Supervisory Protection and Automated Event Diagnosis Using PMU Data

Milan Biswal, *Member, IEEE*, Sukumar M. Brahma, *Senior Member, IEEE*, and Huiping Cao

***Abstract*—This paper presents a new framework for supervisory protection and situational awareness to enhance grid operations and protection using modern wide-area monitoring systems. In contrast to earlier approaches dealing with the combined processing of data from multiple phasor measurement units (PMUs), the proposed approach analyzes only the PMU data with the strongest or the most prominent disturbance signature. The specific contributions of this paper are: a) new criteria for identification of PMU with the strongest signature, b) simplified approach for quick detection of faults, c) early classification of eight other disturbances suitable for near real-time response, d) time-frequency transform-based feature extraction techniques for speedy and reliable classifiers, and e) a promising approach to locate disturbances within narrow geographical constraints. The contributions are verified with exhaustive simulation data from the Western Electricity Coordination Council system model and limited real PMU data.**

***Index Terms*—Classification, feature extraction, phasor measurement units, power system disturbance, wide-area monitoring systems (WAMS).**

I. Introduction

IT was observed after the North American blackout in August 2003 that effective wide area system visualization can make power systems more resilient to large scale outages [1]; therefore utilities have been upgrading their measurement infrastructure by deploying more PMUs. By 2014, over 1100 PMUs have been installed in North America, and the number keeps increasing [2]. It is expected that near real-time detection and identification of disturbances will be used for power system protection and control [3]. Researchers have begun to investigate automatic methods for the detection and identification of disturbance events using PMU data in an effort to increase visualization of power systems [4]–[8].

The current state of the art of protection and control applications can benefit from detection of relay misoperations, a problem that has remained elusive to engineers. *Misoperation* is a term that indicates that a relay has operated when there was no fault in its operating zone. Certain misoperations have historically occurred due to defective relay components which have also been called *hidden faults* [9]. There have been documented instances where a misoperation due to a hidden fault either initiated or contributed to system instability, resulting in a blackout [9], [10]. Another type of misoperation that routinely occurs during early stages of a blackout is *load encroachment*, where a distance relay misinterprets overload as a line fault. Event detection and identification using PMU data can play an important role in verifying a relay operation. If a relay operated, but no fault was identified by the event identification tool, this operation can be flagged as misoperation.

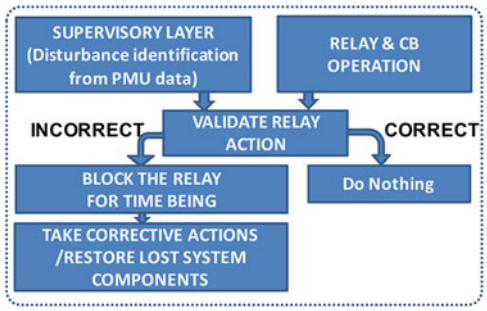

Fig. 1. Supervisory protection system.

Our previous research work elaborates this idea to outline the concept of supervisory protection, where event identification using PMU data can form a *supervisory layer* that verifies every relay operation, and informs the operators if a misoperation is detected [10]. Depending on the state of the system (moderately or heavily loaded), a judicious response can be initiated, either manual or automated. The functional block diagram of a supervisory protection system is shown in Fig. 1. Blackout logs show that the time between misoperations and eventual instability ranged from a few seconds to several minutes or more [10], so speed is an important aspect of the proposed concept.

Thus disturbance detection and classification form the underpinnings of both supervisory protection and visualization. Detection is performed in PMUs based on trigger-thresholds specified in section R8.2 of PRC-002-2 [11], the dynamic disturbance monitoring standard published by North American Reliability Corporation (NERC). The standard recommends the triggering criteria for disturbance detection in WECC system, as either of the following: 1) frequency <59.55 Hz or >61 Hz, 2) rate of change of frequency <-0.05625 Hz/s or >0.124 Hz/s, 3) Under voltage trigger set no lower than 85% of the normal operating voltage for a duration of 5 seconds. However, as this study will reveal, these values are too conservative, and result in very few disturbances being detected. For finer visualization, more sensitive triggers have been proposed [12], [13].

Once detected, the events need to be classified as quickly as possible. The speed of classification will decide if the

Manuscript received June 30, 2015; revised November 21, 2015; accepted January 09, 2016. Date of publication January 22, 2016; date of current version July 21, 2016. This work was supported by the National Science Foundation, CREST under Grant HRD-1345232. Paper no. TPWRD-00834-2015.

M. Biswal and S. M. Brahma are with the Klipsch School of Electrical and Computer Engineering, Las Cruces, NM 88001 USA (e-mail: milanb@nmsu.edu; sbrahma@nmsu.edu).

H. Cao is with the Computer Science Department, New Mexico State University, Las Cruces, NM 88003 USA (e-mail: hcao1@nmsu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPWRD.2016.2520958

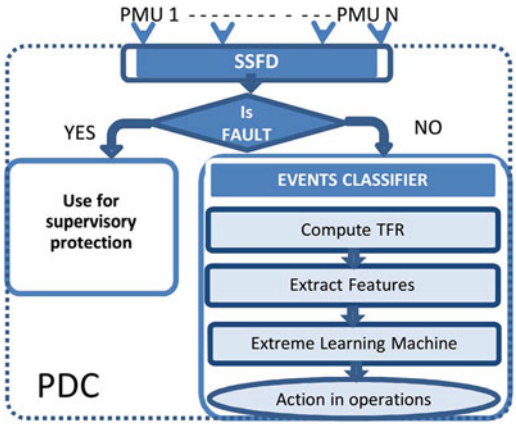

Fig. 2. Flowchart of the proposed PMU data processing function.

visualization can actually be used for supervisory protection, special integrity protection schemes (SIPS), and control applications. Recent studies [4]–[8] have addressed the classification issue using PMU-generated disturbance files, but they simply target major disturbances like faults and/or generation loss. Real power systems have many other types of events, and in order to prove robustness of classification a large variety of events should be considered. None of these studies address speed of classification, which is crucial to evaluate their suitability for protection and control. Lastly most of the papers [4]–[7] use data from multiple PMUs for creating features, thus ignoring the inherent spatial correlation among PMU data.

These drawbacks in the current research provide the motivation for this paper. We propose an efficient scheme for supervisory protection and near real-time automated disturbance diagnosis utilizing voltage and frequency data from the PMU that has the most prominent effect on its measurements due to a particular disturbance, or in other words, has the "strongest" disturbance signal. This scheme is tested with real as well as simulated PMU data. Features extracted from time-frequency representation (TFR) of the voltage and frequency waveforms corresponding to the strongest signal are used to classify *nine* events, including faults, and locate them in narrow geographical regions. A one-versus-many Extreme Learning Machine (ELM) classifier is used as the classification tool because of its computational simplicity, and speedy but accurate classification performance.

II. Overview of the Proposed Methodology

Fig. 2 depicts the flowchart of the proposed methodology. The input data streams from multiple PMUs arrive at a Phasor Data Concentrator (PDC) after some inherent delay in communication. These PMU data are then fed to the strongest signal selection and fault detection (SSFD) module. The SSFD module, after preprocessing the signals by removing bad data, identifies the data stream with the highest strength, which then is used to segregate the faults from all other types of disturbances. If the detected disturbance is not a fault, then the strongest PMU data stream is fed as input to the Events Classifier module. This approach ensures the quick identification of faults for supervisory protection.

A. Strongest Signal Selection and Fault Detection (SSFD)

Although a PMU records multiple variables such as voltage phasors, current phasors, and frequency, we have shown in [14] that the voltage magnitude and frequency data carry sufficient information for event recognition. We use the voltage magnitude alone to find out the strongest signal and the associated PMU, and to detect faults.

Let Pv_i be the discrete series containing the measured voltage-magnitude values recorded by the ith PMU. Also suppose that an event is triggered at the kth sample in this series. The first step of the SSFD module is to identify the PMU with the strongest disturbance signal.

Let n be the number of samples after triggering of a PMU. The positive and negative deviations of the voltage signal n samples after triggering for the ith PMU are calculated according to (1) and (2) respectively, which define the positive $(+Dev_i(n))$ and negative $(-Dev_i(n))$ deviations of the nth sample from the base voltage V_i^{rms} of the ith PMU.

$$+Dev_i(n) = \begin{cases} Pv_i(k+n) - V_i^{rms} & \text{if } Pv_i(k+n) \geq V_i^{rms} \\ 0 & \text{else} \end{cases} \tag{1}$$

$$-Dev_i(n) = \begin{cases} V_i^{rms} - Pv_i(k+n) & \text{if } Pv_i(k+n) < V_i^{rms} \\ 0 & \text{else} \end{cases} \tag{2}$$

V_i^{rms} is calculated as the averaged rms value of the measured voltage magnitude by the ith PMU over the 10 cycles before the triggering point. 10 cycles are generally used in measurement aggregation for power signal analysis [15]. Thus, V_i^{rms} is defined as (3).

$$V_i^{rms} = \frac{1}{10}\sqrt{\sum_{\sigma=k-10}^{k} Pv_i(\sigma)^2} \tag{3}$$

A new metric based on energy deviation, referred to as Cumulative deviation in Energy (CE), is defined for quantifying the strength of the signal recorded by a PMU. The CE for the ith PMU, n samples after triggering is expressed as:

$$CE_i(n) = \sum_{m=1}^{n} (+Dev_i(m))^2 + \sum_{m=1}^{n} (-Dev_i(m))^2 \tag{4}$$

The strongest PMU signal is the signal with the largest CE. CE can also be used to identify faults, but since the voltage was always observed to drop and stay below V_i^{rms} until cleared (due to obvious reasons), we use only the negative deviations to identify faults, and define cumulative Fault Energy (FE) as:

$$FE_i(n) = \sum_{m=1}^{n} (-Dev_i(m))^2 \tag{5}$$

Algorithm 1: Detection of Strongest-signal PMU and Fault.

Input: $W_L, \tau_{fe}, \tau_{ne}$
Input: Areas ids and zone ids of all the PMUs
$W_f = 1 \leftarrow$ *Fault analysis window Length*
$M \leftarrow$ *Number of PMUs*
while $W_f \leq W_L$ **do**
  Compute: $CE_i(W_f), i = 1, 2...M$
  Find $I = \arg\max_i(CE_i(W_f))$
  (i.e., $I \leftarrow$ the index of PMU with the strongest signal)
  if $FE_I(W_f) \geq \tau_{fe}$ **then**
    FLT STATUS $= 1$
    Fault Area $\leftarrow$ Area where ith PMU is placed
    Fault Zone $\leftarrow$ Zone where ith PMU is placed
  else
    FLT STATUS $= 0$
  $W_f = W_f + 1$
  while FLT STATUS $= 1$ **do**
    If $FE_I(W_f) - FE_I(W_f - 1) \leq \tau_{ne}$ **then**
      FLT CLR STATUS $= 1$
    else
      FLT CLR STATUS $= 0$

Algorithm 1 shows the steps to identify the PMU that captures the strongest signal of a disturbance, and to detect whether the disturbance is fault or not. It takes as input a maximum window size containing W_L samples. The fault analysis window contains (W_f) samples, where the sample number is incremented in each iteration until it reaches the maximum value W_L. W_L determines the maximum duration up to which the possibility of occurrence of fault needs to be monitored. The value of W_f defines the number of post triggering samples after which the fault gets detected. The threshold value of FE (τ_{fe}) is fixed as the minimum value of the FEs calculated from a set of strongest PMU signals in the field data obtained from a utility in the United States from 2007 to 2010. If a fault is detected, the information is passed for appropriate supervisory protection. The algorithm also monitors whether the fault is cleared, by comparing the change in FE with another pre-specified threshold value τ_{ne}. This threshold value should not be larger than the square of the maximum deviation of the voltage values allowed in the grid during stable operation [16]. The exact values of W_L, τ_{fe} and τ_{ne} are given in Section VI. The region in which a fault occurs are predicted from the area and zone information of the PMU which captures the strongest signature for the fault.

B. Events Classifier Module

The primary functionality of the Events Classifier Module shown in Fig. 2 is to classify different disturbance types by analyzing both, voltage and frequency signals. There are two key challenges in classifying disturbance events: 1) choosing an appropriate representation of the data in feature space, discussed in Section III, and 2) selecting a suitable classifier, discussed in Section IV.

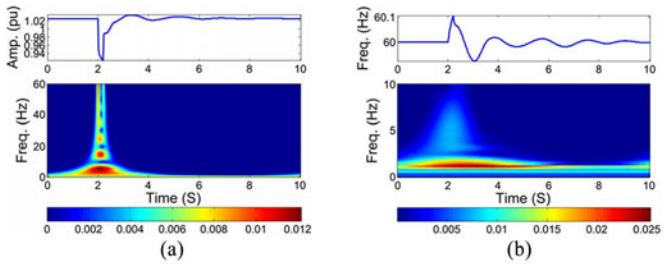

Fig. 3. TFR of voltage and frequency waveforms during a Fault.

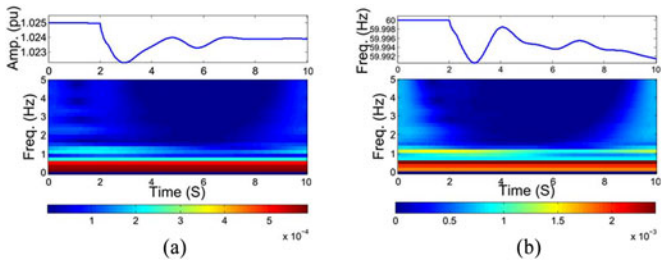

Fig. 4. TFR of voltage and frequency waveforms for loss of generation.

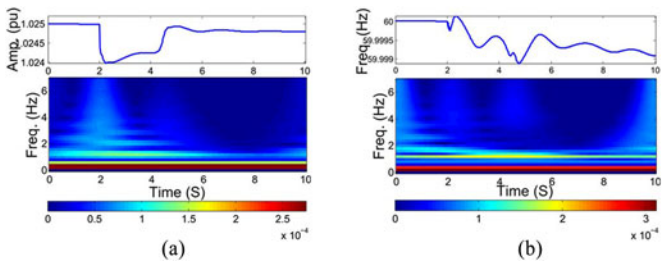

Fig. 5. TFR of voltage and frequency waveforms for load switching ON.

III. Time-Frequency Analysis for Feature Extraction

The variations in voltage and frequency for different power system events often appear similar. For example, loss of generation may create waveforms that look similar to (load switching ON(see Figs. 4 and 5), making classification difficult simply by inspection. However, as the physics behind the events is different, their voltage or frequency waveforms are expected to carry unique information. Time-Frequency analysis (TFA) techniques are sophisticated signal processing methods to decompose any signal into time varying spectral components or modes. These spectral modes capture time varying statistical properties of a signal. Since the differences in the waveforms corresponding to different events can be quantified in terms of statistic-based metrics, TFA is a reasonable approach to create features for this study. However, a fast TFA method is required to minimize the delay. Therefore, a low complexity FDST (Fast variant of Discrete S-Transform) algorithm for calculating time-frequency representation (TFR) of any non-stationary signal proposed in [17] is chosen.

A. Fast Variant of Discrete S-Transform (FDST)

FDST is a linear time-frequency transform which is an extension from the wavelet class of algorithms, with specific merits

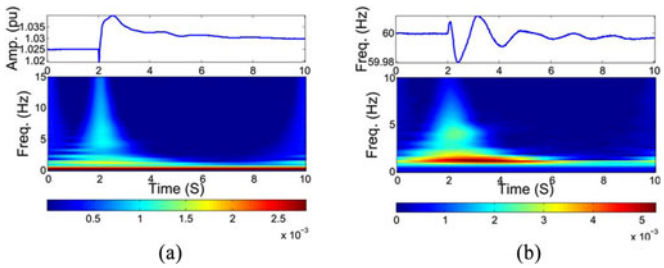

Fig. 6. TFR of voltage and frequency waveforms for shunt capacitor switching ON.

such as fast computation, absolute phase reference and Fourier basis modeling. The FDST is capable of segregating a disturbance waveform in terms of oscillatory modes which correspond to a range of Fourier frequencies.

Let $x(n)(n \in [0, N-1])$ be a value in a waveform with N samples recorded by a PMU, with a sampling frequency of F_s Hz. The frequency domain expression for FDST of $x(n)$ is expressed as

$$S(n,k) = \begin{cases} \frac{1}{N}\sum_{k'=0}^{N-1} & \left(X(k') \times W((k'-n),k)\right) e^{\frac{j2\pi kk'}{N}}, \\ & \text{if } k' = k \text{ and } X(k') \geq \alpha \\ 0 & \text{Otherwise} \end{cases} \quad (6)$$

The meaning of each component in this equation is explained as follows. First, $n = 0, 1, \ldots, (N-1)$ denoting the index of the original sample at time t (in s), thus $n = t.F_s$. And $k = 0, 1, \ldots, N/2$ denoting the index of the sample in the frequency domain for $f \in [0, F_s/2]$ in Hz, thus $k = N \cdot f/F_s$.

Second, $X(k')$ is the Discrete Fourier Transform (DFT) of $x(n)$ and is expressed as

$$X(k') = \sum_{n=0}^{N-1} \left(x(n)\, e^{-\frac{j2\pi k'n}{N}}\right), \; k' = 0, 1, \cdots, (N-1) \quad (7)$$

Third, $W(k', k)$ is a discretized frequency domain normalized Gaussian window function which is expressed as:

$$W(k', k) = e^{-\frac{1}{2}\left(\frac{2\pi k'}{k}\right)^2}, \; k' = 0, 1, \cdots, (N-1) \quad (8)$$

The parameter α is a threshold, whose value is zero in our analysis to preserve all the frequency components. Fig. 3 shows the TFR of the voltage and frequency waveform when a fault happens. Figs. 4 and 5 show the TFR of the voltage and frequency waveforms for generation loss and load switching on (LS-On) events respectively. The distinctive spectral information can be observed from the visual distinction among the plots. Though both the generation loss and LS-On result in the reduction of voltage and frequency values, their spectral variations bear different signatures in the TFR plane. The disturbance waveform and TFR corresponding to a shunt capacitor switching ON event is shown in Fig. 6.

B. Time-Frequency Features

The TFR representation for a length-N waveform is a $N \times (N/2+1)$ matrix S. Entities of S are complex numbers representing amplitude & phase information corresponding to a time and frequency. Because of the big number of dimensions of S, it is not appropriate to feed the matrix S to the classifier. Therefore, relevant features need to be extracted from S, to reduce the number of dimensions and to remove the irrelevant and redundant information, thus achieving better classification performance. In addition, features may carry intuitive information regarding the nature of disturbances which is not perceptible from the raw data. We adopt a feature selection strategy to encode the unique characteristics of the disturbance types in the time-frequency domain. An exhaustive number of candidate features were evaluated. Out of these features, *six* $(F_1 - F_6)$ were found to yield the best classification result when evaluated with *k*-Nearest Neighborhood (kNN) classifier. The reason for employing kNN classifier for selecting the best set of features is as follows. If the patterns belonging to the same class are densely located and lie far apart from the cluster of patterns corresponding to other classes, kNN shows superior performance. kNN classification provides an intuitive insight regarding the performance of features, but it may not generate highest classification accuracy. As it is discussed in Section IV, the performance of ELM classifier is superior to kNN, and is recommended for classification of power system events. We use kNN here for the sole purpose of feature selection.

The final set of selected features are described in the following sub-sections.

1) Cumulative Energy Change (CEC): CEC feature is defined as the difference between the cumulative TFR energy after and before the occurrence of a disturbance. It is calculated using (9) and is used to identify whether a disturbance results in an increase or a decrease in the waveform energy.

$$F_1 = \sum_{n=1}^{N_d}\sum_{k=0}^{N/2}\left[S_d(n,k)\right]^2 - \sum_{n=1}^{N_n}\sum_{k=0}^{N/2}\left[S_n(n,k)\right]^2 \quad (9)$$

where, $S_n(n,k)$ is the absolute value of TFR during normal state (or before triggering of a disturbance) and $S_d(n,k)$ is the absolute value of TFR for the segment after triggering a disturbance. N_d and N_n are the length of the segment before and after the trigger, respectively.

2) Dominant Spectral Component: This feature is the most dominant frequency component in the TFR. The motivation for this feature comes primarily from the nature of frequency waveforms which can be decomposed into multiple oscillatory spectral components. The dominant oscillatory mode or spectral component can be used to distinguish among different types of disturbances. It is calculated according to (10).

$$F_2 = \; k | \forall i \in \{0, 1, \ldots, N/2\} : S_d(n,i) \leq S_d(n,k) \\ \text{and } n = 1, 2, \cdots N_d \quad (10)$$

3) Instantaneous Distortion Factor (IDF): IDF feature is defined as the ratio of the summation of all the non-dominant spectral components to the intensity of the most dominant

component. It quantifies the instantaneous distortion in the voltage and frequency waveforms due to the presence of multiple non-dominant spectral components. IDF is calculated using (11) and (12).

$$S_d^{\max}(n) = S_d(n,p|\,\forall k \in \{0,1,\ldots,N/2\} : S_d(n,k) \le (S_d(n,p)) \quad (11)$$

$$IDF(n) = \frac{\sum_{k=0}^{N/2}(S_d(n,k))^2}{(S_d^{\max}(n))^2}, \quad n \in \{1,\ldots,N_d\} \quad (12)$$

The IDF(n) is a vector representing the IDF at every time sample n. We choose the maximum and minimum value of the IDF vector as the features. Therefore, the features are expressed as: $F_3 = \max(IDF(n))$ and $F_4 = \min(IDF(n))$.

4) Time-Frequency Flux: Spectral Flux is a measure of rate of change of spectral content of a signal over time. The spectral flux was extended in [18] to measure the variation of the signal in both the time and frequency domain, referred to as time-frequency flux. It encodes the information regarding rate of change of the temporal as well as spectral content reflected in the TFR, and is defined in (13).

$$F_5 = \sum_{n=1}^{N_d-l}\sum_{k=0}^{N/2-m} |S_d(n+l,k+m) - S_d(n,k)| \quad (13)$$

The parameters $l < N_d$ and $m < N/2$ define the window length along the time and frequency axis, over which the rate of change in TFR is computed. Large value of these parameters may neglect the fine changes in TFR, and very small values may result in magnification of very small undesired changes. We have found $l = 10$ and $m = 10$ to be a good choice, for analyzing the PMU data.

5) Skewness of TFR: Skewness is a statistical feature that measures the asymmetry of any distribution around the mean value [18]. This feature is chosen because, with high probability, different disturbances have unique skewness values. It is calculated using (14).

$$F_6 = \frac{\frac{1}{N_d}\sum_{n=1}^{N_d}\sum_{k=0}^{N/2}\left(S_d(n,k)-\bar{S}_d(n,k)\right)^3}{\left[\frac{1}{N_d-1}\sum_{n=1}^{N_d}\sum_{k=0}^{N/2}\left(S_d(n,k)-\bar{S}_d(n,k)\right)^2\right]^{3/2}} \quad (14)$$

IV. Classification Algorithm

Features $F_1, \cdots, F_6$ extracted from the TFR are fed to a classifier. We first evaluate the kNN classifier because of its simplicity and wide use in time series analysis. However, the classification results with kNN shown in Section VI, so we have to choose a new classifier.

A performance comparison of several established classifiers for classifying power system events is presented in our previous work [10]. The Support Vector Machine (SVM) was found to be the best, having the highest classification accuracy. However, SVM has high computational cost and has extensive memory requirements, so we propose to use a multi class Extreme Learning Machine (ELM) classifier. ELM is similar to feed forward neural network (FNN), but with a linear learning approach and is capable of achieving generalization bound quickly [19]. In ELM the connection weights between the input and hidden layers (w_i^I) are random values which are fixed in the learning process and the weights between the hidden layer and the output nodes (w_i^H) are adjustable. The number of hidden layer nodes is a parameter that can be adjusted to give the best predictive performance [20]. We use a one versus many approach for classification with ELM. The extracted feature vector $F \subseteq \mathbf{R}^6$, consisting of *six* features derived from voltage waveforms and another *six* features derived from the frequency waveforms, are used as input to the ELM classifier.

The hypothesis space of ELM is expressed as:

$$H_{\psi,V}^j = \sum_{i=1}^{V} w_i^H \psi\left(w_i^I, F\right) : w_i^H \in \mathbf{R}^V \quad (15)$$

Here, the values of $w_i^I \in \mathbf{R}^L$ are random numbers between 0 to 1, drawn from an independent and identically distributed (i.i.d.) distribution. The scalar value L denotes the number of input nodes, which is *six* in our case. The scalar value V denotes the number of nodes in hidden layers. The function $\psi(.) : \mathbf{R}^L \times \mathbf{R}^V \mapsto \mathbf{R}$ is a non-linear activation function. The choice of activation function simplifies the nonlinear classification problem and the linear learning approach significantly minimizes the computational cost [19]. The theoretical justification of ELM in achieving generalization bound while using different types of activation functions (e.g., sigmoidal, polynomial and Nadaraya-Watson) have been extensively discussed in [21]. For all the simulations in this paper, we use a sigmoidal activation function. Let J be the number of output nodes and $y_i \in \mathbf{R}^J$ be the desired output of ELM. The unknown weights $w_i^{H*} \in \mathbf{R}^V$ are estimated by solving a linear optimization problem ((16)).

$$w_i^{H*} = \arg\min_{w_i^H} \sum_{j=1}^{J} |y_j - H_{\psi,V}^j|^2 \quad (16)$$

V. Disturbance Dataset

In order to carry out a comprehensive evaluation, we used two datasets consisting of actual field data and simulated data. The actual field data over four years (2007–2010) were collected from four PMUs located in the 345-kV network of a Western Electricity Coordination Council (WECC) member utility in the United States. The detailed description of preprocessing of the data has been explained in a previous work by this research group [14]. However, this dataset has class labels available for only 58 fault events and 23 generation loss events. The unavailability of class labels for other events is thus the major issue with the field data.

A dataset with sufficient number of disturbance events under various operating conditions is needed for unbiased assessment of the proposed approach. The positive sequence load flow (PSLF) dynamic simulation tool from General Electric (GE) was used for creating a dataset of simulated PMU events belonging to different event classes. In order to simulate actual operating conditions of the grid, we took a load flow base case for the WECC system, tagged as 2008 Heavy Summer, which can

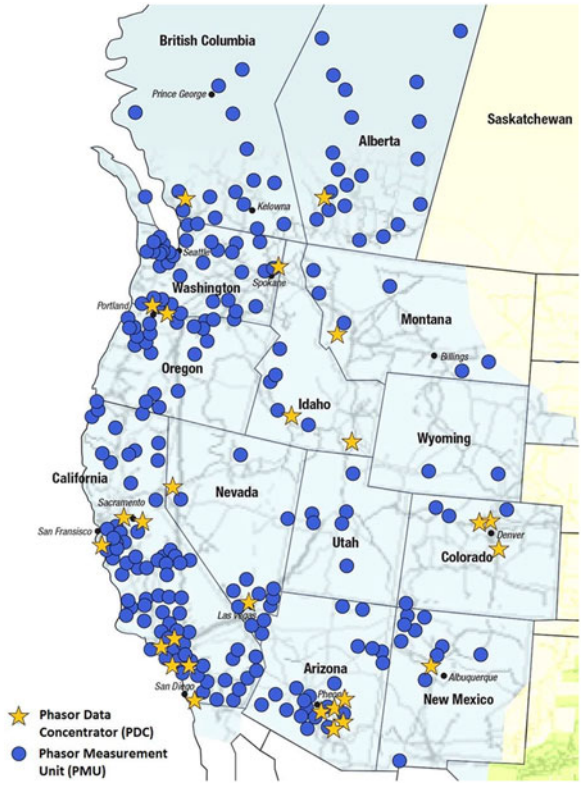

Fig. 7. Map showing PMU locations in WECC up to October 2013 [24].

be loaded into PSLF. A total of 299 time-synchronized voltage and frequency probes were placed at key buses corresponding to simulate the approximate PMUs located in the WECC system. However, since the exact buses where the PMUs are placed in WECC are not available in public domain, we utilized the PMU siting criteria [22], the PMU location map shown in Fig. 7, and information of principal transmission lines in WECC [23] to choose the exact buses for placing PMUs. The PMUs were located at the major generating stations, key substations in high voltage (240 kV and above) transmission system, major load centers, and inter-connections. This reasonably ensures that our measurements correspond to the measurements from actual PMUs in the WECC system.

The simulation time was fixed for 15 s. The disturbance in the system was triggered just after 5 s, resulting in 10 s post disturbance data. The sampling rate of all the PMUs were fixed at 240 samples/sec. The data were down-sampled to 60 samples/s taking into consideration the sampling rate of the PMUs currently available in market. Test cases for fault (FLT) were simulated at the buses with base voltage greater than 120 kV. The fault duration was kept 0.15 s. For simulating the loss in generation (GNL) the generators above 300 MW were switched off at the trigger point. Synchronous motors were switched off one at a time to create events labeled as synchronous motor-off (SM-off). The load switching off (LS-off) events were created by switching off major loads (with both real and reactive power components) with real power above 10 MW. For creating the load switching on (LS-on) event, a particular load was switched off first, and the load flow case was solved. Then the dynamic simulation was initiated and the disturbance was created by switching on the load at the trigger point. The shunt capacitors (SHC) and shunt reactors (SHR) were switched on and off in similar way to create additional data. Two classes were created from these data: 1) reactive power injected into the grid (RIN), and 2) reactive power extracted from the grid (REX). The SHC-on and SHR-off disturbances were grouped under RIN, whereas SHC-off and SHR-on were merged to form REX. The voltage phasors and frequency corresponding to each disturbance event were recorded by all the 299 probes. A total of 3495 events were simulated. However, in real world, not all these events can be detected by all PMUs. It depends on the threshold set inside PMUs. As mentioned in Section I, due to the too conservative thresholds suggested by the NERC PRC-002-2 standard, there have been some attempts to design other detection criteria. A rather complicated Fast Fourier transform, Yule-Walker, and matrix-pencil-based methods to detect power system events have been presented in [12], whereas a method based on voltage phase-angle difference among PMUs is suggested in [13], which requires signals from multiple PMUs. Since we have already used a simple entity CE, which is defined in Section II-A, as a measure to quantify the strength of a disturbance signal, we also use it to create a threshold for disturbance detection. CE is continuously evaluated for the recorded data and an event is registered when CE crosses a specified threshold. The total numbers of events from each type that get detected for different choices of CE-thresholds are summarized in Table I. Clearly the NERC standard based detection is the most conservative criterion. We consider the events detected by different criteria when we test our approach, as described in Section VI.

TABLE I
DISTURBANCE STATISTICS

	Event Detection Criteria					
	All Events	CE $\geq$ 0.0001	CE $\geq$ 0.0005	CE $\geq$ 0.001	CE $\geq$ 0.002	NERC PRC-002
FLT	1260	1260	1260	1260	1260	1032
GNL	558	503	258	205	150	8
LS-off	348	337	149	88	69	0
LS-on	354	314	151	148	107	0
SHC-off	369	346	213	175	133	0
SHC-on	315	306	183	154	117	2
SHR-off	127	123	122	122	118	11
SHR-on	128	128	126	126	120	3
SM-off	36	36	32	26	14	11
TOTAL	3495	3353	2494	2304	2088	1067

VI. RESULTS AND DISCUSSION

The first step (SSFD module) of the proposed method as explained in Section II-A, is to detect faults for supervisory protection. The detection results with both the simulated and field data are presented in Table II. The FE threshold τ_{fe} was fixed at 0.01 (rationale given in Section II-A). The value of W_L was chosen to be 15. Table III shows the time taken to detect faults. Faults were identified within maximum 8 samples

TABLE II
PERFORMANCE IN FAULT DETECTION

	No of fault Instances	No of faults Detected	Fault status Update	Correctly Predicted Zone Count	Correctly Predicted Area Count
SIMULATED	1260	1230	1230	979	1230
ACTUAL	58	55	55	n.a.	n.a.

TABLE III
FAULT DETECTION TIME

	Time in Seconds						
	0.0333	0.05	0.0667	0.0833	0.1	0.1167	0.1333
SIMULATED	1224	2	0	1	0	1	2
ACTUAL	18	18	7	4	6	1	1

$(8/60 = 0.1333$ s). The value of τ_{fn} was chosen to be 0.0025, allowing a maximum of 5% deviation of the voltage during normal operation [16]. As can be seen from Table II, out of 1260 fault events, 1230 events were correctly classified resulting in an accuracy of 97.61%. The status update upon clearance of the fault was correctly updated for all the events. As shown in Table III, out of the correctly detected faults 99.5% took only 0.0333 s for detection, with the worst case detection time being 0.1333 s.

While determining the zone and area where the faults occurred, we simply assumed the disturbance events fall in the zone and area in which the PMU that provided the strongest signal is located. The fault zone and area were correctly identified for 82.03% and 100%, respectively of the detected faults. Here, the zone and area are defined according to the WECC records available in the load flow base case. This approach has a limitation. Sometimes, if the PMU is located near the boundary of a zone/area, it is possible that a disturbance taking place at the other side of the zone/area boundary may provide the strongest signal to this PMU. This will result in misclassification of zone/area. We still adopted this approach because the exact physical locations of the PMUs, say in terms of GPS coordinates, were not available in public domain. With such information, when the distances between PMUs would be known, the fault can be better localized. However, this exercise shows that CE can be used to localize the fault and other disturbance events. Even with the current approach, the error will decrease as the number of PMUs increases.

The cause for undetected faults was found to be the large distance between the faulted bus and the PMU that captured the strongest signal of the event. The distance was so large that even the strongest signal did not have enough energy. Obviously, as the number of PMUs increases, this problem will be resolved. However, for the present study, we adopt another strategy to resolve the issue by feeding unclassified fault data along with other events to the Events Classifier module. Therefore, the fault events that are misclassified at the SSFD module get classified by the Events classifier module. In case of field PMU data, out of 58 fault events only 3 were misclassified by the SSFD module. The zone and area information of the faults were not available for the field data.

TABLE IV
AVERAGE ACCURACY WITH KNN (10 FOLD CROSS VALIDATION)

	All Events	CE $\geq$ 0.0001	CE $\geq$ 0.0005	CE $\geq$ 0.001	CE $\geq$ 0.002	NERC PRC-002
FLT	99.41	99.41	99.42	99.42	99.46	100
GNL	82.90	83.17	83.67	80.4	94.32	100
LS-off	96.13	94.77	97.48	98.17	100	n.a.
LS-on	83.26	85.25	90.74	91.12	92.55	n.a.
REX	87.91	90.30	94.25	96.91	96.59	100
RIN	86.40	90.32	96.18	98.26	99.72	100
SM-off	39.46	39.46	51.22	29.26	58.75	100
TOTAL	82.21	83.24	87.57	84.79	91.63	100

To segregate the other disturbance types including the misclassified faults, the simulated data was subjected to the Event Classifier module, the second step of the proposed framework, where the FDST algorithm was applied to compute the TFR of the disturbance waveforms. For calculating TFR, a data segment of 2 s is considered with 1.5 s before the triggering of an event and 0.5 s after the trigger. Therefore, the classifier has to wait for at least 0.5 s for predicting the disturbance type. This latency is not a concern for the near real-time situational awareness. The six features defined in Section III-B were computed from the TFR of the strongest voltage and frequency signals corresponding to the disturbances. These features were then subjected to the classifier as inputs. Table IV lists the average classification accuracies for kNN classifier with 10-fold cross validation. The classification accuracy obtained with disturbance datasets defined by different event detection thresholds implies that the performance of the classifier improves with more stringent condition for disturbance detection. This is not surprising, since more stringent threshold removes relatively insignificant events with low signal strength from the test dataset. However, the performance of kNN $(\sim 80\%)$ is not acceptable. It should also be mentioned here that the unusually low accuracy of KNN for SM-off events is misleading. It was observed that these events had waveforms that looked similar to waveforms corresponding to REX and RIN events. The *numbers* of misclassified events among these three classes were similar, but the total number of events generated for SM-off were small (36) compared to RIN (442) and REX (497). This is because there were limited synchronous motors in the system.

The classification accuracies with one versus many ELM are summarized in Table V. The number of nodes in the hidden layer (V) of ELM was fixed to be 100. The overall classification accuracy for all the disturbance events was found to be 96.08%, with classification accuracy for fault as 99.41%. Therefore, the small number of misclassified faults from the SSFD module are correctly classified at this stage. The classification accuracy of all the disturbance types are greater than 93%, most of them in the high nineties, which underscores the efficacy of the proposed approach. For the dataset generated by PRC-002-2 - recommended triggers, the fault detection and event

TABLE V
AVERAGE ACCURACY WITH ELM (10 FOLD CROSS VALIDATION)

	All Events	CE $\geq$ 0.0001	CE $\geq$ 0.0005	CE $\geq$ 0.001	CE $\geq$ 0.002	NERC PRC-002
FLT	99.41	99.41	99.42	99.42	99.46	100
GNL	93.20	93.16	97.03	97.93	98.44	100
LS-off	96.70	96.69	99.23	99.47	99.76	n.a.
LS-on	93.87	94.01	96.71	97.8	98.05	n.a.
REX	94.74	95.09	98.54	98.28	98.44	100
RIN	95.52	96.02	97.93	98.46	98.63	100
SM-off	99.13	99.13	99.13	99.12	99.51	100
TOTAL	96.08	96.22	98.28	98.64	98.9	100

TABLE VI
AVERAGE COMPUTATION TIME PER EVENT

	FDST	Feature Extraction	Classification (Testing)	Total
time (s)	0.0034	5.933×10^{-4}	0.734×10^{-4}	0.407×10^{-2}

classification accuracies are 100%, except for load switching events, which were not picked up by this criteria.

The proposed method was implemented in MATLAB [25] software on a Core i7 CPU with 16 GB RAM running windows 7. The average time required for the FDST computation, feature extraction, and event classification for each event is presented in Table VI. The total time required for processing one event is 0.407×10^{-2} s, which is quite small, and justifies the implementation of the proposed approach for near real-time situational awareness. The time can be further reduced through embedded programming.

VII. CONCLUSION

In this paper, a novel method is presented for supervisory protection and automated disturbance diagnosis. Early (two cycles) detection of faults shows promise for the use of this method for supervisory protection and SIPS. The strongest PMU-signal identification, achieved with an efficient and easy to implement algorithm alleviates the need for complex multivariate data analysis methods, and helps in locating a disturbance based on the location of the PMU generating the strongest signal. The TFR based feature extraction technique is shown to be suitable in facilitating the event classification. The easy training and small testing latency of ELM classifier further simplifies the approach. The high average classification accuracy of 96.08% using only 0.5 s of the disturbance record, tested with an exhaustive number of simulated as well as field PMU data corresponding to different event types demonstrates both the effectiveness and the robustness of the approach, making it suitable for near real-time situational awareness. To the best of our knowledge, no prior work has been reported to classify *nine* types of disturbance events with such exhaustive testing. The widespread installation of new PMUs is going to further enhance the capability of the proposed approach.

REFERENCES

[1] N. S. Group *et al.*, *"Technical analysis of the August 14, 2003, blackout: What happened, why, and what did we learn?,"* Report to the NERC Board of Trustees. 2004. [Online]. Available: http://www.nerc.com/docs/docs/blackout/NERC_Final_Blackout_Report_07_13_04.pdf

[2] *Summary of the North American synchrophasor initiative (naspi) activity area.*, 2015. [Online]. Available: http://energy.gov/sites/prod/files/North%20American%20Synchrophasor%20Initiative%20(NASPI)%20Program%20Factsheet.pdf.

[3] PSRC Working Group, *Use of synchrophasor measurements in protective relaying applications*, 2012. [Online]. Available: http://www.pes-psrc.org/Reports/UseofSynchrophasorMeasurementsinProtectiveRelayingApplications_final.pdf.

[4] O. Dahal, S. Brahma, and H. Cao, "Comprehensive clustering of disturbance events recorded by phasor measurement units," *IEEE Trans. Power Del.*, vol. 29, no. 3, pp. 1390–1397, Jun. 2014.

[5] J. Ma, Y. V. Makarov, C. H. Miller, and T. B. Nguyen, "Use multi-dimensional ellipsoid to monitor dynamic behavior of power systems based on PMU measurement," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Jul. 2008, pp. 1–8.

[6] L. Xie, Y. Chen, and P. Kumar, "Dimensionality reduction of synchrophasor data for early event detection: Linearized analysis," *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 2784–2794, Nov. 2014.

[7] L. Fan, R. Kavasseri, Z. Miao, D. Osborn, and T. Bilke, "Identification of system wide disturbances using synchronized phasor data and ellipsoid method," in *Proc. IEEE Power Energy Soc. Gen. Meeting—Convers. Del. Elect. Energy 21st Century*, Jul. 2008, pp. 1–10.

[8] W. Gao and J. Ning, "Wavelet-based disturbance analysis for power system wide-area monitoring," *IEEE Trans. Smart Grid*, vol. 2, no. 1, pp. 121–130, Mar. 2011.

[9] D. C. E. de la Garza, *Hidden failures in protection systems and its impact on power system wide-area disturbances*, M.Sc. dissertation , Virginia Polytechnic Institute and State University, Blacksburg, VA, USA, 2000.

[10] O. P. Dahal, H. Cao, S. Brahma, and R. Kavasseri, "Evaluating performance of classifiers for supervisory protection using disturbance data from phasor measurement units," in *Proc. IEEE Power Energy Soc. Innovative Smart Grid Technol. Conf. Eur.*, Oct. 2014, pp. 1–6.

[11] North American Electric Reliability Corporation (NERC), *"PRC-002-2 disturbance monitoring and reporting requirements,"* Nov. 13 2014.

[12] S. S. Allen, Alicia, and E. Muljadi, *"Algorithm for screening phasor measurement unit data for power system events and categories and common characteristics for events seen in phasor measurement unit relative phase-angle differences and frequency signals,"* National Renewable Energy Laboratory, Tech. Rep. NREL/TP-5500-58611, Aug. 2013.

[13] *PMU Data Event Detection: A User Guide for Power Engineers*, National Renewable Energy Laboratory, Tech. Rep. NREL/TP-5D00-61664, 2014.

[14] O. Dahal and S. Brahma, "Preliminary work to classify the disturbance events recorded by phasor measurement units," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Jul. 2012, pp. 1–8.

[15] *"Electromagnetic compatibility, Part 4: 30: Testing and measurement techniques-power quality measurement methods,"* Tech. Rep. IEC 61000-4-30:2003(E)

[16] *Electric Power Systems and Equipment-Voltage Ratings (60 Hz)* ANSI C84.1-1995, 1995 ..

[17] M. Biswal and P. Dash, "Detection and characterization of multiple power quality disturbances with a fast s-transform and decision tree based classifier," *Digital Signal Process.*, vol. 23, no. 4, pp. 1071–1083, 2013.

[18] B. Boashash, G. Azemi, and N. A. Khan, "Principles of time-frequency feature extraction for change detection in non-stationary signals: Applications to newborn eeg abnormality detection," *Pattern Recogn.*, vol. 48, no. 3, pp. 616–627, 2015.

[19] S. Lin, X. Liu, J. Fang, and Z. Xu, "Is extreme learning machine feasible? A theoretical assessment (part ii)," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 1, pp. 21–34, Jan. 2015.

[20] C. M. Bishop, *Pattern Recognition and Machine Learning*, New York, USA: Springer, 2006.

[21] X. Liu, S. Lin, J. Fang, and Z. Xu, "Is extreme learning machine feasible? A theoretical assessment (part i)," *IEEE Trans. Neural Netw. Learning Syst.*, vol. 26, no. 1, pp. 7–20, Jan. 2015.

[22] J. Chow *et al.*, *Guildelines for siting phasor measurement units,* North American Synchro Phasor Initiative, Research Initiative Task Team (RITT), Tech. Rep., Jun. 15, 2011. [Online]. Available: https://www.naspi.org/.
[23] 2015 wecc map of principal transmission lines., May 2015. [Online]. Available: https://www.wecc.biz/PlanningServices/Pages/Maps.aspx.
[24] North American SynchroPhasor Initiative (NASPI), *PMUs and synchrophasor data flows in North America*, Feb. 2014. [Online]. Available: https://www.smartgrid.gov/document/pmus_and_synchrophasor_data_flows_north_america/.
[25] *Matlab-Simulink, the Mathworks*, 2015. [Online]. Available: http://www.mathworks.com/.

Milan Biswal (M'09) received the Ph.D. degree in electrical engineering from Siksha O Anusandhan University, Bhubaneswar, India, in 2013.

Currently, he is a Postdoctoral Researcher with CREST Interdisciplinary Center of Research Excellence in Design of Intelligent Technologies for Smart Grids (iCREDITS) at New Mexico State University, Las Cruces, NM, USA. His current research interests include signal-processing applications in smart grid, specifically for wide-area monitoring systems.

Sukumar M. Brahma (M'04–SM'07) received the B.E. degree in engineering from L. D. College of Engineering, Ahmedabad, India, in 1989, the M.Tech. degree in electrical engineering from the Indian Institute of Technology, Bombay, in 1997, and the Ph.D. degree in electrical engineering from Clemson University, Clemson, SC, USA, in 2003.

Currently, he is William Kersting Endowed Chair Associate Professor and Associate Director of the Electric Utility Management Program (EUMP) at New Mexico State University, Las Cruces, NM, USA. He is an editor for IEEE TRANSACTIONS ON POWER DELIVERY.

Dr. Brahma is Chair of IEEE's Power and Energy Education Committee and a member of the Power System Relaying Committee.

Huiping Cao received the Ph.D. degree in computer engineering from The University of Hong Kong (HKU), Hong Kong, China, in 2007.

She is an Assistant Professor in Computer Science at New Mexico State University, Las Cruces, NM, USA. Her research expertise is in the general area of data mining, including time-series mining, graph mining, and graph databases.

Dr. Cao has served on the editorial board for *Journal on Data Semantics* (JoDS), as Conference Organizer for SIGMOD, and as a program committee member for many international conferences and workshops.