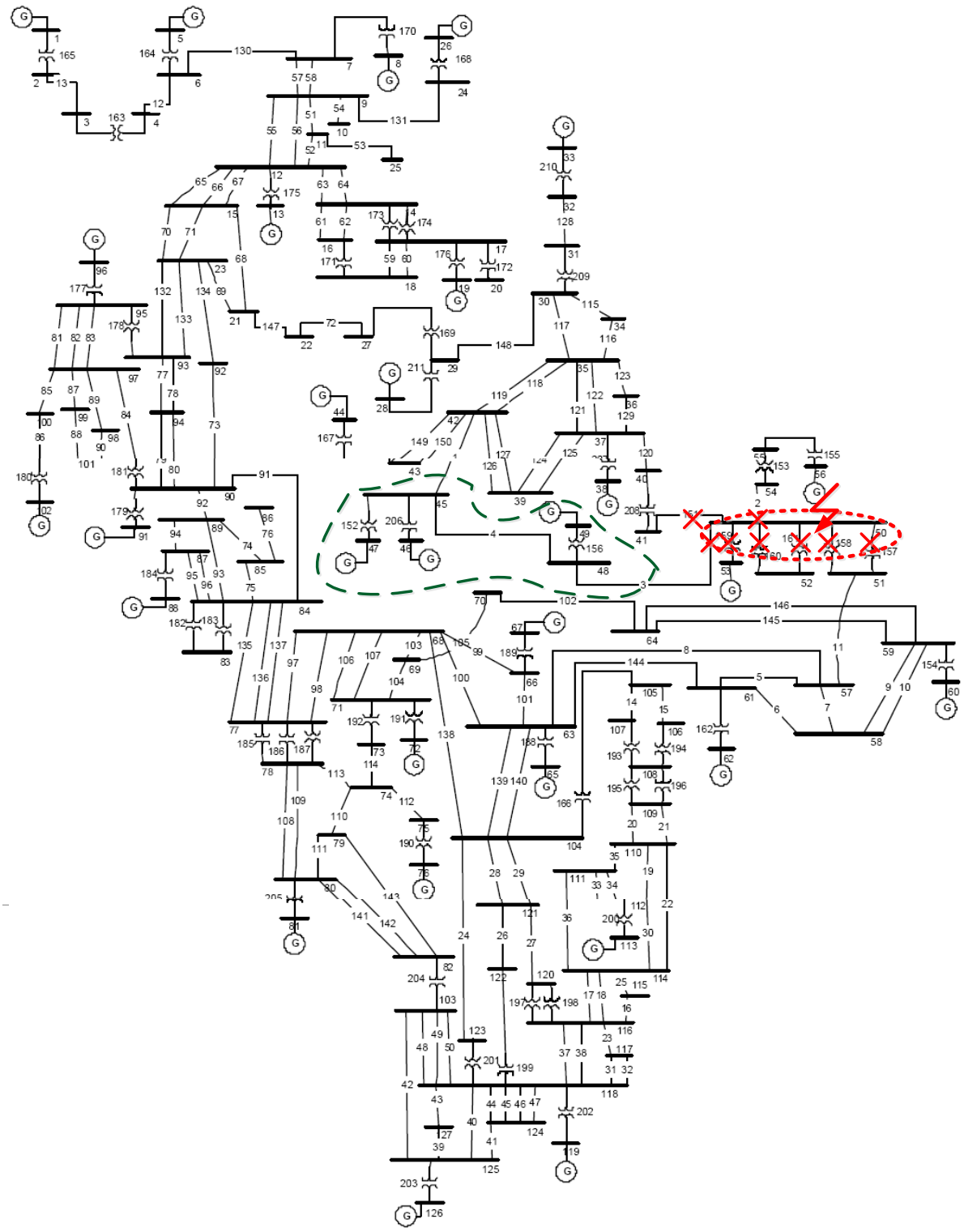


Figure 7. 6: Island form after three phase fault at bus 50

After the intelligent controlled islanding scheme proposed in this this dissertation, each resulting island has a balance between the generation output and the load demand, so that the islands can continue to operate at the nominal frequencies ($\Delta f < 0.1Hz$). All the voltages of the islands remained at the acceptable levels ($\Delta V < \pm 5\%$) and all the generators within each island remained synchronism. The rotor angles of the island and the remaining system are shown in Figure 7.7 and Figure 7.8:

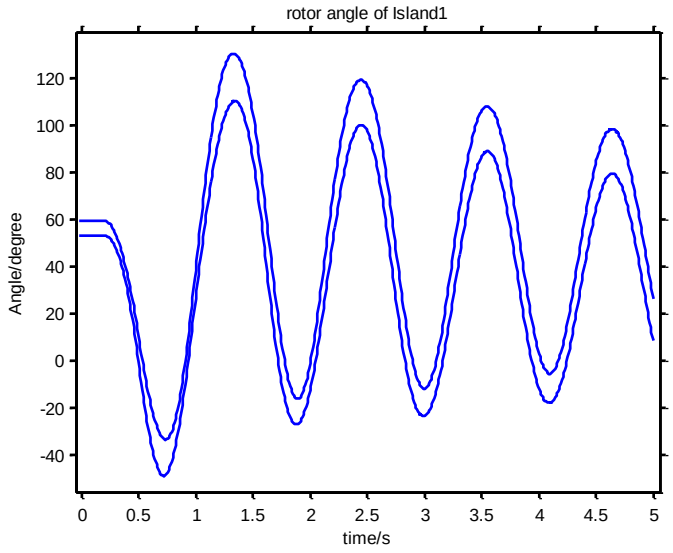


Figure 7. 7: The generator angle of the island

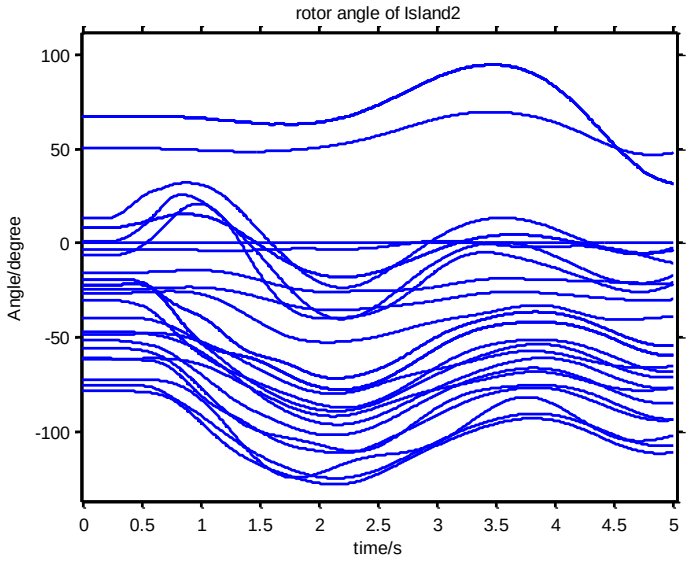


Figure 7. 8: The rotor angle of the remaining system

Figure 7.9 shows the active power flow in the islanding boundary (line 42-45) without intelligent islanding.

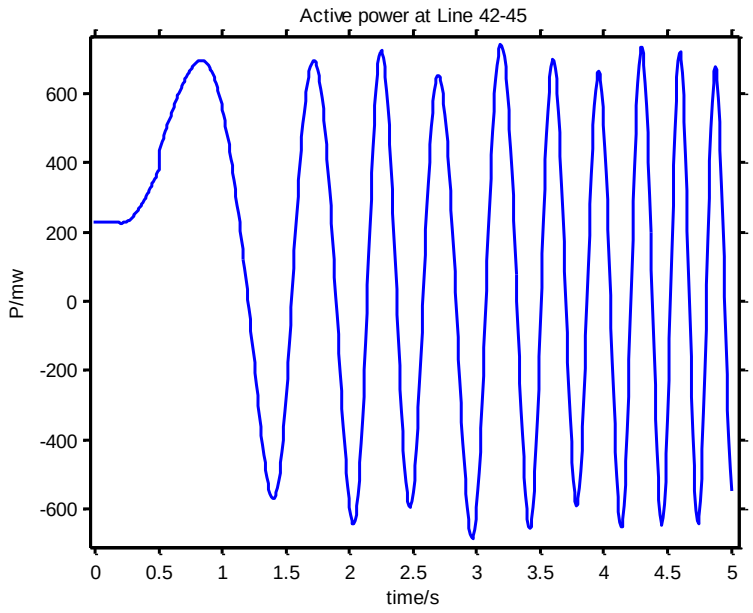


Figure 7. 9: Active power of Line 42-45

It can be seen from Figure 7.9 that the active power of line 42-45 fluctuates dramatically during the contingency, which is identical with the current criteria of islanding boundary's detection in the power system.

Case study 2: Transients2

A three phase fault is created at bus 68 at 0.2 seconds that the fault lasts for 300 milliseconds. The fault is cleared by tripping all the lines connected to bus 68. At t=0.8s, generator 72 is already out of step. At t=1.0s, two coherent groups are oscillating against each other:

Coherent group1: generator at bus 1, 5, 8, 13, 19, 26, 28, 88, 91, 96,102

Coherent group2: generator at bus 33, 38, 44, 46, 47, 49, 53, 60, 62, 65, 67, 76, 81, 113, 119, 126, 5699

The rotor angle plots for all generators without intelligent islanding are shown as Figure 7.10.

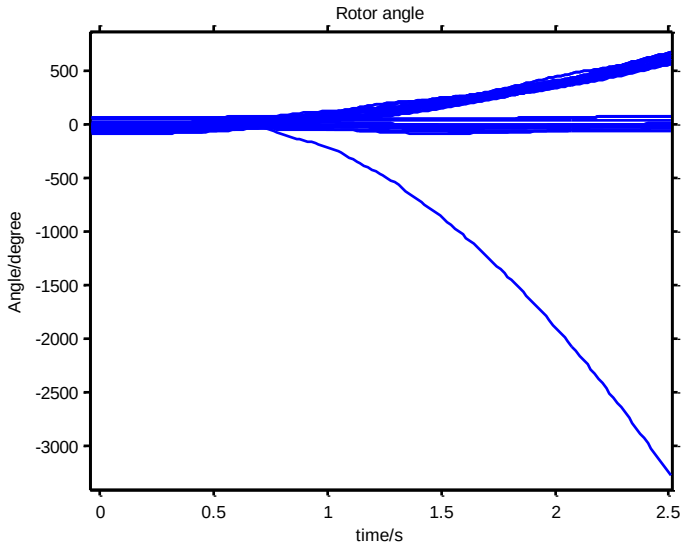


Figure 7. 10: Rotor angles without controlled islanding

The flow chart process of intelligent islanding shown in Figure 7.1 is thus followed.

Step 1: Calculate the relative electrical distance (RED) of every load bus with respect to the generators in coherent group 1. The detailed RED table of this coherent group in island 1 can be seen in Appendix C. At t=1.0s, the generators' capacity in island 1 is 36726.7 MW. Based on the RED index, Table 7.7 lists the top loads to match the generators' capacity in island 1.

Table 7. 7 The top ranking of load with respect to the generators in island 1

GEN in island 1	RED Ranking													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
GEN1	2	6	7	9	25	11	12	14	18	20	17	15	23	93
GEN 5	6	3	7	9	25	11	2	12	14	18	20	17	15	23
GEN 8	7	9	25	11	12	14	18	20	17	15	2	6	93	24
GEN 13	12	14	18	20	17	25	11	15	9	23	93	92	94	7
GEN 19	20	17	18	14	12	25	11	15	9	23	93	92	94	22
GEN 26	24	9	25	11	7	12	14	18	20	17	15	23	93	92
GEN 28	29	27	22	30										
GEN 88	87	83	89	85	86	84	68							
GEN 91	90	92	94	93	101	23	100	98	99	97	95	89	15	85
GEN 96	95	97	99	98	100	93	94	23	92	101	15	90	22	27
GEN 102	101	100	98	99	97	90	95	94	92	23	15	89	85	86

As mentioned in the second requirement of the load searching in Section B, the accumulation of load with respect to the generators in one island will stop when the next load/bus is the terminal buses of the generators which belongs to the other coherent group. Therefore the load searching of GEN 28 is stopped at bus 30, which is the terminal bus of generator 29. The load searching with respect to GEN 88 stops at bus 84 because the next ranked bus is 68, which has been tripped during the contingency. Next, loads with high ranking in the island are accumulated until the imbalance between generation output of coherent group 1 and the load demand is minimized. In this case, fourteen depths are needed for the load selection, which are shown in Table 7.8.

Table 7. 8 Fourteen depths of load searching to balance the capacity of island 1

Depth	Load selected	Load demand (MW)
1st	2,6,7,12,20,24,29,87,90,95,101	27163.45
2nd	2,6,7,12,20,24,29,87,90,95,101,3,9,14,17,27,83,92,97,100	31720.18
3rd	2,6,7,12,20,24,29,87,90,95,101,3,9,14,17,27,83,92,97,100, 25,18,22,89,94,99,98	33812.15
4th	2,6,7,12,20,24,29,87,90,95,101,3,9,14,17,27,83,92,97,100, 25,18,22,89,94,99,98,11,20,85,93,	35972.50
5th	2,6,7,12,20,24,29,87,90,95,101,3,9,14,17,27,83,92,97,100, 25,18,22,89,94,99,98,11,20,85,93,86	36015.76
6th	2,6,7,12,20,24,29,87,90,95,101,3,9,14,17,27,83,92,97,100, 25,18,22,89,94,99,98,11,20,85,93,86,84,23	36679.50
…		
14th	2,6,7,12,20,24,29,87,90,95,101,3,9,14,17,27,83,92,97,100, 25,18,22,89,94,99,98,11,20,85,93,86,84,23,15,19	36709.38

The load searching stops at depth 14th when the load demand (36709.38MW) is maximally match the generator capacity (36726.7MW) in island1. Also the connection between the generators in the island is guarantee by:

$$L_i, L_j \subset island\ ,\ load_i \in L_i, load_j \in L_j$$

$$\forall load_i \cap load_j \neq \phi, i \neq j$$

in which, L_i, L_j stand for the loads set selected by RED method with respect to generator GEN_i and generator GEN_j respectively.

Step 2: Rearrange the admittance matrix with the sequence of $(\mathbf{Y}_{\mathbf{island1}}, \mathbf{Y}_{\mathbf{island2}})$.

(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,83,84,85,86,87,88,89, 90,91,92,93,95,96,97,98,99,100,101,102) $\in \mathbf{Y}_{\mathbf{island1}}$

$(remaining\ \ buses\ \ in\ \ system) \in \mathbf{Y}_{\mathbf{island2}}$

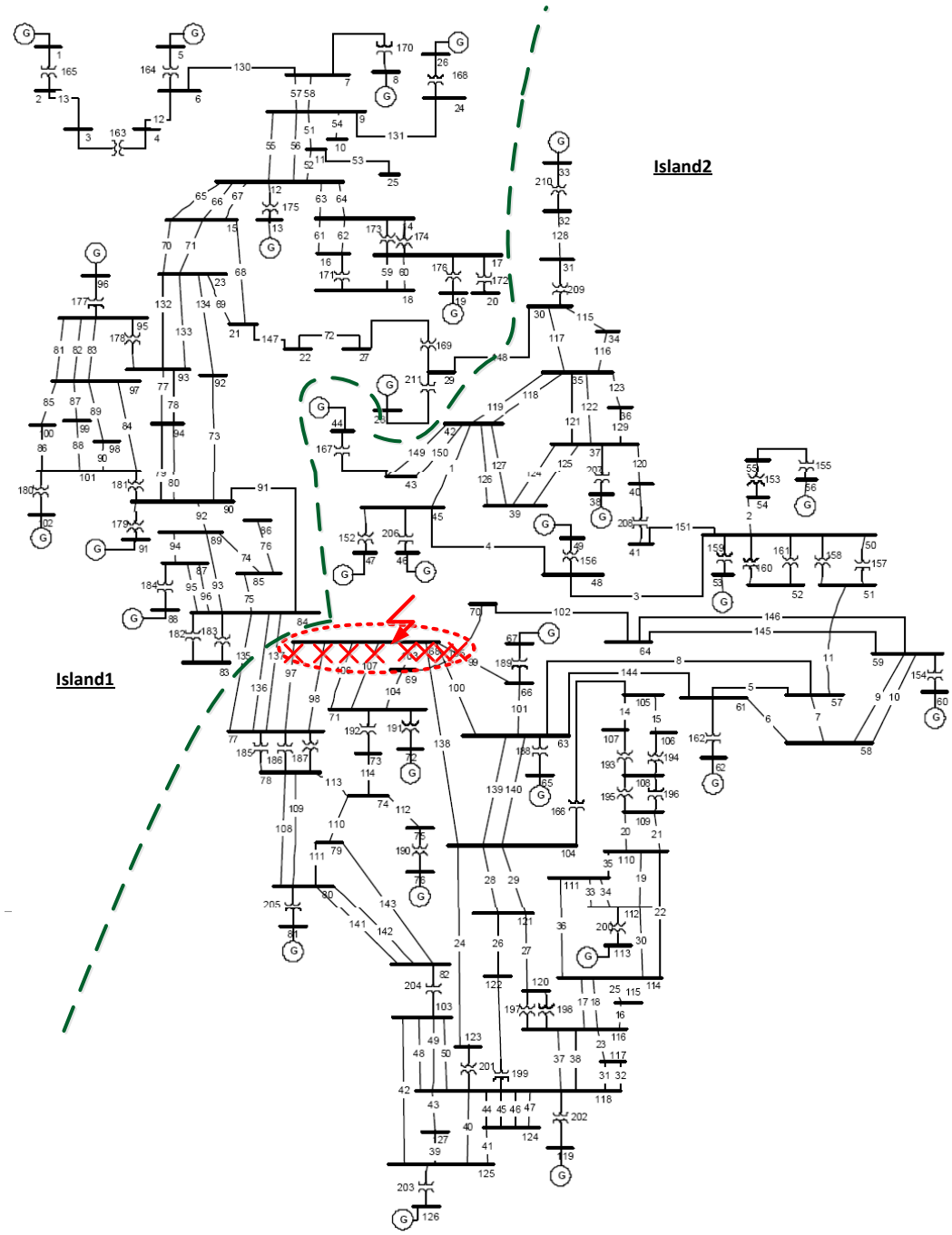


Figure 7. 11: Controlled islanding of WECC 128 buses system

The non-zero elements of the mutual admittance matrix $\mathbf{Y_B}$ are $Y_{29,30}$ and $Y_{77,78}$ therefore, to isolate island 1 and island 2, both line 29-30 and line 77-84 need to be tripped. The islanding boundaries of transient 2 are shown in Figure 7.11.

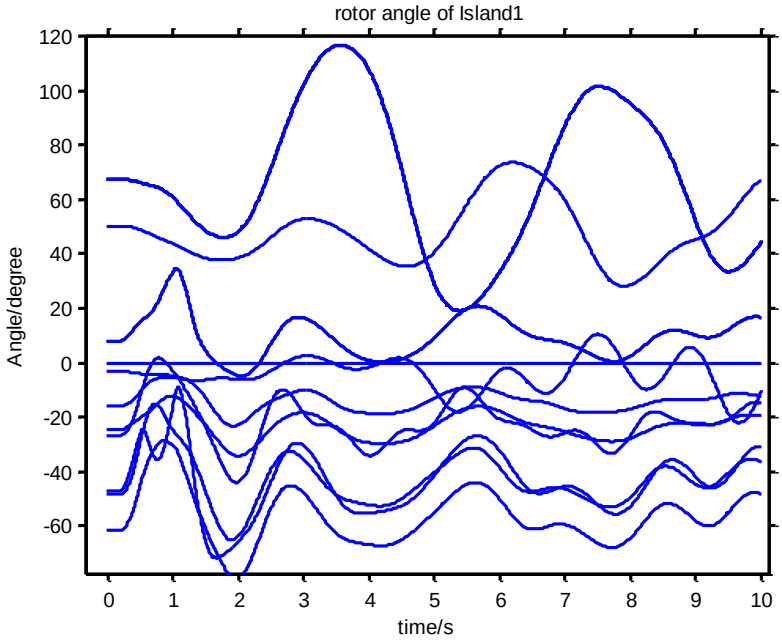


Figure 7. 12: Rotor angles of island 1

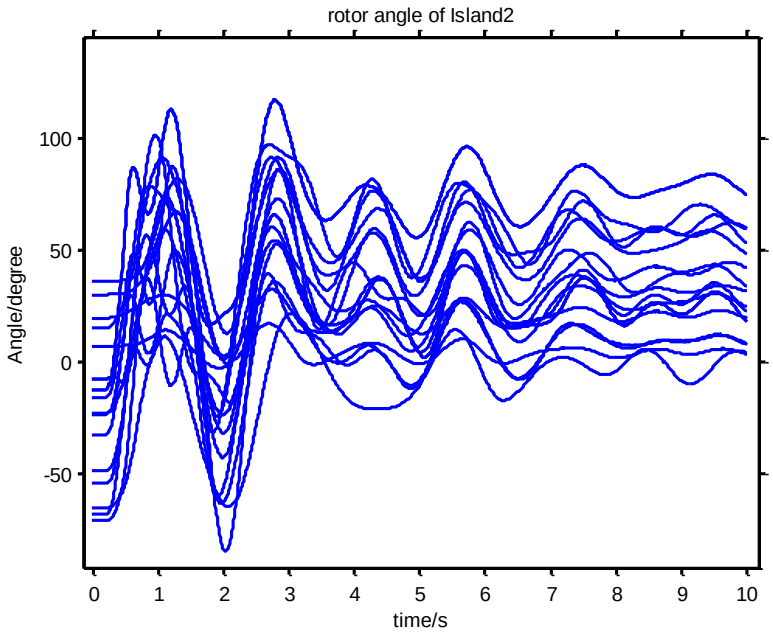


Figure 7. 13: Rotor angles of island 2

It can be seen from Figure 7.12 and Figure7.13 that the oscillations of the rotor angles in island 1 and island 2 are within the stable boundary (less than 180 degree) and they are damping along with time. All the buses' voltages are within the $\pm 5\%$, the frequencies are within 0.5Hz deviation from 60Hz.

To validate the feasibility of this method, the active power of the boundary lines are simulated without the controlled islanding. The active power of line 29-30, line 77-84 are shown as below:

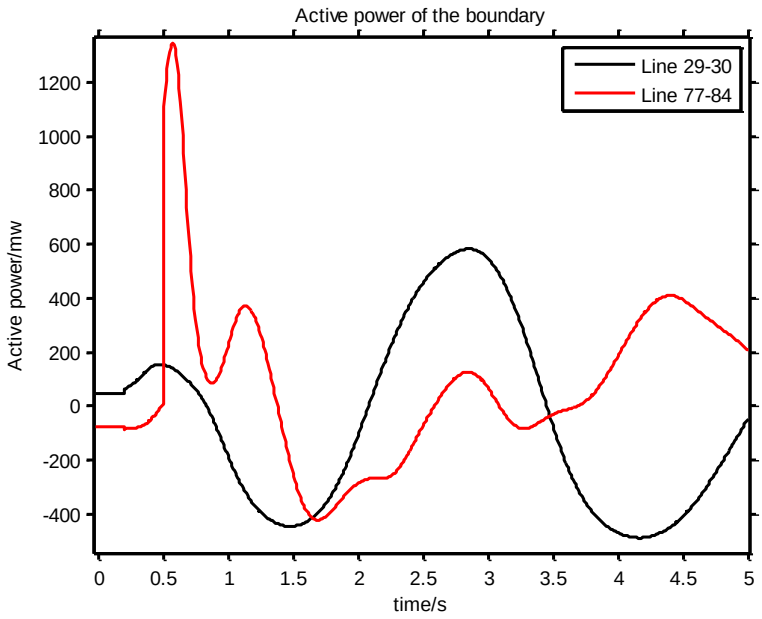


Figure 7. 14: Active power of Line 29-30 and Line 77-84

It can be seen from Figure 7.14, the active power of line 29-30 and line 77-78 fluctuate remarkably during the contingency, which are identical with the current criteria of islanding boundary detection in the power system.

7.3 Conclusion and Discussion

This chapter proposed a novel intelligent controlled islanding method. This method has been tested in the IEEE-39 bus system and WECC 128 bus system, which verified its practicability and validity in the steady state and the transient simulation.

The proposed method, based on algebraic calculation is fast and easy to implemented. There are mainly two main steps in the intelligent controlled islanding. Step1: After the recognition of coherent groups, the load can be chosen within the island based on the RED load index table. Step 2: Detect the islanding boundaries with the mutual admittance matrix. The proposed method saves large initial searching space of the islanding strategies. The RED index table is pre-built prior to the disturbance and applicable during contingency, since the basic configuration of the islands does not change dramatically in a short time frame when some lines are removed during contingency. The loads ranked higher on the list have greater impact on the generator angle performance and should be chosen as the loads to match the generator capacity in the island. Furthermore, the load ranking in this method that based on electrical distance, is insensitive to the random load changes.

Considering multiple coherent groups oscillation, based on the conjecture 3 in reference [61], it is believed that the loss of synchronism does not occur simultaneously according to the experience that the Out of Step relay operates one after another. The proposed islanding scheme allows the stability to be restored within the two islands at the early stage before successive new islands occurrence.

It can be concluded that the proposed method is a very efficient and remarkably high speed method that can be applied to a power system for real time splitting. It not only helps to seek the boundaries of different generator groups in the steady state, but also rapidly detects the controlled islanding boundaries for the operators when the system splitting is unavoidable during the contingencies.

Chapter 8 Conclusion and Future Work

8.1 Conclusion

The main contribution of this dissertation is providing an alternative technology to predict the transient stability. By using the apparent impedance trajectory, the transient stability is predicted in real time with high accuracy. Furthermore, the critical locations for PMUs installations are decided for the transient stability prediction. Based on the information about the state of the power system provided by the phasor measurements, this dissertation uses the apparent impedance trajectory to forecast system stability in real-time.

Three main steps are involved for the transient stability prediction. The first step is to record the apparent impedance trajectory of each branch for a few cycles after the clearance of the fault. Next, a dimension-reduction technique is developed that integrates the high-dimensional complex apparent impedance trajectory recorded by PMUs, to a single dimensional evaluation index for the transient stability prediction. The 24-dimensional apparent impedance trajectories were depicted by a single index value. The distance of the 12-cycle trajectory with respect to the pre-built hyperplane is calculated. The final step involves using data mining techniques to select those critical trajectories that would characterize the system status based on the evaluation index developed in the previous step. By pre-processing thousands of cases and building the decision tree offline, transient stability prediction using the apparent impedance trajectory was found to take about 200 milliseconds, which makes it a very fast prediction technique. The high prediction success of the rotor angle stability is obtained which verifies the validity of the proposed method.

Using the parent nodes of the decision tree, the critical locations for PMUs placements are also determined. Therefore, PMUs would only be required at the critical locations of the network which will then optimize the number of installations. This is followed by an execution of the decision tree to determine the transient stability of the system. With prior knowledge of the distances of the data points from a previously computed hyperplane, a decision tree can be built in CART® to predict the transient stability based on the apparent impedance trajectory in real time. Furthermore, multiple

classes can also be classified by the proposed method. Prediction of the apparent impedance trajectory provides important information for the operator. After the system is predicted to lose synchronism, immediate action must be taken to return the system to stable operation. Pre-determined protection schemes such as load shedding, intelligent islanding, and generator tripping are options for returning the system back to an acceptable security level. Three systems are used to validate the proposed method: the 17 Machine-162 Buses System, the 128 Buses abbreviated WECC System and the 4000 Buses WECC System. Their transient stability was predicted with high accuracy by the decision tree with few nodes, using the 12-cycle apparent impedance trajectories.

Another contribution of this dissertation is developing a real time slow coherency-based controlled islanding technique to reduce the hazards of the large area blackouts, based on the relative electrical distance and the admittance matrix. After the transient instability was predicted by the proposed method and the coherency of generators have been recognized, an intelligent controlled islanding scheme can be applied as the final resort to restore the transient stability of the system. Two steps are required in the intelligent controlled islanding. The first step of the controlled islanding is searching the loads that should remain within the island by the relative electrical distance (RED) method. Then the loads are accumulated according to the RED index from small to large until the imbalance between the generator capacity and the load demand are minimized. The second step of the controlled islanding is by rearranging the admittance matrix, the islanding cut sets can be determined by the non-zero elements in the "mutual admittance matrix." This method is computationally very fast since the RED index can be pre-calculated before contingency and all the calculation are algebraic. Furthermore, the RED index is insensitive to the load variations since it only reflects the relative location of the loads with respect to the generators. The IEEE 10 Machine-39 Buses system is used to test the controlled islanding method in the steady state and the 128 Buses WECC System is used to test the controlled islanding method in the transients. After the islanding, each new formed island is found to remain synchronized and both of the voltage and the frequency stabilized at an acceptable level.

8.2 Future Work

In the future, more considerations will be taken into account with respect to the controlled islanding.

The intelligent controlled islanding can be improved by several aspects:

Develop an improved coherency recognition scheme: The proposed controlled islanding scheme is based on the slow-coherency recognition. Considering the importance of initial coherency detection, a fast and efficient coherency determination algorithm should be developed to assist the controlled islanding scheme.

Optimize the controlled islanding: In this dissertation, the generation and load balance were addressed while determining the cut set. More comprehensive constraints may be included according to different operation requirements, such as the transmission capacity limits within each island, minimizing the economy loss of the power market, etc..

What to do after controlled islanding: In Chapter 7, the "when" and "where" were focused on to implement intelligent controlled islanding. Furthermore, the "What" measures to be taken after the controlled islanding to stabilize the system within a short time should be also considered, such as damping the frequency oscillation and restoring the voltage to an acceptable level rapidly.

Finally, the controlled islanding scheme should be tested in a real power system.

Appendix

Appendix A: RED Index of GEN 31 and GEN 32 in IEEE 39-Bus System

The buses with purple color are load buses.

	GEN31	GEN32
BUS1	0.982797	0.980389
BUS2	0.954702	0.948363
BUS3	0.891137	0.877074
BUS4	0.767608	0.749654
BUS5	0.66219	0.729942
BUS6	0.636197	0.721147
BUS7	0.666766	0.740853
BUS8	0.682141	0.750776
BUS9	0.869689	0.897827
BUS10	0.776917	0.531236
BUS11	0.730154	0.596021
BUS12	0.754517	0.585514
BUS13	0.781804	0.57995
BUS14	0.799389	0.692729
BUS15	0.891313	0.840206
BUS16	0.931306	0.904344
BUS17	0.925992	0.904937
BUS18	0.912664	0.894266
BUS19	0.975043	0.965247
BUS20	0.98674	0.981536
BUS21	0.952634	0.934043
BUS22	0.974824	0.964942
BUS23	0.973969	0.963752
BUS24	0.937425	0.912866
BUS25	0.966181	0.960694
BUS26	0.959302	0.94926
BUS27	0.943946	0.928826
BUS28	0.979437	0.974364
BUS29	0.98593	0.982459

Appendix B: Determination of Controlled Islanding Cut Sets

For display convenience, the nonzero elements in the matrix have been replaced with ones while rearranging the admittance matrix. The sub-square matrix in the lower right corner, colored purpled, indicates island 1, the sub-square-matrix in top left corner indicates island 2. The red colored nonzero elements shown in the matrix is the intersection lines between island 1 and island 2. Those lines (Line 3-4,Line 14-15,and Line 8-9) will be tripped in order to isolate island 1 and island 2.

Island 2 Island 1

BUS	1	2	3	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	33	34	35	36	37	38	39	9	14	4	5	6	7	8	10	11	12	13	31	32
1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
15	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
16	0	0	0	1	1	1	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

0																																						
33	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
36	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
37	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
39	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0
14	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0
4	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	0	0	1	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	1	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	1	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	1	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	1	0	0
31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0
32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1

Appendix C: RED of Load Buses of Coherent Group 1 of 128-Bus WECC System

	GEN1	GEN5	GEN8	GEN13	GEN19	GEN26	GEN28	GEN88	GEN91	GEN96	GEN102
2	0.018122	0.978814	0.998802	0.999841	0.999954	0.999966	0.999992	0.999999	0.999981	0.999995	0.999997
3	0.20833	0.76616	0.986776	0.998245	0.999491	0.999629	0.999906	0.999984	0.99979	0.999943	0.999964
4	0.236666	0.627703	0.978946	0.997206	0.99919	0.99941	0.999851	0.999974	0.999666	0.999909	0.999943
6	0.983316	0.019232	0.944537	0.992639	0.997866	0.998445	0.999608	0.999931	0.99912	0.999759	0.999849
7	0.998404	0.906153	0.136051	0.885345	0.966759	0.975773	0.993888	0.998926	0.986291	0.996253	0.997652
9	0.998784	0.928545	0.342186	0.708478	0.915482	0.9384	0.98446	0.997269	0.965142	0.990472	0.994031
10	0.998761	0.927146	0.329314	0.702774	0.913828	0.937195	0.984156	0.997215	0.96446	0.990285	0.993914
11	0.999076	0.945659	0.499744	0.523084	0.861732	0.953155	0.974578	0.995532	0.942974	0.984412	0.990235
12	0.999673	0.980763	0.822904	0.340191	0.808708	0.983416	0.964829	0.993818	0.921105	0.978434	0.98649
14	0.99971	0.982966	0.843183	0.415746	0.672543	0.985315	0.968857	0.994526	0.93014	0.980904	0.988037
15	0.999783	0.987245	0.882576	0.562514	0.873164	0.989004	0.892432	0.981094	0.758706	0.934042	0.958681
16	0.999712	0.983049	0.843954	0.41862	0.666928	0.985387	0.96901	0.994553	0.930483	0.980998	0.988096
17	0.999751	0.985353	0.86516	0.497626	0.572823	0.987373	0.973221	0.995293	0.93993	0.98358	0.989714
18	0.999743	0.984878	0.860788	0.481339	0.593447	0.986964	0.972353	0.995141	0.937983	0.983048	0.98938
20	0.999748	0.985165	0.863426	0.491164	0.567328	0.987211	0.972877	0.995233	0.939158	0.983369	0.989581
21	0.999825	0.989723	0.905391	0.647516	0.897808	0.991141	0.831553	0.978164	0.721314	0.923821	0.952278
22	0.999936	0.996217	0.965177	0.87026	0.962386	0.996739	0.362558	0.991962	0.897422	0.971961	0.982435
23	0.999856	0.991507	0.921811	0.708691	0.915543	0.992678	0.897324	0.969785	0.614375	0.89459	0.933966
24	0.999908	0.994618	0.950455	0.978043	0.993634	-0.02734	0.99883	0.999794	0.997375	0.999282	0.99955
25	0.999051	0.944218	0.486471	0.51043	0.858064	0.951912	0.973904	0.995413	0.941461	0.983998	0.989976
27	0.999945	0.99675	0.970077	0.888517	0.967679	0.997198	0.270094	0.993093	0.911857	0.975906	0.984906
29	0.999981	0.998854	0.989451	0.960697	0.988605	0.999012	0.163398	0.997564	0.968925	0.991506	0.994679
30	0.999995	0.999692	0.997163	0.989429	0.996935	0.999734	0.775026	0.999322	0.991628	0.997715	0.998567
31	0.999995	0.999734	0.997547	0.990862	0.997351	0.99977	0.805527	0.999414	0.992763	0.998024	0.998761
32	0.999999	0.999963	0.999656	0.998717	0.999628	0.999968	0.972689	0.999918	0.998984	0.999723	0.999826
34	0.999997	0.999809	0.998245	0.993463	0.998105	0.999836	0.860899	0.999566	0.994814	0.998586	0.999112
35	0.999998	0.999865	0.998759	0.995376	0.998659	0.999884	0.90162	0.999682	0.996325	0.998999	0.999371
36	0.999998	0.999895	0.999034	0.996401	0.998957	0.99991	0.923451	0.99974	0.997132	0.999221	0.999509
37	0.999999	0.999965	0.99968	0.998806	0.999654	0.99997	0.974672	0.999877	0.999026	0.99974	0.999834
39	0.999999	0.999945	0.999497	0.998125	0.999456	0.999953	0.960154	0.999847	0.998495	0.999593	0.999743
40	1	0.999982	0.999838	0.999398	0.999825	0.999985	0.988758	0.999013	0.998938	0.999834	0.999832
41	1	0.999986	0.999875	0.999535	0.999865	0.999988	0.991973	0.998833	0.998929	0.999857	0.999833
42	0.999999	0.999928	0.999334	0.997518	0.99928	0.999938	0.94721	0.999821	0.998022	0.999463	0.999662
43	1	0.999977	0.999791	0.999222	0.999774	0.99998	0.983442	0.999944	0.99938	0.999831	0.999894
45	1	0.999997	0.999969	0.999883	0.999966	0.999997	0.997594	0.999944	0.999878	0.999973	0.99998
48	1	0.999998	0.999977	0.999915	0.999975	0.999998	0.999585	0.999151	0.999412	0.99995	0.999911
50	1	0.999996	0.99996	0.999849	0.999956	0.999996	0.999354	0.998444	0.998926	0.999909	0.999838

51	1	0.99999	0.999904	0.999644	0.999897	0.999991	0.99933	0.995801	0.997139	0.999766	0.999569
52	1	0.999996	0.99996	0.999849	0.999956	0.999996	0.999354	0.998444	0.998926	0.999909	0.999838
54	1	0.999997	0.999975	0.999908	0.999973	0.999998	0.999604	0.999046	0.999341	0.999944	0.999901
55	1	0.999999	0.999987	0.999952	0.999986	0.999999	0.999794	0.999505	0.999658	0.999971	0.999948
57	0.999999	0.99996	0.999631	0.998626	0.999602	0.999965	0.999423	0.982603	0.988226	0.999054	0.998229
58	1	0.999972	0.999738	0.999025	0.999717	0.999975	0.999626	0.987631	0.99163	0.999328	0.998741
59	1	0.999974	0.999762	0.999113	0.999743	0.999978	0.999671	0.988736	0.992378	0.999388	0.998854
61	0.999999	0.999968	0.99971	0.998918	0.999686	0.999973	0.999569	0.986281	0.990716	0.999254	0.998604
63	0.999999	0.999915	0.999214	0.99707	0.99915	0.999926	0.998938	0.962793	0.974824	0.997978	0.996214
64	0.999999	0.999919	0.999255	0.997225	0.999195	0.99993	0.999009	0.964753	0.976151	0.998085	0.996414
66	0.999999	0.999933	0.999381	0.997696	0.999332	0.999942	0.999176	0.970732	0.980196	0.99841	0.997022
68	0.999996	0.999781	0.997985	0.992493	0.997824	0.999811	0.997344	0.904637	0.935476	0.994819	0.990297
69	0.999998	0.999856	0.99867	0.995044	0.998563	0.999875	0.998245	0.937051	0.957408	0.99658	0.993595
70	0.999998	0.999887	0.998963	0.996137	0.99888	0.999903	0.998628	0.950938	0.966803	0.997335	0.995008
71	0.999998	0.999867	0.998776	0.99544	0.998678	0.999885	0.998386	0.942078	0.960809	0.996853	0.994107
73	0.999998	0.999868	0.998787	0.995482	0.99869	0.999886	0.998402	0.942614	0.961172	0.996882	0.994161
74	0.999998	0.999882	0.998915	0.995959	0.998828	0.999898	0.998573	0.94866	0.965263	0.997211	0.994776
75	1	0.999974	0.999758	0.9991	0.999739	0.999977	0.999682	0.988564	0.992262	0.999379	0.998836
77	0.999992	0.999516	0.995546	0.983407	0.995189	0.999583	0.994145	0.789214	0.857379	0.988549	0.978554
78	0.999994	0.999624	0.996541	0.987112	0.996264	0.999676	0.995452	0.836279	0.889224	0.991106	0.983342
79	0.999998	0.999872	0.99882	0.995605	0.998726	0.99989	0.998447	0.944174	0.962227	0.996967	0.99432
80	0.999998	0.999864	0.998745	0.995324	0.998644	0.999882	0.998348	0.940599	0.959809	0.996773	0.993956
82	0.999998	0.999865	0.998757	0.995369	0.998657	0.999884	0.998362	0.941176	0.960199	0.996804	0.994015
83	0.999988	0.999272	0.993298	0.975032	0.992761	0.999372	0.991194	0.682816	0.785389	0.982769	0.967728
84	0.999989	0.999342	0.993942	0.97743	0.993457	0.999433	0.99204	0.713286	0.806005	0.984424	0.970828
85	0.999987	0.999258	0.993171	0.974556	0.992623	0.99936	0.991027	0.703187	0.781297	0.98244	0.967113
86	0.999987	0.999258	0.993171	0.974556	0.992623	0.99936	0.991027	0.703187	0.781297	0.98244	0.967113
87	0.999992	0.999549	0.995852	0.984547	0.99552	0.999612	0.994551	0.408355	0.867178	0.989336	0.980027
89	0.999985	0.999145	0.99213	0.97068	0.9915	0.999263	0.989662	0.689603	0.747987	0.979766	0.962104
90	0.99996	0.997634	0.978217	0.918842	0.97647	0.99796	0.971394	0.945342	0.302406	0.943991	0.8951
92	0.99991	0.994696	0.951168	0.818065	0.947253	0.995427	0.935874	0.955778	0.435605	0.919638	0.911323
93	0.999898	0.994031	0.945051	0.795278	0.940647	0.994854	0.927843	0.964073	0.541476	0.848547	0.918664
94	0.999922	0.995394	0.957598	0.842021	0.954199	0.996029	0.944318	0.956713	0.447539	0.884684	0.909276
95	0.999946	0.996804	0.970574	0.890369	0.968216	0.997244	0.961359	0.978313	0.723215	0.442893	0.900838
97	0.999952	0.99716	0.973855	0.902591	0.971759	0.997552	0.965667	0.975497	0.687275	0.553334	0.792779
98	0.99996	0.997658	0.978444	0.919687	0.976716	0.997981	0.971692	0.972084	0.643715	0.702914	0.653399
99	0.999958	0.997526	0.977225	0.915147	0.975399	0.997867	0.970092	0.972845	0.65342	0.664483	0.68713
100	0.999965	0.99793	0.980944	0.929004	0.979417	0.998216	0.974976	0.970483	0.62328	0.782061	0.583338
101	0.999973	0.998421	0.985465	0.945847	0.9843	0.998639	0.980913	0.967973	0.591246	0.92181	0.465216
103	0.999998	0.999863	0.99874	0.995306	0.998639	0.999882	0.998339	0.940375	0.959657	0.996761	0.993933
104	0.999997	0.999823	0.998367	0.993917	0.998237	0.999847	0.997843	0.922735	0.947721	0.995802	0.992139

105	0.999997	0.999839	0.998515	0.994466	0.998395	0.999861	0.998038	0.929701	0.952434	0.996181	0.992847
106	0.999998	0.999868	0.998783	0.995466	0.998685	0.999886	0.998393	0.9424	0.961027	0.996871	0.994139
107	0.999998	0.999868	0.998783	0.995466	0.998685	0.999886	0.998393	0.9424	0.961027	0.996871	0.994139
108	0.999998	0.999874	0.998839	0.995674	0.998746	0.999891	0.998467	0.945043	0.962815	0.997014	0.994408
109	0.999998	0.999879	0.998887	0.995853	0.998798	0.999896	0.998531	0.94732	0.964356	0.997138	0.99464
110	0.999998	0.999883	0.998924	0.995991	0.998838	0.999899	0.99858	0.949079	0.965546	0.997234	0.994819
111	0.999998	0.999884	0.998936	0.996035	0.99885	0.9999	0.998596	0.949633	0.965921	0.997264	0.994875
112	0.999998	0.999911	0.99918	0.996946	0.999115	0.999923	0.998918	0.961208	0.973752	0.997893	0.996053
114	0.999998	0.999878	0.998874	0.995805	0.998784	0.999895	0.998514	0.946712	0.963944	0.997105	0.994578
115	0.999998	0.999873	0.998835	0.995661	0.998742	0.999891	0.998463	0.944885	0.962708	0.997006	0.994392
116	0.999998	0.999868	0.998783	0.995468	0.998686	0.999886	0.998395	0.942425	0.961044	0.996872	0.994142
117	0.999998	0.999866	0.998769	0.995414	0.99867	0.999885	0.998376	0.941739	0.96058	0.996835	0.994072
118	0.999998	0.999865	0.998761	0.995384	0.998662	0.999884	0.998366	0.94136	0.960323	0.996814	0.994034
120	0.999998	0.999853	0.998649	0.994967	0.998541	0.999874	0.998217	0.936071	0.956745	0.996527	0.993496
121	0.999997	0.999824	0.998376	0.993951	0.998246	0.999848	0.997855	0.923158	0.948007	0.995825	0.992182
122	0.999997	0.999848	0.9986	0.994783	0.998488	0.999869	0.998152	0.933733	0.955162	0.9964	0.993258
123	0.999997	0.999843	0.99855	0.994598	0.998434	0.999864	0.998086	0.931383	0.953572	0.996272	0.993018
124	0.999998	0.999867	0.998778	0.995448	0.99868	0.999886	0.998389	0.94218	0.960878	0.996859	0.994117
125	0.999998	0.999904	0.999117	0.996709	0.999046	0.999917	0.998835	0.958196	0.971714	0.997729	0.995747
128	0.999998	0.99987	0.998802	0.995535	0.998706	0.999888	0.99842	0.943284	0.961625	0.996919	0.994229

Reference

[1] C. W. Taylor and D.C. Erickson, "Recording and analyzing the July 2 cascading outage," IEEE Comput. Appl. Power, vol. 10, pp. 26–30, Jan.1997.

[2] X. Vieira et al., "The March 11:th 1999 blackout: Short-term measure to improve system security and overview of the reports prepared by the international experts," presented at the SC 39 Workshop on Large Disturbances, Paris, France, Aug. 29, 2000. CIGRÉ Session.

[3] P. Kundur, J. Paserba, V. Ajjarapu, G. Anderson, A. Bose, C. Canizares, N. Hatziargyriou, D. Hill, A. Stankovic, C. Taylor, T. V. Cutsem, and V. Vittal, "Definition and classification of power system stability," IEEE Trans. on Power Syst., vol. 19, no. 3, pp. 1387–1401, Aug. 2004.

[4] C. W. Taylor and D. C. Erickson, "Recording and analyzing the July 2 cascading outage [Western USA Power System]," IEEE Comput. Appl. Power, vol. 10, no. 1, pp. 26–30, 1997.

[5] H. B. Ross, N. Zhu, J. Giri, and B. Kindel, "An AGC implementation for system islanding and restoration conditions," IEEE Trans. Power Syst., vol. 9, pp. 1399–1410, Aug. 1994.

[6] J. M. Haner, T. D. Laughlin, and C. W. Taylor, "Experience with the R-Rdot out of step relay," IEEE Trans. Power Syst., vol. PWRD-1, pp. 35–39, Apr. 1986.

[7] H. You, V. Vittal, and X. Wang, "Slow coherency-based islanding," IEEE Trans. Power Syst., vol.19, no. 1, pp. 483–491, Feb. 2004.

[8] A. G. Phadke and J. S. Thorp, Synchronized Phasor Measurements and Their Applications. New York: Springer, 2008.

[9] D. Novosel and K. Vu, "Benefits of PMU technology for various applications," International Council on Large Electric Systems Symposium on Power System Management, 7 annual, Croatia, Nov.2006.

[10] Hiskins, I; "Nonlinear Dynamic Model Evaluation From Disturbance Measurements" IEEE Transactions on Power Systems, Vol.16, No.4, November 2001 Page(s) 702-710

[11] P. Kundur, Power System Stability and Control, New York: McGraw-Hill, 1994.

[12] P. M. Anderson, Analysis of faulted power systems. New York: IEEE Press, 1995.

[13] A. Llamas, and J. De La Ree, "Stability and the transient energy method for the classroom," Proc., 25th Southeastern Symposium on System Theory, Tuscaloosa, AL, pp. 79-85, Mar. 1993.

[14] A. Llamas, J. De La Ree, L. Mili, A. G. Phadke, and J. S. Thorp, "Clarifications of the BCU method for transient stability analysis," IEEE Trans. on Power Systems, Vol. 10, No. 1, pp. 210–219, Feb. 1995.

[15] R. C. Bansal, "Bibliography on the fuzzy set theory applications in power systems (1994–2001)," IEEE Trans. on Power Syst., Vol. 18, No. 4, pp. 1291–1299, Nov. 2003.

[16] C. W. Liu, M. C. Su, S. S. Tsay, and Y. J. Wang, "Application of a novel fuzzy neural network to real-time transient stability swings prediction based on synchronized phasor measurements," IEEE Trans. on Power Syst., Vol. 14, No. 2, pp. 685–692, May 1999.

[17] Y. Mansour, E. Vaahedi, and M. A. El-Sharkawi, "Dynamic security contingency screening and ranking using neural networks," IEEE Trans. on Neural Networks, Vol. 8, No. 1, pp. 942–950, Jul. 1997.

[18] C. A. Jensen, M. A. El-Sharkawi, and R. J. Marks II, "Power system security assessment using neural networks: Feature selection using Fisher discrimination," IEEE Trans. on Power Systems, Vol. 16, No. 4, pp. 757-763, Nov. 2001.

[19] L. S. Moulin, A. P. A. da Silva, M. A. El-Sharkawi, and R. J. Marks II, "Support vector machines for transient stability analysis of large scale power systems," IEEE Trans. on Power Syst., Vol. 19, No. 2, pp. 818–825, May 2004.

[20] L. S. Moulin, A. P. A. da Silva, M. A. El-Sharkawi, and R. J. Marks II, "Support vector and multilayer perceptron neural networks applied to power systems transient stability analysis with input dimensionality reduction," Proc. IEEE Power Eng. Soc. Summer Meeting, Chicago, IL, Vol. 3, pp. 1308-1313, Jul. 2002.

[21] H. Mori, "State-of-the-art overview on data mining in power systems," in Proc. IEEE Power Eng. Soc. General Meeting, Montreal, Jun., 2006.

[22] L. Wehenkel, M. Pavella, E. Euxibie, and B. Heilbronn, "Decision tree based transient stability method a case study," IEEE Trans. on Power Syst., Vol. 9, No. 1, pp. 459-469, Feb. 1994.

[23] L. Wehenkel, and M. Pavella, "Decision tree approach to power systems security assessment," International Journal of Electrical Power and Energy Systems, Vol. 15, No. 1, pp. 13-36, Dec. 1993.

[24] S. Rovnyak, S. Kretsinger, J. Thorp, and D. Brown, "Decision trees for real-time transient stability prediction," IEEE Trans. on Power Systems, Vol. 9, No. 3, pp. 1417-1426, Aug.1994.

[25] F. G. Velez-Cedeno, “Multiple Swing Out-of-Step Relaying,” Ph. D. Dissertation, Bradley Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA, 2010.

[26] K. K. Anaparthi, B. Chaudhuri, N. F. Thornhill, and B. C. Pal, “Coherency identification in power systems through principle component analysis,” IEEE Trans. Power Syst., vol. 20, no. 3, pp. 1658–1660, Aug. 2005.

[27] J. H. Chow, R. Galarza, P. Accari, and W. Price, “Inertial and slow coherency aggregation algorithms for power system dynamic model reduction,” IEEE Trans. Power Syst., vol. 10, pp. 680–685, May 1995.

[28] B. Stott, “Power System Dynamic Response Calculations”, Proceedings of the IEEE, vol. 67, no. 2, pp. 219-241, February 1979.

[29] C. K. Tang. C. E. Graham, et al., “Transient stability index from conventional time domain simulation,” IEEE Trans. on Power Syst., Vol. 9, No.3, pp. 1524-1532, Aug.1994.

[30] P. Kundur, N. J. Balu and M. G. Lauby, Power system stability and control. New York: McGraw-Hill, 1994.

[31] Y. Xue, T. Van Custem and M. Ribbens-Pavella, "Extended equal area criterion justifications, generalizations, applications," IEEE Trans. on Power Syst., vol. 4, pp. 44-52, 1989.

[32] Fang Da-zhong, T.S.Chung, and A.K.David, ”Fast transient stability estimation using a novel dynamic equivalent reduction technique” IEEE Trans. on Power Syst., Vol. 9, No. 2, pp. 995-1001, May.1994.

[33] Minghui Yin, C. Y. Chung, K. P. Wong, Yusheng Xue, and Yun Zou. “An improved iterative method for assessment of multi-Swing transient stability limit”, IEEE Trans. on Power Systems, Vol. 26, No. 4, pp. 2023-2029, Nov.2011.

[34] M. Pavella, D. Ernst, and D. Ruiz-Vega, Transient stability of power systems: A unified approach to assessment and control, Kluwer Academic Publisher, Boston, U.S.A, 2000.

[35] M. A. Pai, Energy function analysis for power system stability. Boston: Kluwer Academic Publishers, 1989.

[36] Athay, etal. "Transient Energy Analysis"， System Engineering for Power, Emergency Operating State Control, Section IV , US Department of Energy, Publication No. CONF-790904 P I .

[37] Kakimoto, etal. "Transient Stability Analysis of Electric Power Systems via Lure Type lyapunov Function, Parts I and II," Trans IEE of Japan, Vol 98. No. 5/6, May/June 1978.

[38] M.A. El-Kady, C.K. lang, etal. "Dynamic Security Assessment Utilizing the Transient Energy Function Method," Proc of the 1985 PlCA Conference, San Francisco, Californis, May 1985,pp. 132-139.

[39] G. A. Maria, C. Tang, and J. Kim, "Hybrid transient stability analysis," IEEE Trans. Power Syst., vol. 5, no. 2, pp. 384–393, May 1990.

[40] D. Ernst, D. Ruiz-Vega, M. Pavella, P. M. Hirsch, and D. Sobajic, "A unified approach to transient stability contingency filtering, ranking and assessment," IEEE Transactions on Power Systems, vol. 16, no. 3, pp. 435–443, 2001.

[41] G. Li and S. M. Rovnyak, "Integral square generator angle index for stability ranking and control," IEEE Transactions on Power Systems, vol. 20, no. 2, pp. 926–934, 2005.

[42] Y. Mansour, E. Vaahedi, A. Y. Chang, B. R. Corns, B. W. Garrett, K. Demaree, T. Athay, and K. Cheung, "B. C. Hydro's on-line transient stability Assessment (TSA) model development, analysis, and post-processing," IEEE Trans. on Power Syst., vol. 10, no. 1, pp. 241–253, Feb. 1995.

[43] D. Z. Fang, T. S. Chung, Y. Zhang, and W. N. Song, "Transient stability limit conditions analysis using a corrected transient energy function approach," IEEE Trans. on Power Syst., vol. 15, no. 2, pp. 804–810, May2000.

[44] X. Wu, J. Zhao, A. Xu, H. Deng, and P. Xu, "Review on transient stability prediction methods based on real time wide-area phasor measurements," in Proc. IEEE- DRPT 4th Int. Conf., Shandong, China, Jul. 2011, pp. 320–326.

[45] D. J. Sobajic and Y. H. Pao, "Artificial neural-net based dynamic security assessment for electric power systems," IEEE Trans. Power Systems, vol. 4, no. 1, pp. 220–228, Feb. 1989.

[46] Y. Mansour, E. Vaahedi, and M. A. El-Sharkawi, "Dynamic security contingency screening and ranking using neural networks," IEEE Trans. Neural Networks, vol. 8, no. 1, pp. 942–950, July 1997.

[47] C. Jensen, M. El-Sharkawi and R. Marks, "Power System Security Assessment Using Neural Networks: Feature Selection Using Fisher Discrimination," Power Engineering Review, IEEE, vol. 21, pp. 62-62, Nov 2001.

[48] E. H. Mamdani and S. Assilian, "Applications of fuzzy algorithms for control of simple dynamic plant", Proc. Inst. Elec. Eng., vol. 121, pp.1585-1588, 1974.

[49] S. G. Cao, N. W. Rees, and G. Feng, “Analysis and design for a class of complex control systems, part I: Fuzzy modeling and identification,” Automatica, vol. 33, no. 6, pp. 1017–1028, 1997

[50] D. Driankov, H. Hellendoorn and M. Reinfrank, “An introduction to fuzzy control”, 2nd, rev. ed. Berlin ; New York: Springer, 1996.

[51] Francisco Ramon Gomez Lezama, “Prediction and Control of Transient Instability Using Wide Area Phasor Measurements Ph. D. Dissertation, Electrical and Computer Engineering Department, University of Manitoba, 2011

[52] M. H. Haque, A. H. M. A. Rahim. “Determination of first swing stability limit of multimachine power systems through Taylor series expansions.” IEE Proc. Gener. Transm. Distrib. vol. 136, no.6, pp. 373~379, 1989.

[53] J. Sun and K. L. Lo. “Transient stability real-time prediction for multimachine power systems by using observation.” Proceedings of IEEE TECON 93. Beijing, 1993, pp. 217-221.

[54] Hoff, J.C., A practical guide to Box-Jenkins forecasting. 1983, Belmont, Calif.: Lifetime Learning Publications. xii, 316.

[55] C. W. Liu, J. Thorp. “Application of synchronized phasor measurements to real-time transient stability prediction.” IEE Proc. Gener. Transm. Distrib. 1995, vol. 142, no.4, pp.355~360.

[56] F. Song, T. Bi and Q. Yang. “Perturbed trajectory prediction method based on wide area measurement systems.” Automation of Electric Power Systems. vol. 30, no.23, pp.27-31, 2006.

[57] Bernabeu, E.E., Methodology for a security-dependability adaptive protection scheme based on data mining. 2009, Blacksburg, Va.: University Libraries Virginia Polytechnic Institute and State University.

[58] C. Cortes, and V. Vapnik, “Support vector network,” Machine learning, Vol. 20, pp. 273-297, 1995.

[59] A. E. Gavoyiannis, D. G. Vogiatzis, D. P. Georgiadis, and N. D. Hatziargyriou, “Combined support vector classifiers using fuzzy clustering for dynamic security assessment,” in Proc. Power Eng. Soc. Summer Meeting, vol. 2, 2001, pp. 1281–1286. 2001.

[60] S. Shahnawaz Ahmed, Narayan Chandra Sarker, Azhar B Khairuddin, Mohd Ruddin B Abd Ghani, and Hussein Ahmad. “A Scheme for Controlled Islanding to Prevent Subsequent Blackout” IEEE Trans. Power Sys., vol. 18, No. 1, pp. 136-143, Feb. 2003

[61] M. M. Adibi, R. J. Kafka, Sandeep Maram, et al. “On power system controlled separation”, IEEE Trans on Power System, vol. 21, pp. 1894- 1902, Nov. 2006.

[62] N. Senroy, G. T. Heydt, and V. Vittal, "Closure to Discussion of 'Decision Tree Assisted Controlled Islanding'," IEEE Trans. Power Sys., vol. 22, No. 4, pp. 2293-2293, Nov. 2007.

[63] X.Wang and V.Vittal, "System islanding using minimal cutsets with minimum net flow," Proceedings of the 2004 IEEE PES Power System Conference and Exposition, New York, October 2004.

[64] H.You, V.Vittal, and Z.Yang, "Self-healing in power systems: an approach using islanding and rate of frequency decline-based load shedding," IEEE Trans. on Power Systems, vol.18,No.1,pp174-181,Feb.2003.

[65] Guangyue Xu, and Vijay Vittal, "Slow Coherency Based Cutset Determination Algorithm for Large Power Systems" IEEE Trans. on Power Systems, vol.25,No.2,pp877-884,Feb.2010.

[66] Qianchuan Zhao, Kai Sun, Da-Zhong Zheng, Jin Ma, and Qiang Lu." A Study of System Splitting Strategies for Island Operation of Power System: A Two-Phase Method Based on OBDDs" IEEE Trans. Power Sys., vol. 18, No. 4, pp. 1556-1565, Feb. 2003

[67] C. W. Taylor, J. M. Haner, L. A. Hill, W. A. Mittelstadt, and R. L. Cresap, "A new out-of-step relay with rate of change of apparent resistance augmentation," IEEE Trans. Power App. Syst., vol. PAS-102, no. 3, pp. 631–639, Mar. 1983.

[68] Arun G. Phadke, and James S. Thorp, Computer Relaying for Power Systems, John. Wiley & Sons Inc., New York 2nd edition, 2009, pp. 294-296.

[69] V. Centeno, J. De La Ree, A. G. Phadke, G. Michel, R. J. Murphy, and R. O. Burnett Jr., "Adaptive out-of-step relaying using phasor measurement techniques," Computer Applications in Power, IEEE, Vol. 6, No. 4, pp. 12-17, 1993.

[70] M. C. Bozchalui, and M. Sanaye-Pasand, "Out-of-step relaying using phasor measurement unit and equal area criterion," IEEE Power India Conference, April 10-12, 2006, 6 p.

[71] Xue,Y., Wehenkel, L., Belhomme, R. et al. Extended equal area criterion revisited [EHV power systems], IEEE Trans.on Power Systems,vol.7, issue 3, pp.1012-1022.

[72] V. Centeno, A. G. Phadke, and A. Edris, "Adaptive Out-Of-Step relay with phasor measurement," IEEE. Developments in Power System Protection, Sixth International Conference on (Conf. Publ. No. 434) 25-27 March 1997, pp. 210-213, 1997.

[73] CART® Tree-Structured Non-parametric Data Analysis, Classification and Regression Trees by Salford Systems, www.salford-systems.com

[74] X. Lu , Z. Y. Dong and X. Li "Electricity market price spike forecast with data mining techniques", Int. J. Elect. Power Syst. Res., vol. 73, pp.19 -29 200

[75] S.-J. Huang and J.-M. Lin "Enhancement of anomalous data mining in power system predicting-aided state estimation", IEEE Trans. Power Syst., vol. 19, no. 1, pp.610 -619 2004

[76] LI Meiyan, MA Jin, An Analytical Method to Study the Impact of Load Model Uncertainty on the Power System Dynamic Simulations[J]. Automation of Electric Power Systems. 2010, 34(7): 16-20.

[77] LI Meiyan, MA Jin. Uncertainty Analysis of Load Models in Small Signal Stability[C]. International Conference on Sustainable Power Generation and Supply, 2009. SUPERGEN'09.

[78] Ralf Mikut, Markus Reischl: Data mining tools. Wiley Interdisc. Rew.: Data Mining and Knowledge Discovery 1(5): 431-443 (2011)

[79] S. Madan, W. K. Son, and K. E. Bollinger, "Applications of data mining for power systems", Proc. IEEE Can. Conf. Elect. Comput. Eng., pp.403 -406 1997

[80] Jefferson Morais, Yomara Pires, Claudomir Cardoso and Aldebaro Klautau. "An Overview of Data Mining Techniques Applied to Power Systems". Signal Processing Laboratory (LaPS) Federal University of Pará (UFPA), Belém PA Brazil.

[81] Emanuel E. Bernabeu, "Methodology for a Security-Dependability Adaptive Protection Scheme based on Data Mining", PhD Dissertation, Virginia Tech, 2009

[82] John F. Elder IV & Dean W. Abbott, "A Comparison of Leading Data Mining Tools", Fourth International Conference on Knowledge Discovery & Data Mining, 1998, New York

[83] Salford Systems, CART. [Online]. Available: http://www.salford-systems.com/cart.php.

[84] L. Breiman, J. H. Friedman, R. Olshen, and C. J. Stone, Classification and Regression Tree, Wadsworth & Brooks/Cole Advanced Books & Software, Pacific California.

[85] M. Pavella, D. Ernst, and D. Ruiz-Vega, Transient stability of power systems: A unified approach to assessment and control, Kluwer Academic Publisher, Boston, U.S.A, 2000.

[86] Santosh Veda, "WAMS-based Intelligent Load Shedding Scheme for Preventing Cascading Blackouts " PhD Dissertation, Virginia Tech, 2012.

[87] S. Balakrishnama and A. Ganapathiraju, "Linear Discriminant Analysis—A Brief Tutorial," Mississippi State Univ., StarKVille, MS, 1998.

[88] Jieping Ye, Ravi Janardan,Qi Li, Two-Dimensional Linear Discriminant Analysis, Advances in neural information processing system, 2005, vol. 17, pp. 1569-1576

[89] P.N. Belhumeour, J.P. Hespanha, and D.J. Kriegman. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7):711–720, 1997.

[90] R.Gopinath," Maximum likelihood modeling with Gaussian distributions for classification," in Proceeding of ICASSP, pp. II-161-II-664,Seattle,1998.

[91] R. Fisher, "The use of multiple measurements in taxonomic problems," Annals of Eugenics, 7, pp. 179-188, 1936.

[92] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, pp. 711–720, 1997.

[93] S. Dudoit, J. Fridlyand, and T. P. Speed, "Comparison of discrimination methods for the classification of tumors using gene expression data," Journal of the American Statistical Association, Vol. 97, No. 457, pp. 77–87, 2002.

[94] A. Pal, J. S. Thorp, T. Khan, and S. S. Young, "Classification Trees for Complex Synchrophasor Data", submitted to Electric Power Components and Systems.

[95] Elizondo, D., "Hidden Failures in Protection Systems and its Impact on Power SystemWide Area Disturbances", PhD Dissertation, Virginia Tech, 2000

[96] K. Sun, S. Likhate, V. Vittal, V. Sharma Kolluri, and S. Mandal, "An Online Dynamic Security Assessment Scheme Using Phasor Measurements and Decision Trees", IEEE Trans. on Power Systems, Vol. 22, No. 4, pp. 1935-1943, Nov. 2007.

[97] Vishaka K., Thukaram D., Jenkins L.: "Transmission charges of power contracts based on relative electrical distances in open access", Electr. Power Syst. Res., 2004, 70, pp. 153–161