Automatic power system event classification using quadratic discriminant analysis on PMU data


Published in:

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Open Access
This research has been made openly available by Queen's academics and its Open Research team. We would love to hear how access to this research benefits you. – Share your feedback with us: http://go.qub.ac.uk/oa-feedback
Automatic Power System Event Classification Using Quadratic Discriminant Analysis on PMU Data

Mark Rafferty*, Xueqin (Amy) Liu†
*System Operator for Northern Ireland (SONI Ltd.), Belfast, UK
mark.rafferty@soni.ltd.uk
†School of Electronics, Electrical Engineering and Computer Science, Queen’s University, Belfast, UK
x.liu@qub.ac.uk

Abstract—Rapid detection and diagnosis of events in power system wide area monitoring is of great interest to system operators, with event classification being a major aspect of diagnosing an event. Other event diagnostic aspects include the time the event occurred, location of the event, root cause of the event and magnitude of the event. Automatic event classification enhances the operators’ ability to identify the types of events occurring in a system quickly, which helps to assist fast decision making when restoring power to the system. This paper proposes an approach for classifying power system events, namely Generation Dip, Loss of Load and Line Trip Events, by employing Quadratic Discriminant Analysis (QDA) on Phasor Measurement Unit (PMU) data in combination with a forward selection technique. QDA is a commonly used supervised, statistical technique for data classification, and works by finding a combination of features that separates the data into different classes by modelling the difference between them. Historical power system event data is used to construct an event database, and as new events are detected the methodology automatically classifies the event based on the effect on power system variables. The reliability of the proposed method is demonstrated using simulated case studies, constructed using DigSilent PowerFactory, and real data case studies, acquired from the UK and Irish Power System.

Index Terms—Power system event classification, quadratic discriminant analysis, PMU data, power system monitoring, machine learning

I. INTRODUCTION

Increasing levels of renewable energy resources world-wide as a supplement or alternative to traditional, non-renewable energy resources, has led to a greater level of distributed generation (DG) on the system. This increase in DG is mainly facilitated by small capacity generating units not owned by the main system operator, but by independent power producers.

These renewable DGs have many benefits to the system, firstly they allow the supply of energy to areas far from load centres, relieving the requirement for transmission congestion and lowering capacity investment in transmission lines and distribution substations. This is particularly important as visible sites are normally in geographically isolated area where the construction of large substations can be challenging and expensive. Also, due to DGs serving the local area, the need for long-distance power transmission lines is made redundant and can decrease the power loss in the overall system.

However, integration of a high quantity of DGs into the power system has raised considerable concerns regarding power system operation and protection. Traditionally power flow in the system was unidirectional, from high (power stations) to low (loads), but with the introduction of DGs to the system, network topology has changed.

Another issue with connecting DG to the power system is that utility and DG protection systems must be coordinated in order to respond correctly to system events while avoiding the occurrence of nuisance trips.

As power system events are inevitable a scheme to detect and diagnose them as quickly as possible is required to lessen the damage done and restore full power to the system with minimum delay. Considering this, many countries have invested in smart grid technologies, employing a network of Phasor Measurement Units (PMUs) in Wide Area Monitoring Systems (WAMS). This allows power system dynamics to be monitored; providing a way to observe power system events and carry out retrospective analysis.

On-line power system monitoring is of great interest to system operators, and with the growth of installed PMUs in the system there is vast amounts of data to be analysed. Motivating the need for intelligent, statistical tools to extract useful information. After the occurrence of an event, it is extremely important to firstly detect the event before diagnosis can occur, a task which has been the focus of our previous work [1]–[4]. Once an event has been detected fast classification is of significant importance as it allows operators to quickly determine the nature of the event occurring, so that appropriate counter measures can be implemented to fix the problem.

In machine learning, classification [5] is the assignment of data to known groups within a system of categories distinguished by some metric or characteristic in the data. In literature many methods have been investigated for the classification of data in different domains, including support vector machines (SVM) [6], decision trees [7], artificial neural networks (ANNs) [8], kernel density estimation [9] and discriminant analysis (DA) [10].

In [3] a classification metric for 'net' system over/under frequency was presented, however this method could not classify multiple loss of load or generation events occurring in the system simultaneously; it was only able to classify the consequent 'net' over/under generation. Also, line trips were not considered. Similar work has been investigated with the focus on event classification in [11]–[13], where [11], [13] employed a supervised learning approach, such as SVM. [11]
only considers application on the frequency variable, and [13] on both frequency and voltage. [12] made use of a strong signal based approach, could be problematic - where there is a large disturbance, the PMU measurements on transient events close to the source of the disturbance could be inaccurate. In addition, both [12] and [13] require gathering event signals 0.5 seconds before and 1.5 seconds after the event. This pre-defined 2-second window length approach may not reflect the true event length.

In contrast to existing literature, this paper proposes a systematic application to frequency, voltage and phase angle signals as well as the difference and the rate of change of the corresponding variables from various locations, and employs a forward selection technique to select the variables which contribute most to event classification accuracy. The window length used in this paper is determined by an event detection algorithm described in our previous work [2], [3]. The results demonstrate that the phase angle difference, the frequency difference, and the rate of change of voltage has most contribute most to improving classification accuracy.

To achieve high classification accuracy, we have (1) developed a historical database of power system events using DigSilent Power Factory and previously event detection tools [3] to allow new detected events to be classified; (2) employed a sequential forward selection technique to systematically select the most important features for classification accuracy; (3) adopted Quadratic Discriminant Analysis which is capable of handling non-linear boundaries in comparison to Linear Discriminant Analysis for event classification; (4) validated the proposed approach with numerical cased studies based on simulated and real power system data.

II. METHODOLOGY

A. Discriminant Analysis

Discriminant analysis (DA) is a supervised machine learning technique, used to find a linear combination of features in a dataset that best discriminate between two or more mutually exclusive groups on the basis of predefined features in the data. These combinations are commonly used for dimensionality reduction before classification.

Within discriminant analysis there are two main stages, separation and allocation. Firstly, a dataset of labelled training data, X is required for the separation stage. Here the objective is to find discriminant functions that maximise the separation between the classes in the dataset. The most common way of generating the discriminant functions is by using linear methods, such as Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA), both of which are presented in this section. During the allocation stage, the objective is to assign an unclassified sample into one of the known classes, k using the discriminant functions.

A.1. Linear Discriminant Analysis - LDA

In LDA the objective is to find the posterior probability that a sample, x belongs to a class, k for classification, G. If π_k and π_ℓ denote the prior probability that a randomly selected sample comes from the k-th and l-th class respectively, and is calculated from, π_k = \frac{N_k}{\text{Total No. samples in class } k}, and let f_k(x) denote the density function of x in class G = k. From the Bayes theorem the following can be stated [14]:

$$Pr(G = k|X = x) = \frac{f_k(x)\pi_k}{\sum_{\ell = 1}^{K} f_\ell(x)\pi_\ell}$$

where X is the training dataset and is assumed to follow a multivariate normal distribution [15]. The class-conditional density, f_k(x), is then given by:

$$f_k(x) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma_k|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x - \mu_k)^T\Sigma_k^{-1}(x - \mu_k)\right)$$

where \mu_k is a class specific mean vector and \Sigma_k is the corresponding covariance matrix for class k. In the case of LDA it is assumed that the classes all have a common covariance matrix, \Sigma_k = \Sigma, \forall k, and thus the linear discriminant function can be expressed as [15]:

$$\delta_k(x) = x^T \Sigma^{-1} \mu_k - \frac{1}{2} \mu_k^T \Sigma^{-1} \mu_k + \log \pi_k$$

Classification, \hat{G}, that sample x belongs to class k is when:

$$\hat{G}(x) = \arg \max_k \delta_k(x)$$

A.2. Quadratic Discriminant Analysis - QDA

QDA is an extension of LDA, and like LDA, assumes multivariate data, following a normal distribution and again uses the Bayes theorem to calculate the probability a sample, x belongs to class k. However a separate covariance matrix, \Sigma_k, is assumed for each class, k = 1, 2, ..., K, (n = number of classes), yielding the quadratic discriminant function as [15]:

$$\delta_k(x) = -\frac{1}{2} \log |\Sigma_k| - \frac{1}{2}(x - \mu_k)^T\Sigma_k^{-1}(x - \mu_k) + \log \pi_k$$

with optimal classification again given by Eq. 4.
B. Power System Event Classifier - PSEC

The process for the proposed classification scheme is presented in Fig 1, and involves two main stages with a third stage, event detection and isolation (presented in [3], [4]) included. The two main components of the classification scheme are: 1) off-line training (event database construction); and 2) on-line automatic classification.

To construct the event database, historic PMU data of past events is required to train the model and calculate the discriminant functions. This enables the boundaries between the classes to be determined so that new events can be classified. If \( f_i, \phi_i \) and \( v_i \) denote a measurement of frequency, \( F \), phase angle difference, \( \Phi \), and voltage, \( V \), from a PMU respectively, then the change in each, \( \Delta F, \Delta \Phi \) and \( \Delta V \) is calculated from the previous \( (f_{i-1}, \phi_{i-1} \text{ and } v_{i-1}) \) values of \( F, \Phi \) and \( V \) respectively. Also calculated was the rate of change (ROC) for each variable over time, \( \frac{\Delta f}{\Delta t}, \frac{\Delta \phi}{\Delta t} \) and \( \frac{\Delta v}{\Delta t} \) respectively. A time of 0.5 seconds was chosen as \( \Delta t \). Event data is isolated from non-event data, using an appropriate isolation technique, and labelled by the user. Once an adequate number of the different events has been gathered and labelled, the boundaries between adjacent classes can be calculated using discriminant analysis.

PSEC can now be used to classify new events detected on the power system. During the monitoring of the power system, the change and rate of change for each variable are continuously calculated, and at the occurrence of an event all data available at the PMU is isolated. For each sample of the event recorded at the PMU, the posterior probability that the samples belongs to each class is calculated, with the sample classified by solving the discriminant function for each class, \( k \) and applying the classification calculation, Eq. 4. this allows PSEC to begin classification as soon as an event is detected.

III. Evaluation with Power System Data

To evaluate PSEC, dynamically simulated power system scenarios (generated on the IEEE-39 bus system [16]) and real power system data (acquired from UK and Irish power systems) were used. To generate simulated system data a PMU was placed at each bus in the system not directly connected to a generator. In reality this is neither cost effective or efficient but is used to illustrate the concept of the methodology. Real power system data was acquired through a network of PMUs implemented on the UK and Irish power system, [17].

Simulated case studies consist of 33 single line trip, 132 generation dip and 112 loss of load events of varying severity. Fig 2 shows typical frequency, phase angle difference and voltage waveforms for the aforementioned events. A number of real cases, including 4 line trip, 4 generation dip and 4 loss of load events were selected for test in this paper.

A. Construction of the PSEC Event Database

Construction of the event database requires an adequate number of cases for each event. Therefore for each type of event two sub groups were constructed, training and testing, with a random 50-50 split between training and testing observed. As previously highlighted, in Fig 2, the locality of the PMU to the event dictates the severity of event on system variable measurements and therefore, it is important to use all of the system data during an event for the event database.

Evaluation of PSEC Accuracy

Once the event database has been constructed it is beneficial to evaluate it on the training dataset used in its construction. This allows the true error rate to be measured and the classifier error rate on the entire population to be calculated.

The \( K \)-fold method [18] was used to cross validate the PSEC classifier, checking how well the classifier will generalise to new data. The \( K \)-fold method works by dividing the training datasets into \( k \) subsets (folds), with all bar 1 of the folds used for training the classifier and the remaining folds used for testing. This process is repeated until all the folds have been used for testing. The cross validation true error, \( CV_E \), is calculated by finding the mean of the errors calculated for each trial, \( E_k \), and is given by:

\[
CV_E = \frac{1}{K} \sum_{k=1}^{K} E_k
\]

An advantage of this method is that it allows all the samples in the dataset to be used for both training and validation, with each sample used for validation only once. A common value for the number of folds used, \( k \), is 10.

Selection of Variables for PSEC Construction

In machine learning variable selection is a process where subsets of available variables are chosen which contribute most to model accuracy, with the remaining, unimportant features, which have little or no predictive information, discarded [19].
Inclusion of too many variables can decrease a classifier's performance with respect to computation and recognition accuracy. However, including too few will also result in lower accuracy.

Sequential forward selection (SFS) [19] was used on the training dataset, yielding the results displayed in Fig 3. SFS works by selecting at each step the variable that yields the lowest cross validation error. The process of adding variables to the previously selected ones is repeated until either a predefined number have been selected, or the cross-validation error begins to increase. From Fig 3, it can be observed that $\Delta \Phi$ was selected first, followed by $\Delta f$ and finally $\Delta v$. To illustrate the concept, a 2-dimensional representation of 2 variables ($\Delta f$ and $\Delta \phi$) is shown in Fig 4 (an extra class of normal data is also included).

### B. Classification of Detected Events

After construction of the event database, PSEC was evaluated using the remaining case studies to gauge it’s performance with data it has no prior experience with. Total classification results for each type of event are displayed in TABLE I and II. Also included in TABLE I are results for a LDA based classifier. This will be discussed in Section IV. Fig 5 shows, conceptually: (a) results for a number of the simulated case studies; and (b) results for real power system data events; for each type of event, in relation to the class boundaries calculated in Section III-A and depicted in Fig 4.

From Fig 5 (a) it can be seen that during line trip events, some samples are located inside more than one class, this is due to the amount and direction of power flow during a line trip. If a line trip occurs when a large amount of power is flowing through it, the bus exporting power experiences a form of load loss (increasing the frequency). However, using the large line trip in Fig 5, as an example it can be seen that the majority of samples are located inside the line trip class. Also from Fig 5, the affect large system events have on successful classification can be seen, with large events very easily classified. If a small loss (load or generation) is experienced the closer to normal power system conditions the system is currently experiencing and thus the more challenging it is to successfully classify the event. TABLE I provides full classification results for the simulated case studies.

Full testing results for the real power system event case studies, captured on the UK and Irish power systems, are portrayed in TABLE II, whilst Fig 5 (b) depicts three of these (one of each type) to show conceptual results for the classifier. The events depicted in Fig 5 (b) occurred when (a) an interconnector between one of the Northern Isles of Scotland and the Scottish mainland was lost, whilst it was exporting power, leading to a shortfall in generation being witnessed at the PMU situated on the Isle; (b) a planned disconnection of a DC inter-connector between Ireland and the UK was carried out whilst it was exporting 246 MW of power to the UK leading to the PMU located in Ireland recording a loss of load; and (c) an islanding event occurred at a site in the Northern Isle, whilst the other PMUs in the system recognise the line trip event occurring that leads to the island. TABLE II summarizes the classification results for the real data case studies used to test the classifier. As can be seen each event tested was successfully classified by PSEC.

### IV. Discussion

The accuracy of LDA and QDA based classification for simulated data is portrayed in TABLE I, with accuracy defined
as the number of successful classifications expressed as a percentage of the total number of classifications made. This illustrates that for classification of generation dip and loss of load events both methods are quite accurate (100%). However, LDA fails short when classifying line trips, with a successful classification rate of only 17.85%, compared to 96.77% when using QDA. Results for the real system data used to test PSEC (presented in TABLE II) indicate that the proposed method shows good promise for real world implementation. As mentioned previously, PSEC is to be implemented at the PMU, meaning for each event all PMUs in the system will experience some form of event; with those closer experiencing it to a greater extent. Using event data from all PMUs in the system for the event database allows event locality to be taken into consideration and makes classification possible even if the event occurs at the other side of the system. Finally, implementation at the PMU level also allows automatic classification as the data is measured at the PMU, with reduced time delay.

Through the work carried out in this paper, our previous work [3], [4] and trends in the industry for automated control, the potential of PMUs for Wide-Area Monitoring and Control is highlighted. Through the detection and diagnosis techniques proposed, it is illustrated that PMU data applications can move away from historical, off-line analysis, which they are currently utilized for. Specifically, the output of the detection and classification schemes can be used as input knowledge for intelligent control system to assist with real-time decision-making in power system operation, and thus methods to restore full power to the system based on the current power system conditions. Once such example application could be reducing the output power from generators once a loss of load event has been successfully detected and classified or discharging a large battery energy storage system (BESS) when a generation dip has occurred.

The case studies also highlighted limitations of the PSEC methodology. Firstly, a small loss (load or generation) leads to the power system measurements being closer to normal power system conditions and making successful event classification more challenging. However, as the methodology is intended to be used after an event has been detected the risk of classifying an event as normal system operation is mitigated. Also this methodology adopts a static database, however it would be beneficial, and relatively straightforward, to include an adaptive methodology to allow the retraining of the database to include new, successfully classified events in the system.

V. CONCLUSIONS AND FUTURE WORK

In this paper an automatic power system event classifier (PSEC) was proposed for a number of regularly occurring power system events. The methodology was based on quadratic discriminant analysis and used wide-area synchronized frequency, phase angle difference and voltage measurements collected from a network of PMUs located on a power system. The methodology is extensively tested for a large number of simulated and real data case studies, and results are presented.

From the results it can be seen that this method is able to classify generation dip and loss of load events with high accuracy, however this is not the case for line trip events (96.77% accurate) and for future work other techniques will be investigated to increase the accuracy of the classifier for these type of event.

More specifically, future work will look at enhancing a number of aspects of the classifier. Firstly as mentioned previously the intended location for this methodology is at PMU level in the system, hence modifications need to be incorporated to allow the methodology to work on a wide area level. Secondly, a variable selection technique (SFS) was used to determine the inputs to the classifiers, but other dimensionality reduction methods may improve classification accuracy and reduce the training and testing time complexity of the method. Also, the proposed method utilizes a static training scheme, therefore a potential improvement to classification accuracy could occur when an adaptive event database is implemented. This would allow self training of PSEC after the successful classification of an event. Finally, improvements and expansion to the simulated power system event database will be conducted. These enhancements will include the generation of more complex scenarios including events under different operating conditions (i.e. a number of different load profiles will be generated for each load and different wind penetration levels) and the addition of other event types (e.g. BESS charging/discharging, capacitor switching and motor start events).

ACKNOWLEDGMENT

The research is supported by a British Council Newton Institutional Links Programme grant with Helwan University, Egypt and a PhD studentship from the Department of Education and Learning, Northern Ireland.

REFERENCES


