DATA-DRIVEN DIAGNOSTICS OF ISSUES RELATED TO POWER SYSTEM DYNAMICS USING PMU MEASUREMENT

by

Mahmoodreza Arefi

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Electrical Engineering

Charlotte

2018

Approved by:

_____
Dr. Badrul Chowdhury

_____
Dr. Yogendra Kakad

_____
Dr. Valentina Cecchi

_____
Dr. Tao Hong

ABSTRACT

MAHMOODREZA AREFI. Data-Driven diagnostics of issues related to power system dynamics using PMU measurement (Under the direction of Dr. BADRUL CHOWDHURY)

This work investigates the use of data driven techniques to diagnose issues related to power system dynamics. The source behind the data are simulated Phasor Measurement Unit (PMU) measurements. First, this study examines the application of Weighted Support Vector Machine (WSVM), ensemble WSVM and Adaptive Neuro Fuzzy Inference System (ANFIS) for prediction of post-fault transient stability condition. The performance of the ensemble classifier is compared with other methods for accuracy of prediction. The method is tested on the IEEE 39-bus test system. Second two methods are introduced for predicting the voltage and rotor angle stability status of a power system instantly after a large disturbance. Generator voltages and angles gathered instantly after clearing a fault are used as inputs of Feedforward Neural Networks to classify stability status of a system. Grey Wolf Optimizer (GWO) and Particle Swarm Optimization (PSO) are applied as training methods for Feedforward Neural Networks. The performance of two methods are determined by applying the two methods on the IEEE 39 bus test system. Results show that the ensemble method achieves the highest accuracy using rotor angles of generators and voltage magnitudes after fault clearance.

The resulting accuracy of GWO and PSO are compared. Examination showed that the applied methods can predict the stability status 30 cycles after fault clearance.

This study also presents a model-free method for detecting coherency of generators in power systems by means of the Wavelet Packet Decomposition and the Recurrence

Quantification Analysis (RQA). The time-series rotor angles of generators are used in this approach to detect coherent group of generators. Noise is a critical issue that needs to be considered when evaluating the performance of the coherency detection methods based on monitored data. This work also focuses on a method to detect low-frequency oscillations and identify frequency modes by applying measured data from PMUs. An algorithm based on reduced dimensionality of the data for detection of oscillations is proposed. The Slow Feature Analysis is applied to extract low-dimensional features from the PMU data. Based on the RQA, two thresholds are applied to identify low-frequency oscillations by using the slow features. The recurrence-derived Fourier Transform is applied to determine frequency modes. A 29-machine 179-bus system is considered for the study. Studies on the system shows the effectiveness of the proposed methods.

Finally, a method is proposed for disturbance event detection and classification by applying the RQA. To evaluate the characteristics of location and type of disturbance events, nonlinear measures of the power system, such as voltage and frequency, are examined based on recurrence plots. For dimensionality reduction, the Principle Components Analysis is used. An unsupervised clustering method is applied to determine two types of disturbance events, which are fault and generator tripping. Simulations conducted on the 29-machine 179-bus system model and results reveal that the RQA might be an effective tool to identify the location and type of disturbance events.

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor Prof. Badrul Chowdhury. I appreciate all his contributions of time, ideas to make my Ph.D. experience fruitful and motivating. His energy, inspiration, support, and advice were a valuable assist to accomplish this work.

I would also to thank Dr. Yogendra Kakad, Dr. Valentina Cecchi and Dr. Tao Hong for being part of my graduate committee and for letting me learn from their experience.

I would like to thank my family for all their love and encouragement. For my parents who supported me in all my pursuits. And most of all for my loving, kind, encouraging, and patient wife Ronak Ghandriz whose true support during this journey is so appreciated.

Finally, I thank the EPIC, ECE department, and graduate school for providing me the financial assistance to accomplish this dissertation.

DEDICATION

To my beloved Family

My Parents: Firouz Arefi, Mahvash Javidi

My sibling: Golnar

And my beloved wife: Ronak Ghandriz

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1 : OVERVIEW

1.1 Introduction

Future smart grids may run closer to their stability limits due to market conditions provoked by newer generation types and loads. To empower safer operation, corrective control and stabilization are likely feasible options. Ensuring the stability of a power system is critical in averting cascading failures. During major disturbances, rapid detection of unstable system behavior is essential in providing enough time for taking necessary automatic corrective control actions [1].

Research in this field shows the utilization of machine learning methods like decision trees, fuzzy logic, artificial neural networks (ANN), and support vector machines (SVM) [2]- [3] are useful methods for solving some complicated power system problems.

Meta-heuristic optimization methods have become prominent during the last two decades. Some of them such as Genetic Algorithm (GA) [4], Ant Colony Optimization (ACO) [5], and Particle Swarm Optimization (PSO) [6] have been used in computer science as well as in other fields. The reason for the popularity of these methods are: simplicity, flexibility, derivation free system, and local optima prevention.

Electricity demand is increasing steadily, albeit at a slow pace, while transmission capacity does not always keep pace [7]. Exacerbating the situation is the rapid growth [8] in utility-scale renewable energy capacity, which can create operational uncertainties because of its variability. These factors have forced the power grid to occasionally run

closer to its maximum stability limit, which makes it more susceptible to any possible disturbances. In the worst-case scenario, severe blackouts can result from small-signal instability associated with low frequency oscillations [9], [10]. After the Northeast blackout of 1965, several studies were done in order to prevent blackouts in North America. Nevertheless, cascading events causing widespread blackouts have occurred [11] in North America and other countries around the world. Consequently, there is a need for improved methods to quickly detect the onset of instability such that preventive or remedial actions can be taken to avoid blackouts. The detection of unstable behavior begins with the identification of one or more generators that tend to lose coherency or synchronism.

Wide-area measurement systems have the important duty to monitor the dynamic characteristics of a power system. There are several measurement units that can furnish important parameters about a system such as generator voltages, frequencies and rotor angles [12]. Applying this measurement is a dilemma to examine instinctive changes in dynamic characterization of a system. Determining the coherency of generators after fault clearance, on the other hand, could be indicative of the state of stability of multi-machine power system.

Generator coherency can be determined by examining and comparing rotor angles through synchronized measurement. This can eventually help operators to take corrective actions if the system exhibits symptoms of instability.

Model based approaches are used in [13], [14] and are grounded on the eigenvalue study of a power system. The accuracy of these methods is good when predominant system condition and different parameters are accurate. Nevertheless, changes in topology and

serious load alterations can force certain generators with low synchronizing torque to switch from one group to another [15].

One of the main concerns in stable power system operations is low frequency oscillations. Some power quality problems could be a result of oscillations and might damage equipment in the power system. In some serious cases, increasing oscillations may cause serious blackouts [16]. A 0.25 Hz oscillation lead to the 1996 Western Electricity Coordination Council (WECC) [17] blackout. For detecting low-frequency oscillations traditional methods need detailed dynamical demonstration of the system to perform modal analysis [18]. Nevertheless, the penetration of renewable resources causes difficulties for precise dynamic modeling of the system.

Synchrophasors have demonstrated the capability for successful wide-area monitoring and protection [19]. Analyzing low-frequency oscillations by online data is possible since PMUs can provide 30 Hz or higher sampling rates [20]. Estimation of eigenvalues for monitoring inter-area oscillation was done by applying Fourier spectral analysis to synchrophasor data in [21]. PMUs have been used in power systems to deliver system states and dynamics in real time [22]. A monitoring system using PMU measurements that has different capabilities such as detecting phase angle stability, voltage stability and power system oscillations was introduced by ABB [23]. By considering the growing number of PMU data, there is a need to find a good solution to manage and use the huge amount of data provided by synchrophasors in real-time.

Data for real time wide-area monitoring, control, and protection applications are mostly provided by Phasor Measurement Units (PMU) [24]. Parameters of the power system are measured by each PMU and Global Positioning System provides a time reference. The

time stamped data from different PMUs are stored in Phasor Data Concentrators (PDC) which includes disturbance data that can be applied for secure operation of power systems by improved monitoring, and customized remedial actions.

There are many discussions in the pertinent literature about applying PMUs to improve wide-are monitoring, protection and control [25], [26] , [27]. The Lyapunov exponents of the voltage phasors were used for monitoring short-term voltage stability [28]. The discrete Fourier transform was used in [29], [30] for transmission line fault detection and location. The system topology with phasor angle measurements were applied in [31] to identify line outages.

Specific criteria for dynamic disturbance recording which includes placement of recorders, recording durations and electrical quantities to be recorded are provided by the North American Electric Reliability Corporation (NERC) standards [32]. A PMU measures voltage phasor, frequency and the rate of change of frequency. When the records change from a pre-specified value, the disturbance recording feature activates [33]. There are different reasons for the change, such as changes in generations, or loads, faults, capacitor switchings, etc. A disturbance may propagate from one utility to another in a large system like the Western Electricity Coordinating Council (WECC). Finding the exact location of the event can be a stiff challenge for large multi-area systems.

In this dissertation an adaptive neuro fuzzy weighted combined SVM is used to establish a technique that can predict the transient stability of a system. Also, two metaheuristic methods: Grey Wolf Optimization and Particle Swarm Optimization are applied to train weights of FNN for prediction of stability status of power systems after a large disturbance.

This work mainly focuses on: 1) finding coherent groups of generators from data collected by PMU throughout a given network; 2) detecting low-frequency oscillations, as well as identifying frequency modes by applying the RQA method; 3) finding the type and location of disturbance events by implementing the dimensionality reduction of the RQA measures and applying a clustering method.

The main contributions of this dissertation are applying the RQA method for diagnosing anomalies in system dynamics, including:

(a) post-fault coherency detection of generators

(b) post-fault detection of sustained oscillations, and

(c) identification of the class of disturbance event.

The definition of blocks (light blue rectangles in Fig. 1.1) is explained in related chapters.

For coherency detection, Fig. 1(a) shows the nonstationary time series obtained by PMU is decomposed into different scales by applying wavelet packet decomposition. Measures of the RQA method are derived, and since these features may lead to an extremely high dimensionality, a feature selection method is applied for feature reduction. Finally, a clustering technique is used to detect the coherent groups of generators.

Fig. 1(b) shows the overview of the proposed method for detection of sustained oscillation, where a Slow Feature Analysis (SFA) is carried out on the frequency of generators to extract the slowly varying features. One of the RQA measures is applied on the selected features to detect low-frequency oscillations. Also, the recurrence-derived Fourier Transform (FT) is applied to determine the frequency modes.

The method for identifying the class of disturbance events is shown in Fig. 1(c). For determining the fault location, two measures of the RQA are used on each bus. To identify the type of disturbance events, thirteen measures of the RQA are applied on the PMU data, and then the Principle Component Analysis (PCA) method is used on the RQA measures for feature reduction. Finally, a clustering method is considered to detect the event type.

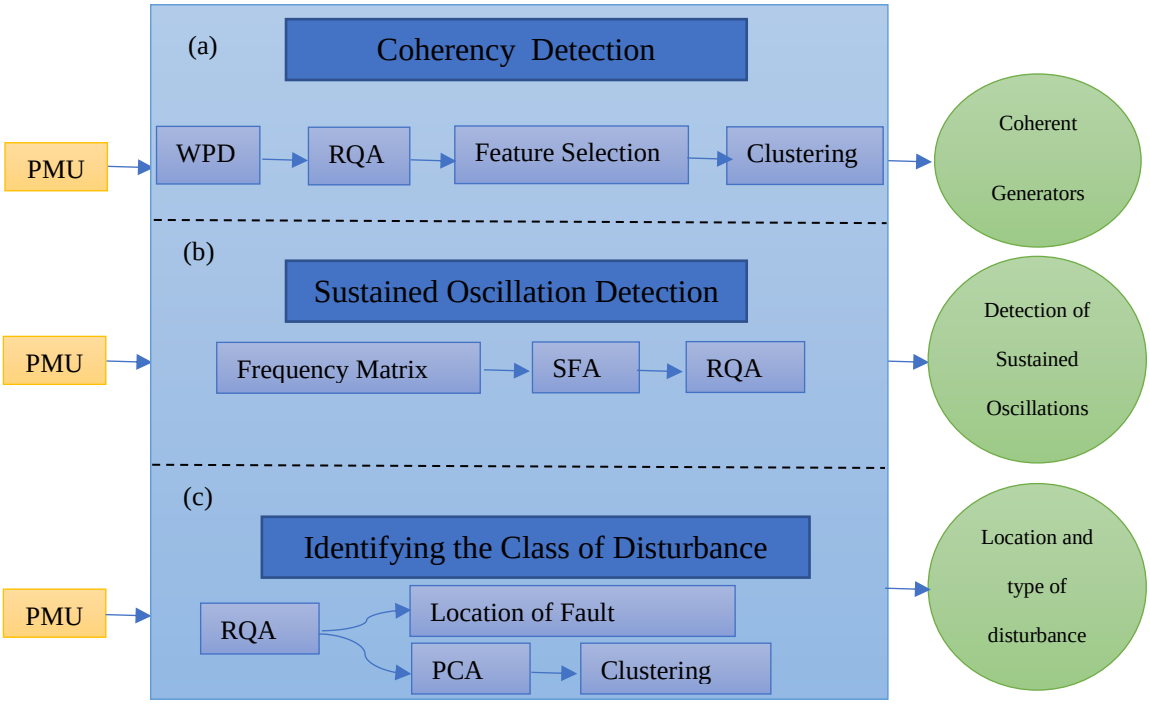

Figure 1-1: Flowchart of proposed algorithms based on the RQA

1.2 Contribution

A summary of the contributions are as follow:

- In terms of formulation, two methods are proposed for prediction of transient stability. The first method considers the SVM and ANFIS, and the second method studies training of a feedforward neural network by the PSO and the GWO.

- For detecting coherent generators this dissertation examines recurrence dynamics in multiple wavelet scales and improves the performance of recurrence attributes. The application of wavelet transform allows one the capability to consider the recurrence property in various wavelet scales and the correlation of recurrence behaviors across scales.

- A novel method is proposed for early detection of low-frequency oscillations by applying synchrophasor data. This research considers the problem of detecting the oscillations quickly. The proposed method is based on measured data from PMUs. The SFA is used to extract the slow varying features, SFs, from the raw measurements. Then, two thresholds based on RQA is used by applying the SFs for oscillation detection. The recurrence-derived FT is applied to identify frequency modes.

- A method is proposed to identify location and type of disturbance events based on the RQA measures. The RQA captures the characteristics of obtained data such as voltage, phase angle and frequency. Then, the PCA and an unsupervised clustering method are applied to determine two disturbance events which are generator tripping and fault.

- In term of analysis, this thesis establishes useful measures. Rotor-angles are applied to predict transient stability and to detect coherency. Also, measures applied on frequency and active power of transmission lines to identify low-frequency oscillation and frequency modes are discussed. Voltages, phase angles and frequency are used to find location of fault and to determine type of

disturbance events. These analyses are performed in several case studies such as natural oscillations, cascading failures, trained neural networks.

- In term of testing, this dissertation uses two power system case studies including, the IEEE 39-bus system and a reduced Western Electricity Coordinating Council (WECC) ac transmission system that includes 29 generators, 179 buses, and 263 branches. These case studies are tested by using Powerworld and PSSE software.

1.3 Dissertation Organization

Smart grids may evolve into a stressed condition as a result of network complication, uncertainties, renewable generation and market requirement. The incident of transient disturbances is more probable to drive a system over its stability restraint. In this framework, to evaluate and predict system stability, new and more effective methods are needed so that remedial control actions can be taken to prevent instability of the system or reduce its consequences. Detecting coherent group of generators based on the RQA and WPD is a way to take preventive actions and to avoid possible blackouts. We are interested in detecting low-frequency oscillations, identifying frequency modes, determining location and type of disturbance events.

Chapter 2 provides reviews of previous work for prediction of transient stability, coherency detection, detecting low-frequency oscillation, identifying frequency mode and determining location of disturbance events as well as type of disturbance events.

Chapter 3 focuses on prediction of transient stability and proposes an algorithm grounded on the Adaptive Neuro-Fuzzy Inference Systems and the Support Vector Machine.

Chapter 4 investigates the prediction of transient stability by applying two metaheuristic methods namely PSO and GWO to train weights of neural network.

Chapter 5 develops a flexible and reliable algorithm for distinguishing coherency to consider noisy measurement data. A method for detecting wide-area coherency employing wavelet packet decomposition and recurrence quantification analysis is proposed.

Chapter 6 proceeds with a novel method for early detection of low-frequency oscillations and identifying frequency mode based on the RQA method.

Chapter 7 concentrates on disturbance event problem and proposes an algorithm for determining the location of fault and identifying two type of disturbance events based on the RQA and PCA methods.

Chapter 8 concludes the dissertation and discusses possible future work.

Chapter 2 : REVIEW OF LITERATURE

2.1 Transient Stability

The simplest approach to identify unstable behavior is by solving differential-algebraic equations (DAE) that are related to a power system dynamic behavior [34], [18]. This method needs precise configuration of the system all along and after a fault. The Transient Energy Function (TEF) is a direct approach for detecting stability after a disturbance [35]. In this case, evaluation of stability can be performed by analyzing the difference between the kinetic energy (KE) and potential energy (PE) for a specific disturbance. Determining the KE and PE amounts for certain cases has many complications [35]. The prominent equal area criterion (EAC) presents another direct means of evaluating transient stability of multi machine systems by considering all generators of a power system as a single generator that are linked to an infinite bus and no need for solving the DAE. In this method, an equivalent machine is created from all the generators present, and at the same time, just the dynamic parameters which are related to mechanic part of equivalent generator are considered [36], [37]. The combination of energy functions and time domain simulations are used in what has become known as the extended equal area criterion (EEAC) method [38]. This method is more efficient in computing the end result of a disturbance, but is less accurate than time-domain simulations.

The support vector machine introduced in [39] has its primary application in the identification of optical characters. Support Vector Classifiers (SVC) have been applied to transient stability assessment in [40] where only post-fault voltages are considered, while

generator rotor angles are not included. A linear SVM using scaled variables to study stability of power system is applied in [41]. To improve the performance of SVM in the presence of outliers and noise, the weighted SVM is introduced in [42], and results showed an improvement in accuracy.

A core vector machine is used in [43] to assess transient stability and results exhibited the speed and performance of the applied method. Machine learning techniques that predict transient stability of a system has been used earlier in [44]. The decision tree method shown in [45] needed about 2 seconds to predict the stability of the Hydro-Quebec system after fault clearance, which does not provide enough time for automatic corrective actions. In [46], only out of–step conditions in a tie line is predicted by using the decision tree method. An SVM classifier in association with a several pre-defined voltage curve templates is used in [40] to predict transient stability condition. To evaluate fuzzy memberships, the voltages obtained from busses are compared with voltage patterns. This process identifies similarities between the measured voltages and the templates. For training purposes, the input of the SVM has similarity values.

Swarm intelligence (SI) is one of population-based metaheuristics, class of methods that are mainly created from natural colonies. Some of the most well-known techniques are ACO [5], PSO [6] and Artificial Bee Colony (ABC) [47]. Some advantage of SI algorithms to other metaheuristic methods, such as evolutionary and physics based methods, are as follows [48]: 1) In SI, information about search space is maintained during the iteration, though the information of previous generation is neglected in evolutionary algorithms. 2) SI has fewer parameters to adjust. 3) Compared to evolutionary methods, SI methods have fewer operators.

The ACO algorithm [5] is stimulated by the social behavior of ants in an ant colony. The main inspiration of ACO is their social intelligence that seeks to find the shortest path between the nest and source of food. The ABC algorithm imitates the collective behavior of bees in finding food sources. The Bat algorithm [49] inspired by the location of objects which is the result of reflection by sound and has been suggested lately. The literature shows that stochastic training methods have received more attention as a result of high local optima prevention.

There are other SI techniques like: Artificial Fish-Swarm Algorithm [50], Monkey search [51], Cuckoo Search [52], Firefly Algorithm [53], Krill Herd [54], etc. Their convergence rates are slow, and so they are not preferable for transient stability prediction. The GWO introduced in [48], has several advantages over other techniques. These are: easier to implement due to simpler structure; less storage requirement; faster convergence due to continuous reduction of search space; and fewer number of parameters to adjust. Different heuristic optimization methods have been used to train FNNs like Simulated Annealing (SA) [55], [56], Genetic Algorithm (GA) [57], Particle Swarm Optimization (PSO) [58], and Magnetic Optimization Algorithm (MOA). According to [59], some of these methods could avoid local minima, but still have slow convergence rates. The PSO is one of the most efficient [60] optimization algorithms which has the ability of avoiding local minima and has a fast convergence rate.

2.2 Coherency of Generators

Wide-area measurement systems have the important duty to monitor the dynamic characteristics of a power system. There are several measurement units that can furnish important parameters about a system such as generator voltages, frequencies and rotor

angles [61]. Applying this measurement is a dilemma to examine instinctive changes in dynamic characterization of a system. Determining the coherency of generators after fault clearance, on the other hand, could be indicative of the state of stability of multi-machine power system. Generator coherency can be determined by examining and comparing rotor angles through synchronized measurement. This can eventually help operators to take corrective actions if the system exhibits symptoms of instability. Model based approaches are used [13], [14] and are grounded on the eigenvalue study of a power system. The accuracy of these methods is good when predominant system condition and different parameters are accurate. Nevertheless, changes in topology and serious load alterations can force certain generators with low synchronizing torque to switch from one group to another [15]. With rapidly increasing placement of PMUs in power systems, approaches based on measurements have been broadly examined in recent literature to observe the coherency of power systems. The methods using these measurements can detect coherent group of generators without depending on dynamic models of generators, prime-movers or even the topology of the system. In order to detect a coherent group of generators, the Artificial Neural Network (ANN) is used in [62].

Particle Swarm Optimization and K-means clustering (PSO-KM) method are applied in [63] for coherency detection. Partitioning Around Medoids (PAM) is a clustering algorithm used in [64] to identify coherent groups, the representative objects of a data set are called Medoids. A graph theory-based method is proposed in [65] to determine coherent groups of generators. System coherency is identified in [66] by applying Hilbert-Huang Transform (HHT). In [67] and [68], Fast-Fourier Transform (FFT) was applied for detection of coherent generators. The Principle Component Analysis (PCA) was investigated in [69] to

find the coherent generators, and Independent Component Analysis (ICA) was used in [70] to eliminate noise in the obtained measurement to identify coherent generators. Wavelet phase difference was used in [71] to determine coherent generators. Magnitude and phases of the Koopman modes was used in [72] to identify coherent groups.

In [73] the RQA method is used to detect coherent group of generators. Wavelet transform method was used in [74], [75] to detect low frequency oscillations and local oscillations. The empirical mode decomposition and wavelet shrinkage analysis is used in [76] to specify nonlinear trends of large interconnected power systems. For detecting oscillation modes, wavelet based support vector data description is applied in [77]. Continuous wavelet transform is used in [78] to study power system low-frequency oscillations. Coherency detection was also studied using the Projection Pursuit [79] method and the Spectral Clustering [80] method. Even though the methods mentioned above demonstrated satisfactory results to identify coherent groups of generators, each of them has a number of constraints.

ANN algorithms [62] need enormous offline training data, while the PSO-KM and PAM [63], [64] engender excessive computational burden to detect coherent generators. The graph theory method [65], in some cases needs subjective decisions to detect coherency. The HHT [66] is unable to determine the frequency mode if closely spaced. The Fast Fourier Transform [21], [22] considers stationary data and linearized systems to detect coherent groups of generators.

By obtaining measured data, the PCA [23] and the ICA [1] can directly detect coherent groups of generators, but to detect precise coherent groups, a priori information of the system is needed. In addition, the PCA is sensitive to noisy measurements. The wavelet

phase difference [2] and Koopman [3] methods need to decompose the measurement data, which will carry high computational burden. Only two measures of the RQA are considered in [4], and noisy data was not considered. The wavelet based methods decompose the signal and need more computation time.

2.3 PMU and data reduction methods

By considering the growing amount of PMU data, there is a need to find a good solution to manage and use the huge amount of data provided by synchrophasors in real-time. A Phasor Data Concentrator (PDC) that collects data from 100 PMUs with sampling rate of 30-Hz creates more than 50 GB of data every day [81].

There are different methods for phenomenon detection, and they can be classified into two categories: statistical algorithms and signal processing algorithms. For statistical algorithms [82], the challenging task is to set a threshold for detecting the phenomenon. There are different techniques applied in signal processing such as waveform analysis and Fourier [83], nonlinear adaptive filter [84] and Kalman filters [85]. The signal processing methods have a proven track record as powerful tools in power systems. The drawback of these methods is applying complex matrix calculation or high sampling rates that raises computational burden.

The data reduction methods are compressing techniques like wavelet analysis [86] and linear-feedback shift register reseeding [87]. Adaptive fuzzy logic method has been used to detect power-quality issues from waveforms [88]. Nevertheless, a few compressing techniques have been considered to real power systems and less attention is drawn to the compressing methods that contain the critical phenomenon such as sustained oscillations.

SFA [89] is essentially parameter-free and only two parameters dimension and time delay of the embedding vectors should be selected. Also, this method is not sensitive to noise.

2.4 Frequency Mode

Different reasons can cause oscillations in the power systems. Most of the oscillations are damped by the system but a few of them may remain undamped and lead to system collapse. So, investigation of low-frequency oscillations in the power system is necessary for power system operation and control. The Prony analysis [90] of the signals gives the low- frequency modes but when the signal has noise it is difficult to estimate actual frequencies existing in the signal. Fast Fourier Transform [91] gives the dominant low frequency modes with the related amplitudes. The Stockwell transform [92] of the signals provides a contour plot that helps to find the zone of disturbance, therefore this method helps to detect the low- frequency modes of oscillations related with the zone of disturbances. The Wigner distribution [92] helps in finding the contour plot of the signals, so it would be easy to find out low- frequency modes of oscillation. The Estimation of Signal Parameters by Rotational Invariance Technique [92] has negligible bias in the presence of the colored Gaussian noise.

2.5 Disturbance Events

Some studies have been done for identification and classification of disturbance events. In [93], frequency deviation before and after a disturbance, and the peak initial swing frequency derivative was applied as features for each disturbance signal. The difference between the normal and the post-disturbance values of voltage, current, and frequency is used in [94] as features to identify and classify disturbance events. A decision tree is used

in [95] to identify and classify three types of faults which are single phase to ground, double phase to ground, and three phase to ground.

To identify disturbance events, frequency and voltage indicators have been studied in [96], [97], and [98]. Event identification based on wavelet and the PCA are presented and examined in [99], [100], [101]. The unsupervised clustering method is examined in [102]. To identify patterns of events, machine learning methods have been used to analyze PMU data [103] , [104].

Chapter 3 : ENSEMBLE ADAPTIVE NEURO FUZZY SUPPORT VECTOR MACHINE FOR PREDICTION OF TRANSIENT STABILITY

3.1 Introduction

In power systems, transient instability has been the foremost problem because, if the power system becomes unstable, the consequences will be catastrophic and lead to huge financial losses [105]. Hence, for the security operation of power systems, studies on transient stabilityare essential.

Since machine learning techniques have the capability of mining the useful information from PMU data as well as high speed calculation and accuracy, they have been used to predict transient stability based on PMU data [106], [107]. Transient stability prediction can be considered as a two-class classification problem. For offline training relevant features should be selected under different operation conditions and then suitable classification method should be applied to predict the stability status.

For prediction of transient stability, the features used as input for the machine learning techniques include rotor angle and speeds [106], [107] generator terminal voltage [108]. The combination of different variables and time series are considered as input features [109], [110]. The transient energy functions were considered in [111] as input features and prediction accuracy of applied classifier has shown enhanced results.

For predicting transient stability, different machine learning methods, such as artificial neural networks [106], decision trees (DTs) [109], support vector machines (SVMs) [40], and extreme learning machines [112] have been examined. The DT algorithm can deliver the prediction results and make known the rules it learned [113]. Also, key features through the training process can be filtered.

The main contributions of this chapter are: 1) Three separate weighted support vector machine are considered as base classifiers; 2) a combined WSVM and ANFIS method is proposed to build the classifier, where the WSVM is applied to obtain better classification accuracy, and the ANFIS is applied to find the most proper membership functions to predict stability; the proposed method can predict the transient stability with a very high accuracy.

The rest of this chapter is organized as follows. Section 3.2 introduces support vector machine. Section 3.3 introduces test system, data generation and cross validation. Section 3.4 includes the proposed method based on the WSVM and ANFIS methods. Results and conclusions are provided in Sections 3.5 and 3.6 respectively.

3.2 Support Vector Machine

Support Vector Machines in machine learning are supervised learning models which are related to learning algorithms that examine data applied for classification. Having a set of training instances, each labeled as fitting to one of two categories, the SVM shapes a model based on its training algorithm that allocates new data set to one of two categories. The points in the space can be represented by SVM model, which are mapped to distinct categories with the maximum distance. Same space will be considered for mapping new data and projected to fit in a category based on which side of border they fall. Besides linear classification, SVMs are capable of performing nonlinear classification by kernel function which maps their inputs to high dimensional feature space.

In SVM technique the main purpose is to construct a hyperplane that considers maximum distance for classifying input data. Given *n* training data points, Fig. 3.1 shows the hyperplane:

$$\{(x_i, y_i)\}_{i=1}^{n},\ x_i \in R^n,\quad y_i \in \{-1,1\} \tag{3.1}$$

$$w.x_i + b \geq +1 \ \ when\ y_i = +1 \tag{3.2}$$

$$w.x_i + b \geq -1 \ \ when\ y_i = -1 \tag{3.3}$$

Where x_i are the input vectors and y_i are class labels [114]. N is the number of inputs. Hyperplane H is defined such that:

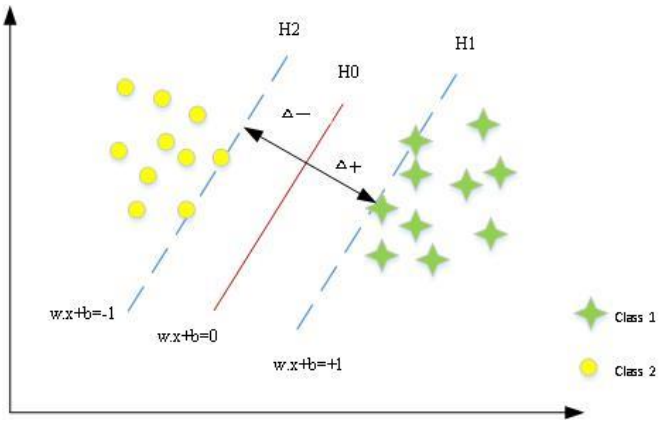

Figure 3-1:SVM hyperplane

H_1 and H_2 are the planes. The plane H_0 is the median in between, where $w.x_i + b = 0$. $\Delta+$ is the shortest distance to the closest positive point and $\Delta-$ is the shortest distance to the closest negative point. w is a weight vector, b is bias. The distance between H_0 and H_1 is calculated by

$$d(positive) = \frac{|w.x + b|}{\|w\|} \tag{3.4}$$

The objective function, shown below, is the solution of the support vector method:

$$Minimize \ \ \phi(w) = \frac{1}{2} w^T w + C \sum_{i=1}^{n} \xi_i \tag{3.5}$$

Subject to:

$$y_i(w, \phi(x_i) + b) \geq 1 - \xi_i, \quad i = 1, ..., n \tag{3.6}$$

$$\xi_i \geq 0, \qquad i = 1, ..., n \tag{3.7}$$

The function maps the training vector x_i into a higher-dimensional space, parameter C plays the role to make adjustment between classification errors and having maximum distance specified by user, the norm of a vector which is perpendicular to isolated hyperplane is called w. The slack variable ξ_i assess the rate of data are not classified correctly. The kernel function [115] is defined by the product $(\phi(x_i).\phi(x_j))$ and designated by $K(x_i, x_j)$. There are different kernel functions like polynomials, radial basis function (RBF) and sigmoid function. The RBF is used in this paper because it gives better results when compared to other alternatives. The width of the Gaussian is shown by δ and x_j are the support vectors.

$$K(x_i, x_j) = e^{(-\|x_i - x_j\|/2\sigma^2)} \tag{3.8}$$

The theory of SVM is extended in [116] since each data point in the real world is not exactly classified into one of two classes. Each data point in the weighted support vector machine (WSVM) has a different influence to create the hyper-plane, so weighted

memberships are allocated to each data point to show the importance of each point in building the SVM hyperplane, so we have same objective function as equation 3.5 and the only added term is fuzzy membership s_i .

$$Minimize \ \ \phi(w) = \frac{1}{2} w^T w + C \sum_{i=1}^{n} s_i \xi_i \tag{3.9}$$

The term is the major difference between SVM and WSVM. Considering as the error in the SVM, the fuzzy membership can increase or reduce the impact of each training data point. The constraints are similar to those for the SVM.

3.3 Transient stability predictors

Transient instability happens in a system when at least one rotor angle becomes unbounded in comparison to the rest of the system under the condition that the system experiences high power flows, and therefore is under stress. Angle measurements are made with respect to a reference commonly known as the center of inertia angle (COI) [34], [18]. However, a certain time after fault clearance is needed to observe the evolution of generators angles that determines stability. Thus, both rotor angles and voltage magnitudes are investigated for predicting transient instability.

3.3.1 Test System

To assess the ANFIS training and the performance of the classifier, the IEEE 39-bus test system [117] is used. This system consists of 39 buses, ten generating units, 46 transmission lines and 19 loads. The system diagram is shown in Fig. 3.2.

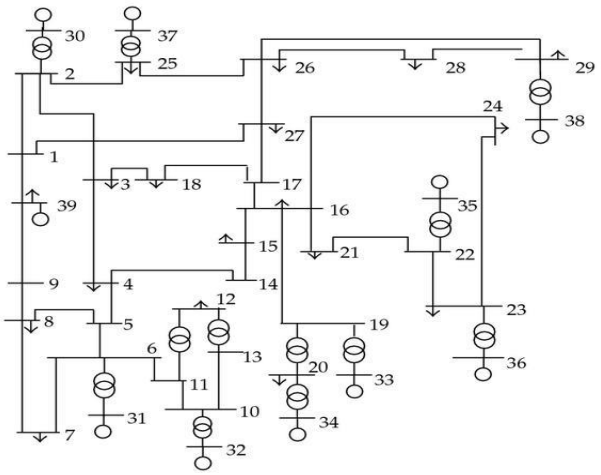

Figure 3-2: IEEE39-bus test system

3.3.2 Data generation and cross validation

For the training classifier, the PowerWorld software was used to generate data for the system. Three-phase to ground faults were created on each transmission line at 25%, 50% and 75%. For each contingency, the fault clearing time is assumed to be 5 cycles, and the fault is removed by opening circuit breakers at both ends of the line. Load levels are considered base load and overloaded for 5%, 10%, 15% and 20%. Single phase to ground faults are not considered in this paper. A total of 300 simulations were created, and for each case, variations of rotor angles and voltage magnitudes are reported. For each simulation, a class label was defined according to transient stability index which is based on the angle [108], [118]. If the system falls in the ‘stable’ class, label “1” is used for the simulation; alternatively, if the system is unstable, then a class label “-1” is assigned for the simulation.

Due to the large number of samples, a 10-fold cross validation sample was chosen. Almost 80% of the generated data was applied for training purpose, and the other 20% was used for testing the performance. To obtain better accuracy of the weighted SVM, the

parameters C and δ are optimized through a grid search. C=64 and δ =32 gives the best classification accuracy which is 93.05%. Table 3.1 provides the classification accuracies of the SVM and the WSVM methods; as seen, the classification accuracies of the WSVM are higher than the SVM.

Table 3-1: Classification performance

Classifiers	*C* Parameter	Kernel Parameter (δ)	Classification Accuracy (%)
SVM	64	32	93.05
		10	90.12
		1	87.35
WSVM	100	32	95.08
		10	93.58
		1	89.66

3.4 Adaptive neuro fuzzy inference combined WSVM

An ANFIS is a fuzzy system that, by using neuro-adaptive learning techniques, tunes its membership function parameters. In Madani's [119] fuzzy system, fixed membership functions are selected arbitrarily, and the user specifies rules based on characteristic of variables in the model.

The membership function parameters are fine-tuned by the back-propagation method only, or a hybrid method, which is a combination of backpropagation and least squares type. Fuzzy systems learn from the data they are modeling by applying the necessary adjustments. Fig. 3.3 shows ANFIS inference which has 5 layers. Inputs Δ_1 and Δ_2 distances are obtained from weighted SVMs and *f* is the output that indicates system transient stability status. The detailed structure of the ANFIS is [119] shown in Fig. 3.5.

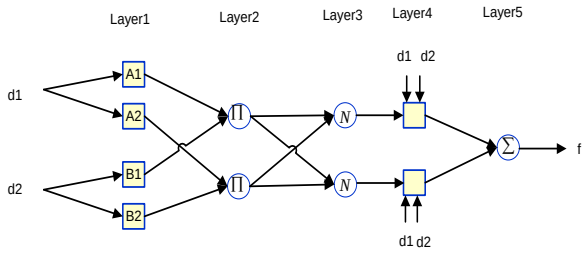

Figure 3-3:Adaptive neuro-fuzzy architecture

In order to increase the performance of WSVM, three different classifiers are combined. Combination of base classifiers determine the final decision, and this is shown in Fig. 3.4. For classification two stages are considered. Identical training data are used for train of all WSVMs in first stage. ANFIS is applied to combine the result of three WSVM in second stage. The output of ANFIS determines transient stability status.

The WSVM classifier and the ANFIS process are described in sections 3.4.1 and 3.4.2

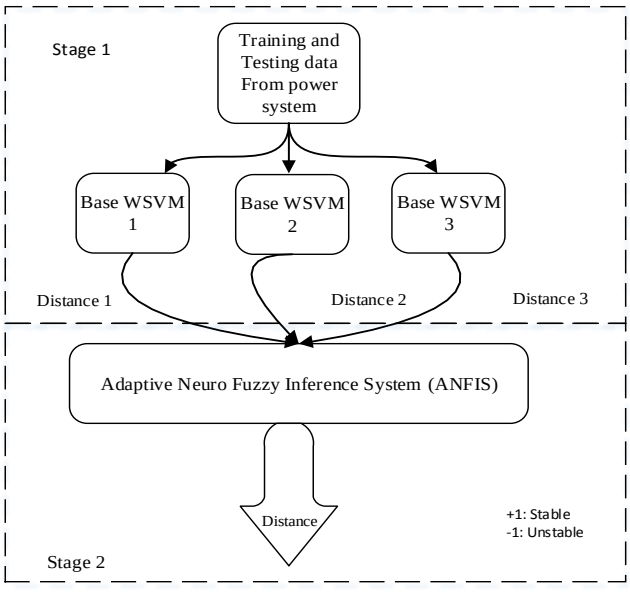

Figure 3-4: Two stage scheme for prediction

3.4.1 Base classifiers

Development of the WSVM with RBF kernel using different values of δ and C leads to six different base WSVMs. Base classifiers, kernel parameters with different values, classification accuracy and validation accuracy are shown in Table 3.2. Table 3.3 gives the distances $\Delta_1, \Delta_2, \Delta_3$ of six different operating conditions from separating hyperplanes of base classifiers A1, A4 and A6. A few selected cases are shown; the first five conditions are related to unstable cases, and the last one is related to a stable case.

Table 3-2:Base classifiers

Kernel Function		Base SVM No.	Kernel Parameter	Classification Accuracy %	Validation Accuracy %
RBF	C=100	A1	δ=32	95.08	93.05
		A2	δ=10	93.58	91.12
		A3	δ=1	89.66	87.25
	C=120	A4	δ=32	93.41	91.32
		A5	δ=10	91.75	89.78
		A6	δ=1	85.14	85.01

The load at three different buses are increased by 10% and 50%. Distances for each classifier and for different operating condition are computed.

Table 3-3: Distances of testing data

Operating Conditions	Δ_1	Δ_2	Δ_3
Load increased 150% at bus 16	-0.34	-0.338	-0.378
Load increased 110% at bus 16	-0.163	-0.158	-0.181
Load increased 150% at bus 18	-0.293	-0.291	-0.248
Load increased 110% at bus 18	-0.156	-0.153	-0.106
Load increased 150% at bus 8	-0.305	-0.296	-0.383
Load increased 110% at bus 8	0.153	0.142	0.1811

3.4.2 Fuzzy Integration (Stage 2)

To study transient stability entire contingencies are applied for training. Distances of for all operation conditions are used from Table 3.3. The ANFIS used here has 8 rules, with two membership functions being assigned to each input variable. 250 training data and 50 checking data were sampled from inputs. The training data is applied for the training of ANFIS; at the same time checking data is applied for validation of the established ANFIS. The structure of the ANFIS after training is shown in Fig. 3.5. Two membership functions are assigned to each input and there are 8 rules that define the output of system. The logical operation for rules are the AND operators.

The input distances from WSVM1 and WSVM 2 (input 1 and input 2) and the output are illustrated in Fig. 3.6. The surface has two important regions, yellow and dark blue. When the inputs are mapped to the yellow region, the output of the system is "1", which indicates a stable system, and when inputs are mapped to the dark blue region, the output is "-1" which indicates that the system is unstable. The hybrid method for membership optimization is used (mixed least squares and backpropagation). For training purpose, error tolerance is set to zero. For comparison, the average percentage error is used.

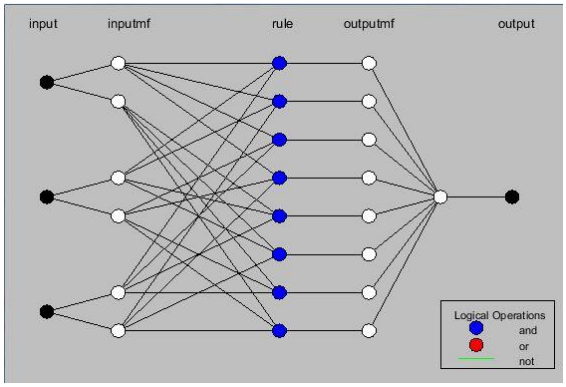

Figure 3-5:Three input ANFIS with 8 rules

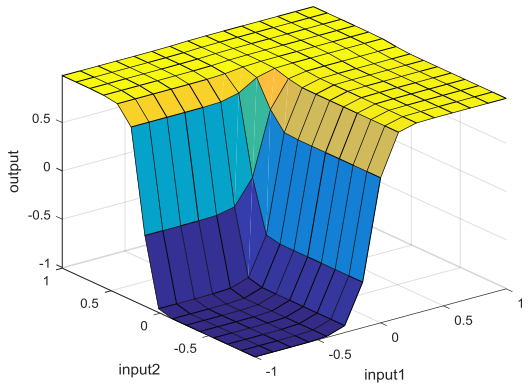

Figure 3-6: Surface at 40 epochs

The training error is about 1.1%. By increasing the number of epochs, the error decreases. The average percentage error is 2.1% for testing data.

3.5 Results

The combined ANFIS consists of classifies A1, A4, A6 and result of classification are shown in Table 3.4. After 40 epochs, the training average percentage error (APE) is 1.1%, which has high accuracy when compared to accuracy of base weighted classifiers.

Table 3-4: Classification result

Model	APE Train (%)	APE chk(%)	Training set size
Combined ANFIS	1.1	2.1	250

In this dissertation, we assume the structure of ANFIS to be fixed and tuning of membership function to be solved through the hybrid learning rule. To demonstrate reliability of method, process examined under different conditions by changing the topology of the system. Assessment is based on three cases: 1) the transmission line interconnecting buses 16 and 24 was taken out of service by opening both circuit breakers at the end of the line; 2) the transmission line interconnecting busses 4 and 14 was taken

out of service; and 3) the transmission line interconnecting busses 25 and 26 was taken out of service. Such network alteration, although unusual, could happen on a given day because of line maintenance or forced outages. When the topology of power system changes consequently power flows and voltages of system alters. So, after a fault happens, the recovery characteristic would be new and different. The classifier introduced in section 3.4 was assumed to have been trained with the fully intact system, and then used to predict the transient stability of the altered system. A three phase fault was created in each transmission line, and fault clearance time was about 5 cycles. A three-phase-to-ground fault was created on each of the remaining transmission lines at 25%, 50% and 75% of capacity. For each scenario, 200 cases were created. The results of stability status for classification are illustrated in Table 3.5. Results reveal that the applied ensemble method can predict stability status with high precision for a power system under different contingencies and topology changes. Comparison of result Table 3.5 with results Table 3.1 shows that proposed method has better accuracy.

Table 3-5:Stability prediction under topology changes

Scenarios	Prediction Accuracy		
		Classified as Stable (%)	Classified as Unstable (%)
Line 16-24 open	Stable	99.16% (119/120)	0.83% (1/120)
	Unstable	0% (0/80)	100% (80/80)
Line 4-14 open	Stable	100% (120/120)	0% (0/120)
	Unstable	1.25% (1/80)	98.75 (79/80)
Line 25-26 open	Stable	97.5(117/120)	2.5% (3/120)
	Unstable	1.25% (1/80)	98.75 (79/80)

3.6 Conclusion

Although the SVM can create classifiers with high prediction accuracy, in order to decrease the effect of outliers in the training model, and to contribute more weight to the

important training data points, a weighted support vector machine (WSVM) was built. The fuzzy membership related training data helps to provide different degrees of membership in the training of the classifier. Hence, the WSVM obtains superior classification accuracy compared to the SVM. Numerical tests show that the applied approach is adequate to determine the stability status of a power system. Multiple SVM classifiers were combined with an adaptive neural fuzzy system that considers distances of the data points from the optimal hyperplane. The improvement of this method is seen when the WSVMs are combined with an ANFIS that is able to tune its membership function parameters to achieve higher accuracy. The proposed adaptive neuro fuzzy combined WSVM persistently exceeds the individual base SVM classifiers.

Chapter 4 : POST-FAULT TRANSIENT STABILITY STATUS PREDICTION USING GREY WOLF AND PARTICLE SWARM OPTIMIZATION

4.1 Introduction

The power system stability can be considered as the capability of a power system to recover from a large disturbance such as a fault, and settle back to an operating equilibrium state. When most of the system variables are constrained which means variables such as voltages and angle of busses are within their limits, the entire system remains in steady-state condition [105].

The power system deviates from the pre-fault equilibrium and experiences a change in the direction of the post-fault dynamics when a disturbance occurs on the system. Rotor angle stability denotes the capability of synchronous machines to be in synchronism with the rest of the system after a fault occurrence. The rotor angle increase of some generators causes the loss of their synchronism with other generators and, in this case, instability happens. Transient instability could happen when the system faces a severe disturbance (N-1 contingency). The system response includes a large increase of generator rotor angles, which is described by the nonlinear power-angle relationship. The reason for this kind of instability is inadequate synchronizing torque [105].

Lately, Feedforward Neural Networks (FNNs) with two layers have been used extensively [120] and are appropriate for classification of nonlinearly separable patterns [121]. Any continuous and non-continuous function can be estimated by two-layer FNNs [122]. Learning is a necessary part of neural networks and the standard [123] or improved [124] back-propagation (BP) algorithm are mostly used for training of FNNs. The drawbacks of BP algorithm are slow convergence [59] and trapping in local minima.

Throughout the learning process of FNNs, the aim is to find finest combination of weights and biases to obtain the minimum error. Nevertheless, in most cases the convergence of FNNs lead to best locally solutions and not globally. The initial values of weights, biases, learning rate and momentum [59] has high impact on the convergence of the BP algorithm. Based on the literature, to improve the algorithms founded on BP algorithm, applying novel heuristic optimization methods or evolutionary algorithm is prevalent alternative.

For training of FNNs different heuristic optimization methods have been applied, such as Simulated Annealing (SA) [55], Particle Swarm Optimization [58], Magnetic Optimization Algorithm (MOA) [125], and Differential Evolution (DE) [126]. Some methods such as SA and GA [59] can decrease the probability of being stocked in local minima, but slow convergence rates are a problem yet. For reducing both mentioned drawbacks PSO is one of the most [60] efficient algorithms. In this chapter, the performance of GWO is examined for training FNNs in comparison with PSO. The contribution of this chapter is applying the PSO and GWO methods on FNNs to train their weights for classification of stable and unstable cases, post-fault rotor angles and voltages magnitudes are used as inputs to feedforward neural network classifiers.

The reminder of this chapter is organized as follows. Section 4.2 presents transient stability prediction approaches, as well as the GWO and PSO algorithms. The test system and data generation are discussed in Section 4.3. Section 4.4 shows how the FNN is trained by the GWO and PSO. Results and conclusions are provided in Sections 4.5 and 4.6 respectively.

4.2 Transient stability prediction approaches

4.2.1 Grey Wolf Optimization

The main idea of GWO includes social hierarchy and hunting behavior of wolves:

4.2.1.1 Social hierarchy and encircling prey

For designing the GWO, the social hierarchy of wolves can be considered; the leaders are called alpha; the second level is beta; and the lowest ranking wolves is named omega. The alpha wolf is dominant since his commands should be followed by others. The beta is second level in the hierarchy of grey wolves; betas are subordinates' wolves that help the alpha in decision-making. Delta wolves dominate the omega and should submit to alphas and betas. To design the GWO, the social hierarchy of wolves is mathematically modeled such that the fittest solution is considered alpha (α); beta (β) and delta (δ) are specified the second and third best solution respectively.

The last candidate solutions are considered as omega (ω). The optimization in the GWO algorithm is directed by α, β and δ, and they are pursued by ω wolves. The following equation shows the encircling behavior of wolves:

$$\vec{D} = |\vec{C} \,.\, \vec{X}_p(t) - \vec{X}(t)| \tag{4.1}$$

$$X(\vec{t}+1) = \vec{X}_p(t) - \vec{A}.\vec{D} \tag{4.2}$$

$\vec{X}$ shows the position vector of a grey wolf and $\vec{X}_p(t)$ is the position vector of the prey, and $\vec{D}$ is distance between the grey wolf and prey. Current iteration is shown by *t*; $\vec{A}$ and $\vec{C}$ are coefficient vectors, and calculated as follow:

$$\vec{A} = 2\vec{a}.\vec{r_1} - \vec{a} \tag{4.3}$$

$$\vec{C} = 2.\vec{r_2} \tag{4.4}$$

The value of $\vec{a}$ linearly decreases from 2 to 0 in the process of iterations, r_1 and r_2 are random vectors in the range [0, 1]. These values are considered based on [48]; other values result in reduced training accuracy. To observe the effects of equations (4.1) and (4.2), a position vector with two dimensions are depicted in Fig. 4.1. As illustrated in the figure, a grey wolf in the position (M, N) can change its position according to the position of the prey (M*, N*). By adjusting the value of $\vec{A}$ and $\vec{C}$, each wolf (agent) has good flexibility to reach the prey.

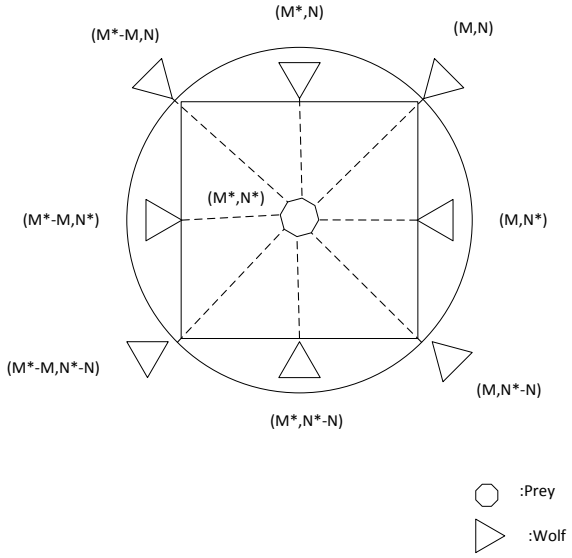

Figure 4-1: Search agents position updating mechanism

The same approach can be extended to higher dimensions, for instance, *n* dimensions, and the grey wolves will move in hyper-cubes around the best solution achieved earlier.

4.2.1.2 Hunting

Generally, we know about the location of the optimum (prey), and so, it is assumed alpha, beta, and delta have better knowledge about the possible location of the optimum. As a result, the first three best solutions achieved until now is memorized and require other search agents to adjust their positions in proportion to the position of the best agents. The related formulas are shown in [48] as follows:

$$\vec{D}\alpha = |\vec{C}_1 . \vec{X}_\alpha - \vec{X}| \tag{4.6}$$

$$\vec{D}_\beta = |\vec{C}_2 . \vec{X}_\beta - \vec{X}| \tag{4.7}$$

$$\vec{D}_\delta = |\vec{C}_3 . \vec{X}_\delta - \vec{X}| \tag{4.8}$$

$$\vec{X}_1 = \vec{X}_\alpha - \vec{A}_1 . (\vec{D}_\alpha) \tag{4.9}$$

$$\vec{X}_2 = \vec{X}_\beta - \vec{A}_2 . (\vec{D}_\beta) \tag{4.10}$$

$$\vec{X}_3 = \vec{X}_\delta - \vec{A}_3 . (\vec{D}_\delta) \tag{4.11}$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \tag{4.12}$$

a) Exploitation: To find the optimum, the value of $\vec{a}$ decreases [48]. It should be considered that variation range of $\vec{A}$ is decreased by $\vec{a}$. $\vec{A}$ is a random value in the interval (-2*a*, 2*a*) where *a* decreases from 2 to 0 [48] during the iterations. The next position of a search agent can be in any position between its current position and the position of optimum when the random values of $\vec{A}$ are in range [-1, 1].

b) Exploration: Grey wolves' pattern search is based on the position of the alpha, beta and delta parameters. First they spread to search for prey (optimum) and then they get together to attack the prey. $\vec{A}$ is assigned with random values greater than 1 or less than -1 to compel the search agent to diverge from the prey. It forces the GWO algorithm to search globally. The condition $|\vec{A}|>1$ makes the grey wolves to diverge from the prey in order to try and find a better and fitter prey.

In the GWO algorithm, the search starts with creating a random population of grey wolves. During iteration, alpha, beta and delta wolves estimate the possible position of the prey. Each possible solution adjusts its distance from the prey. The parameter decreases from 2 to 0 to accentuate exploration and exploitation. Possible solutions are likely to diverge from the prey when $|\vec{A}|>1$ and converge close to the prey when $|\vec{A}|<1$. Finally, by meeting the convergence criteria, GWO algorithm terminates. The GWO algorithm is as follow:

GWO Algorithm

Set the initial values of the population size parameter *a*, coefficient vectors *A, C* and max no. of iteration.
Set t: =0
For (i=1 : i<n) do
Generate an initial population X_i *(t)* randomly.
Evaluate the fitness function of each search agent $f(X_i)$
End
Assign the values of the first, second and the third best solution X_α, X_β and X_δ respectively.
For (*i=1: i<n*) do
Update each agent in the population
Decrease the parameter a from 2 to 0.
Update the coefficients $\vec{A}$ and C .
Evaluate the fitness function of each search agent $f(X_i)$
End
Update the vectors X_α, X_β and X_δ
Set *t=t+1*
Until ($t<Max_{it}$)
Produce the best solution X_α.

4.2.2 Particle Swarm Optimization

The PSO is formed based on the swarming nature of particles seeking food in a cooperative style. The algorithm has become well known by researchers because of its faster convergence rate, limited selection of parameters, and straight forward application. The algorithm includes a swarm in a search space that has *D* dimensions in which each particle's position $xi = [ni1,ni2,...,nik]$ consists of *K* cluster vectors. The centroid cluster of c_{ik} is n_{ik}. Each particle's position is related to a velocity $Vi=[vi1,vi2,...,vik]$ where they are initialized as random numbers in the search space. An appropriate fitness function assesses the fitness of particles. According to fitness values, the best previous positions obtained by the particles express the local solutions given by $Pi=[pi1,pi2,....pik]$. For the first run $Pi=xi$ is considered. The best position obtained by the swarm in a generation given by $Pg=[pg1, pg2, ...,pgt]$ is the global solution, *t* represents the number of iteration (generation).

$$V_{ik}(t+1) = w_p * v_{ik}(t) + c_1 * r_1 * (p_{ik}(t) - x_{ik}(t)) + c_2 * r_2 * (p_g(t) - x_{ik}(t)) \tag{4.12}$$

$$X_{ik}(t+1) = x_{ik}(t) + v_{ik}(t+1) \tag{4.13}$$

Where *w* is the inertia weight which is considered 4.12. The acceleration constants are c_1 and c_2 both with a value of 2, r_1 and r_2 are random numbers between 0 and 1. These values are considered based on [17]; other values result in less training accuracy. The PSO algorithm as follow:

PSO Algorithm

Initialize particle
While maximum iteration is not obtained
For each particle
Calculate fitness value (f)
If the fitness value is better than the best fitness value (p_{ik}) until now
set current value as the new p_{ik}
End

Choose the particle with the best fitness value of all the particles as the p_g (global solution)
For each particle
Calculate particle velocity based on (6.12)
Update particle position based on (6.13)
End

4.3 Test System and data generation

Although, all power systems are different, faults have to be cleared within a specific amount of time (usually, a few cycles) to maintain stability. Voltage magnitude and rotor angle trajectories should be monitored for several more cycles after the fault clearing to determine whether the system maintains stability after the fault inception and fault clearing events. We consider 30 cycles (or ½ second of a 60 Hz cycle) from fault inception for training the FNN. If the stability condition can be determined within this timeframe, certain corrective actions can be taken.

4.3.1 Test system

The IEEE39-bus test system is considered to determine the performance of the FNN classifier that is trained by the GWO and the PSO algorithms. This system consists of 39 buses, ten generator units, 46 transmission lines and 19 loads. The single line diagram of the test system is shown in Fig. 4.2.

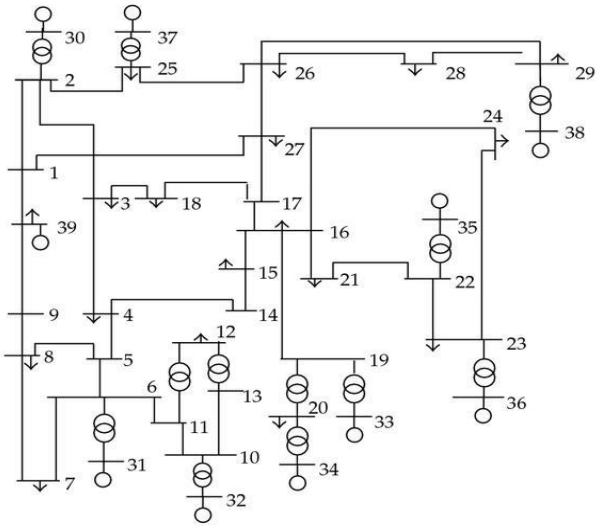

Figure 4-2: New England 39-bus system

4.3.2 Data generation

For training the FNN-classifier, several cases are simulated in PowerWorld software to generate data. Three phase to ground faults were simulated at 25%, 50% and 75% length of transmission lines. Fault clearing time is considered to be 5 cycles when the fault is removed by opening the line at both ends. A total of 300 contingencies were created, and for each case, the data for trajectory of bus voltage magnitudes and generator rotor angle recorded. According to transient stability index, which is based on rotor angle [18] a class label is assigned for each case. If system is considered as stable, class label "1" is selected, and for unstable cases, the class label "-1" is used.

4.4 The GWO and PSO Algorithms for training a feedforward neural network

The GWO and PSO are applied to train the weights of a feedforward neural network (FNN) with two layers. Fig. 4.3 illustrates the structure of the FNN, in which, the input layer has 2 nodes; the hidden layer has H nodes and output layer has O output nodes. H is the number of nodes in the hidden layer, and O is the number of nodes in the output layer.

Considering that the hidden transfer function is sigmoid, and the output transfer function is a linear activation, the output of the hidden node would be:

$$f(s_j) = 1/(1+\exp(-(\sum_{i=1}^{n} w_{ij}.x_i - \theta_j))), \quad \mathrm{j} = 1,2,...,\mathrm{H} \tag{4.14}$$

Where w_{ij} is the connection weight from the i^{th} node of input layer to the j^{th} node of hidden layer, n is the number of input node, θ_j is the bias of the j^{th} hidden node (threshold) and x_i is the i^{th} input. The weight input sum in the hidden layer is indicated by s_j and is as follow:

$$s_j = \sum_{i=1}^{n} w_{ij}.\mathrm{x}_i - \theta_j \tag{4.15}$$

$$y_r = \sum_{j=1}^{H} w_{rj}.f(s_j) - \theta_r \qquad r = 1,2,...,\mathrm{O} \tag{4.16}$$

$$E = \sum_{r=1}^{u} E_r/(u*O) \qquad where \qquad E_r = \sum_{i=1}^{O} (y_i^r - d_i^r) \tag{4.17}$$

The connection weight from j^{th} hidden node to the r^{th} output node is w_{rj} and threshold of the r^{th} output unit is shown by θ_r. The learning error E can be calculated by equation (4.17): Where u is the total number of training sample, $y_i^r - d_i^r$ is the error of the desired and actual output of the i^{th} output unit when r^{th} training sample is used for training. So the fitness function of the i^{th} training sample is defined as follows: Fitness (Xi)=E(Xi)

For the GWO algorithm, a set of weights are presented by each member of population and for training weights of the FNN by PSO algorithm, a set of weights are expressed by a particle.

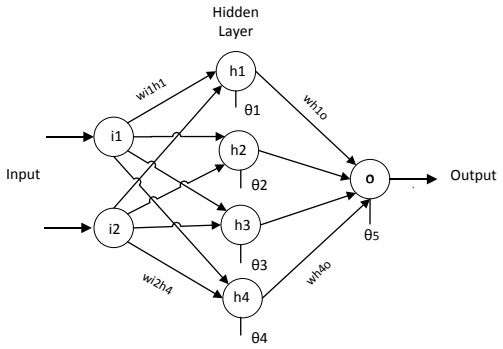

Figure 4-3:Feedforward Neural Network

In this approach, each population in the GWO method and every particle in the PSO method are encoded for a matrix. As an example, an FNN with the structure of 2-4-2 can be written as follows according to equations (4.15) and (4.16):

$$W_1 = \begin{bmatrix} w_{i1h1} & w_{i2h1} \\ w_{i1h2} & w_{i2h2} \\ w_{i1h3} & w_{i2h3} \\ w_{i1h4} & w_{i2h4} \end{bmatrix} \qquad T_1 = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{bmatrix}$$

$$W_2 = \begin{bmatrix} w_{h1o} \\ w_{h2o} \\ w_{h3o} \\ w_{h4o} \end{bmatrix} \qquad T_2 = \begin{bmatrix} \theta_5 \end{bmatrix}$$

Where W_1 is the input layer weight matrix, T_1 is the hidden layer bias matrix, W_2 is the output layer weight matrix and T_2 is the output layer bias matrix. Structure 2-4-2 indicates that feedforward neural network has 2 inputs, 4 nodes in the hidden layer and two outputs.

4.5 Results and discussions

We compare the performance of the GWO algorithm and the PSO algorithm in training the FNN for predicting post-fault transient stability of the IEEE 39-bus test system. For GWO based on [48], the population size is considered 100; the value of $\vec{a}$ linearly decreases

from 2 to 0 in the process of iterations; r_1 and r_2 are random vectors in the range [0,1]. For PSO, it is assumed [59] that the weights generated at random in the range of [0,1] are allocated to every initial particle. The values of acceleration constants c_1 and c_2 are 2, the initial weight w is 1.8 and r_1 and r_2 are two random numbers in the range of [0,1]. The population size is 30, and initial velocities of the initial particles were generated randomly in the range of [0,1]. Other values for the above-mentioned parameters results in less training accuracy.

Three cases for inputs are considered 1) voltage magnitudes 2) rotor angles 3) both voltage magnitudes and rotor angles. 30 cycles of data are considered after the clearance of fault. For instance, FNN is trained with the structure of 30-H5-2 meaning 30 inputs, five nodes in the hidden later, and 2 outputs. The input can be voltage magnitudes, rotor angles or both. When the output is 1, it indicates that the system is stable, and if it is -1, system is considered to be unstable.

For training purpose, cross validation is used. Assuming the number of samples is M, we train the FNN with M-1 samples. One sample is used to test generalization performance of the FNN. This process is repeated M times to ensure the generalization performance. For every fixed hidden unit number, two algorithms were run and maximum training iterations is set to 300. Comparison of the two algorithms when the inputs are voltage magnitudes are given in Table 4.1. Results show that training error when training by GWO is less than when trained by PSO. We also see that when the hidden layer has 10 neurons, the FNN has the lowest training error, which is 0.057977. Table 4.2, shows the results when the inputs of FNN are only rotor angles. The GWO algorithm has better performance

compared to the PSO algorithm. It is also clear that training error is less when compared to Table 4.1.

When using both rotor angles and voltage magnitudes, we consider 60 inputs - 30 for voltage magnitude samples and another 30 for rotor angles. Table 4.3 shows that the results when the FNN is trained by GWO are better than those from the FNN when trained by PSO. Generally, using both rotor angles and voltage magnitudes yields better training errors.

Table 4-1: Comparison of the performance of the two training algorithms when inputs are voltage magnitudes only

Hidden Neurons	Training Error by GWO	Testing Error By PSO
5	0.063298	0.071995
10	0.057977	0.080121
15	0.060096	0.078075

Table 4-2: Comparison of the performance of the two training algorithms when inputs are rotor angles only

Hidden Neurons	Training Error by GWO	Training Error By PSO
5	0.0028239	0.015411
10	0.0027488	0.045658
15	0.0051686	0.024059

Table 4-3: Comparison of the performance of the two training algorithms when inputs are both voltage magnitudes and rotor angles

Hidden Neurons	Training Error by GWO	Training Error By PSO
5	0.0018413	0.0029389
10	0.0014557	0.0043846
15	0.0038142	0.0099064

The training errors for FNN-GWO and FNN-PSO are shown in Figs. 4.4 through 4.6 for the three types of inputs. In all three cases, the FNN-GWO has a smaller training error. Figure 4.6 indicates that when both rotor angles and voltage magnitudes are considered as inputs to the FNN, it converges in the fewest number of iterations.

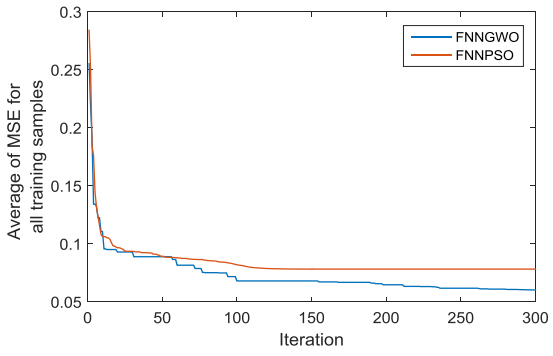

Figure 4-4:Training error with 15-nodes FNN with voltage magnitudes as inputs

The average Mean squared Error (MSE) for all training samples are based on 10 independent run. These figures confirm that the FNN-GWO achieves the best classification accuracy for all values of hidden neurons. Also it can be inferred that the FNN-GWO has a better training accuracy than FNN-PSO, and the FNN-PSO has a faster convergence rate than FNN-GWO.

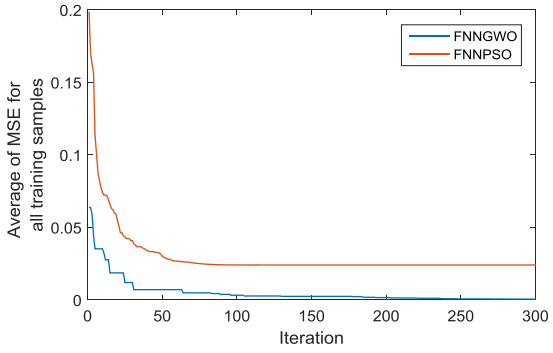

Figure 4-5: Training error for the 15-node FNN with rotor angles as inputs

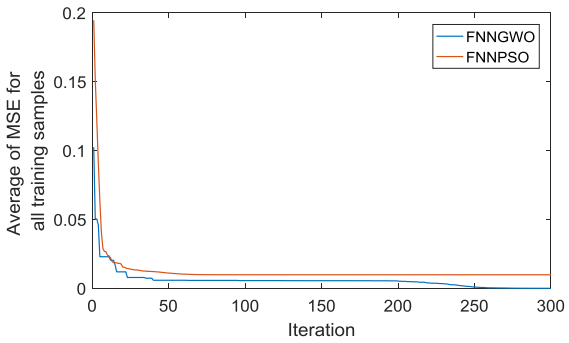

Figure 4-6: Training error for the 15-node FNN with both voltage magnitudes and rotor angles as inputs

To determine the robustness of the methods, the transient stability prediction procedure was checked by changing the topology of the test system. Transmission line 4-14 is opened, and a three-phase fault simulated with fault clearing time set to about 5 cycles. Three-phase to ground faults were initiated on each transmission line at 25%, 50% and 75% length of the line. A total of 100 cases were created for each scenario. The testing errors are illustrated in Table 4.4. According to the obtained results both classifiers predict transient stability status with good accuracy even if the topology of system is changed even though, the GWO-trained FNN has better accuracy. Figures 4.7-4.9 illustrates that FNN-GWO has better training accuracy for all three cases and FNN-PSO has faster convergence rates.

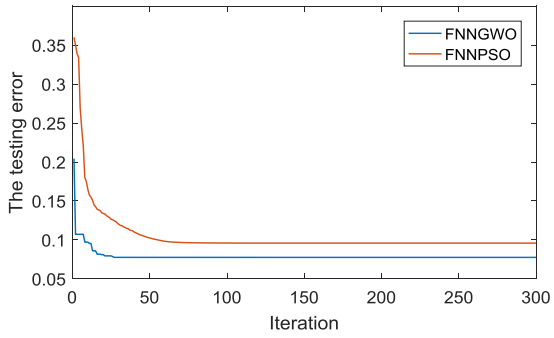

Figure 4-7: Testing error for 15-node FNN with voltage magnitudes as inputs when system topology is changed

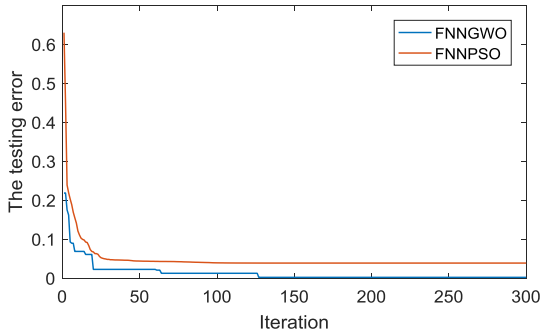

Figure 4-8: Testing error for 15-node FNN rotor angles as inputs when system topology is changed

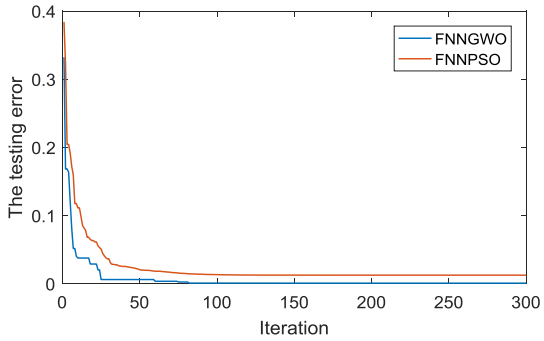

Figure 4-9: Testing error for 15-node FNN with both voltage magnitudes and rotor angles as input when system topology is changed

Results show that the FNN-GWO has the smallest training and testing errors when both voltage magnitudes and rotor angles are used. The reasons that the FNN-GWO has better training accuracy may be explained as follows: 1) Adaptive values of $\vec{a}$ and $\vec{A}$ guarantees the exploration and exploitation. By decreasing $\vec{A}$, half of the iterations are committed to exploration, and the rest of the iterations are allocated to exploitation. It should be mentioned here that $\vec{C}$ does not decrease linearly and it provides stochastic values regularly to point out exploration both on initial and final iterations. This is useful in case

of local optima inactivity which means that final solution is a global solution among all possible solutions and not just restricted to particular local neighborhood of values.

Table 4-4: Comparison of the performance of the two algorithms for a 15-node FNN

Input Data Type	Testing Error by GWO	Testing Error By PSO
Voltage Magnitudes	0.077571546	0.095881884
Rotor Angles	0.0081111439	0.038471153
Voltage Magnitudes and Rotor Angles	0.005424561	0.012661489

4.6 Conclusion

Accurate prediction of power system instability can lead to prevention of blackouts. Two metaheuristic methods for training the weights of an FNN to predict the stability status of a power system using post-disturbance voltage magnitudes and generator rotor angles are studied. The FNN trained by the grey wolf optimization method can predict system stability status better than the particle swarm optimization method, even though the latter method converged faster. This method could predict transient stability of the system with good accuracy even when the network topology was changed.

It is possible to use the adopted intelligent methods for real time transient stability assessment and control of larger systems. However, some effort will be needed to train the FNN with credible contingencies under varying operating conditions.

Chapter 5 : COHERENCY DETECTION OF GENERATORS USING RECURRENCE QUANTIFICATION ANALYSIS

5.1 Introduction

With the increasing installations of PMUs, enormous data related with operation of power systems can be obtained more quickly. Therefore, the situational awareness [127] should be improved in the smart grid with the attained PMU data. Applying data-driven methods to determine the coherency among the generators in power system is one of the important tasks of situational awareness [128].

Methods for examining the hidden patterns of nonlinear systems have recently improved and they bring more opportunities to better understand nonlinear behavior of power systems from new perspectives. For example, an effective time-frequency study for nonstationary data can be accomplished by the use of wavelets. The RQA was demonstrated as an effective method to examine and quantify the nonlinear and nonstationary signals [129]. It should be mentioned here that power system measurements, such as voltage and current, are nonstationary in nature. The combination of WPD and RQA can bring more accurate results because the RQA statistics are computed from multiple wavelet scales, and these statistics are more sensitive to path variation and less sensitive to noise. By applying the WPD and RQA, the measured rotor angles, which are actually measured bus angles, are first decomposed into several wavelets by discrete wavelet packet transformation. Then the rotor angles are quantified inside every wavelet scale. Hence, the RQA measures consider both the main scale and several wavelet scales, whereby the transient and steady-state nature of the signal can be obtained more effectively

since they are dependent on local scales. The scales at each level of wavelet decomposition are known as local scales.

The chapter is organized as follows: the WPD and RQA theory are studied in section 5.2. The proposed method is presented in section 5.3. Simulation results and case studies are discussed in section 5.4. Comparison with other approaches are presented in section 5.5. Finally, conclusions are drawn in section 5.6.

5.2 Review of the WPD and RQA theory

5.2.1 Wavelet Packet Decomposition

The wavelet decomposition (WD) has an extended version called WPD which offers enhanced signal examination [130]. The main signal is divided into two subspaces V and W by wavelet decomposition. One of the subspaces V carries low frequency information and the other subspace W carries high frequency information related to the main signal. The subspace V which contains low frequencies is decomposed and shown in Fig. 5.1. The WD only divides the frequency

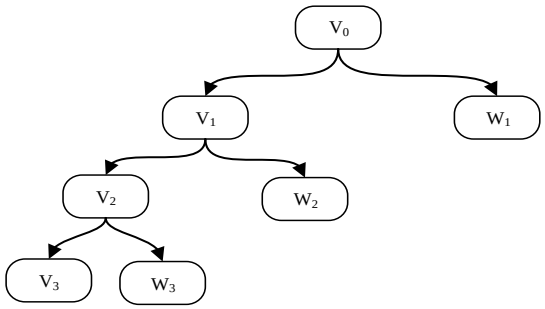

Figure 5-1:Wavelet decomposition structure

axis precisely approaching low frequencies. While the high frequency information is ignored in WD, the WPD [30] considers the high frequency information and decomposes them into different scales. The WPD considers whole packet tree of a wavelet as illustrated

in Fig. 5.2, where $U_{j,n}$ is the *n*th subspace (*n* is frequency factor) of wavelet packet at the *j*th scale, and $U_{j,k}^{n}(t) = 2^{-j/2}u^{n}(2^{-j}t - k)$ (*k* can be interpreted as time localization parameter), which are defined as:

$$u_{j,0}^{n}(t) = \sum_{k} h_0(k) u_{j-1,k}^{i} \quad (n \ \ is \ \ even) \tag{5.1}$$

$$u_{j,0}^{n}(t) = \sum_{k} h_1(k) u_{j-1,k}^{j} \quad (n \ \ is \ \ odd) \tag{5.2}$$

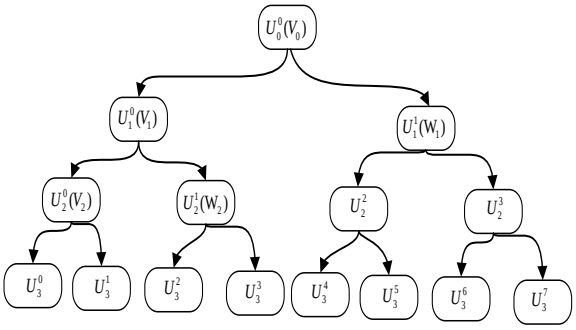

Figure 5-2: Wavelet packet decomposition structure

Where $j,k \in Z, n = 0,1,2,...,2^{j} - 1, h_0(\mathrm{k}), \mathrm{h}_1(k)$ stand as pairs of multiply mirror filters that are unrelated to scales, and described by means of:

$$h_1(k) = (-1)^{1-k} h_0(1-k) \tag{5.3}$$

The coefficient of WPD at the *j*th level and *k*th samples can be defined by quadrature mirror filters for both low and high frequency information.

5.2.2 Recurrence Quantification Analysis

Recurrence is an equity of dynamical systems that demonstrates whether the states revisit points in the phase space trajectory. Briefly, two points on the phase-space trajectory are deliberated to be recurrent if the distance between them is less than a threshold value

[131]. R is a two-dimensional squared matrix which visualize recurrence of a state at time *i* at a different time *j* with dots; both axes here are time axes.

$$R_{i,j}^{m,\varepsilon_i} = \Theta(\varepsilon_i - \|x_i - x_j\|) \quad i, j = 1...N \tag{5.4}$$

Where *N* is the number of considered states *xi*, ε_i is the threshold distance, *m* is the dimensional phase space trajectory, and $\Theta(.)$ is the Heaviside function [131].

Fig. 5.3 and Fig. 5.4 demonstrates a Recurrent Plot (RP) for a sinusoidal and random noise. RPs are specified by line-of-identity (LOI) which is the solid main diagonal. The periodicity of the signal can be shown by large diagonals repeating themselves periodically. Short diagonals are representation of non-deterministic or stochastic behavior. Stationary data leads to homogenous plot while non-stationary data leads to RP with the upper side of the plot on left and the lower side of the plot on right side disappearing.

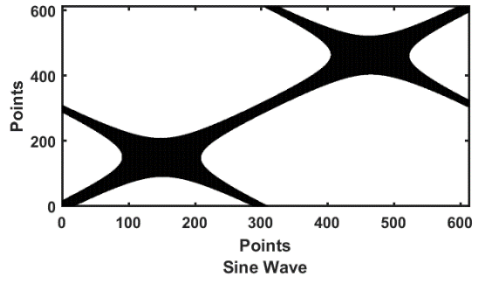

Figure 5-3: Sine wave recurrence plot

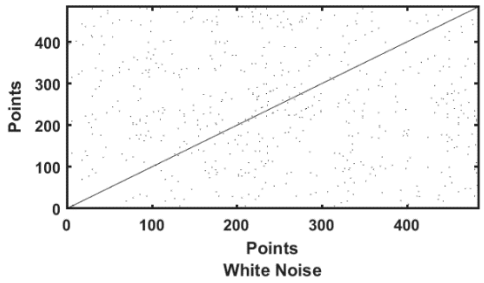

Figure 5-4: White noise recurrence plot

Recurrence Rate (RR): RR sums the total number of the black points in the RP without considering the LOI. It represents the relative frequency of recurrence points.

$$RR(\varepsilon, N) = \frac{1}{N^2 - N} \sum_{i \neq j=1}^{N} R_{i,j}^{m,e} \tag{5.5}$$

Where *N* is the length of data, and *R* is the recurrence matrix.

Determinism (DET): The portion of recurrence points which shape diagonal lines with minimum length are denoted by DET. Definition of lines in the RP are specified by d_{min} parameter which sets the lower bounds and normally is considered as 2.

$$DET = \frac{\sum_{l=d_{\min}}^{N} lH_D(l)}{\sum_{i,j=1}^{N} R_{i,j}} \tag{5.6}$$

Where $H_D(l)$ represents how many times a diagonal with length l exits.

Maximal line length in the diagonal direction (Dmax): D_{max} represents the length of longest diagonal in RP.

$$D_{\max} = \arg\max_{l} H_D(l) \tag{5.7}$$

Shanon Entropy (ENT): ENT is Shanon entropy of the frequency distribution of the diagonal line lengths. It shows the complexity of the deterministic structure in the system. The higher ENT means that the dynamics are more complex. For instance, the value of ENT is small for uncorrelated noise or oscillations, which means low complication in the dynamics of the system.

$$ENT = -\sum_{l=d_{\min}}^{N} p(l)\ln p(l) \tag{5.8}$$

Where

$$p(l) = \frac{H_D(l)}{\sum_{l=d_{\min}}^{N} H_D(l)} \tag{5.9}$$

Laminarity (LAM): LAM indicates the percentage of recurrent points in vertical structures though the fraction of recurrent points in diagonal structures are represented by DET.

$$LAM = \frac{\sum_{l=v_{\min}}^{N} lH_V(l)}{\sum_{i,j=1}^{N} R_{i,j}} \tag{5.10}$$

Trapping Time (TT): TT is the average length of vertical structures; it applies the minimum length. It indicates how long the state is trapped.

$$TT = \frac{\sum_{l=v_{\min}}^{N} vH_V(l)}{\sum_{l=v_{\min}}^{N} H_V^e(l)} \tag{5.11}$$

5.3 Proposed Methodology

The proposed method advances the performance of recurrence attributes by inspecting recurrence dynamics in multiple wavelet scales. The recurrence properties are considered in various wavelet scales and the correlation of recurrence behaviors across scales.

As illustrated in Fig. 5.2, the nonlinear and nonstationary time series is decomposed into short-term subseries in different scales by applying wavelet packets. After this stage, the recurrence dynamics are quantified in the wavelet subseries rather than the main time series. Hence, the nonstationary presence will be transformed into non-overlapped wavelet scales to simplify the characterization of nonlinear dynamics. Additionally, the subseries will make the computation of recurrence plots more effective. Lastly, feature selection and clustering will be applied to determine coherent groups of generators.

5.3.1 Feature Selection

For each of the wavelet subseries RQA measures, RR, DET, LMAX, ENT, LAM, and TT are derived based on the flowchart shown in Fig 5.5. The *kth* level packet decomposition will create a high-dimensional feature space ($6 \times 2^k = 48$). Respectively, this may cause "curse of dimensionality" for the clustering techniques. When the number of features increases overfitting problem typically happens. Here, a feature selection method is applied to optimally choose a subset of features that are correlated with variation.

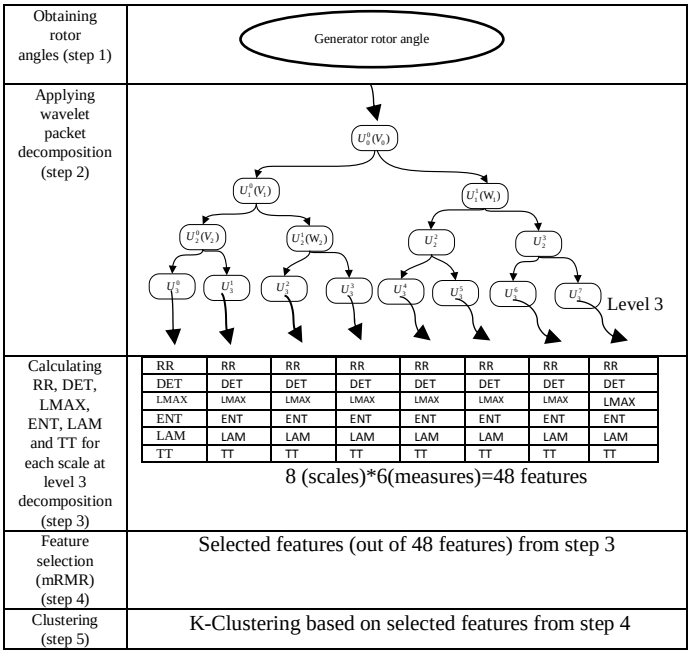

Figure 5-5: Flow chart of applied methodology

The minimum-redundancy-maximum-relevance (mRMR) is a feature selection method [132] where the mutual information between the joint distribution of the designated features and classification variable are maximized and a feature's relevancy by its repetition in the existence of the other elected features are penalized. It could be written as an optimization problem:

$$mRMR = \max_{x\in\{0,1\}^{ng}} \left[\frac{\sum_{i=1}^{ng} c_i (xg)_i}{\sum_{i=1}^{ng} (xg_i)} - \frac{\sum_{i,j=1}^{ng} a_{ij} (xg_i)(xg_j)}{(\sum_{i=1}^{ng} (xg_i))^2} \right] \tag{5.12}$$

$$c_i = I(f_i; c) \qquad a_{ij} = I(f_i, f_j)$$

Where *ng* is the full-set features, *xgi* denotes the established membership index function for feature *fi* (*xgi*=1 shows existence and *xgi*=0 shows lack of the feature *fi* in the most relevant features). The individual feature *fi* and the class *c* has mutual information and its average value is shown by *ci* and indicates the importance of a feature for that class. The redundancy of features is shown by a_{ij} and it specifies average value of mutual information between two features *fi* and f_j. The mRMR method is a supervised learning method, while the task of this research is an unsupervised learning task. The test system has 29 generators, and each generator is assumed to have a PMU installed at the terminal. Thus, PMU1, for example, is related to generator 1. So, in the mRMR method, each rotor angle is labeled the same as its associated generator. Now, to avoid overfitting, the most relevant features are selected, which are correlated with variation, meaning that the dependency between the joint distribution of the selected feature and classification variable can be maximized by mRMR. After this stage, k-clustering method will be performed on selected features to determine the coherent group of generators.

5.3.2 Clustering

For a specific contingency, coherent generators that demonstrate similar oscillatory nature should be determined. The k-means clustering method is used considering features selected by mRMR method. The k-means clustering is based on vector quantization and originated from signal processing, and it is used for cluster examination in data mining.

The purpose of k-means clustering is dividing the number of observations into a defined number of clusters, and each observation fits into the cluster with the nearest mean. It uses an iterative algorithm which has two phases that minimizes the total distances of point-to-centroid, which are summed over all defined clusters. For finding the proper number of clusters in a dataset, the elbow method [133] is used. This method considers the percentage of variance described as a function of the number of clusters.

5.4 Case Studies

In this section, the proposed method is applied to the 179-bus system shown in Fig. 5.6. This system is a reduced Western Electricity Coordinating Council (WECC) ac transmission system that includes 29 generators, 179 buses, and 263 branches. The data for 23 simulated contingencies [134] are used as test cases. For the base instance, synchronous generators are represented as classical machines with damping parameter D set to 4 for all generators except generators at buses 45 and 159. The value of damping parameter D for generators at 45 and 159 are set at -2 and 1 respectively [134]. This makes generator 45 the source of the disturbance. All loads in the system are modeled as constant MVA.

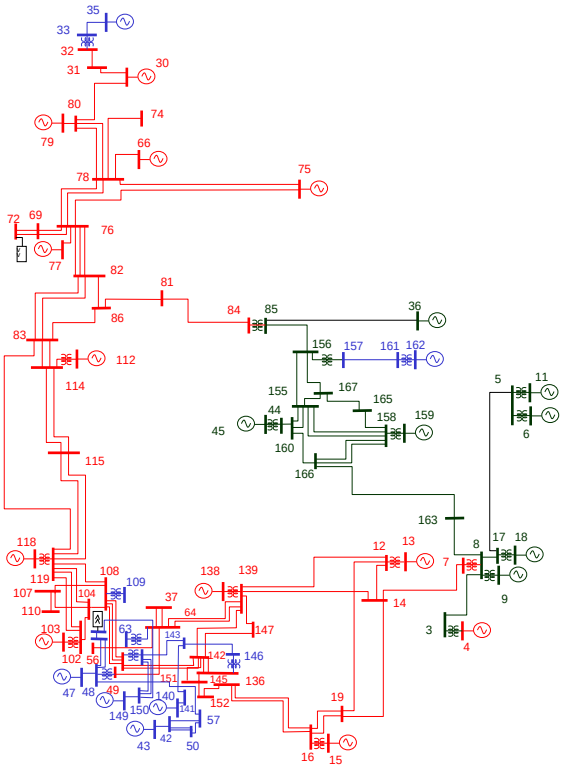

Figure 5-6: The reduced WECC 179-bus power system

The algorithm for one of the contingencies is described as follows. The results of a small signal study show a low damping 0.01% for the local mode at 1.41 Hz. This mode can be excited by creating a disturbance at bus 159. In order to meaningfully excite the 1.41 Hz mode, a three-phase short circuit is applied on bus 159 at 0.5s and cleared after 0.05s. It should be mentioned here that the generator which has the negative damping value is the source of the oscillation. The rotor angles of all generators in the system are investigated to see the impact of the disturbance. Fig. 5.7 shows plots of a subset of these rotor angles, specifically for G36, G159, G45, G162, G9, G18, G35 and G47.

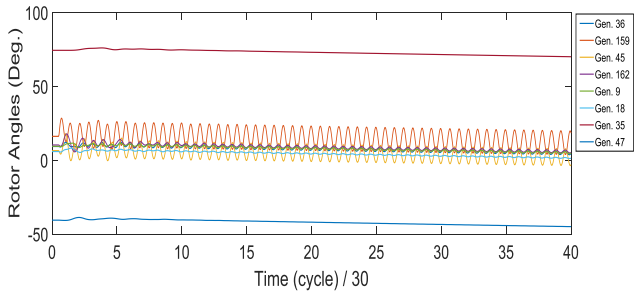

Figure 5-7: Rotor angles of some generators

Results of 23 simulated contingencies are used in the study. Generator coherency can be studied based on fluctuations of generator rotor angles after the fault is cleared. It expands further beyond the transient duration in large wide area power systems that exhibit poor damping. In order to monitor the system performance after a fault, mostly a 20-s (1200 cycles) time window is considered as acceptable [135]. Based on the IEEE recommended practice for synchrophasor measurements in power systems, a 100 Hz sampling frequency is used [136]. The time trends in Fig. 5.7 are used to create recurrent plots.

The recurrence analysis breaks down and represents the repeated dynamics of the system at distinct levels that captures the transient and steady-state behavior of the system. The rotor angle of generator 159 is examined in several wavelet scales $U_3^0, U_3^1, U_3^2, U_3^3, U_3^4, U_3^5, U_3^6$ and U_3^7. The recurrence plots for scales U_3^0 to U_3^7 are calculated from the wavelet detail series. The three wavelet scales with related recurrence plots for generator 159 are shown in Fig. 5.8. The different frequency components in the rotor angle of generator 159 are separated for the investigation of recurrence dynamics that facilitates exploring the transient and steady-state behavior of a system. All rotor angles are considered after fault clearance.

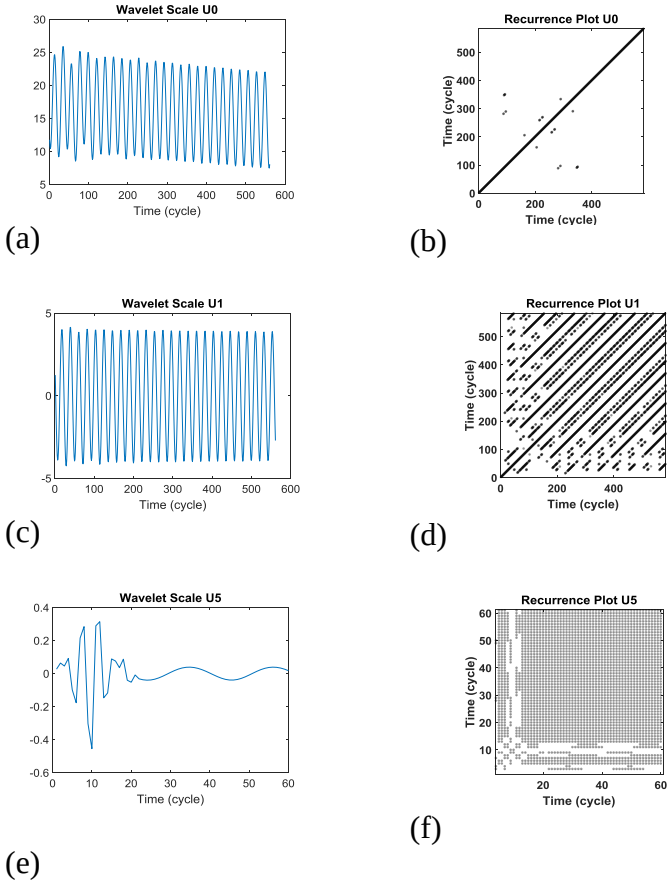

Figure 5-8: Recurrence plot of generator 159 (rotor) in wavelet scales

In this examination, the decomposition level is elected to be 8 for the rotor angles. The wavelet scale and recurrence plot of wavelet scales U_3^0, U_3^1 and U_3^5 are shown in Fig. 5.8. For wavelet scale U_3^0 recurrence plots, there is a long diagonal which shows the deterministic behavior as seen in Fig. 5.8(b). Some single points are away from the main diagonal which shows states are rare and they are present only for a short time. For wavelet scale U_3^1, the periodicity of the signal can be shown by large diagonals repeating themselves periodically as seen in Fig. 5.8(d). For wavelet scale U_3^5, the darker areas consist of several hundred dots, which shows the presence of noise, as seen in Fig. 5.8(f). In the latter figure, there is also a very short diagonal at around 10 cycles which represents the transient behavior. The

lighter area for wavelet scale U_3^5 indicates sudden changes in the dynamic behavior. So recurrence plots can give detailed quantified measures about rotor angles.

Six measures of the RQA, considering RR, DET, LMAX, ENT, LAM and TT are derived from the recurrent plots for each wavelet scale. Totally, 48 features are pulled in for recurrence examination in order to determine generator coherency indication.

Since each rotor angle has 48 features and it may lead to the "curse of dimensionality" in the clustering process, the mRMR feature selection method is applied to choose a subgroup of features that are correlated with the process variation. Additional analysis of selected feature subset will detect which measures of RQA play more significant roles in determining coherent groups of generators. The clustering error is shown in Fig. 5.9. This error can be decreased by adding optimal features into the clustering method. It should be mentioned here that when there is no noise in the measured data, there is no error in the clustering. The number of selected features by the mRMR method has direct effect on the clustering error. More details are provided in [132].

The error rate does not decrease when the number of features is greater than 16. Hence, the optimal number of selected features is considered to be 16 for the study, which also brings a consistent number of features for building the clustering model, which is shown in Table 5.1.

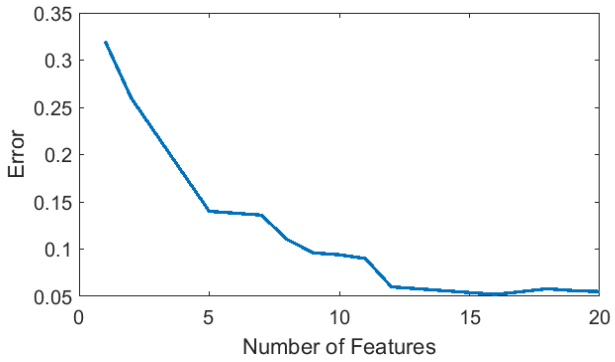

Figure 5-9: Cluster error against the number of features

The feature selection is based on the mRMR method as mentioned in section 5.3. Table 5.1 shows that for scale U_3^0, measures RR, LMAX and TT are the selected features. The black dots represent selected features for each scale. The meaning of RQA measures such as RR and DET is explained in section 5.3.2.

Table 5-1: Selected features of RQA measures

	RR	DET	LMAX	ENT	LAM	TT
U_3^0	●		●			●
U_3^1	●			●		●
U_3^2						●
U_3^3						●
U_3^4	●					●
U_3^5	●					●
U_3^6	●					●
U_3^7				●		●

In all, 23 cases are considered and for each case, the selected features are different. The number of most relevant features varies based on different cases. For example, for case 2, the number of selected features is 14 and for case 3 the number of selected features is 18.

After applying clustering, it is observed that the four generators - G36, G162, G159, G45 - create four different groups namely group 1, group 2, group 3 and group 4. Generators G9 and G18 swing together and form a separate cluster of generators, specifically group 5. The other generators G4-G149, form group 6. Table 5.2 shows the coherent areas of the test system.

Table 5-2: Coherent groups for 179 bus system

Group	Coherent Generators
1	36
2	162
3	159
4	45
5	9,18
6	4,6,70,103,138,11,77,144,148,13,79,15,112,116,118,30 ,40,35,43,140,65,47,149

The six measures of RQA at level 3 scale for some of the generators are shown in Table 5.3.

Table 5-3: Measures of RQA at level 3 for scale U0

Generator	RR	DET	LMAX	ENT	LAM	TT
36	0.31	93.82	128	3.84	0	0
30	9.17	99.68	1132	5.61	99.40	24.32
40	8.78	99.69	1132	5.23	99.74	27.98
35	10.1	99.68	1132	5.88	99.77	28.14
159	0.00	53.33	5	1	0	0
9	1.95	95.60	1121	3.47	16.12	3.00
45	0.00	50	4	1	0	0
162	0.39	89.60	139	3.42	0.53	3.5
18	1.91	95.782	1122	3.55	69.46	3.89

According to Table 5.1, the measures RR, LMAX and TT play important roles for this scale, in determining the coherent groups of generators. Generators G30, G40 and G35 are in the same group and their RR value are close to each other. LMAX is the same for all of

them, and trapping times are very close to each other in comparison to other coherent groups.

The generators G9 and G18 are in the same coherent group, as their RR are close to each other, the LMAX is almost same for both generators, and trapping times are close in comparison to other coherent groups. The four generators G36, G159, G45 and G162 each form distinct groups, and their RR values are zero or close to zero, LMAX is different for each of them, and the trapping time is zero except for generator G162.

5.5 Comparison with other Methods

The accuracy of the determination of coherent generators, achieved by applying the RQA technique, is examined in contrast with the PCA method [69] with and without noise in the measurement on the same test system.

5.5.1 Comparison without noise

The PCA method [69] is applied to detect coherent groups of generators when there is no noise, and results reveal that the PCA method identified the coherent groups of generators in 6 clusters. This indicates that, applying both the proposed RQA method and the PCA method in this paper reveal similar coherent groups of generators.

5.5.2 Comparison under noise

The Signal to Noise ratio (SNR) compares the level of a signal such as voltage to the noise level. The performance of methods based on measurements can be affected by the presence of noise in the acquired data. The performance of the RQA in the presence of noise is evaluated against the PCA. For this investigation, identical datasets are applied by considering white noise with varying levels of SNR ranging from 20 to 100.

White Gaussian noise is added to rotor angles using Matlab. The trajectories of rotor angles are the same except that noise is added, and it is independent and identically distributed from a zero-mean normal distribution with variance N. Different signal to noise ratios are considered such as SNR=80 dB and SNR=20 dB. After fault clearance, the two methods - the RQA and the PCA - are applied to detect coherent groups of generators.

For both methods, the rotor angles of all generators of the test system are obtained at each cycle for 20s after fault clearance. A three-phase short circuit is created at bus 159 at 0.5s and cleared after 0.05s. The results achieved by means of the RQA technique is shown in Fig. 5.10 and the coherent groups identified applying the PCA method are shown in Fig. 5.11, both for SNR=20.

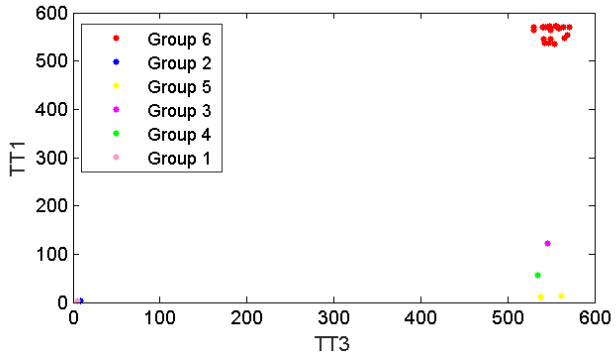

Figure 5-10: RQA method (Comparison of coherency methods)

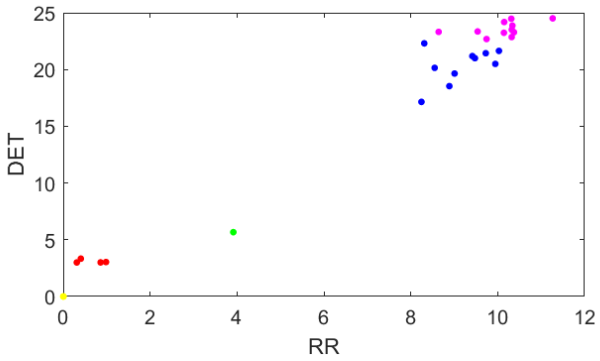

Figure 5-11: Comparison of coherency methods (PCA method)

From Fig. 5.10, it is clear that the power system comprised of 6 coherent groups as determined by the RQA method. A group of coherent generators (group5) is formed by generators G09 and G18. TT1 and TT3 are the trapping times at scales 1 and 3 at level three of the wavelet packet decomposition. Generators G4, G6, G70, G103, G138, G11, G77, G144, G148, G13, G79, G15, G112, G116, G118 form group 6. The other four generators, G36, G162, G159, and G45 each forms a distinct coherent group, which are respectively, groups 1, 2, 3 and 4.

Fig. 5.11 illustrates the coherency of generators achieved by applying the PCA technique when the SNR is considered to be 20. With an SNR of 20, the PCA method was not able to determine coherent groups correctly. However, when the SNR was 60, the

method yielded more accurate classification when compared with an SNR of 20. The RQA method, on the other hand, could detect coherency even when the SNR was 20.

When the SNR is less than 20, the RQA method is unable to detect coherent generators with good accuracy. This shows that the presence of noise in the measured rotor angles has an important role in accurately determining the outcome of techniques based on measurement.

5.5.3 Robust PCA

One of the most widely used statistical tool [137] for data analysis and dimensionality reduction is the PCA. However, it is sensitive to noisy observations. For a given matrix the PCA tries to find the best L2-norm low-rank approximation of that matrix. Nonetheless, the L2-norm is sensitive to outliers and noise, which is often observed in measurement data.

Hence, much attention has been drawn to the robust PCA. The robust PCA has been introduced in [138] and it decomposes a given data matrix into a low-rank matrix and a sparse matrix. It is represented by M - the original data matrix, L the low-rank component and N the sparse component. The robust PCA can be formulated by convex optimization problem as follow:

$$\min_{L,E} \|L\|_* + \lambda \|N\|_1 \tag{5.13}$$

$$\text{subject to } M=L+N$$

Where $\|L\|_*$ represents the nuclear norm of L, $\|N\|_1$ represents the L_1-norm of N. It has been proven that L and N can be recovered from M with high probability under certain noise sparsity. The robust PCA has certain limitations. As indicated in equation 5.13, it

uses L_1-norm to describe N and it is only optimal for Laplacian noise. While the, L_1-norm can better fit sparse noise than the L_2-norm, the real noise has much more complex structures in statistical terms, and is often not Gaussian or Laplacian.

5.5.4 Coherency detection under noise (Robust PCA)

The performance of the robust PCA in the presence of noise is assessed against the PCA. For this investigation, identical datasets as in the previous section are applied by considering white noise with varying levels of SNR.

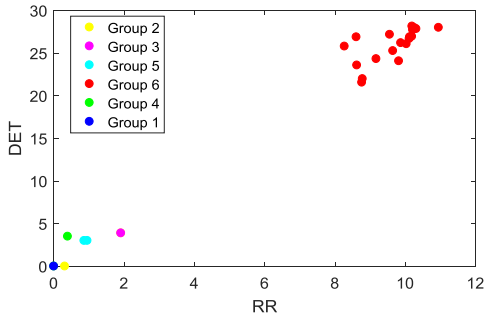

Figure 5-12: Comparison of coherency methods (robust PCA)

The robust PCA is applied on the rotor angles when the SNR is considered to be 40 and the method was able to identify coherent groups correctly as shown in Fig. 5.12. For SNRs in the range of 40-100, the accuracy of the robust PCA is investigated and results show good accuracy, which means the method can determine the coherent generators in six groups like the proposed method which applies the WPD and the RQA. Group 6 is represented by red color, Group 5 which includes generator 9 and 18 is shown by light blue, and the other four distinct groups are illustrated by green, blue, yellow and pink. However, when the SNR was 20, the method produced less accurate classification when

compared with SNR 40-100. The RQA method is capable to detect coherency even when the SNR was 20.

5.5.5 Empirical Mode Decomposition

The Empirical Mode Decomposition (EMD) is a signal processing method introduced in [137]. This method is suitable for nonlinear and nonstationary signal processing and it is widely used in different applications such as process control and biology.

This method identifies a baseline signal within a distorted waveform, and then extracts it from the main waveform. The mean of an upper envelope and a lower envelope determines the baseline signal. By interpolating between the local maxima, the upper envelope is formed, and by interpolating the local minima, the lower envelope is attained. The original distorted signal is continually examined till no further baseline signals can be identified. This new attained signal is called an IMF. The IMFs have two features: a) the number of local minima equals (or varies by one) from the number of local maxima. b) All maxima are positive, and all minima are negative. The following steps show the process for obtaining IMF from a distorted waveform $d_s(t)$.

1) Determine local maxima and minima of distorted signal, $d_s(t)$.

2) To obtain upper $e_u(t)$ and lower envelopes $e_l(t)$, implement cubic spline interpolation between the maxima and the minima.

3) Calculate the mean of the two envelopes.

$$e_m(t) = (e_u(t) + e_l(t))/2 \tag{5.13}$$

4) Extract

$$imf(t)=d_s(t)-e_m(t) \tag{5.14}$$

5) If the number of local extrema of *imf(t)* is equal to, or differs from the number of zero crossings by one, and also the average of *imf(t)* is zero, then *imf(t)* is an IMF. Otherwise, steps 1-4 should be repeated, and $d_s(t)$ should be replaced by *imf(t)*, until the conditions of being an IMF is met by the new *imf(t)*.

6) Calculate the residue:

$$r(t)=d_s(t)-imf(t) \tag{5.15}$$

To prevent "over-improving" the IMF, which can cause loss of modal information, proper stopping criteria should be used in step 5.

When all IMFs are extracted from a distorted waveform, the final residue has one or zero extremum points. The original signal can be obtained as follows:

$$d_s(t) = \sum_n imf(t) + r(t) \tag{5.16}$$

5.5.6 Coherency detection under noise

The performance of the WPD in the presence of noise is evaluated against the EMD, For this examination, identical datasets are used by considering white noise with varying levels of SNR which were applied in the previous section.

The EMD method is applied on rotor angles and all IMFs extracted; then the measures of the RQA are applied on the IMFs. Since each rotor angle is decomposed into three IMFs and 6 measures of the RQA are applied on these three IMFs, the total number of features for each rotor angle would be 18 (3x6=18). To avoid the overfitting problem, a feature reduction method is applied, and then K-clustering method is used to identify the coherent

groups of generators. The EMD of the rotor angles are processed to give 3 IMFs for each rotor angle. For example, three IMFs of rotor angle of generator 9 are shown in Fig. 5.12. The results achieved by means of the EMD technique are shown in Fig. 5.13 for SNR=50.

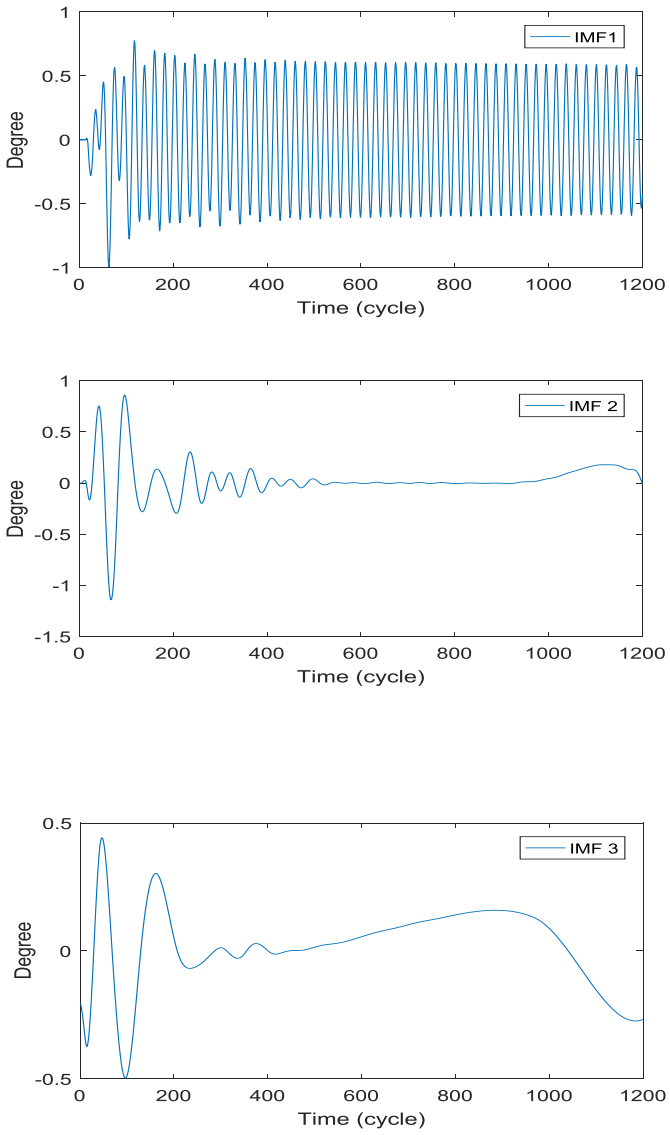

Figure 5-13: IMFs extracted from the rotor angle of gen. 9

Fig. 5.13 shows the coherency of generators obtained by applying the EMD technique when the SNR is 50. Results show that under the noise condition, the EMD method cannot perform accurately when compared with the WPD method.

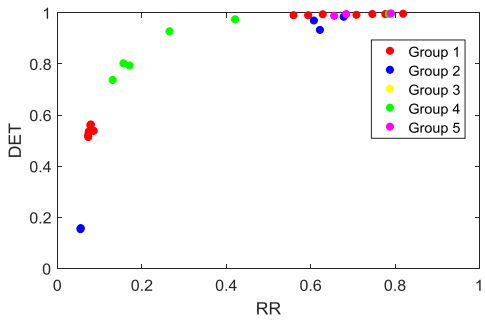

Figure 5-14: Comparison of coherency methods (EMD)

Fig. 5.13 illustrates that the power system contains 5 coherent groups as specified by the EMD method. These groups are identified in Table 5.4. and when these results are compared with the WPD method (Table 5.3), it is clear that by applying the EMD in the proposed method, accurate result cannot be obtained.

Table 5-4: Coherent group for 179 bus system (EMD)

Group	Coherent Generators
1	36,40,35,4,103,159,11,43,45,140,47,112,18,116
2	138,77,144,79,149
3	9,148,65,162,13,118
4	30,70,15
5	6

5.6 Conclusion

The RQA method based on data acquisition is proposed to detect generator coherency in a power system. The rotor angles decomposed by wavelet decomposition and the RQA measures are used to capture characteristics of rotor angles, and then coherent groups of generators are detected by applying a clustering technique.

A wavelet transform decomposes the nonstationary time series into different frequency bands to efficiently keep apart the system transient and steady-state behaviors. Additionally, wavelet subsampling facilitates the computation of recurrence plots from the

time series data. Further, recurrence analysis becomes more compelling in the multiple wavelet scales under a stationary premise.

The results obtained by the RQA method are compared with results obtained by the PCA and the robust PCA methods. When there is noise in the measured data, the PCA method is unable to accurately detect the coherent group of generators. The robust PCA has better performance in the presence of noise when the SNR is in range 40-100. However, when the SNR is set to 20, the accuracy of the method decreases. Since the RQA method investigates recurrence dynamics in multiple wavelet scales, and advances the performance of recurrence characteristics, it can determine the coherent groups of generators precisely even in the presence of noise.

To evaluate the performance of the proposed method the WPD was replaced by the EMD. When the EMD was used to decompose the rotor angles, the results showed that the algorithm can not detect coherent group of generators accurately. This indicates that the WPD method has better computing efficiency than the EMD method.

Chapter 6 : DATA-DRIVEN DETECTION OF SUSTAINED OSCILLATIONS AND FREQUENCY MODE IN POWER SYSTEMS USING RQA

6.1 Introduction

Different events may cause oscillations in power systems and the problem of low-frequency oscillation in 0.1-2 Hz range is observed in many power systems. The instability could happen if the oscillation is not damped adequately. Therefore, it is necessary to investigate the low-frequency oscillation modes for dynamic system security. Rich information are obtainable by ringdown portion of the signal about the low-frequency oscillatory modes and can be used for this purpose.

A few linear analysis methods have been considered to exactly examine inter-area oscillations. The Kalman filter and the Prony analysis methods are implemented in national grids [137]. The Prony analysis technique considers a signal to identify low frequency. Combination of WAMS with the Kalman filter is used in [138] to determine poorly damped oscillations. In [139] an enhanced the Prony analysis is introduced for detection of small frequency oscillations. For the parameter detection of low-frequency oscillation the Discrete Fourier Transform is applied in [140]. Identification of low-frequency oscillation is investigated by applying wavelet method and result give better understanding for low frequency approximation from PMU data, still the instantaneous frequencies cannot be identified from the contour plot.

In this chapter, we propose a new method to detect low-frequency oscillation and identify frequency modes. Chapter focuses on the SFA and RQA method. The main

contribution of this chapter is the theoretical and application of the proposed algorithm in the power system. The proposed method is evaluated by using different created cases on the WECC 179 bus system.

This chapter is organized as follows: Section 6.2 introduces the SFA and the RQA methods. The proposed method is presented in Section 6.3. Case studies and simulation results are presented in section 6.4. Conclusions are discussed in section 6.5.

6.2 Review of SFA and RQA

Non-linear dynamics for investigating the unobserved patterns have enhanced recently and yields a convenient opportunity to realize non-linear behavior of the power systems. The SFA is a method for getting features that change slowly from a rapidly altering signal. A single underlying driving force with high precision can be projected by applying SFA to nonstationary time series such as frequencies of generators in the power system. For example, an effective slowly varying features study for nonstationary data can be shaped by SFA. The RQA was shown as a useful method to investigate dynamics of a system, and can provide a quantitative description for the nonlinear and nonstationary signals [141]. By applying the SFA and the RQA on the acquired frequencies of generators, first slowly varying features will be extracted; then related slowly varying features are quantified.

6.2.1 Slow Feature Analysis

The main purpose of the SFA is to get slowly varying features from a signal that varies swiftly. If $x(t)$ is considered as input signal with M dimension where t indicates time, the objective is to try to find a function $g(x)$ that creates an output signal $y(t)= g(x(t))$ that varies very slowly by considering the following optimization problem:

$$\min \sum_{i=1}^{J} \Phi[\dot{g}_j^2(x(t))] \tag{6.1}$$

$$s.t. \quad \Phi[g_j(x(t))] = 0, \quad \Phi[g_j^2(x(t))] = 1 \tag{6.2}$$

$$\Phi[g_i(x(t))g_i(x(t))] = 0 \tag{6.3}$$

$$i = 1, ..., J, \qquad j \neq i \tag{6.4}$$

Where the first order derivative of $g_j\ (x(t))$ is shown by $\dot{g}$ and Φ represents the sample mean over time. The objective is to decrease the temporal variation of $y_j(t)$ by minimizing the objective function. The output $y_j(t)$ should have zero mean and unit variance which is shown by constraint (2). The constraint (3) indicates that J outputs $y_1(t)$ through $y_j\ (t)$ should be mutually uncorrelated. If g_j is linear in $x(t)$, then the objective function has a closed form solution such as $g_j(x(t)) = w_j^T\ x(t)$ where w_j is a weighting vector. Then, SFA can be solved as a generalized eigenvalue problem [89]:

$$\Phi[\dot{x}(t)\,\dot{x}(t)^T]W = \Phi\,[x(t)\,x(t)^T]W\,E \tag{6.5}$$

Where the weighting vectors are the same as the generalized eigenvectors and E is a diagonal matrix of the generalized eigenvalues. The eigenvalues specify the order of slow features and the lowest index represents the most slowly varying signal. Reference [89] provides more details. The SFA is not sensitive to noise since it focuses on small eigenvalues after sphering and determining the time derivative; so it rejects small but rapidly changing noise elements that naturally have large eigenvalues. It is necessary to mention these eigenvalues are different from the eigenvalues of the power system and related to the SFA optimization problem.

The SFA comprises of four steps: 1) use some fixed nonlinear functions to expand the input signal 2) To achieve components with zero mean and unit covariance matrix, sphere the expanded signal 3) calculate the derivative of sphered expanded signal to identify the normalized eigenvector of its covariance matrix 4) In order to get the output signal, project the sphered expanded signal onto this eigenvector.

Fig. 6.1 shows that frequency of generators is considered as time varying input signals *x(t)*. Non-linear functions *g(x)* transform the input signals to slowly-varying signals *y(t)*. Left side of the figure shows the generators frequency and the right-side demonstrations slowly-varying output signals. It is necessary to mention that since the transformations must be instantaneous, solution such as low-pass filtering is not possible.

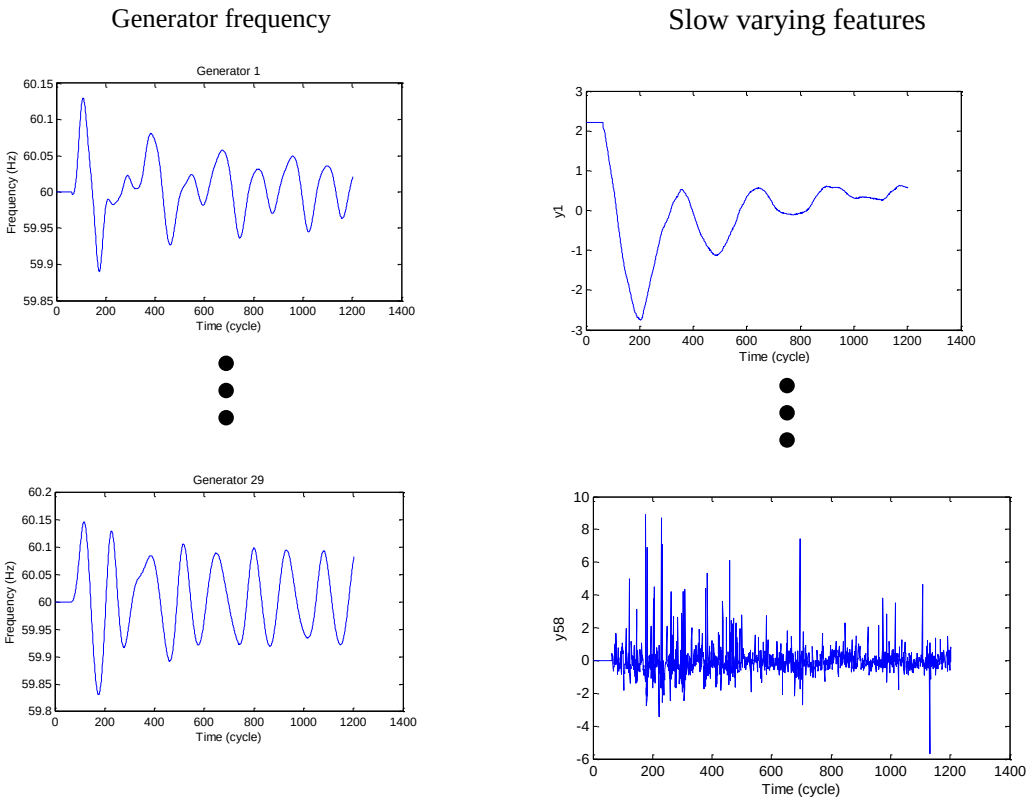

Figure 6-1: Schematics of the optimization problem solved by slow feature analysis

6.2.2 Recurrence Quantification Analysis

A vital property of a deterministic dynamical system is the recurrence of states [131]. Recurrence Plots (RP) introduced in [131] can explain the recurrence property of a deterministic dynamical system, and the main step of RP is to calculate the subsequent matrix:

$$R_{i,j}^{m,\varepsilon_i} = \Theta(\varepsilon_i - \|x_i - x_j\|) \quad i, j = 1...N \tag{6.6}$$

Where N is the number of considered states x_i, threshold distance is shown by ε, m is a dimensional phase space trajectory, and $\Theta(.)$ is the Heaviside function. The main advantage of RP is its application to analyze nonstationary data, such as power system data, and RP can explain the attributes of both large and small-scale patterns of a dynamical system [131].

Recurrence rate (RR): The RR is one of the measures of RQA that counts the black dots in the RP without considering the line of identity which is a solid main diagonal [131]. It is a measure of the recurrence point's density. Further details are explained in [131].

$$RR(\varepsilon, N) = \frac{1}{N^2 - N} \sum_{i \neq j=1}^{N} R_{i,j}^{m,e} \tag{6.7}$$

Where N is the length of data and R is the recurrence matrix. The probability of occurrence of the same states in a system is shown by the recurrence rate. This measure responds to the correlation sum.

The power system is considered as a discrete-time dynamical system, where the sampling rate of PMUs are measured as time points. Fig. 6.2 shows the frequency of a

generator connected to a transmission system, when at time 0.5 seconds (30 cycles), a three-phase fault occurs at its terminals, and then is cleared after 0.05 second (33 cycles). Low-frequency oscillations are observed after the fault clearance. Fig. 6.3 illustrates the RR plot, when the system exhibits low-frequency oscillations. From the plot, the RR decreases from 1.0 to 0.1. In the steady state condition, the value of RR is equal to 1.0.

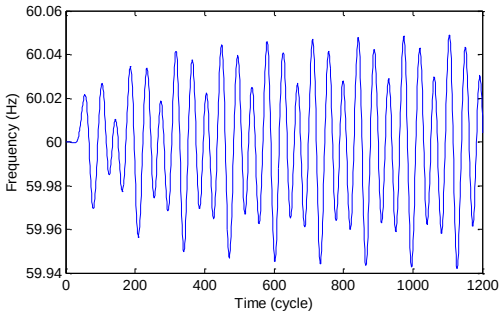

Figure 6-2: Generator Frequency

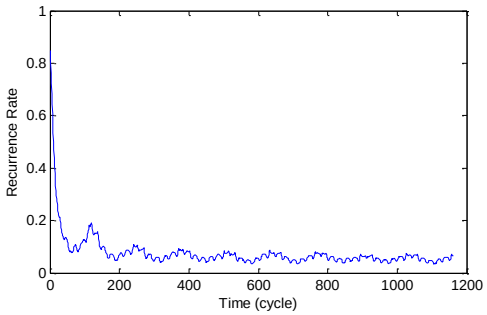

Figure 6-3: Recurrence rate of generator frequency

In case of high damping the system returns to the steady state condition. The value of RR starts to increase from zero after fault clearance and would eventually reach 1.0 after a short period. Details are discussed in section 6.4.

Fig. 6.4 shows the RR plot of the generator frequency for the same fault condition: a three-phase fault occurs at time 0.5 second (30 cycles), and then cleared after 0.05 second (33 cycles). When the fault occurs, the RR decreases from 1.0 to almost zero almost immediately. After the clearance of the fault, the frequency returns close to 60 Hz at t=13 seconds (780 cycles), while the RR increases from zero to 1.0, and the system returns to steady-state condition.

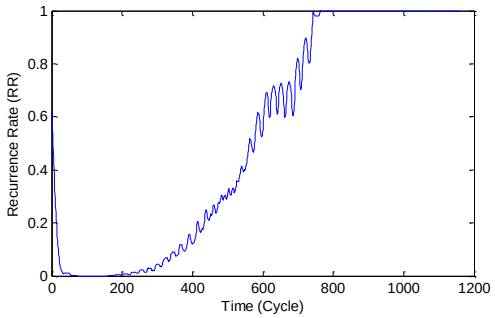

Figure 6-4: Recurrence rate for frequency of a generator

6.3 Proposed Methodology

6.3.1 Low-Frequency oscillations detection

For detecting low-frequency oscillations, we propose an algorithm in this section. The flowchart is presented in Fig. 6.5. For offline learning, historical phenomena from PMU data is considered to form X_{off}, which is the frequency of generators, and then SFA is carried out on X_{off} to extract the slowly varying features.

When the slowly varying features have been determined by the SFA method, the question that should be addressed is which components represent the oscillation sources. The pattern of extracted features should be similar to the sources of oscillations. Considering that slowly varying components are linear combinations of monitored generator frequencies, the following similarity index [142] is applied:

$$SI(i,j) = \hat{u}_j^T u_j / (u_j^T u_j) \ , \ \hat{u}_j = \overline{X}_i (\overline{X}_i^T \overline{X}_i)^{-1} \overline{X}_i^T u_j \tag{6.8}$$

Where slowly varying j is denoted by u_i, the frequency of generators is shown by $\overline{X_i}$. The similarity index SI (i,j) has a value between 0 and 1. If the generator is disturbed and the source of disturbance is u_j, then u_j can be described by to some extent, and has an

obvious connection with $\overline{X_i}$. The strong connection will be near to 1, and weak connection will be near to zero.

The method of specifying the sources of oscillation is as follows: a) Obtaining the frequency of generators and forming matrix X. b) Extracting slowly varying components by applying the SFA. c) Applying cross-correlation analysis to eliminate the slowly varying components which have same periods in comparison with previous ones. d) Calculating the similarity index (SI) for all slowly varying components, the component which has the highest values is the probable source of oscillation.

Then, the RR is calculated by applying the selected slow varying feature, and two thresholds for detection of low-frequency oscillations are specified. When the system is in the steady state condition, the value of RR is equal to one (first threshold). After the occurrence of a phenomenon the value decreases. The value of |RR| is calculated for three consecutive window sizes dt=120, and the difference between the first and the second window sizes, as well as the second and the third window sizes is determined, the higher value will be considered as the second threshold. For online monitoring, the SFA method is applied on real-time PMU data on X, and then the RR is calculated using the current selected slow varying features. If the value of RR is less than 1 (first threshold), then |d(RR)/dt| for window size dt=120 is calculated. These values are stored for 3 window sizes. If the changes for these three values from one window size to the next window size are less than 0.02 (threshold 2), then low-frequency oscillation is detected. Otherwise, the algorithm indicates steady-state behavior in the system. A modal analysis can be initiated at time t, else X will be updated and iterated.

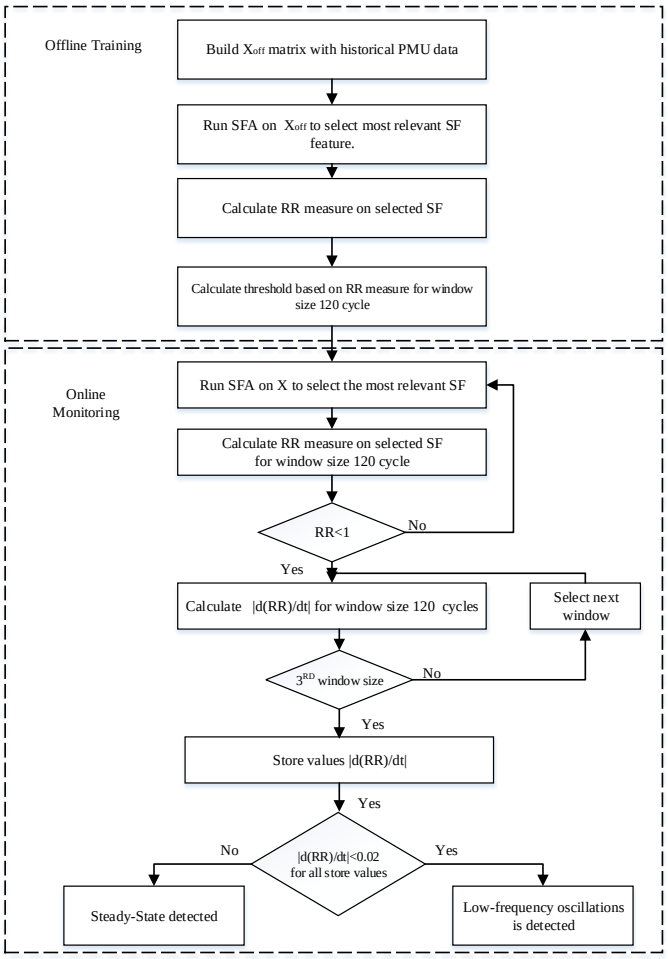

Figure 6-5: Flow chart of Low-frequency oscillation

The mathematical explanation for RR is as follows: the probability of a recurrence plot (RP) can be obtained by [143]:

$$P_b(\xi)=\lim_{N\to\infty}\frac{1}{N^2}\sum_{i,j=1}^{N}R_{i,j} \tag{6.9}$$

For a straight line with magnitude of "b" the probability distribution function is

$$\rho(x)=\begin{cases}1, & if \quad x=b\\ 0 & else.\end{cases} \tag{6.10}$$

A recurrent point in the RP has the probability of

$$P_b(\xi)=\int_{-\xi}^{\xi} R(x).d(x)=2\int_{0}^{\xi} \rho(x)*\rho(x).d(x) \tag{6.11}$$

Where $P_b(\xi)$ is the probability of a recurrence plot, $\rho(x)$ is the probability density function and *R(x)* is the distance distribution by applying the self-convolution on $\rho(x)$. Two points are called recurrent if the distance between them is less than a threshold ξ. All the points in the RP will be recurrent when the value of RR=1. The recurrence rate is determined by the probability density function of the generator frequency and threshold ξ. Since RR calculation is based on convolution integral of probability density function in the range of ξ to -ξ, when the density function decreases, the RR decreases, and when the density function increases, the RR increases consequently.

6.3.2 Recurrence-derived FT

The Wiener-Khinchin theorem mentions that the Fourier Transform of the autocovariance function is the power spectrum. One way of getting autocovariance function could be applying recurrence quantification. The probable advantages of applying a recurrence-derived FT [144] are use of the embedding theorem to catch dynamics of higher dimensional spaces, as well as its relaxed assumptions of stationary and non-linearity. Therefore, periodicities are not determined in the regular FT periodogram, or the autocovariance- derived Fourier Transform will be detected in this method. The magnitude squared of the signal's Fourier Transform represents the power spectrum of a deterministic, discrete-time signal, *x(i):*

$$S_x(\omega) = \frac{1}{N}\left|\sum_{i=0}^{N-1} x(i)e^{-j\omega i}\right|^2 \tag{6.12}$$

The power spectrum of a signal is equal to the Fourier Transform of the autocovariance function *Cx* of the signal when the Wiener-Khinchin theorem is applied:

$$S_x(\omega) = \sum_{\tau=-\infty}^{\infty} C_x(\tau)e^{-j\omega\tau} \tag{6.13}$$

Where the autocovariance function of a stochastic time series *x(n)* is described as:

$$C_x(\tau) = \frac{1}{N}\sum_{i=0}^{N-1-\tau} x(i)x^*(i+\tau) \tag{6.14}$$

The distance matrix D between all data points is the base for calculation of a RP from phase space vectors $\vec{x} \in R^m$ (*m* the dimension of the system):

$$D(i,j) = \left\|\vec{x}(i) - \vec{x}(j)\right\| \quad \vec{x} \in R \tag{6.15}$$

The average of the distance values $d(\tau)$ for a specific τ described as:

$$d(\tau) = \frac{1}{N}\sum_i D(i, i+\tau) \tag{6.16}$$

It is discussed in [144] that by applying a threshold to distance matrix D, the matrix will be limited to periodic orbits:

$$R = \Theta(\varepsilon - D) \tag{6.17}$$

Where R is the recurrence matrix. At that point, it is considered that the probability the system recurs after time:

$$RR(\tau) = \frac{1}{N-\tau}\sum_{i=1}^{N-\tau} R(i, i+\tau) \tag{6.18}$$

and replace $C_x(\tau)$ in equation (6.13) by $RR(\tau)$. Actually, the expectation values are replaced here. More details are discussed in [144].

$$E\left\{\Theta(\varepsilon - \|\vec{x}(i) - \vec{x}(i+\tau)\|\right\} \tag{6.19}$$

6.4 Case Study

The proposed method is applied on the 179-bus system illustrated in Fig. 6.6 This is the reduced Western Electricity Coordination Council (WECC) ac transmission system that consists of 29 generators, 179 buses, and 263 branches. There are 8 simulated contingency cases, which are created by Powerworld software, and applied as test cases. These cases include poorly damped oscillations, which include single sources, and local or inter-area modes. The threshold 2 for rate of change RR is specified 0.02 based on an analysis of these phenomena.

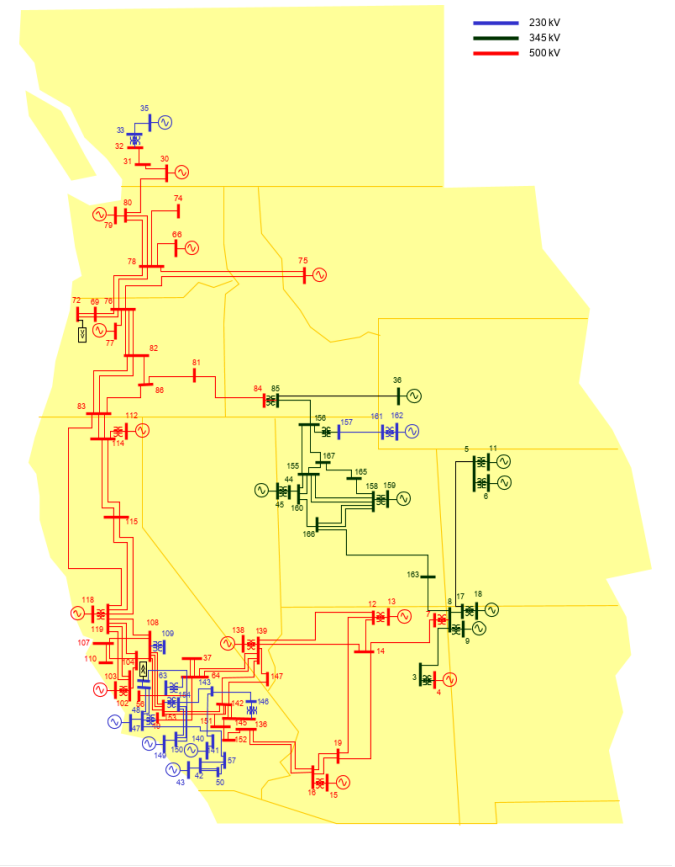

Figure 6-6: A 179-bus power system

The classical model is considered for all generators. A value of 4 is selected as the damping parameter for all generators, except generators 65 and 30. All loads are modeled as constant MVA. The algorithm for one of the contingencies is described as follows: The value of damping parameter D for generator 65 is considered to be -1.5 and for generator 30 set as 0.5; so generator 65 would be the source of the oscillation. The outcome of a small signal study indicates a low damping ratio for the inter-area mode at 0.73 Hz. There is presence of a large excitability at bus 30 for this mode. A three-phase short circuit is applied on bus 30 at 0.5s and cleared after 0.05s to implicitly excite the 0.73 Hz. The source of oscillation is the generator with the negative damping value.

6.4.1 Synchrophasor data dimension reduction

The effectiveness decreasing dimension for synchrophasor data will be shown in this section. Assuming that for each generator, a PMU is installed, and the sampling rate is 60 Hz. Also, noise is added to the data with signal-to-noise ratio (SNR) of 92 dB. The two slowest features *f1-f2* are illustrated in Fig. 6.7, and a maximum similarity index that shows the *jth* slowly changing feature is the result of source influence, and is more

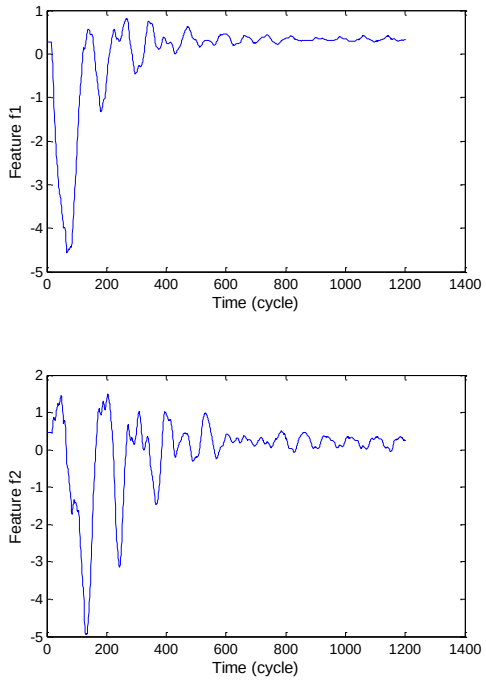

Figure 6-7: Slowly varying component

probable to be the source of oscillation as shown in Fig. 6.8. It can be determined that f_4 has the most similarity with the behavior of the source generator than the rest of the features. For the generator frequency ω, the similarity index calculated from SFA is 96.24%.

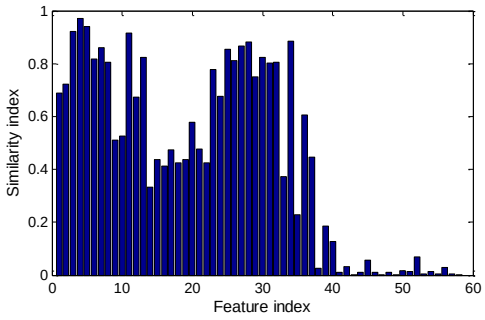

Figure 6-8: Similarity index of 58 features

Comparison between the most relevant varying component f_4 (red) and generator 65 (blue) normalized frequency is shown in Fig. 6.9. It can be discovered the source signal can be represented by the slowly varying signals. The latter are a good estimation of source signals.

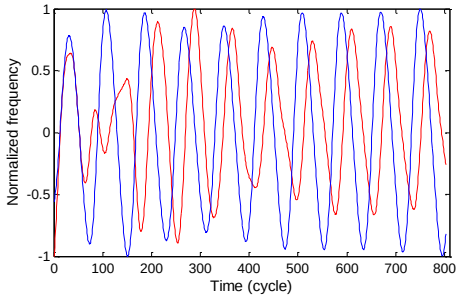

Figure 6-9: Normalized generator frequency vs slowly varying feature

6.4.2 Recurrence Quantification Analysis

The initial stage in the RQA method is to reconstruct the phase space trajectory with *m*-dimension. The embedding dimension *m* and delay time τ of the selected feature f_4 are determined by using the mutual information (MI) and Cao's method [145]. By applying the MI method, the delay time was chosen to be 4 at the first local minimum to reconstruct the phase space. For specifying the embedding dimension, the Cao method was used, and more details are explained in [145]. The embedding dimension does not have major

changes at *m*=3, and so, the minimum embedding dimension was selected as 3. Fig. 6.10 shows the RR of the most relevant feature f_4.

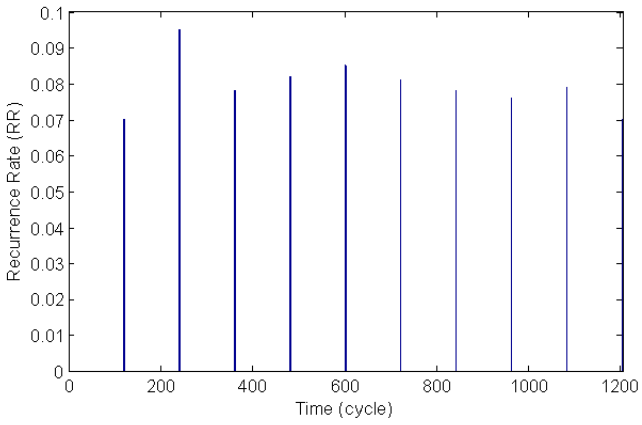

Figure 6-10: Recurrence rate for feature f_{24} (dt=120)

It is seen that for every 120 cycle (dt=120), the change in RR is less than 0.02 as time elapses, and it is a representation of oscillations. If RR increases continuously at every dt, it is an indication of steady-state in the power system. An RR equal to one indicates a steady state condition. Fig. 6.10 shows the RR values are not the same, and the reason is that in Fig. 6.9, the magnitude of the slowly varying feature, which is shown in orange, is not constant and varies with time.

6.4.3 Frequency mode detection

The signals attained from the simulated cases have two zones that are the ambient zone which is related to the pre-disturbance zone and the ring down zone, which is related to after-fault conditions. The ringdown portions are considered for analysis in this paper because it gives richer information about the low-frequency oscillatory modes. Detrending is a statistical operation of removing the trend from the signal and considered for pre-processing of the PMU signals.

The first case has one lightly damped local mode of 0.73 Hz, and the source of oscillation is the generator at bus 65. A three-phase short circuit is created at bus 30 at 0.5s and cleared at 0.55s. Fig. 6.11 shows the active power in line 84-99.

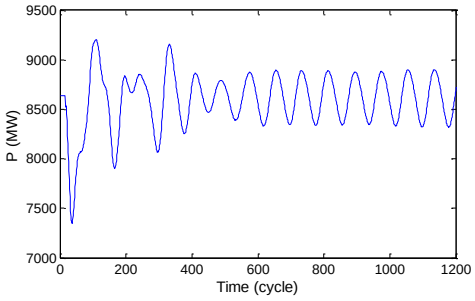

Figure 6-11: Active power in transmission line 84-99

To specify if the recurrence-derived FT can identify frequency modes, the active power in transmission line 84-99 was considered. The embedding (dimension 3, and the delay 4), and a recurrence threshold ε of 0.1 is applied. Also, for comparison, a FT spectrum is calculated. Fig. 6.12 shows the FT spectrum and Fig. 6.13 shows the recurrence-derived FT.

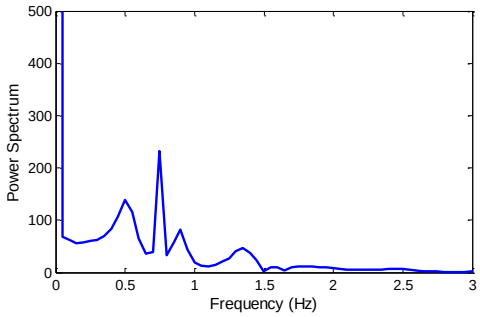

Figure 6-12: FT Spectrum

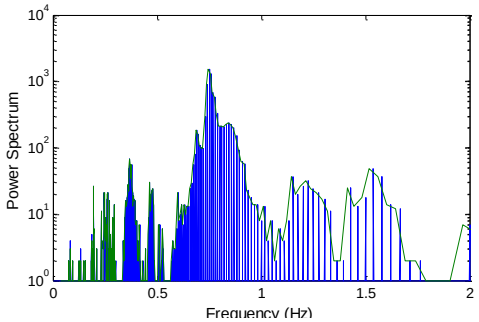

Figure 6-13: Recurrence-derived FT spectrum

The FT method detects a frequency mode of 0.73 Hz. The recurrence-derived FT is able to specify a frequency mode of 0.73 Hz, and some other frequencies in the signal

which are not detectable by the FT method. These frequencies are in the range 0.1- 1.70 Hz.

The second case has a single source, one local mode and two poorly damped inter-area modes. The source of oscillation is generator 11. A three-phase short circuit is created at bus 79 at 0.5s and cleared after 0.05s. Fig. 6.14 shows the active power in line 84-89.

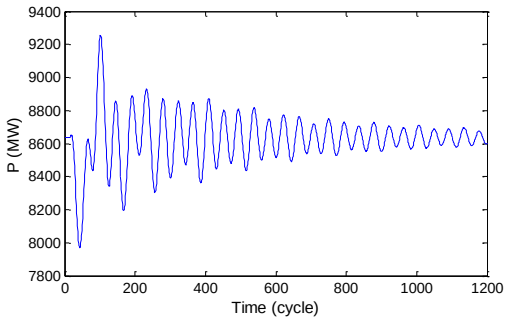

Figure 6-14: Active power in transmission line 84-99

The Fourier method detects one local and two inter-area modes when applied on active power transmission line 81-84. Fig. 6.15 shows the frequency modes. The recurrence-derived FT applied on the same line to determine frequency modes. The local mode at 1.42 Hz is detected and also inter-area modes of 0.88 Hz and 0.52 Hz are identified. Fig. 6.16 indicates that results are consistent with the FT method.

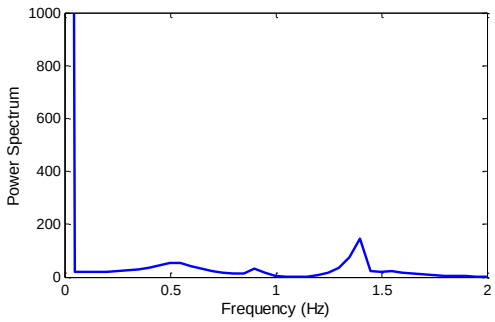

Figure 6-15: FT Spectrum (line 81-84)

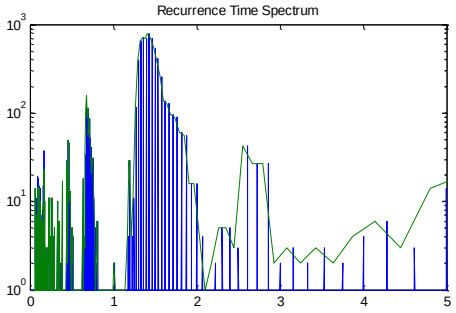

Figure 6-16: Recurrence-derived FT (line 81-84)

In the recurrence-derived FT method, some other frequencies which less than 0.52 Hz and more than 2 Hz are detected, which are not identified by the FT method.

6.5 Comparison with other methods

The accuracy of the determination of frequency modes, achieved by applying the recurrence-derived FT technique, is examined in contrast with the Welch's method on the same test system.

6.5.1 Periodgram

A nonparametric estimate of the Power Spectral Density (PSD) of a stationary random process is called the periodogram [146]. It is the Fourier Transform of the biased estimate of the autocorrelation sequence. The periodogram of a signal x_n , with frequency sample f_s can be described as :

$$\hat{S}(f) = \frac{\Delta t}{N}\left|\sum_{n=0}^{N-1} x_n e^{-i2\pi fn}\right|^2 \qquad (4.20)$$

Where Δt is the sampling interval. The periodgram has several drawbacks. The spectral leakage causes high variance and severe bias which is one of the major drawbacks. In order

to get an estimator with better variance and bias properties there are some methods to modify periodogram.

6.5.2 Welch's method

A modified version of the periodogram is called Welch's estimator [147] and the following steps are considered for this method. The original N samples are divided into N_B overlapping blocks with N_S samples each. A data taper (h_t) is applied on each block. Data taper is a method for smoothing the periodogram in the frequency domain. The most advantage of tapering is an improved convergence of periodogram estimates to spectral density. In order to reduce the bias because of spectral leakage, it is suggested that each block is windowed by a Hamming data taper. For each block obtain a periodogram and average the individual periodograms together to shape an overall spectral estimate. The Welch's estimator will be given by

$$\hat{S}(f) = \frac{\Delta t}{N_B} \left| \sum_{j=0}^{N_B-1} \sum_{t=1}^{N_S} h_t X_{t+jn} e^{-i2\pi fn} \right|^2 \tag{4.21}$$

Where n represents the amount of overlap between each block. The nonparametric spectral estimators such as the Welch are likely to yield high resolution PSDs with low variance and low bias. This method is applied on two cases in this section. The first case has one lightly damped local mode and the second case has one local mode and two poorly damped inter-area modes. The figure 6.17 and 6.18 show the frequency modes detected by the Welch's method on the first and the second cases accordingly.

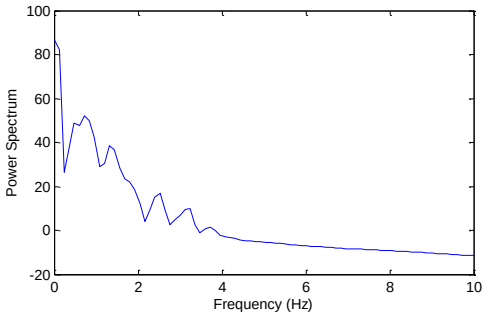

Figure 6-17: Power spectra the Welch's method line 81-84

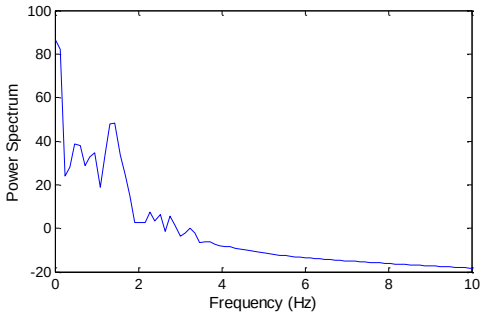

Figure 6-18: Power spectra the Welch's method line 81-84

The system modes identified by Welch's method, and the visible peaks are representation of modes in the spectrum estimate. Fig. 6.17 shows the frequency mode at 0.73 Hz for case 1, which has the highest magnitude. Fig. 6.18 shows the three peaks of high magnitude that happen at 0.52, 0.88 and 1.42 Hz, which represents the inter-area modes and local mode in case 2.

6.6 Conclusion

A new data-driven method to detect low-frequency oscillations early at inception is proposed by considering the dimensionality reduction of PMU data. The linear SFA dimensionality reduction method for the synchrophasor data is considered to extract slow features (SFs) features. The similarity index is applied on slowly varying components to specify a probable oscillation source. Two threshold values are used. The first threshold is applied to detect the power system is not in steady state condition, and the second threshold is applied to distinguish low-frequency oscillation and steady-state. By applying the RQA method on extracted features, oscillation detection is achieved. Case studies on the test system show the effectiveness of the oscillation detection algorithm in on-line applications.

The bus frequency recurrence rate of change is applied to detect low-frequency oscillation of the power system. This method does not depend on ubiquitous network

topology and operating conditions. A recurrence-derived FT method is applied to detect frequency mode of low-frequency oscillation. This method has two advantages. First it relaxes assumptions of stationary and non-linearity and second it applies embedding theorem to obtain dynamics of higher dimensional spaces. So, frequencies not observed in the regular FT periodogram are specified by this method.

Chapter 7 : CLUSTERING OF DISTURBANCE EVENTS USING RECURRENCE QUANTIFICATION ANALYSIS

7.1 Introduction

As more intelligence trickles into the power grid, the better situational awareness can be attained by identification of disturbance events through converting data to knowledge. This awareness can be useful when the system is under stress, or is in the initial phases of a probable blackout. For example, the study of the telephone log of a blackout [148] shows the system operators report all disturbance events to control centers. The implemented emergency responses are formulated based on this information. Real-time identification of disturbance events can perform the same task in an automated way, and in much faster speed. This paper introduces a novel method for developing a real-time event identification.

The main features of the proposed algorithm are: 1) it is data-driven and there is no need to know the system topology in advance; 2) it implements the dimensionality reduction of RQA features to extract key characteristics of the PMU data; 3) the lengthy buffering data, which is essential for frequency-domain analysis, is not needed here.

This chapter is organized as follows: Section 7.2 introduces the RQA and PCA methods. The proposed method is presented in Section 7.3. Case studies and simulation results are presented in section 7.4. Finally, conclusions are provided in section 7.5.

7.2 RQA and PCA

To investigate the hidden patterns of nonlinear systems, methods have been recently enhanced and they help to better understand nonlinear behavior of power systems from new standpoints. For examining and quantifying the nonlinear and nonstationary signals, the RQA was verified as a productive method [131]. It is necessary to mention that power system measurements, such as voltage and frequency, are nonstationary.

The goal of this section is to define and give possible explanations of all the thirteen RQA features, as reported in [131]. The features are RR, DET, D, Dmax, ENT, LAM, TT, Vmax, recurrence time of 1st type, recurrence of 2nd type, Trec, clust and Γ. These features are explained in more detail in the next section.

7.2.1 Recurrence Quantification Analysis

Recurrence is the representation of dynamical systems that determines whether the states revisit points in the phase space trajectory.

In summary, two points are called recurrent on the phase-space trajectory if the distance between them is less than a threshold value [149]. The recurrence of a state at time i at a different time j are visualized by R which is a two-dimensional squared matrix.

$$R_{i,j}^{m,\varepsilon_i} = \Theta(\varepsilon_i - \|x_i - x_j\|) \quad i, j = 1...N \tag{7.1}$$

Where N is the number of considered states xi, εi is the threshold distance, m is the dimensional phase space trajectory, and Θ(.) is the Heaviside function [150].

Fig. 7.1 shows a recurrence plot (RP) for voltage (sine signal) when there is no fault and Fig. 7.2 shows the voltage when a fault happens at t=5 seconds. The line-of-identity (LOI)

which is the solid main diagonal specifies RPs. Generally, large diagonals repeating themselves periodically represent periodicity of the signal which are shown in Fig. 7.1. Short diagonals are indication of non-deterministic or stochastic behavior which are observed in Fig. 7.2. Homogenous plot is result of stationary data while the RP with the upper side of the plot on left and the lower side of the plot disappearing is a result of non-stationary data.

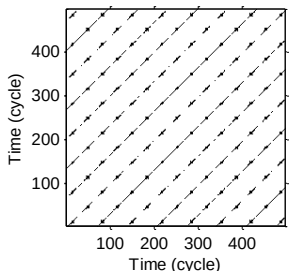

Figure 7-1: Recurrence plot (voltage) with no fault

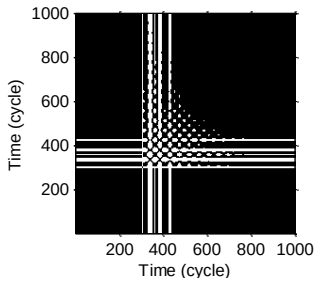

Figure 7-2: Recurrence plot (voltage) with fault

The explanation and equation six measures of the RQA which are RR, DET, Dmax, ENT, LAM and TT were discussed in chapter 5 section 5.2.2. The rest of measures will be discussed here.

Average (D) line length in the diagonal direction: D denotes the length of the average diagonal line in RP.

$$D = \frac{\sum_{l=d_{\min}}^{N} lH_D(l)}{\sum_{l=d_{\min}}^{N} H_D(l)} \tag{7.2}$$

Longest vertical line: V_{max} is the representation of the longest vertical line.

$$V_{\max} = \arg\max H_v(l) \tag{7.3}$$

Recurrence times of 1st and 2nd type: These are the mean times between consecutive points.

Recurrence period density entropy (Trec): The entropy delivers information about the aperiodicity.

$$T_{rec} = -\frac{1}{\ln(T_{\max})}.\sum_{t=1}^{T_{\max}} P(t).\ln(p(t)) \tag{7.4}$$

$$P(T) = R(T) / \sum_{k=1}^{T_{\max}} R(k) \tag{7.5}$$

Where P(T) is the recurrence time probability. The histogram of recurrence times are shown by R(T) and the maximum recurrence time is represented by $T_{max.}$

Clustering coefficient (clus): Among the recurrent states clus indicates the mean number of very closely related states.

$$clust = \sum_{i=1}^{N} \frac{\sum_{j,k=1}^{N} R_{i,j} R_{j,k} R_{k,i}}{\sum_{j=1}^{N} R_{i,j}} \tag{7.6}$$

Transitivity (Γ) : The probability that two nodes connected to a third node are also connected directly is called transitivity. Regular dynamics like a sine wave are related with a higher transitivity and disordered systems demonstrate low values of transitivity. It is an indication of how strong the recurrent states are.

$$\Gamma = \frac{\sum_{i,j,k} R_{i,j} R_{j,k} R_{k,i}}{\sum_{i,j,k} R_{k,i} R_{k,j}} \tag{7.7}$$

7.2.2 Principle Component Analysis

The main application of the principal components analysis is the examination of data to specify patterns to reduce the dimensions of the dataset with insignificant loss of information.

Since the number of PMU's can be quite large, and the data size can be massive, dimensionality analysis and reduction methods are being studied in the literature [151], [152] . The PCA is one the most common methods used to reduce dimensionality by maintaining the most variance of the main data [153]. It is used for coherency detection [154], extraction of fault features [155], and determining fault locations [156].

The steps for the PCA method are as follow:

1- Consider the dataset which includes d-dimensional without class labels

2- Compute the d-dimensional mean vector

3- Compute the covariance matrix of the data set

4- Compute eigenvectors and related eigenvalues

5- Sort the eigenvectors based on eigenvalues to shape *dxk* dimensional matrix *W*

6- Transform samples to new subspace by using *dxk* eigenvector matrix. The equation $y=W^{T}\times x$ can represent this, where x is a d×1 dimensional vector which shows one sample and y is the transformed k×1-dimensional sample in the new subspace.

7.3 Proposed Methodology

The proposed method is illustrated in Fig. 7.3. The main steps to identify different events are: (i) reading PMU data, (ii) constructing the RQA measures in each window, (iii) dimensionality reduction and (iv) applying k-means clustering.

The ratio of DET to RR is defined as variable ratio (RATIO) [150], and it is the frequency distribution of the lengths of the diagonal lines. The study of physiological systems indicated that the RATIO is suitable in identifying dynamic transitions [149]. For certain types of transitions, the DET does not change, while the RR decreases; thus, the ratio increases.

$$RATIO = N^2 \frac{\sum_{l=l\min}^{N} lH_D(l)}{\left(\sum_{l=1}^{N} lH_D(l)\right)^2} = \frac{DET}{RR} \tag{7.8}$$

As illustrated in Fig. 7.3, the two RQA measures of the nonlinear and nonstationary time series obtained by PMU data will be calculated at every 120 cycles. These measures are RR and DET. If the RATIO is 1, no action will be taken. When the RATIO is less than 1, the location of the event and the type of the event will be identified. The explanation of these steps is mentioned in the following sections as feature extraction and clustering.

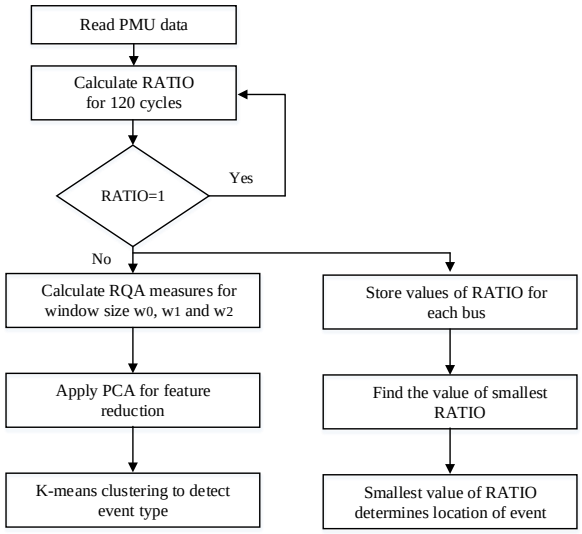

Figure 7-3: Flow chart for location detection and type of disturbance event

7.3.1 Feature extraction from disturbance events

Since the RQA is a successful method in obtaining dynamic behavior of different systems [149], the method was applied to extract features from disturbance events. Three data windows are considered in this study as shown in Fig. 7.4. To cover the pre-disturbance and post disturbance w_0 is considered before triggering, and both w_1 and w_2 are considered after triggering, for disturbance events. The interval is chosen 2-s for each data window. Then, to encircle every data window, the RQA measures are calculated for each bus. Voltage, phase angle, frequency were separately applied to calculate measures of the RQA. Characteristic features such as RR, DET, explained in section 7.2.1 were calculated. For constructing a feature vector, the total number of 117 features was gathered from each disturbance. By combining feature vectors for all disturbances, the feature matrix is formed. Fig. 7.5 shows the steps for shaping the feature matrix. This feature matrix is used as input to the PCA method for dimensionality reduction.

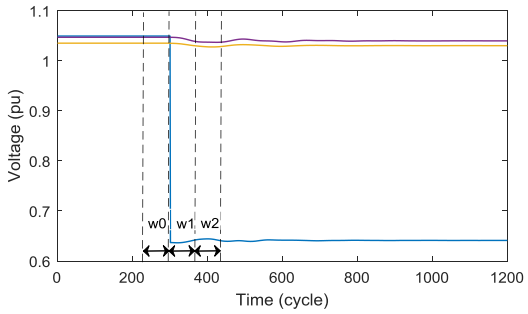

Figure 7-4: Plots of a sample event

Each row is the representation of one bus. Since the test system has 179 buses the feature matrix has 179 rows. Feature $f_{1,w0}$ consists of the RQA measures of voltage, phase angle and frequency before an event happens (w_0) for bus 1. Features $f_{1,w1}$ and $f_{1,w2}$ include same measures of the RQA after an event happens for the same bus.

$$\begin{bmatrix} f_{1,w0} & f_{1,w1} & f_{1,w2} \\ f_{2,w0} & f_{2,w1} & f_{2,w2} \\ f_{3,w0} & f_{3,w1} & f_{3,w2} \\ \vdots & \vdots & \vdots \\ f_{179,w0} & f_{179,w1} & f_{179,w2} \end{bmatrix}$$

Figure 7-5: Feature Matrix

7.3.2 K-means Clustering

Different disturbance events that illustrate similar behavior should be identified. Features build by the PCA are applied to k-means clustering method to identify distinct categories of events. The k-means clustering is a partitioning method. In this method, data is partitioned into *k* mutually exclusive clusters, which represent different category of events.

K-means method applies the centroids of clusters to describe the data. They are specified by minimizing the sum of squared errors,

$$J_k = \sum_{k=1}^{K} \sum_{i \in C_k} (x_i - m_k)^2 \tag{7.9}$$

Where the data matrix is $(x_{1,\ldots},x_n) = X$ and $m_k = \sum_{i \in c_k} x_i / n_k$ is the centroid cluster C_k and n_k is the number of points in C_k. The elbow method [157] is used to find the proper number of clusters. In this method the percentage of variance is defined as a function of the number of clusters.

7.4 Case Study

In this section, the effectiveness of the event detection algorithm, including the RQA and the dimensionality reduction method are shown. The Powerworld® software is used to produce the synthetic PMU data.

The proposed method is applied to the 179-bus system shown in Fig. 7.6. This system is a reduced Western Electricity Coordinating Council (WECC) ac transmission system that includes 29 generators, 179 buses, and 263 branches. We assume that each bus in the system has a PMU. For emulating field condition, we added noise to the simulated data with a signal-to-noise ratio (SNR) of 92 dB.

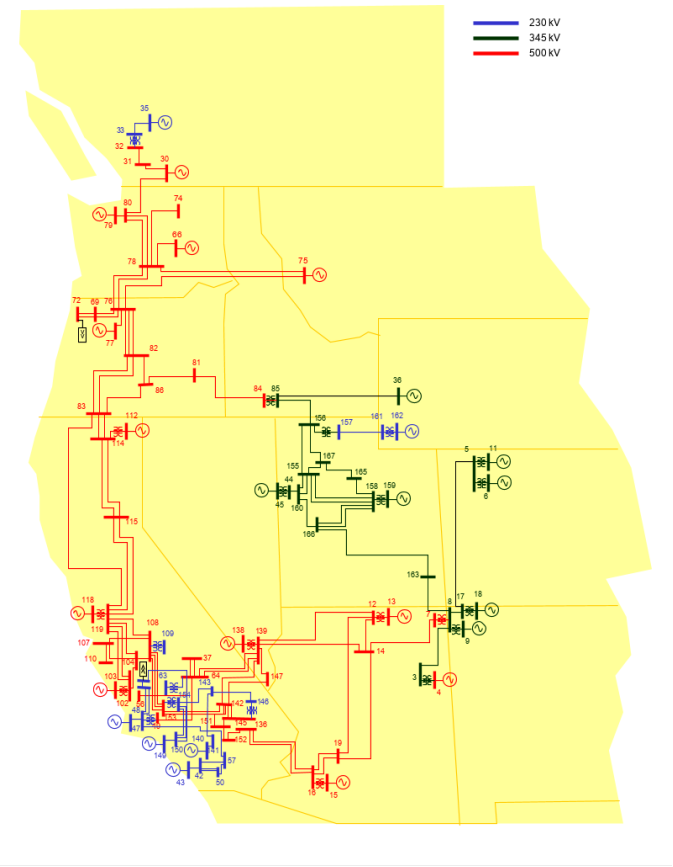

Figure 7-6: The reduced WECC 179-bus power system

7.4.1 Detecting location line tripping event

The proposed algorithm in Fig. 7.3 shows that when RATIO is equal to one, the system is in a steady state, and when it is less than 1, an event is detected. The test system has 179 buses and all voltages, phase angles, and frequencies are measured. As shown in Fig. 7.7, a fault happens at line 86-88 at t=5 s (300 cycles), and after 6 cycles, the fault clears.

Figs 7.9 and 7.11 show that before the fault occurs on line 86-88, the RR and DET values are one, and so, RATIO is one. This represents the fact that the system is in steady state condition. When the fault happens, the voltage drops from 1.04 p.u. to zero. Fig 7.9 indicates that the RR of bus 86 is 0.5 from 300 to 420 cycles, and Fig. 7.11 shows that the

relevant DET did not change from its value of 1. The RATIO value for the fault event is 1.98.

Bus 88 is a neighboring bus to 86 and Fig. 7.8 shows that the voltage during the fault decreases from 1.06 p.u. to 0.73 p.u. Fig. 7.10 shows the RR of bus 88 decreases from 1 to 0.4 and Fig. 7.12 indicates that the DET value is 1, and so RATIO would be 2.47.

Table 7.1 shows the RATIO value for some of the busses. Since the fault happens at line 86-88, bus 86 has the lowest RATIO value, and is therefore, the location of the fault. Bus 73 is far away from the other buses, and has the lowest voltage drop; so it has a minimum RR among the other busses, and highest value of RATIO, which is 3.19. It is necessary to mention here that the distance of a bus from the location of the fault does not always mean a lower value of RR. In this test system, a bus might be far from the fault location, but close to another generator; so it would not have a large voltage drop, and consequently, a lower value of RR.

Table 7-1: Measures of RATIO for some busses

BUS	RR	DET	RATIO
86	0.5	0.99	1.98
88	0.40	0.99	2.47
82	0.44	0.99	2.25
73	0.31	0.99	3.19

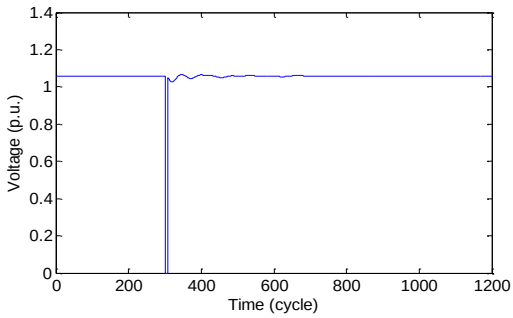

Figure 7-7: Fault at bus 86

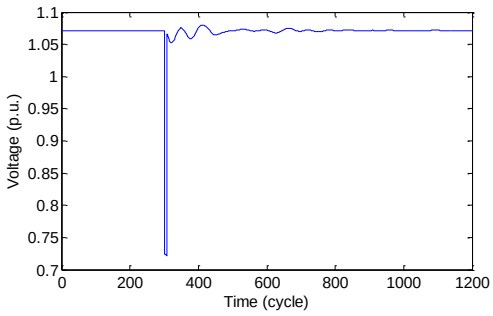

Figure 7-8: Fault at bus 88

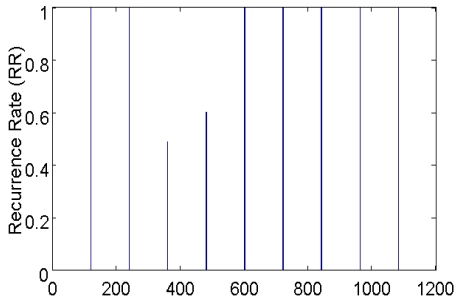

Figure 7-9: RR of bus 86

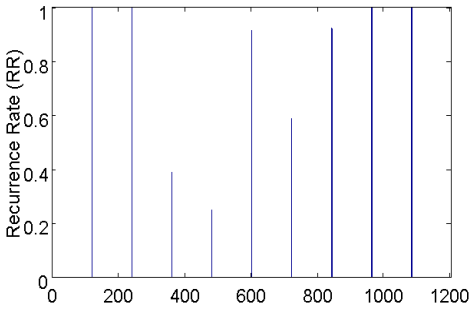

Figure 7-10: RR of bus 88

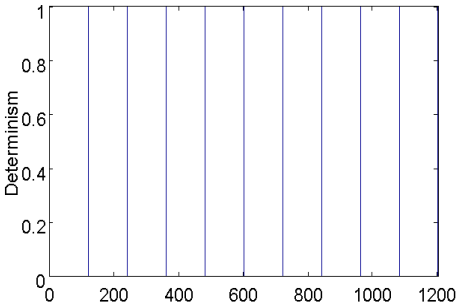

Figure 7-11: DET of bus 86

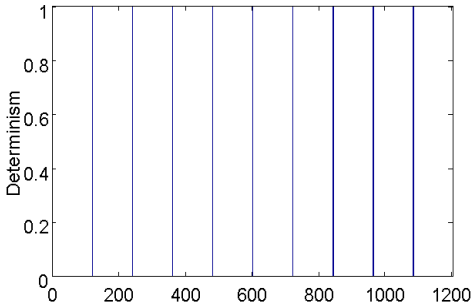

Figure 7-12: DET of bus88

Based on the proposed algorithm, RATIO will be calculated for all busses at the same time, and these values will be compared whereby, the bus with the lowest value would be identified as the location of an event.

7.4.2 Detecting location unit tripping event

As shown in Fig. 7.13, a generator at bus 66 (Montana) is tripped at t=5s, and the voltage decreases from 1.04 p.u. to 0.65 p.u. The voltage of two other substations which are in different areas are also shown. These voltages are measured at busses 46, 22. The voltage drop for these busses are low since they are not close to the location of the fault.

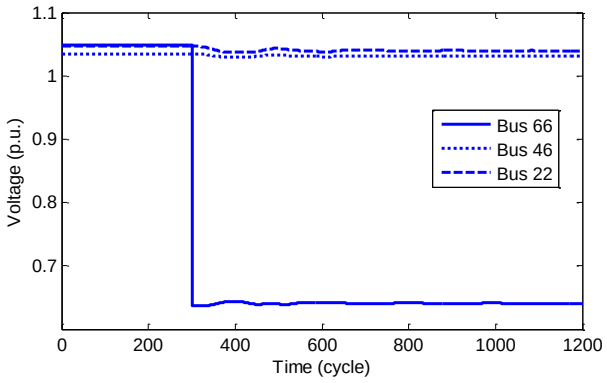

Figure 7-13:Voltage profile of some generators (generator trip at bus 66)

The RR and DET of generator at bus 66 are shown in Fig. 7.14 and 7.16. Before the generator trips at t=5 second (300 cycles), their values are 1, but after trip, the RR value changes from 1 to 0.51, but the DET remains unchanged. The RR and DET values for bus 46, which is in a different area, and far away, are shown in Figs. 7.15 and 7.17. The values of RR and DET were constant before the generator trip, while after the DET still remains constant, but RR changes from 1 to 0.32. Since the voltage pattern of busses 46 and 22 are similar, the RQA measures of bus 22 are not considered here.

The RATIO value for some of the generators and busses are shown in Table 7.2. At bus 66, RATIO is 1.94, and at bus 46, it is 3.09. The lowest value of RATIO identifies the location of the event, and in this case, it is bus 66 (Montana) where the generator trips. The RR has the highest value at bus 66, which has the maximum voltage drop. Busses 22, 46,

99 and 164 are not close to the location of the generator trip, and as expected, they have lower values of RR, and as a result, a higher values of RATIO. As seen, the value of DET for all busses are 0.99, except for busses 99 and 164, which have lower values. The impact of these lower values in determining values of RATIO is negligible though.

Table 7-2: Measures of RATIO for some busses

BUS	RR	DET	RATIO
66	0.51	0.99	1.94
46	0.32	0.99	3.09
22	0.22	0.99	4.50
99	0.06	0.92	15.33
164	0.08	0.97	12.12

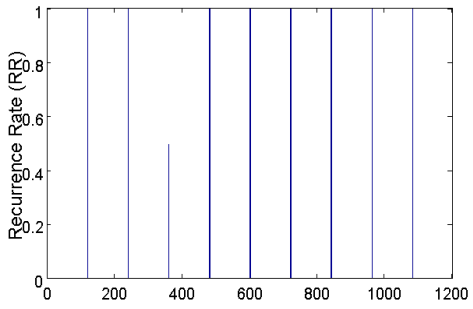

Figure 7-14: RR of bus 66

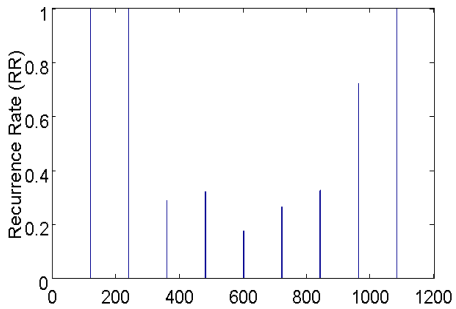

Figure 7-15: RR of bus 46

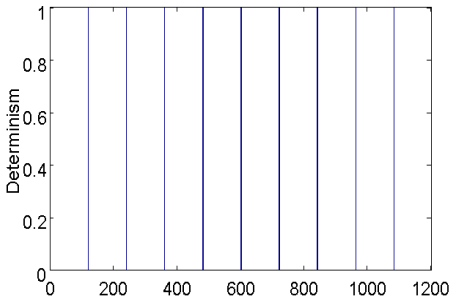

Figure 7-16: DET of bus 66

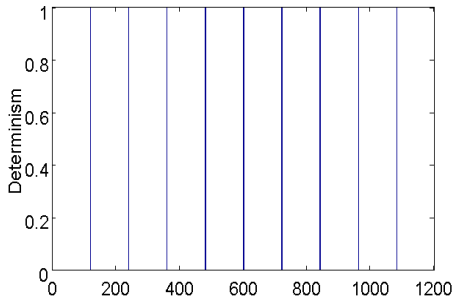

Figure 7-17: DET of bus 46

Three different windows, shown in Fig. 7.4, i.e., window w_0 before an event happens and windows w_1 and w_2 after the disturbance event, were captured to build the thirteen measures of the RQA. Some of these measures related to the time before the event are

shown in Table 7.3. These measures are related to pre-disturbance for busses 172, 124 and 26 and all measures have the same values as expected.

Since most of the post-event variation happens right after the fault, we show just the data for window w_1 in Tables 7.4 and 7.5.

Some measures of the RQA are presented in Table 7.3. These measures are related to the pre-disturbance period for busses 172, 124 and 26, and all measures have the same values for each bus, as expected. Some measures of the RQA for windows w_0 are not applicable when compared with Tables IV and V; so they are not shown here. These variables are laminarity, trapping time, recurrence time of the 2nd type, and recurrence period density entropy. For laminarity, $H_v(l)$ is the number of times a vertical line with length l exists. Since for window w_0 which is a flat line, there is no diagonal line, the numerator, as well as the denominator, are both zero; therefore, the value is indeterminate.

Also, for trapping time, both the numerator and the denominator are zero, and therefore, the output is indeterminate. The recurrence time of the 2nd type is defined as the mean time between consecutive points [158], and since the voltage does not change for window time w_0, it is not applicable here.

In order to investigate recurrence period density entropy we consider recurrence distance tends to be zero for a trajectory. The recurrence time probability [159] density is :

$$P(T) = \begin{cases} 1 & if \quad T = kk \\ 0 & otherwise \end{cases} \tag{7.10}$$

Where *kk* is the period of the trajectory. Since the recurrence time in the phase space is not equal to the period of trajectory, the value of *P(T)* is zero, and when substituted in (7.4), results T_{rec} to be zero.

Table 7-3:Some measures of RQA for window w_0

Fault Location	RR	DET	D	D_{max}	ENT
Bus 172	1	0.99	31	60	4.07
Bus 124	1	0.99	31	60	4.07
Bus 26	1	0.99	31	60	4.07

Table 7.4 shows some measures of the RQA for faults at busses 172, 124 and 26 and Table V shows the same measures for generator trips at buses 35, 30 and 79. The RR values are higher for fault events, and the DET is almost 0.99 for both events as expected. For periodic behaviors, DET can be considered as the predictability of the system. The periodic behavior here refers to systems that do not show a chaotic process.

The TT, V_{max} and recurrence time of the 2nd type (the mean times between consecutive points) are higher for fault events. The recurrence period density entropy, or T_{rec}, for the generator trip events are higher than that for the fault events.

Table 7-4: Some measures of RQA for window w_1 (Fault)

Fault Location	RR	DET	TT	Vmax	Recu. time 2nd type	T_{rec}
Bus 172	0.79	0.99	17.67	52	24.16	0.25
Bus 124	0.73	0.99	14.25	51	19.42	0.40
Bus 26	0.68	0.99	14.00	50	18.57	0.30

Table 7-5: Some measures of RQA for window w_1 (Generator trip)

Generator trip	RR	DET	TT	Vmax	Recu. Time 2nd type	T_{rec}
Gen. 35	0.11	0.97	4.53	12	9.13	0.77
Gen. 30	0.16	0.99	6.08	18	10.64	0.75
Gen. 79	0.17	0.97	5.32	19	12.27	0.72

7.4.3 Identifying disturbance events

Three phase to ground faults were created on 100 transmission lines in the 179-bus test system. Once again, the PowerWorld simulator was used to generate the data. The clearing time for faults is considered to be 6 cycles. A generator trip is also considered for all 29 generators. Thus, a total of 129 simulations were conducted, and for each case, voltages, angles and frequencies were recorded. For each simulation, w_0, w_1 and w_2 were considered. Fig. 7.18 shows the cumulative variance calculated from the PCA method. The first four PCs contains about 92.62% of the variance and the first 7 PCs preserves about 98.45% variance. It is observable from Fig. 7.18 that the dimensionality reduction can indeed be obtained by the PCA method.

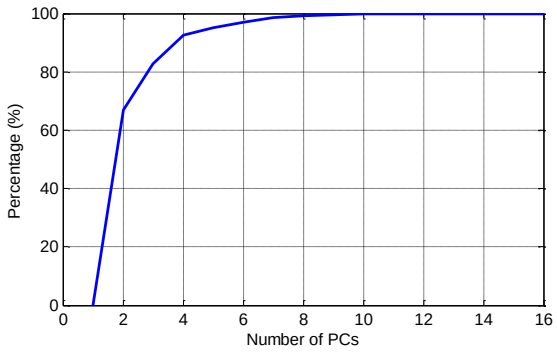

Figure 7-18: Cumulative variance preserved by PCs

Principle component coefficients for four variables *D*, D_{max}, V_{max} and *clust* as well as the principal component scores for 20 events are visualized in Fig. 7.19. This figure shows

all four variables by a vector. The contribution of each variable to the two-principal components are illustrated by the direction and length of the vector. For instance, the first principal component, PC1, has positive coefficients for the D_{max}, *clust* and V_{max} variables. Thus, the direction of these vectors are toward the right half of the plot. For PC1, the largest coefficient is V_{max}. The vertical axis represents the second principal component, or PC2. It has positive coefficients for the variables *D*, D_{max}, and *clust*, and a negative coefficient for the measure V_{max}.

Each of the 20 events is shown in the biplot (red circles), and the coordinates indicate the score of each event for the first two principal component. For instance, the highest scores for PC1 are located on right side of the plot. It is necessary to mention that all the RQA measures illustrated in Fig. 7.19 are related to voltages.

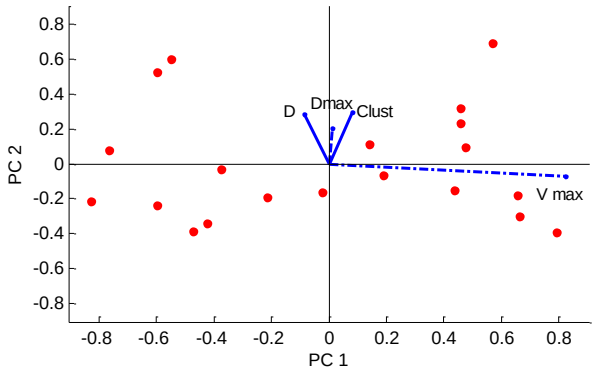

Figure 7-19: Orthonormal principal component coefficients

Thirteen measures of the RQA, including RR, DET, LMAX, ENT, LAM, TT are computed from voltage, angle and frequency. These features are pulled in for recurrence

examination for each window, and since three windows are considered, the number of features would be 39. Subsequently, we are measuring voltage, angle and frequency, and so the total number of features would be 117 (39×3=117).

Fig. 7.20 illustrates the fact that disturbance events are divided into two groups. The red cluster (circles) shows the generator trip events, and the blue cluster (triangles) illustrates the fault events. PC1 and PC2 are the first two components of the PCA. Out of 129 simulations, only 49 events are shown for clarity, twenty-nine events are related to generator trips, and the other 20 are related to the three phase to ground faults.

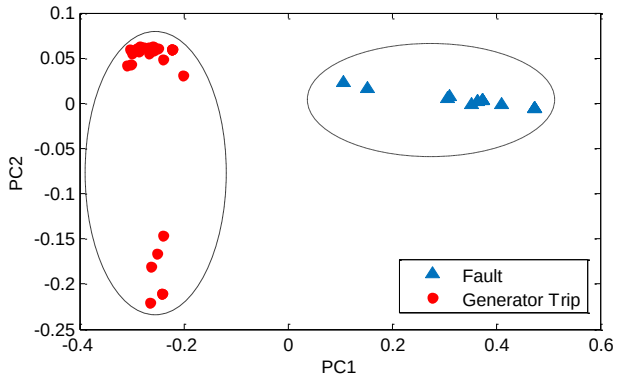

Figure 7-20: Disturbance event clusters for generator trip and fault

7.4.4 Missing data

Time stamped data are continuously sent by PMUs to the PDC. Nonetheless, in some cases, there may be missing data because of lapses in communication between the PMUs and the PDC. If a long interval of communication failure happens, or the related equipment breaks down, all the data, or a substantial portion of the data may go unrecorded.

To process the amount of missing data, we examine the voltage, phase angle and frequency recordings to identify the total number of missing data. Each event file has 360 samples (3 seconds) of data, out of which, 60 samples (1 second) are from before the disturbance event, and 120 samples (2 seconds) are from after the event. Since there are 179 PMUs and the total number of samples are 360, we expect a total of 360×179 =64440 samples to make the classification of event type. A threshold of 5% (3222 samples) is considered for the disturbance file. Thus, if the total number of missing data is more than 5% of the total number of data samples, then that disturbance event cannot be considered. However, if the missing data amounted to less than the defined threshold, then the average of the two previous samples can be considered to fill the missing data.

To evaluate the effectiveness of the proposed approach in the presence of missing data, we consider three cases which are shown in Table 7.6. For the first case, only voltages are considered, and it is observed that the error in correctly identifying faults is 20% that for correctly identifying generator trips is 24.13%. This would mean that 24.13% of generator trips were misclassified as fault events. In the second case, voltage magnitudes and phase angles were considered. The error for fault events was 10% and that for generator trips was 17.24%. When voltage and frequency were considered in the third case, we observed that the error for correctly identifying faults was 5% and that for generator trips is 13.79%.

Table 7-6: Clustering error for faults and generator trip

Scenarios	Clustering error for faults (%)	Clustering error for Generator trips (%)
Volt	20.00%	24.13%
Volt and Phaseangle	10.00%	17.24%
Volt and Frequency	5.00%	13.79%

7.5 Conclusion

A novel method based on data acquisition is proposed in this study for determining the location and types of disturbance events based on the RQA method. The voltages, phase angles and frequencies are obtained for each bus, and RQA measures are used to capture specific characteristics of these system quantities. The PCA method is applied for dimensionality reduction of the data, along with a clustering method to identify different types of disturbance events.

The proposed method identifies the location and differentiates between two types of disturbance events in the system: faults and generator trips. The method is applicable on large systems since it does not depend on training data. Also, there is no need for the system topology because it is entirely data-driven. Results for a 179-bus test system reveal the accuracy of the proposed method in locating and identifying disturbance events.

Missing data was considered in evaluating the performance of the proposed method. When the missing data is less than a defined threshold, the averaging method is applied, and results of the clustering method reveal that the accuracy of the proposed method in identifying disturbance events decreases somewhat. However, the misclassification errors in identifying fault events were less than that for generator trips.

Chapter 8 : CONCLUSIONS AND FUTURE WORKS

Some disturbances in power systems can lead to sustained interruptions causing extensive blackouts.. Data-driven methods can play an important role to fully understand the consequence of serious incidences. The goal of applying technologies such as synchrophasor data in power systems is to reach better management of the system security.

In this dissertation, novel data-driven methods for predicting transient stability, detecting of generators coherency, identifying low-frequency oscillation, recognizing frequency mode of the power system and determining location and type of disturbances are proposed.

8.1. Concluding remarks

In chapter 3, a method based on WSVM and ANFIS was proposed. Although the SVM can create classifiers with high prediction accuracy, in order to decrease the effect of outliers in the training model, and to contribute more weight to the important training data points, a weighted support vector machine (WSVM) was built. The fuzzy membership related training data helps to provide different degrees of membership in the training of the classifier. Hence, the WSVM obtains superior classification accuracy compared to the SVM. Numerical tests show that the applied approach was adequate to determine the stability status of a power system. Multiple SVM classifiers were combined with an adaptive neural fuzzy system that considers distances of the data points from the optimal hyperplane. The improvement of this method was seen when the WSVMs were combined with an ANFIS that was able to tune its membership function parameters to achieve higher

accuracy. The proposed adaptive neuro fuzzy combined WSVM persistently exceeded the individual base SVM classifiers.

In chapter 4, two metaheuristic methods for training the weights of an FNN to predict the stability status of a power system using post-disturbance voltage magnitudes and generator rotor angles were studied. The FNN trained by the grey wolf optimization method can predict system stability status better than the particle swarm optimization method, even though the latter method converged faster. This method could predict transient stability of the system with good accuracy even when the network topology was changed.

It is possible to use the adopted intelligent methods for real time transient stability assessment and control of larger systems. However, some effort will be needed to train the FNN with credible contingencies under varying operating conditions.

In chapter 5, the RQA method based on data acquisition was proposed to detect generator coherency in a power system. The rotor angles decomposed by wavelet decomposition and the RQA measures were used to capture characteristics of rotor angles, and then coherent groups of generators were detected by applying a clustering technique.

A wavelet transform decomposes the nonstationary time series into different frequency bands to efficiently keep apart the system transient and steady-state behaviors. Additionally, wavelet subsampling facilitates the computation of recurrence plots from the time series data. Further, recurrence analysis becomes more compelling in the multiple wavelet scales under a stationary premise.

The results obtained by the RQA method were compared with results obtained by the PCA method. When there is noise in the measured data, the PCA method is unable to accurately detect the coherent group of generators. Since the RQA method investigates recurrence dynamics in multiple wavelet scales, and advances the performance of recurrence characteristics, it can determine the coherent groups of generators precisely even in the presence of noise.

In chapter 6, a new data-driven method to detect low-frequency oscillations early at inception was proposed by considering the dimensionality reduction of PMU data. The linear SFA dimensionality reduction method for the synchrophasor data was considered to extract slow features (SFs) features. The similarity index was applied on slowly varying components to specify a probable oscillation source. Two threshold values were used. The first threshold was applied to detect that the power system was not in steady state condition, and the second threshold was applied to distinguish low-frequency oscillation and steady-state. By applying the RQA method on extracted features, oscillation detection was achieved. Case studies on the test system show the effectiveness of the oscillation detection algorithm in on-line applications.

The bus frequency recurrence rate of change was applied to detect low-frequency oscillation of the power system. This method does not depend on ubiquitous network topology and operating conditions.

A recurrence-derived FT method was applied to detect frequency mode of low-frequency oscillation. This method has two advantages. First it relaxes assumptions of stationary and non-linearity and second it applies embedding theorem to obtain dynamics

of higher dimensional spaces. So, frequencies not observed in the regular FT periodogram were specified by this method.

In chapter 7, a novel data-driven method based on the RQA was proposed to determine location of disturbance events and identify two types of disturbance events. For each bus the voltages, phase angles and frequencies were considered, the RQA measures were applied to capture the features of obtained data. For dimensionality reduction of PMU data, the PCA method was applied and then k-means clustering was used to identify disturbance events.

The location and two types of disturbance events (faults and generator tripping) were determined by the proposed method. Since the method does not depend on training data, it can be applied on large systems. The proposed method is data-driven and does not depend on topology of the system. Results indicates the accuracy of the proposed method by applying it on the 179-bus system.

8.2. Direction for Future Work

Renewable Energy:

Wind energy is growing very fast, and with penetration of these wind turbines, power systems will include many asynchronous machines. Therefore, dynamics and operational characteristics of power systems will experience alteration. Hence, the influence of these technologies on the system stability needs to be considered. For future work, the impact that enormous wind farms may have on the stability of the power system is recommended. The following items are needed to be addressed:

1- The determination of the coherent group of generators in a power system when there is a large penetration of wind energy.
2- The effect of wind power intermittency on the identification of coherent group of generators.

PMU:

The data presented in dissertation is generated by Powerworld software. Real data provided by filed PMUs can be applied on proposed techniques for evaluating the performance of the algorithms.

REFERENCES

[1] D. Novosel, M. Begovic, C. Taylor, J. Clerfeuillie, N. Rhatt and D. Karlsson, "Wide area protection and emergency control," *IEEE Power Engineering Society*, pp. 917-923, 1999.

[2] L. S. Moulin, A. P. A. da Silva, M. A. El-Sharkawi and R. J. Marks, "Support vector machines for transient stability analysis of large-scale power systems," *IEEE Transactions on Power Systems*, vol. 19, no. 2, pp. 818-825, 2004.

[3] C. A. Jensen, M. A. El-Sharkawi and R. J. Marks, "Power system security assessment using neural networks: feature selection using Fisher discrimnation," *IEEE Transaction on Power Systems*, vol. 16, no. 4, pp. 757-763, 2001.

[4] E. Bonabeau, M. Dorigo and G. Theraulaz, "Swarm Intelligence from natural to artificial systems," London, Oxford University Press, 1999, pp. 68-73.

[5] M. Dorigo, M. Birattari and T. Stutzle, "Ant colony optimization," *IEEE Computational Intelligence Magazine*, vol. 1, no. 5, pp. 28-39, 2006.

[6] J. Kennedy and R. Eberhart, "Particle swarm optimization," *IEEE International Conference on Neural Networks*, p. 1942.1948, 1955.

[7] A. group, "Annual Energy Outlook 2017 with projections to 2050," U.S. energy information administration department of energy, 2017.

[8] A. W. E. Association, "AWEA U.S. Wind Industry Annual Market Report Year Ending 2016," American wind energy association, 2016.

[9] D. N. Kosterev, C. W. Taylor and W. A. Mittelstadt, "Model validation for the August 10, 1996 WSCC system outage," *IEEE Transactions on Power Systems*, vol. 14, no. 3, pp. 967-979, 1999.

[10] V. Venkatasubramanian and Y. Li, "Analysis of 1996 western American electric blackouts," in *Proc. Bulk Power System Phenomena-Stability and Control*, Italy, 2004.

[11] A. Atputharajah and T. K. Saha, "Power system blackouts- literature review," in *International Conference on Industrail and Information Systems (ICIIS)*, Sri Lanka, 2009.

[12] I. Kamwa and R. Grondin, "PMU configuration for system dynamic performance measurement in large, multiarea power systems," *IEEE Transactions on Power Systems*, vol. 17, no. 2, pp. 385-394, 2002.

[13] G. Rogers, Power Systems Oscillations, Norwell,MA: Kluwer, 2000.

[14] S. B. Yusof, G. J. Rogers and R. T. H. Alden, "Slow coherency based network partitionig including load buses," *IEEE Transactions on Power Systems,* vol. 8, no. 3, pp. 1375-1382, 1993.

[15] G. N. Ramaswamy, C. Evrard, G. C. Verghese, O. Fillatre and B. C. Lesieutre, "Extensions, simplifications, and tests of synchronic modal equivalencing (SME)," *IEEE Transactions on Power Systems,* vol. 12, no. 2, pp. 896-905, 1997.

[16] D. N. Kosterev, C. W. Taylor and W. A. Mittelstadt, "Model validation for the August 10, 1996 WSCC system outage," *IEEE Transactions on Power Systems*, vol. 14, no. 3, pp. 967-979, 1999.

[17] N. A. E. R. Council, *1996 system disturbances,* August 2002.

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

[19] P. Gao, M. Wang, S. G. Ghiocel, J. H. Chow, B. Fardanesh, G. Stefopoulos, "Misssing data recovery by exploiting low-dimensionality in power system

synchrophasor measurements," *IEEE Transactions in Power Systems,* vol. 11, no. 99, pp. 1-8, 2015.

[20] J. Ma, T. Wang, Z. Wang, J. S. Thorp, "Adaptive damping control of inter-area oscillations based on federated kalman filter using wide area signals," *IEEE Transactions on Power Systems,* vol. 28, no. 2, pp. 1627-1635, 2013.

[21] N. Kakimoto, M. Sugumi, T. Makino and K. Tomiyama, "Monitoring of interarea oscillation mode by sychronized phasor measuremetn," *Power Systems IEEE Transactions on,* vol. 21, no. 1, pp. 260-268, 2006.

[22] R. Yousefian, A. Sahami and S. Kamalasadan, "Hybrid Transient Energy Function-Based Real-Time Optimal Wide-Area Damping Controller," *IEEE Transactions on Industry Applications,* vol. 53, no. 2, pp. 1506-1516, 2017.

[23] D. Santos, L. Fabiano, G. Antonova and M. Larsson, "The use of synchrophasors for wide area monitoring of electrical power grids," *Actual Trends in Developing Power system protection,* 2013.

[24] A. G. Phadke, "Synchronized phasor measurements in power systems," *Computer Applications in Power,* vol. 2, no. 6, pp. 10-15, 1993.

[25] A. Phadke and R. Moraes, "The wide world of wide-area measurement," *IEEE Power Energy Magazine,* vol. 6, no. 5, pp. 52-65, 2008.

[26] V. Terzija, G. Valverde, D. Cai, P. Regulski, V. Madani and A. Phadke, "Wide-area monitoring, protection, control of future electric power networks," *Proceeding IEEE,* vol. 99, no. 1, pp. 80-93, 2011.

[27] L. Xie, Y. Chen and H. Liao, "Distributed online monitoring of quasi-static voltage collapse in multi-area power systems," *IEEE Transactions on Power Systems,* vol. 27, no. 4, pp. 2271-2279, 2012.

[28] S. Dasgupta, M. Paramasivam, U. Vaidya and V. Ajjarapu, "Real-time monitoring of short-term voltage stability using PMU data," *IEEE Transaction on Power Systems*, vol. 28, no. 4, pp. 3702-3711, 2013.

[29] J. Jiang, J. Yang, Y. Lin, C. Liu and J. Ma, "An adaptive PMU based fault detection/location technique for transmission lines. I. theory and algorithms," *IEEE Transaction on Power Delivery*, vol. 15, no. 2, pp. 486-493, 2000.

[30] J. Jiang, Y. Lin, J. Yang, T. Too and C. Liu, "An adaptive PMU based fault detection/location technique for transmission lines. II. PMU implementation and performance evaluation," *IEEE Transaction on Power Delivery*, vol. 15, no. 4, pp. 1136-1146, 2000.

[31] J. Tate and T. Overbye, "Line outage detection using phasor angle measurements," *IEEE Transaction on Power Systems*, vol. 23, no. 4, pp. 1644-1652, 2008.

[32] Regional disturbance monitoring and reporting requirements, "North American Electric Reliability Corporation, NERC Standard PRC-002-1," Washington, 2006.

[33] M. Adamiak and R. Hunt, "Applcation of phasor measurement units for disturbance recording," in *10th Annual Georgia Tech Fault and Disturbance Analysis Conference*, Atlanta, GA, 2007.

[34] P. M. Anderson and A. A. Fouad (Abdel-Aziz A.), Power system control and stability (2nd ed), Piscataway: Wiley-Interscience, 2003.

[35] M. A. Pai, Energy Function Analysis for Power System Stability, Boston: Kluwer Academic, 1989.

[36] D. Ruiz-Vega and M. Pavella, "A compprehensive aproach to transient stability control I. Near optimal preventive control," *IEEE Transactions on Power Systems*, vol. 18, no. 4, pp. 1446-1453, 2003.

[37] Y. Xue,L. Wehenkel, R. Belhomme, P. Rousseaux, M. Pavella, E. Eu-xibie,B. Heliborn and J. F. Lesigne, "Extended equal area criterion revisited (EHV power systems)," *IEEE Transaction on Power System,* vol. 7, no. 3, pp. 1012-1022, 1992.

[38] M. Pavella, D. Ernst and D. Ruiz-Vega, Transient Stability of Power Systems: A unified Approach to Assessment and control, Boston,MA: Kluwer, 2000.

[39] H. Drucker, C.J. Burges, L. Kaufman, A. Smola and V. Vapnick, "Support vector regression machines," *Advances in Neural Information Processing Systems,* pp. 155-161, 1997.

[40] A. D. Rajapakse, F. Gomez, K. Nanayakkara, P. A. Crossley and V. V. Terzija, "Rotor Angle instability prediction using post-disturbance voltage trajectories," *IEEE Transactions on Power Systems,* vol. 25, no. 2, pp. 947-956, 2010.

[41] C. Andersson and J. E. Solem, "Improvements in classification of power system security," in *2006 IEEE Power Engineering Society General Meeting*, Montreal, 2006.

[42] C. W. Hsu and C. J. Lin, "A comparision of methods for multiclass support vector machines," *IEEE Transactions on Neural Networks*, vol. 13, no. 2, pp. 415-425, 2002.

[43] B. Wang, B. Fang, Y. Wang, H. Liu and Y. Liu, "Power system transient stability assessment based on big data and the core vector machine," *IEEE Transactions on Smart Grid*, vol. 7, no. 5, pp. 2561-2570, 2016.

[44] S. Ye, X. Li, X. Wang and Q. Qian, "Power system transient stability assessment based on Adaboost and support vector machines," in *Asia-Pacific Power and Energy Engineering Conference*, Shanghai, 2012.

[45] V. Krishnan, J. D. McCalley, S. Henry and S. Issad, "Efficient database generation for decision tree based power system security assessment," *IEEE Transactions on Power Systems,* vol. 26, no. 4, pp. 2319-2327, 2011.

[46] S. M. Roynyak, C. W. Taylor and Y. Sheng, "Decision trees using apparent resistance to detect impending loss of sychronism," *IEEE Transactions on Power Delivery,* vol. 15, no. 4, pp. 1157-1162, 2000.

[47] B. Basturk and D. Karaboga, "An artificial bee colony (ABC) algorithm for numeric function optimization," in *IEEE swarm intelligence symposium,* 2006.

[48] S. Mirjalili and A. Lewis, "Grey wolf optimizer," *Advances in Engineering Software,* vol. 69, pp. 46-61, 2014.

[49] XS. Yang, "A new metaheuristic bat-inspired algorithm," *Springer Journal,* vol. 284, pp. 65-74, 2010.

[50] X. Li, "A new intelligent optimization artificial fish swarm algorithm," Thesis, Zhejianfg University of Zhejiang, China, 2003.

[51] A. Mucherino and O. Seref, "Monkey search: a novel metaheuristic search for global optimization," *American Institute of Physics (AIP) conference proceedings,* vol. 953, no. 1, pp. 162-166, 2007.

[52] XS. Yang and S. Deb, "Cuckoo search via Levy flights," *Nature and biologically inspired computing world conference,* pp. 210-214, 2009.

[53] XS. Yang, "Firefly algorithm stochastic test functions and design optimization," *International Journal Bio-Inspired computer,* vol. 2, pp. 78-84, 2010.

[54] A. H. Gandomi and A. H. Alavi, "Krill herd: a new bio-inspired optimization algorithm," *Communications in Nonlinear Science and Numerical Simulation Journal,* vol. 17, no. 12, pp. 4831-4845, 2012.

[55] S. Shaw and W. Kinsner, "Chaotic simulated annealing in multilayer feedforward networks," *Canadian conference on electrical and computer engineering,* pp. 265-269, 1996.

[56] S. K. Chang, O. A. Mohammed and S. Y. Hahn, "Detection of magnetic body using article neural network with modified simulated annealing," *IEEE Transactions Mag.,* vol. 30, pp. 3644-3647, 1994.

[57] D. J. Montana and L. Davis, "Training feedforward neural networks using genetic algorithms," *Eleventh International joint conference on artificial intelligence,* pp. 762-767, 1989.

[58] S. Kiranyaz, T. Ince, A. Yildrim and M. Gabbouj, "Evolutionary artificial neural networks by multi-dimensional particle swarm optimization," *Neural Networks Journal,* vol. 22, pp. 1448-1462, 2009.

[59] J. R. Zhang, J. Zhang, T.M. Lock and M.R. Lyu, "A hybrid particle swarm optimization-back-propagation algorithm for feedforward neural network training," *Applied mathematics and Computational Journal,* vol. 185, pp. 1026-1037, 2007.

[60] M. Cells and B. Rylander, "Neural network learning using particle swarm optimization," *The International Conference on Advances in Information Science and Soft Computing*, pp. 224-226, 2002.

[61] I. Kamwa and R. Grondin, "PMU configutation for system dynamic performance measurement in large, multi area power systems," *IEEE Transactions on Power Systems,* vol. 17, no. 2, pp. 385-394, 2002.

[62] M. A. E. A. Abd-EI-Rehim, I. D. Helal and M. A. Hassan Omar, "Multi-machine power system dynamic equivalents using artificail intelligence (ANN)," in *2006 Eleventh International Middle East Power Systems Conference*, El-Minia, 2006.

[63] M. Davodi, H. Modares, E. Reihani and A. Sarikhani, "Coherency approach by hybrid PSO, K-Means clustering method in power system," in *2008 IEEE 2nd International Power and Energy Conference*, Johor Bahru, 2008.

[64] G. C. Pyo, J. W. Park and S. I. Moon, "A new method for dynamic reduction of power system using PAM algorithm," in *IEEE PES General Meeting*, Minneapolis,MN, 2010.

[65] M. A. Rios and O. Gomez, "Identification of coherent groups and PMU placement for inter-are monitoring based on graph theory," in *2011 IEEE PES congerence on innovative smart grid technologies latin america (ISGTLA)*, Medellin, 2011.

[66] N. Senroy, "Generator coherency using the Hilbert-Huang transform," *IEEE Transactions on Power Systems,* vol. 23, no. 4, pp. 1701-1708, 2008.

[67] K. L. Lo, Z. Z. Qi and D. Xiao, "Identification of coherent generators by spectrum analysis," *IEEE Proceedings-Generation, Transmission and Distribution,* vol. 142, no. 4, pp. 367-371, 1995.

[68] M. Jonsson, M. Begovic and J. Daalder, "A new method suitable for real-time generator coherency determination," *IEEE Transactions on Power Systems,* vol. 19, no. 3, pp. 1473-1482, 2004.

[69] K. K. Anaparthi, B. Chaudhuri, N. F. Thornhill and B. C. Pal, "Coherency Identification in power systems through Principle Component Analysis," *IEEE Transactions on Power Systems,* vol. 20, no. 3, pp. 1658-1662, 2005.

[70] M. A. M. Ariff and B. C. Pal, "Coherenct identification in interconnected power system - an independent component analysis approach," in *2013 IEEE Power & Energy Society General Meeting*, Vancouver, BC, 2013.

[71] S. Avdakovic, E. Becirovic, A. Nuhanovic and M. Kusljugic, "Generator coherency using the wavelet phase difference approach," *IEEE Transactions on Power Systems,* vol. 29, no. 1, pp. 271-278, 2014.

[72] Y. Susuki and I. Mezic, "Nonlinear koopman modes and coherency identification of coupled swing dynamics," *IEEE Transactions on Power Systems,* vol. 26, no. 4, pp. 1894-1904, 2011.

[73] P. N. Papadopoulos, J. V. Milanovic, P. Bhui and N. Senroy, "Feasibility study of applicability of recurrence quantification analysis for clustering of power system dynamic responses," *2016 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, pp. 1-6, 2016.

[74] T. Hashiguchi, H. Ukai, Y. Mitani, M. Watanabe, O. Saeki and M. Hojo, "Power system dynamic performance measured by phasor measuremnt unit," *IEEE Lausanne Power Tech*, pp. 1694-1699, 2007.

[75] M. Hojo, T. Ohnishi, Y. Mitani, O. Saeki and H. Ukai, "Observation of frequency oscillation in western Japan 60 Hz power system based on multiple sychronized phasor measurements," *2003 IEEE Bologna Power Tech Conference Proceedings*, pp. 6-12, 2003.

[76] A. R. Messina, V. Vittal, G. T. Heydt and T. J. Browne, "Nonstationary approaches to trend Identification and denoising of measured power system oscillations," *IEEE Transactions on Power Systems,* vol. 24, no. 4, pp. 1798-1807, 2009.

[77] X. Liu, D. Mcswiggan, T. B. Littler and J. Kennedy, "Measurement-based method for wind farm power system oscillations monitoring," *IET Renewable Power Generation*, vol. 4, no. 2, pp. 198-209, 2010.

[78] J. L. Rueda, C. A. Juarez and I. Erlich, "Wavelet-based analysis of power system low-frequency electromechanical oscillations," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1733-1743, 2011.

[79] T. Jiang, H. Jia, H. Yuan, N. Zhou and F. Li, "Projection Pursuit: A general methodology of wide-area coherency detection in bulk power grid," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2776-2786, 2016.

[80] Z. Lin, F. Wen, Y. Ding and Y. Xue, "Data-driven coherency Identification for generators based on spectral clustering," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 99, pp. 125-129, 2014.

[81] M. Patel, S. Aivaliotis and E. Ellen, Real-Time application of synchrophasors for improving reliability, Princeton, NJ: North American Electricity Rel. Corp., 2010.

[82] M. J. Fuente, D. Garcia-Alvarez, G. I. Saint-Palmero and T. Villegas, "Fault detection and identification method based on multivariate statistical techniques," in *2009 IEEE Conference on Emerging Technologies & Factory Automation*, Mallorca, 2009.

[83] S. Santoso, W. M. Grady, E. J. Powers, J. Lamoree and S. C. Bhatt, "Characterization of distribution power quality events with fourier and wavelet transform," *IEEE Transaction Power Delivery*, vol. 15, no. 1, pp. 247-254, 2000.

[84] S. Santoso, E. J. Powers, W. M. Grady and P. Hofmann, "Power quality assessment via wavelet transform analysis," *IEEE Transaction on Power Delivery*, vol. 11, no. 2, pp. 924-930, 1996.

[85] J. Barros and E. Perez, "Automatic detection and analysis of voltage events in power systems," *IEEE Transaction Instrument Measurement*, vol. 55, no. 5, pp. 1487-1493, 2006.

[86] P. K. Dash, "Power quality disturbance data compression detection and classification using integrated spline wavelet and S-transform," *IEEE Transaction on Power Delivery*, vol. 18, no. 2, pp. 595-600, 2003.

[87] J. Lee and N. A. Touba, "Low power test data compression based on LFSR reseeding," in *Proc. IEEE International Conference Computing Design (ICCD)*, 2004.

[88] W. R. A. Ibrahim, R. Anis and M. Morcos, "Novel data compression technique for power waveforms using adaptive fuzzy logic," *IEEE Transaction Power Delivery*, vol. 20, no. 3, pp. 2136-2143, 2005.

[89] L. Wiskott and T. Seinowski, "Slow feature analysis: Unsupervised learning of invariance," *Neural Computation*, vol. 14, no. 4, pp. 715-770, 2002.

[90] J. F. Hauer, "Application of Prony analysis to the determination of modal content and equivalent models for measured power system response," *IEEE Transactions on Power Systems,* vol. 6, no. 3, pp. 1062-1068, 1991.

[91] J. Choi, K. Shim, H. Nam, Y. Lim and S. Nam, "Parameter estimation method of low frequency oscillating signals using discrete Fourier Transforms," *Electrical Eng. Technology,* vol. 7, no. 5, pp. 163-170, 2011.

[92] P. Ray, "Power system low frequency oscillation mode estimation using wide area measurement systems," *Engineering science and technology, an international journal,* vol. 20, pp. 598-615, 2017.

[93] A. Bykhovsky and J. Chow, "Power system disturbance identification from recorded dynamic data at the northfield substation," *International Journal of Electrical Power Energy Systems,* vol. 25, no. 10, pp. 787-795, 2003.

[94] T. Bi, X. Song, J. Wu and Q. Yang, "Novel method for disturbance identification in power systems," in *Power Engineering Society General Meeting*, 2006.

[95] J. Ma, R. Diao and E. Dagle, "Building decision trees for characteristic ellipsoid method to monitor power system transient behaviors," in *Power and Energy Society General Meeting*, 2010.

[96] W. Wang and P. Markham, "Multiple event detection and recognition through sparse unmixing for high-resolution situational awarness in power grid," *IEEE Transaction on Smart Grid,* vol. 5, no. 4, pp. 1654-1664, 2014.

[97] Y. Zhang, Z. Wang, J. Zhang and J. Ma, "PCA fault feature extraction in complex electric power systems," *Advances in Electrical and Computer Engineering,* vol. 10, no. 3, pp. 102-107, 2010.

[98] R. Diao, K. Sun and V. Vittal, "Decision tree-based online voltage security assessment using PMU measurements," *IEEE Transaction on Power Systems,* vol. 24, no. 2, pp. 832-839, 2009.

[99] M. Rafferty, X. Liu and D. Laverty, "Real time multiple event detection and classification using moving window PCA," *IEEE Transactions on Smart Grid*, vol. 7, no. 5, pp. 2537-2548, 2016.

[100] X. Liu, D. Laverty, "Principle component analysis of wide-area phasor measurements for islanding detection-a geometric view," *IEEE Transaction on Power Delivery*, vol. 30, no. 2, pp. 976-985, 2015.

[101] D. Kim, T. Chun and S. Yoon, "Wavelet-based event detection method using PMU data," *IEEE Transaction on Smart Grid*, vol. 8, no. 3, pp. 1154-1162, 2017.

[102] O. Dahal, S. Brahma and H. Cao, "Comprehensive clustering of disturbance events recorded by phasor measurement units," *IEEE Transaction on Power Delivery*, vol. 29, no. 3, pp. 1390-1397, 2014.

[103] F. Gomez, A. Rajapakse and U. Annakkage, "Support vector machine-based algorithm for post-fault transient stability status prediction using synchronized measurements," *IEEE Transactions on Power System*, vol. 26, no. 3, pp. 1474-1483, 2011.

[104] C. Zheng, V. Malbasa and M. Kezunovic, "Regression tree for stability margin prediction using synchrophasor measurements," *IEEE Transactions on Power System*, vol. 28, no. 2, pp. 1978-1987, 2013.

[105] P. Kundur, J. Paserba and V. Vittal, "Definition and classification of power system stability ieee/cigre joint task force on stability terms and definitions," *IEEE Transactions on power systems*, vol. 19, no. 3, pp. 1387-1401, 2004.

[106] C. Liu and Y. Wang, "Application of a novel fuzzy neural network to real-time transient stability swings prediction based on synchronized phasor measurements," *IEEE Transaction on Power System*, vol. 14, no. 2, pp. 685-692, 1999.

[107] H. Sawhney and B. Jeyasurya, "A feed-forward artificial neural network with enhanced feature selection for power system transient stability assessment," *Electrical Power System Research,* vol. 76, no. 12, pp. 1047-1054, 2006.

[108] F. Gomez, A. Rajapakse, U. Annakkage and I. Fernando, "Support vector-based algorithm for post-fault transient stability status prediction using synchronized measurements," in *2011 IEEE Power and Energy Society General Meeting*, San Diego, 2011.

[109] T. Y. Guo and J. V. Milanovic, "Probabilistic framework for assessing the accuracy of data mining tool for online prediction of transient stability," *IEEE Transaction on Power System,* vol. 29, no. 1, pp. 377-384, 2014.

[110] B. Selvi and N. Kamaraj, "Transient stability assessment using fuzzy combined support vector machines," *Electric. Power Compon Syst.,* vol. 38, no. 6, pp. 737-751, 2010.

[111] J. Geeganage, U.D. Annakkage and B. A. Archer, "Application of energy-based power system features for dynamic security assessment," *IEEE transaction on Power System,* vol. 30, no. 4, pp. 1957-1965, 2015.

[112] Y. Xu and K. P. Wong, "Real-time transient stability assessment model using extreme learning machine," *IET Generation Transmission Distribution,* vol. 5, no. 3, pp. 314-322, 2011.

[113] C. X. Liu and K. Sun, "A systematic approach for dynamic security assessment and the corresponding preventive control scheme based on decision trees," *IEEE Transaction on Power System,* vol. 29, no. 2, pp. 717-730, 2014.

[114] C. J. C. Burges, A tutorial on support vector machines for pattern recognition, Norwell, MA: Kluwer, 1998.

[115] X. Yuesheng and H. Zhang, "Universal kernels," *Journal of Machine Learning Research* , vol. 7, pp. 2651-2667, 2006.

[116] C.F. Lin and S.D. Wang, "Fuzzy support vector machines," *IEEE Transactions on Neural Networks,* vol. 13, no. 2, pp. 464-471, 2002.

[117] T. Athay, R. Podmore and S. Virmani, "A practical method for the direct analysis of transient stability," *IEEE Transactions on Power Apparatus and Systems,* Vols. PAS-98, no. 2, pp. 573-584, 1979.

[118] T. Guo, J. He, Z. Li and J. V. Milanovic, "Evaluation of classification methods for on-line identification of power system dynamic signature," in *2014 Power Systems Computation Conference*, Wroclaw, 2014.

[119] E. H. Mamdani, "Application of fuzzy algorithms for control of simple dynamic plant," *Electrical Engineers, Proceedings of the Institution,* vol. 121, no. 12, pp. 1585-1588, 1974.

[120] B. Irie and S. Miyake, "Capability of three-layered perceptrons," *Proceedings of IEEE International Conference on Neural Networks,* pp. 641-648, 1998.

[121] N. Mat, "Clustered-hybrid multilayer perceptron network for pattern recognition application," *Applied Soft Computing,* vol. 11, no. 1, 2011.

[122] K. Homik, M. Stinchcombe and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks,* vol. 2, pp. 359-366, 1989.

[123] D.R. Hush and N.G. Horne, "Progress in supervised neural networks," *IEEE Signal Process Mag.,* vol. 10, pp. 8-39, 1993.

[124] M.T. Hagar and M.B. Menhaj, "Training feedforward networks with the Marquardt algorithm," *IEEE Transaction Network,* vol. 5, no. 6, pp. 989-993, 1994.

[125] S. Mirjalili and A. Safa Sadiq, "Magnetic optimization algorithm for training multi layer perceptron," *IEEE International Conference on Industrial and Intelligent Information (ICIII 2011),* vol. 2, pp. 42-46, 2011.

[126] T. Si, S. Hazra and N. Jana, "Artificial neural network training using differential evolutionary algorithm for classification," *Proceedings of the International Conference on Information Systems Design and Intelligent Applications,* pp. 769-778, 2012.

[127] Z. Z. Lin, T. Xia and Y. Z. Ye, "Application of wide area measurement systems to islanding detection of bulk power system," *IEEE Transactions on Power Systems,* vol. 28, no. 2, pp. 2006-2015, 2013.

[128] N. Zhou, "A cross-coherence method for detecting oscillations," *IEEE Transactions on Power Systems,* vol. 31, no. 1, pp. 623-631, 2016.

[129] H. Yang, "Multiscale recurrence quantification analysis of spatial cardiac vectorcardiogram signals," *IEEE Transactions on Biomedical Engineeering* , vol. 58, no. 2, pp. 339-347, 2011.

[130] W. Ting, Y. Guo-zheng, Y. Bang-hua and S. Hong, "EEG feature extraction based on wavelet packet decomposition for brain computer interface," *Journal of the International Measurement Confederation* , vol. 41, no. 6, pp. 618-625, 2008.

[131] C. Webber and N. Marwan, in *Recurrence Quantification Analysis*, Germany , Springer, 2015, pp. 3-20.

[132] H. Peng, F. Long and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 27, no. 8, pp. 1226-1238, 2005.

[133] D. Ketchen, C. Shook, "The application of cluster analysis in strategic management research: An analysis and critique," *Strategic Management Journal,* vol. 17, no. 6, pp. 441-458, 1996.

[134] S. Maslennikov, "A test cases library for methods locating the sources of sustained oscillations," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, Boston, MA, 2016.

[135] P. Kundur, "Definition and classification of power system stability IEEE/CIGRE joint task force on stability terms and definitions," *IEEE Transactions on Power Systems*, vol. 19, no. 3, pp. 1387-1401, 2004.

[136] "IEEE Standard for Synchrophasor Measurements for Power Systems," *IEEE Std C37.118.1-2011 (Revision of IEEE Std C37.118-2005*, pp. 1-61, 2011.

[137] I. T. Jolliffe, Principal Component Analysis, New York: Springer, 2002.

[138] J. Wright, Y. Peng, Y. MA, A. Ganesh and S. Rao, "Robust principal component analysis: Exact recovery of corrupted low-rank matrices by convex optimization," *Neural Information Processing Systems*, 2009.

[139] N. E. Huang, Z. Shen and S. R. Long, "The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Nonstationary Time Series Analysis," *Mathematical, Physical and Engineering Sciences*, 1998.

[140] J. Xiao, Y. Han and J. Wu, "Dynamic tracking of low-frequency oscillations with improved prony method in wide-area measurement system," *Proc. 2004 IEEE Power Engineering Society General meeting*, vol. 1, pp. 1104-1109, 2004.

[141] J.C.H. Peng and N.K.C. Nair, "Comparative assessment of kalman filter and prony methods for power system oscillation monitoring," *IEEE Conf. Power and Energy Society General Meeting*, pp. 1-8, 2009.

[142] S. Zhang, X. Xie and J. Wu, "WAMS-based detection and early warning of low frequency oscillations in large scale power systems," *Electric Power Systems Research*, vol. 78, no. 5, pp. 897-906, 2008.

[143] J. H. Choi and S.R. Nam, "Parameter estimation method of low frequency oscillating signals using discrete Fourier Transforms," *Electrical Engineering Technology,* vol. 7, no. 2, pp. 163-170, 2011.

[144] L. L. Trulla, A. Giuliani, J. P. Zbilut and C.L. Webber, "Recurrence quantification analysis of the logistics equation with transients," *Physics Letters A,* vol. 223, no. 4, pp. 255-260, 1996.

[145] X. Gao, C. Shang, F. Yang and D. Huang, "Detecting and isolating plant-wide oscillations via slow feature analysis," in *2015 American Control Conference (ACC)*, Chicago, IL, 2015.

[146] J. W. Pierre, D. J. Trudnowski and M. K. Donnelly, "Initial results in electromechanical mode identification from ambient data," *IEEE Transactions on Power Systems,* vol. 12, no. 3, pp. 1245-1251, 1997.

[147] J. Zbilut and N. Marwan, "The Wiener-Khinchin theorem and recurrence quantification," *Physics Letter A,* vol. 372, pp. 6622-6626, 2008.

[148] L. Y. Cao, "Practical method for determining the minimum embedding dimension of a scalar time series," *Physiaca D,* vol. 110, pp. 43-50, 1997.

[149] A. Sean and K. Fitz, "Algorithms for computing the time-corrected instantaneous frequency(reassigned) spectrogram with applications," *Journal of the Acoustical Society of America,* vol. 119, pp. 360-371, 2006.

[150] P. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Transactions on Audio and Electroacoustics,* vol. 15, no. 2, pp. 70-73, 1967.

[151] Om Dahal, S. Brahma and H. Cao, "Comprehensive clustering of disturbance events recorded by phasor measurement units," *IEEE Transactions on Power Delivery,* vol. 29, no. 3, pp. 1390-1397, 2014.

[152] C. Webber and N. Marwan, Recurrence Quantification Analysis, Germany: Springer, 2015.

[153] C. Webber and N. Marwan, Recurrence Quantification Analysis, Germany: Springer, 2015.

[154] Y. Chen, L. Xie and P. Kumar, "Dimensionality reduction and early event detection using online synchrophasor data," in *IEEE Power Engineering Society General Meeting*, 2013.

[155] N. Dahal, R. King and V. Madani, "Online dimension reduction of synchrophasor data," in *IEEE PES Transmission and Distribution Conference (T&D)* , 2012.

[156] I. Fodor, *A Survey of dimension reduction techniques,* 2002.

[157] K. Anaparthi, B. Chaudhuri, N. Thornhill and B. Pal, "Coherency identification in power systems through principal component analysis," *IEEE Transactions on Power Systems,* vol. 20, no. 3, pp. 1658-1660, 2005.

[158] Y. Zhang, Z. Wang, J. Zhang and J. Ma, "PCA fault feature extraction in complex electric power systems," *Advances in Electrical and Computer Engineering,* vol. 10, no. 3, pp. 102-107, 2010.

[159] Z. Wang, Y. Zhang and J. Zhang, "Principal components fault location based on WAMS/PMU measure system," in *IEEE Power Engineering Society General Meeting*, 2011.

[160] D. Ketchen and C. Shook, "The application of cluster analysis in strategic management research: An analysis and critique," *Strategic Management Journal*, vol. 17, no. 6, pp. 441-458, 1996.

[161] J. Gao, "Recurrence time statistics for chaotic systems and their applications," *Physical review letters,* vol. 83, no. 16, pp. 3178-3181, 1999.

[162] M.A. Little, P.E. McSharry,S. J. Roberts, D.A.E Costello and I.M. Moroz, "Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection," *Biomedinal Engineering* , vol. 6, no. 23, 2007.