High Impedance Fault Detection in Distribution Network using Time-Frequency Based Algorithm

A. Ghaderi, Student Member, IEEE, H. A. Mohammadpour, Student Member, IEEE, H. Ginn, Senior Member, IEEE, Yong-June Shin, Senior Member, IEEE

Abstract

A new high impedance fault (HIF) detection method using time-frequency analysis for feature extraction is proposed. A pattern classifier is trained whose feature set consists of current waveform energy and normalized joint time-frequency moments. The proposed method shows high efficacy in all the detection criteria defined in this paper. The method is verified using the real-world data, acquired from HIF tests on three different materials (concrete, grass, and tree branch) and under two different conditions (wet, and dry). Several non-fault events, which often confuse HIF detection systems, were simulated, such as capacitor switching, transformer inrush current, non-linear loads, and power electronics sources. A new set of criteria for fault detection is proposed. Using these criteria the proposed method is evaluated and its performance is compared with the existing methods. These criteria are accuracy, dependability, security, safety, sensibility, cost, objectivity, completeness, and speed. The proposed method is compared with the existing methods, and it is shown to be more reliable, and efficient than its existing counterparts. The effect of choice of pattern classifier on method efficacy is also investigated.

Index Terms

High impedance fault, power distribution faults, time-frequency analysis, statistical joint moment, principal component analysis, and protection.

I. INTRODUCTION

HIGH impedance faults (HIF) introduce a unique challenge to power systems, which is their invisibility to the traditional protection devices. HIF occurs when distribution network conductors break and touch the ground surface, or lean and touch a tree branch. This results in fault current less than the load current level. Since it is difficult to detect HIFs, the exposed energized conductor may result in fire, tremendous release of energy, and threat to human life [1]. Although reports state 5% to 10% of all distribution network faults are HIF [2], the actual percentage has to be higher, since only the HIFs ending in bolted faults are included in the event recorder statistics.

The nature and physical properties of the HIF are well documented [3]. A comprehensive survey of HIF detection methods is presented in [4]. The low current magnitude, associated arc, and current asymmetry are the dominant characteristics of these faults. Although the current magnitude of HIFs fall into non-detection zones of network protection systems, their non-stationarity and random behavior have shown promising results in defining signatures for HIF [5]. However, a method that can perfectly detect HIF in different conditions and networks has not been introduced yet.

HIF detection methods can be classified into four classes based on the analysis domain they exploit to define HIF signature. These analysis domains are time domain analysis, frequency domain analysis, time-scale analysis, and time-frequency analysis [4]. The time domain analysis emphasizes irregularity of the HIF waveform. Methods based on power dissipation factor [6] and fractal techniques [7] are considered simple time domain based methods. In [8] and [9], Sarlak et al. introduce a Mathematical Morphology Gradient based method to extract time domain features of HIF.

Methods based on frequency domain analysis take advantage of the fact that HIF creates wide-spectrum voltage and current arcs. Low frequency spectral behavior of magnitude and phase [10] and the energy content of even and odd harmonics, and inter-harmonics [11] have been utilized for HIF detection. The methods based on time domain and frequency domain analysis benefit from low computational burden and simplicity. However, their assumption of the signal periodicity and infinite time (or frequency) window length hinder them in localizing time-frequency information, simultaneously.

The third class of feature extraction method is time-scale analysis (wavelet transform). In [12], discrete wavelet transform of residual current and voltage are used as feature to detect HIF. In [13], Sedighi et al. propose a method to extracts the first two levels of rbio3.1 mother wavelet as the signature to detect HIF. Michalik et al. detection technique utilizes the phase displacement between wavelet coefficients [14]. Wavelet transform provides a time and scale localization. However it is a challenging task to design a systematic detection technique based on wavelet transform. This is due to its narrow high frequency support, subjectivity to the choice of the mother wavelet, and its loss of feature resolution (less interpretable).

Time-frequency analysis (TFA) is the fourth set of analysis domain. TFA has shown high efficacy for identifying discontinuities, repeating patterns, and non-stationarity aspects of signals [17]-[18]. TFA has been successfully applied to...
different power system applications, such as power quality assessment [19]-[20], and direction finding for capacitor switching disturbances [21]. Because of advantages that TFA presents over other methods, namely time- and frequency-localization, objectivity to the signal choice, coherent time-frequency support, and interpretable features, a HIF detection method utilizing TFA for feature extraction can potentially reveal hidden information of the HIF waveforms.

In this paper, a novel method based on TFA and joint moment calculation is proposed for HIF detection. A new set of criteria for the evaluation of the fault detection algorithms is defined. The detection method gives comparable results with the existing methods in accuracy and security, while showing improvement in the rest of the criteria. The features extracted based on TFA are interpretable, objective, and their mathematical description based on joint moment calculation are complete.

II. THEORETICAL FOUNDATION OF THE METHOD

A. Time-Frequency Analysis

Any protection technique can be projected to a pattern classification task [16], which consists of three steps, i.e., measurement, feature extraction and classifier. In order to detect HIF, it is necessary to develop a more precise feature extraction, and classification technique.

In the method presented in this paper, TFA is deployed as the method for feature extraction. TFA was motivated by the need to describe non-stationary signals, where Fourier transform proves ineffective. Several time-frequency representation methods are designed to mathematically describe these signals, namely short time Fourier transform (STFT), Wigner-Ville distribution (WVD), and Choi-Williams distribution (CWD) [22]-[23]. To systematically design proper time-frequency distribution (TFD), Cohen generalization of the quadratic TFDs is used [22]. Cohen proved that one can relate the desirable properties of TFDs to constraints on its kernel:

$$TFD_s(t, \omega) = \frac{1}{4\pi^2} \int \int \int IAC_s(u, \tau) \phi(\theta, \tau) \times e^{-j\theta u - j\tau \omega + j\tau u} d\theta d\tau du$$  (1)

where $\phi(\theta, \tau)$ is a two dimensional function (in Doppler-lag domain), called kernel. $TFD_s$ is the time-frequency distribution of the signal. The constant time cross section of time frequency distribution ($TFD(t_0, \omega)$) represent the frequencies available at time $t_0$, and its frequency cross section ($TFD(t, \omega_0)$) represents the times when frequency $\omega_0$ occurred. And $IAC_s$ is the instantaneous auto-correlation of the signal $s(t)$, respectively. $IAC_s$ is defined in Eq.(2).

$$IAC_s(t, \tau) = s^*(t - \tau/2)s(t + \tau/2)$$  (2)

The desirable properties of TFD and their constraints are defined as follows:

**Time Marginal:** Integration of the TFD over frequency gives the “Instantaneous Power” ($|s(t)|^2$):

$$\int_{-\infty}^{\infty} TFD_s(t, \omega)d\omega = |s(t)|^2 \iff \phi(\theta, 0) = 1$$  (3)

**Frequency Marginal:** Integration of the TFD over time gives the “Energy Spectrum” ($|S(\omega)|^2$):

$$\int_{-\infty}^{\infty} TFD_s(t, \omega)dt = |S(\omega)|^2 \iff \phi(0, \tau) = 1$$  (4)

**Global Energy:** Integration of the TFD over the entire time-frequency plane yields the “Signal Energy” ($E_S$):

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} TFD_s(t, \omega)dt\omega = E_S \iff \phi(0, 0) = 1$$  (5)

**Reduced Interference:** Due to the bi-linearity of the IAC, introducing “artifacts” (interference) is inevitable in generating TFDs. If the kernel has low pass filter characteristic in Doppler-lag domain ($\theta, \tau$), the interference could be reduced. Since the disturbances in power systems are characterized by the presence of multiple frequency components over short duration of time, keeping high time-frequency resolution, while avoiding artifacts is of great significance in their analysis [21].

Although WVD satisfies the first three constraints ($\phi(\theta, \tau) = 1$), large proportion of interference could result in a poor interpretation of a signal. In 1989, Choi and Williams [25] proposed a Gaussian kernel, called Choi-Williams distribution (CWD) that satisfies all four of the constraints. Therefore, the CWD is chosen as the TFD in this paper. This 2D distribution of the non-stationary signal reveals hidden time-frequency behavior, which later will be mathematically described using statistical joint moments. CWD kernel is defined as follows:

$$\phi(\theta, \tau) = e^{-\theta^2 \tau^2/\sigma}$$  (6)

where $\sigma$ is a parameter that is handled to achieve proper characteristics of the distribution.
III. DATA ACQUISITION

A. High Impedance Fault

In order to collect data for HIF, a test was conducted in the a high current research laboratory. The test circuit is shown in Fig. 1. Three different materials (tree branch, grass surface, and concrete surface) in two different conditions (wet and dry) were tested, and their current and voltage magnitude were recorded with a digital data recorder (sampled at 20 kHz). The voltage level of 7.2 kV line to ground is chosen for the test, since this voltage level is one of the dominant voltage levels in distribution networks. The circuit specifications are explained in Table I. A high speed camera (with 100:1 time scale) was used to the detailed behavior of arc initialization, tracking, extinction, and possible explosion. Since HIF test is not an experiment that utilities perform regularly, the research group designed and tested the following configuration for data acquisition on different materials.

1) HIF on the Tree Branch: To conduct the tree branch test, a wooden pallet was set up on four porcelain insulators to establish an “insulated table”. The pallet was covered with rubber blankets, and the tree branch placed on them. This configuration is shown in Fig. 2(a). When the correct output voltage was reached, the test was initiated by random closing of the circuit breaker. The test current was measured by a CT, which supplied current to a precision 0.1 ohm shunt. For tree branch HIF, several specimens including elm, walnut, and willow tree were set on the insulated table. Additionally, to test the wet tree branch, trees were soaked into the water for an hour before the test. The test circuit was disconnected when (1) no remarkable event occurs in 60 sec., (2) the arcing results in bolted fault, or (3) sixty cycles (1 sec.) of HIF fault current is acquired.

2) HIF on the Surfaces: In addition to tree branches, two surfaces that are prone to make HIF are tested to acquire a comprehensive real-world data for HIF, i.e., concrete and grass. For this purpose, a down conductor was dropped across the surface, in a way that it can bounce on the ground and form an arcing fault. Procedure and experimental setup are the same as HIF on tree branches. Fig. 2(b) demonstrates HIF on grass surface, and Fig. 2(c) demonstrates the configuration for HIF on concrete surface, respectively.

B. Distribution Network Events

A protection scheme should not only detect the fault once it happens (dependability), but also it should be able to distinguish it from other conditions (security). In the case of high impedance fault, there are several events that could be nuisance to the
TABLE I
TEST CIRCUIT SPECIFICATION

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voltage Level</td>
<td>7.2 kV</td>
</tr>
<tr>
<td>System Fusing</td>
<td>250 A</td>
</tr>
<tr>
<td>Fault Current</td>
<td>990 A</td>
</tr>
<tr>
<td>Source Impedance</td>
<td>0.1 – 0.2</td>
</tr>
<tr>
<td>Load Impedance</td>
<td>xs Rs</td>
</tr>
<tr>
<td>System Loading</td>
<td>0.4 – 0.6</td>
</tr>
<tr>
<td>Load Configuration</td>
<td>Parallel RL</td>
</tr>
<tr>
<td>Load Power Factor</td>
<td>0.15</td>
</tr>
<tr>
<td>Sampling Frequency</td>
<td>20 kHz</td>
</tr>
</tbody>
</table>

TABLE II
DESCRIPTION OF SIMULATED EVENTS

<table>
<thead>
<tr>
<th>Event</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal Condition</td>
<td>24 (20%)</td>
<td>normal power system loading</td>
</tr>
<tr>
<td>Capacitor Switching</td>
<td>24 (20%)</td>
<td>4 places, 3 natural frequencies, and 2 compensation settings</td>
</tr>
<tr>
<td>Inrush Current</td>
<td>24 (20%)</td>
<td>8 fault inception angles, and 3 saturation characteristics</td>
</tr>
<tr>
<td>Non-linear Loads</td>
<td>24 (20%)</td>
<td>24 harmonic load characteristics</td>
</tr>
<tr>
<td>PE Converters</td>
<td>24 (20%)</td>
<td>24 PE converter configurations</td>
</tr>
</tbody>
</table>

detection method. The method immunity to non-HIF (NHIF) events has to be verified. For this purpose the following events are simulated using PSCAD/EMTDC in IEEE 13 node test feeder: capacitor switching (single, and back to back), transformer inrush current, non-linear loads, and power electronics sources. Table II depicts what is their description, and Fig. 3 demonstrates an example of one cycle of each event. Each figure consists of (1) a current waveform, (2) waveforms of the superimposed disturbance (SID), which is the non-fundamental part of the current defined in section IV, and (3) the CWD of the SID. It is evident that each event provides unique 2D signature in time-frequency domain.

![Fig. 3. Current waveform, normalized SIDs, and their TFDs for different events in distribution network.](image)

IV. PROPOSED HIF DETECTION TECHNIQUE
In order to extract a meaningful, informative, yet minimal set of features that delineates HIF characteristics, a five-step method is proposed. In this section, each step is discussed.
Step 1: Extracting Disturbance using Curve Fitting Routine

In order to find a high resolution TFD, it is necessary to eliminate the non-informative fundamental harmonic of the fault, which is the largest energy density part of the TFD. The rest of the fault current is called super-imposed disturbance (SID), which is more informative and has higher time and frequency resolution. In other word, the two dimensional time-frequency signature of any waveform \( S(t) \) consists of a disturbance current \( D(t) \) that is superimposed on the fundamental frequency current \( F(t) \). Assuming power system frequency \( (\omega_0/2\pi) \) is fixed at 50 Hz, or 60 Hz, one can extract the SID, \( \hat{D}(t) \), by subtracting the estimated fundamental current \( \hat{F}(t) \) from the original current \( S(t) \), \( \hat{D}(t) = S(t) - \hat{F}(t) \). In this curve fitting optimization routine, the objective function is \( \hat{F}(t; a, \phi) = a_0, \cos(\omega_0 t + \phi) \), and fitting parameters are \( a \) and \( \phi \), which can be found using least square optimization in the moving window curve fitting routine [27]:

\[
\{ a_0, \phi_0 \} = \text{arg} \min \left| S(t) - \hat{F}(t) \right|^2
\]  

(7)

Step 2: Time-Frequency Distribution

Next, time-frequency distributions are generated from the quarter-cycle SID signals. In high impedance arcing fault, each cycle contains two dominant arc ignitions and two dominant arc extinctions [28]. Therefore, by dividing each cycle into quarters, key information is preserved at lower computational cost. TFDs are generated using Eq. (1) and (2) with Choi-Williams kernel given by Eq. (6). Since the resulting TFD is a 2D matrix whose size is equal to signal length squared [17], four quarter-cycles of 60 Hz SID sampled at 20 kHz consist of 27556 \( (4 \times 83 \times 83) \) samples, which is one order smaller than 110889 \( (333 \times 333) \) samples, which would be created without division of quarter cycles.

Although time-frequency distributions are useful representations for non-stationary signals, these representations contain large amount of redundant information. Therefore, the purpose of the next two steps is to reduce this dimension by eliminating the redundancy of the time-frequency information through feature extraction and feature selection.

Step 3: Joint-Time-Frequency Moment Calculation

Several articles discuss the feature extraction of a 2D pattern (e.g. time-frequency distribution) [17] and [29]. Singular value decomposition, principal component analysis, non-negative matrix factorization, and independent component analysis are the most common decomposition techniques to decrease the dimension of time-frequency matrix (TFM) [31]. The efficiency of these methods decreases remarkably for large 2D matrices such as TFM.

"Joint Time-Frequency Moment (JTFM)\(^*\), inspired by image processing techniques, has also been used successfully as a feature extraction technique [32]. The JTFM \( \text{MOM}_{p,q} \) of a time-frequency distribution \( \text{TFD}_S(t, \omega) \) is defined as:

\[
\text{MOM}_{p,q} = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (t - t_m)^p (\omega - \omega_m)^q \text{TFD}_S(t, \omega) dt d\omega}{E_S}
\]

(8)

where \( p \) and \( q \) are non-negative integers. Also \( E_S \), is signal energy defined in Eq. (5), which is used to normalize the moments. \( t_m \) and \( \omega_m \) are called mean time, and mean frequency, and are formulated as follows:

\[
t_m = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t \text{TFD}_S(t, \omega) dt d\omega
\]

(9)

\[
\omega_m = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \omega \text{TFD}_S(t, \omega) dt d\omega
\]

(10)

It should be noted that signal time duration \( (T) \) and signal frequency bandwidth \( (B) \) can be found with the calculation of \( \text{MOM}_{2,0} \), and \( \text{MOM}_{0,2} \), respectively.

By keeping all the joint moments of a time-frequency distribution, one can preserve all of its information. This is due to the fact that any time-frequency distribution can be reconstructed using a unique set of joint time-frequency moments and vice versa [22]. This is proved using the mathematical definition of characteristics function, i.e. moment generating function. The characteristic function is a double Fourier transform of the TFD from time-frequency \( (t, \omega) \) domain to Doppler-lag \( (\theta, \tau) \) domain and is defined as follows:

\[
M(\theta, \tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{TFD}_S(t, \omega) e^{j\theta t} e^{j\tau \omega} dt d\omega
\]

(11)

Taylor expansion coefficients of CF are the JTFM:

\[
M(\theta, \tau) = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \left( \frac{\text{MOM}_{p,q}}{p!q!} \right) (j\theta)^p (j\tau)^q
\]

(12)

This means the TFD is uniquely reconstructed from JTFM and vice versa, due to correspondence between Fourier couples and between Taylor series couples [16] and [33]. In other words, the two dimensional time-frequency signature of any waveform
could be uniquely reconstructed from its set of joint moments. Therefore, we utilized JTFM as one-dimensional and faithful representation of the two-dimensional TFD with the following definition:

\[ J_i = \left\{ \frac{|\text{MOM}_{p,q}|}{p!q!} \right\} \tag{13} \]

where \( J_i \) is the joint time-frequency feature set. \( p \) and \( q \) are nonnegative integers.

In the next step, the feature set will be normalized using its statistical descriptions as given in Eq. (14). This normalization prevents features with large magnitudes from dominating the total variance.

\[ J_i = \frac{J_i - \mu}{\delta} \tag{14} \]

where \( J_i \) is the normalized feature set, and \( \mu \) and \( \delta \) are the average and standard deviation of the features over sampling set, respectively.

The unique characteristics of this feature set is that it is capable of keeping all the 2D time-frequency patterns of the signal in a 1D stream of data. As shown in Fig. 3, each event in power systems provide a distinguishable 2D signatures, which is mathematically described in feature set of Eq. (14).

**Step 4: Principal Component Analysis (PCA)**

The objective of PCA is to find a set of orthogonal components that minimizes the mean square error of the reconstructed data and represent the original data with fewer components, which reduces the dimension of the data. The ratio of an eigenvalue \( \lambda_i \) to sum of eigenvalues expresses the loss of information (LOI) due to the elimination of the \( i \) – th eigenvector, as given in Eq. (15). In other words, the algorithm keeps the features that more strongly represent the signature of different events.

\[ \text{LOI}_i = \frac{\lambda_i}{\sum_{j=1}^{n} \lambda_j} \tag{15} \]

where \( n \) is the length of the feature set.

To determine the optimum number of principal components (PC), \( \text{LOI} \) is calculated as the number of PCs is sequentially increased, starting from the one with the largest eigenvalue.

**Step 5: Pattern Classifier**

In this work, support vector machine (SVM) shows the highest reliability compared to other classifiers. SVM is a linear discriminant classifier, which relies on the preprocessing of the data in higher dimension [16]. A Gaussian non-linear \( \varphi(x) \) function is deployed to map the feature set to a higher dimension. It can be proved that in a sufficiently higher dimension, data from two-class classifier can always be separated [33]. The detection algorithm separates two conditions, namely \( H \) (for HIF), and \( NH \) (for NHIF). Assume there are \( n \) samples from \( H \) class, and \( m \) samples from \( NH \) class. After the SVM mapping, a linear discriminant function \( w^T y_i \) is trained, in a way that, maps the feature vectors to a linear class boundary by maximizing the criterion function. The boundary solution \( (w_S) \) of this maximization routine is:

\[ w_S = (S_H + S_{NH})^{-1}(\mu_H - \mu_{NH}) \tag{16} \]

where \( \mu_H \) and \( \mu_{NH} \) are the sample mean vectors and \( S_H \) and \( S_{NH} \) are the scatter matrices of H-class and NH-class respectively. \( \mu_H \), and \( S_H \) are calculated as follows:

\[ \mu_H = \frac{1}{n} \sum_{i=1}^{n} y_{H,i} \tag{17} \]

\[ S_H = \sum_{i=1}^{n} (y_{H,i} - \mu_H)(y_{H,i} - \mu_H)^T \tag{18} \]

where \( y_{H,i} \) is the \( i \) – th mapped sample in the H-class, and \( n \) is the number of samples in the H-class.

\( \mu_{NH} \), and \( S_{NH} \) are calculated the same way as \( \mu_H \), and \( S_H \), respectively, with two differences: \( y_{H,i} \) is replaced with \( y_{NH,i} \), which is the \( i \) – th sample in the NH-class, and \( n \) is number of samples in NH-class.

**V. Algorithm Performance Evaluation**

In this section the efficiency of the proposed technique on real world data is discussed. 240 HIF cases and 120 NHIF cases were used to train and test the algorithm. The criteria for performance evaluation and result will be discussed in the following subsections.
A. Fault Detection Evaluation Criteria (FDEC)

The complexity of the HIF phenomena and their detection techniques necessitate a complete set of criteria for their evaluation. Although tripping the HIF assures public safety, it would disconnect important loads like traffic lights, elevators, and hospitals. In order for utilities to conduct a risk analysis on this trade-off, two criteria are proposed - safety, and sensibility. Due to the fact that HIF detection techniques have higher complexity, some criteria are also needed to compare their efficiency.

Defining power system protection algorithms as a pattern classification task provides the opportunity to evaluate the protection algorithm using the well-defined assessment methods for classifiers, e.g., “Confusion Matrix” (CM). For a two-class classifiers, as of the fault detection algorithms, CM is a two-by-two matrix defined in Eq. (19).

\[
CM = \begin{bmatrix}
DF & ND \\
MF & HC
\end{bmatrix}
\]  

(19)

where “DF” is the number of detected faults, “ND” is the number of not detected faults, “MF” is the number of non-faults that mistakenly activate the relay, and “HC” is the number of healthy condition that do not mistakenly activate the relay.

Based on the confusion matrix five criteria are defined for the evaluation of the reliability of HIF detection algorithm. These criteria are (1) accuracy, (2) dependability, (3) security, (4) safety, and (5) sensibility. In addition, three criteria for method efficiency and practicality evaluation are proposed. These criteria are (6) cost, (7) objectivity, and (8) completeness of the detection method. These parameters form comprehensive fault detection evaluation criteria (FDEC), which are defined as follows:

1. **Accuracy, A**, emphasizes the all-inclusive performance of the method, by means of dividing the number of right decisions over the total number of decisions.

\[
A = \frac{DF + HC}{DF + ND + MF + HC} \%
\]

(20)

2. **Dependability, D**, is defined as the proneness of protection scheme to detect and isolate the fault, which can be formulated as the fraction of number of detected faults to actual number of faults.

\[
D = \frac{DF}{DF + ND} \%
\]

(21)

3. **Security, S**, illustrates the method ability to selectively detect fault, and not to trip for non-fault conditions. It is the fraction of the number of predicted non-fault over the actual number of non-faults.

\[
S = \frac{HC}{MF + HC} \%
\]

(22)

4. **Safety, SF**, is needed to reflect the method ability to isolate the faults, that are hazardous to the public, e.g. HIF in crowded area and islanding [24]. It is the ability of the method to be vigilant for fault occurrence and to make sure that faults are not mistaken for non-faults. This criterion can be calculated by dividing the number of right prediction of non-faults over the total number of predicted non-faults.

\[
SF = \frac{HC}{ND + HC} \%
\]

(23)

5. **Sensibility, SN**, is required to emphasize the risk of tripping sensitive loads, e.g. hospitals. Sensibility (SN) can be formulated as the fraction of the number of right prediction of faults, over the total number of predicted faults.

\[
SN = \frac{DF}{DF + MF} \%
\]

(24)

These five measures are named ‘reliability criteria’, and are desired to be as high as possible. The next four measures weigh the general performance of the method and emphasize their robustness and invest-ability.

6. **Cost, C**, takes into account the computational complexity (operational burden, training complexity) and hardware costs (additional measurement devices and communication infrastructure).

7. **Objectivity, OBJ**, highlights whether the technique is objective to the type of fault, and the network topology. This criterion guarantees that the same method is able to be applied to different applications and network topologies. In other words, this criterion is rather a qualitative philosophy behind the protection scheme than a criterion that ought to be verified quantitatively.

8. **Completeness, COM**, demonstrates if the technique preserves the information during its feature extraction/feature selection routine. This criterion guarantees that no important information is discarded during the process, rather it compresses it. If the completeness criterion is not met, a new observed sample might necessitate the modification of the method.

9. **Speed, V**, the time window of the data needed for the method to make the decision is crucial for HIF, because of its possible safety hazards and intermittent nature. The higher the number of power frequency cycles it takes, the higher the
number of possible arc ignition. As a result the number of cycles of data needed for the algorithm to perform can be a measure of speed of protection technique. This can be formulated, as the fraction of the duration of one cycle of power frequency current ($T_0$) over the time it takes for the method to detect the fault ($T_d$).

$$v = \frac{T_0}{T_d} \quad (25)$$

B. Feature Extraction

JTFM features were extracted utilizing Eq. (13). For each one fourth of the cycle the signal length is $N = 83$ samples, and frequency bins also set to $M = 83$ levels. JTFM size ($p, q = 6$) is set one order of magnitude smaller than $N, M$, in order to obtain an acceptable Taylor expansion estimation of characteristic function. This results in 36 features in each fourth cycle and 144 features in the whole cycle. Due to the possible loss of information through the windowing and normalization process, several common signal processing features are extracted before processing. These five features are energy ($E_c$), mean time ($t_{mc}$), mean frequency ($\omega_{mc}$), time duration ($T_c$), and bandwidth ($B_c$) of the total cycle of the current. These are calculated based on Eq. (5), Eq. (9), Eq. (10), ($MOM_{2,0}$), and ($MOM_{0,2}$), respectively. Thus, the final feature set contains 149 features. In the final step, data is normalized as described by Eq. (14).

$$feature \ set = \{E_c, t_{mc}, \omega_{mc}, T_c, B_c, J_i\} \quad (26)$$

C. Feature Selection

Although the use of JTFM for PCA reduces the dimension of unique features, maximum robustness is achieved through further reduction of features. In PCA, smaller eigenvalues of covariance matrix provide less pattern information as shown in Eq.(15). To determine the optimum number of principal components (PC), loss of information (LOI) is evaluated against the number of PCs. The optimal point occurs when LOI is negligible despite additional PC inclusion. Table III demonstrates these diminishing returns by use of the LOI and reconstruction error (RE). Keeping the first nine PCs almost keeps all the information in the covariance matrix and minimizes the reconstruction error (RE). This is also evident from Fig. 4, where almost all the information is embedded in the first nine principal components. Thus, the first nine PCs form the new dimensionally reduced feature set.

<table>
<thead>
<tr>
<th>PCs</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>6</th>
<th>9</th>
<th>149</th>
</tr>
</thead>
<tbody>
<tr>
<td>RE (%)</td>
<td>24.8</td>
<td>13.8</td>
<td>3.74</td>
<td>1.01</td>
<td>$10^{-3}$</td>
<td>0</td>
</tr>
<tr>
<td>LOI (%)</td>
<td>41.1</td>
<td>28</td>
<td>15.9</td>
<td>5</td>
<td>0.1</td>
<td>0</td>
</tr>
</tbody>
</table>

Table III: PCA efficiency for different number of PCs

Fig. 4. First fifteen principal components.

D. Pattern Classifier

In this step of the algorithm, a hyperplane boundary has to be found in 9-dimensional feature plane, using a pattern classifier technique. Since it is not possible to represent the feature set in nine dimensions, they are represented in a two dimensional feature plane as in Fig. 5. The magnitude of each feature for fault class (HIF) and non-fault class (NHIF) are demonstrated. It can be observed that a linear discriminant function cannot successfully separate these classes.

In this paper SVM classifier is utilized to separate two classes by means of a linear hyperplane in higher dimension feature domain. Training an SVM has low computational burden, with high reliability and stability. Fig. 6 shows the efficiency of
several classifiers with lower complexity (linear discriminant function) and higher complexity (quadratic discriminant function, naive Bayes, and artificial neural network) than SVM. This comparison is performed using the detection criteria (FDEC) introduced in section V-A. For the purpose of training the classifier, 80% of the feature set are randomly selected and the rest are dedicated to the test set. As shown in Fig. 6, SVM classifier with nine principal components has the highest reliability.

![Fig. 5. 2D representation of HIF, and NHIF in feature domain.](image)

<table>
<thead>
<tr>
<th>Method</th>
<th>A</th>
<th>D</th>
<th>S</th>
<th>SF</th>
<th>SN</th>
<th>OBJ</th>
<th>COM</th>
<th>v</th>
</tr>
</thead>
<tbody>
<tr>
<td>JTF</td>
<td>93.6</td>
<td>100</td>
<td>81.5</td>
<td>100</td>
<td>100</td>
<td>YES</td>
<td>YES</td>
<td>1</td>
</tr>
<tr>
<td>Ref[5]</td>
<td>N.A</td>
<td>97.2</td>
<td>96.9</td>
<td>N.A</td>
<td>N.A</td>
<td>NO</td>
<td>NO</td>
<td>0.66</td>
</tr>
<tr>
<td>Ref[8]</td>
<td>97.3</td>
<td>98.3</td>
<td>96.3</td>
<td>N.A</td>
<td>N.A</td>
<td>YES</td>
<td>NO</td>
<td>0.04</td>
</tr>
<tr>
<td>Ref[13]</td>
<td>96</td>
<td>90</td>
<td>100</td>
<td>95</td>
<td>100</td>
<td>NO</td>
<td>NO</td>
<td>0.25</td>
</tr>
<tr>
<td>Ref[15]</td>
<td>98</td>
<td>N.A</td>
<td>N.A</td>
<td>N.A</td>
<td>N.A</td>
<td>NO</td>
<td>NO</td>
<td>1</td>
</tr>
<tr>
<td>Ref[34]</td>
<td>N.A</td>
<td>100</td>
<td>100</td>
<td>N.A</td>
<td>N.A</td>
<td>YES</td>
<td>NO</td>
<td>0.016</td>
</tr>
</tbody>
</table>
Fig. 6. Comparison of different pattern classifiers for high impedance fault detection.

E. Comparison

Table IV demonstrates the comparison between the proposed time-frequency based technique, and recent HIF detection methods, based on FDEC. If the information for extracting of the criterion was not provided by the paper, not available (NA) is mentioned. The method proposed in this paper preserve the objectivity criterion (OBJ), since it is not subjective to the type of the fault and network topology. It also holds the completeness criterion (COM), because it keeps the whole information throughout the steps. In several recent methods which are based on wavelet transform, choosing the mother wavelet is an application-dependent trial and error procedure that is subjective to the network topology and the type of fault. As a result these methods do not preserve objectivity and completeness.

The security of the method proposed in this paper is less than other criteria. This means the percentage of false alarm is increased, but all the faults are detected. This small shortcoming in method efficiency is inevitable, due to the fact that huge amount of information is lost from non-ideal CT, and environment noise in practical cases. For this reason this is not a fair comparison. However, the method shows higher performance in rest of the criteria, e.g. dependability, safety, sensibility, objectivity, and completeness.
VI. CONCLUSION

In this paper, a simple, yet reliable high impedance fault detection technique, using time-frequency analysis, is proposed. In order to verify the method efficiency, an HIF test is conducted in practical conditions, where Hall effect transformers are not available and a low level of noise is not guaranteed. However, for real-world actualization of the method on microprocessor based relays, a more comprehensive test setup with higher variety of fault surfaces and more realistic scenarios should be implemented. A complete set of criteria is introduced for evaluation of protection algorithms. The proposed method provides high performance in all of these criteria, i.e., accuracy (A), dependability (D), safety (SF), and sensibility (SN). The method is also as fast as one cycle of power system frequency ($v = 1$), and it preserves objectivity (OBJ) and completeness (COM) of the information. Low complexity in both hardware and software is provided by the independence of method to VT, and communication structures, and also small computational burden.

Increased number of false alarms (negligible imperfection in security) is the inevitable consequence of the current waveform distortion in practical conditions. It should be noted that the salient shortcoming of this method is that, similar to most of the existing techniques, it is only applicable to radial distribution systems. Therefore, one of the main future goals of this research team is to comply with the fast-growing distributed generation (DG) paradigm, and to extend the proposed HIF detection technique to the bidirectional distribution systems.

ACKNOWLEDGMENT

This work was supported by NSF I/UCRC (Industry/University Collaborative Research Center) Grid-connected Advanced Power Electronic Systems (GRAPES) under #0934378, and Korea National Research Foundation under # NRF-2012M2A8A4055236 and # NRF-2014R1A2A1A01004780. Also, this research was supported by Korea Electric Power Corporation Research Institute through Korea Electrical Engineering and Science Research Institute (Grant Number # R13TA20). The authors would like to acknowledge Mr. Jeff Kester with Hubbell Power Systems Inc. for providing the facilities for real-world data acquisition.

REFERENCES


Amin Ghaderi (S’12) received the B.Sc. degree in electrical engineering from the University of Tehran in 2008. He received the M.Sc. degree in power systems engineering from Iran University of Science and Technology in 2011. Amin is now pursuing his Ph.D degree in Electrical Engineering Department at the University of South Carolina. His main research interests includes power quality, power system protection, and power electronics. He is now working on non-periodic load current compensation using power electronics based devices.

Hossein Ali Mohammadpour (S’10) received the B.Sc. and M.Sc. degrees in electrical engineering and power systems from Iran University of Science and Technology, Tehran, Iran, in 2006 and 2009, respectively. He is currently working toward the Ph.D. degree in electrical engineering at the University of South Carolina, Columbia, SC, USA. His primary research interests include power systems stability, power electronics, renewable energy, flexible ac transmission system technologies, and electric ship system modeling and analysis.

Herbert L. Ginn III (M’96) received the M.S. and Ph.D. degrees in electrical engineering from Louisiana State University, Baton Rouge, respectively. He is currently with the Electrical Engineering Department at the University of South Carolina as an Associate Professor. His areas of specialization are power electronics applications in energy systems, power phenomena and compensation in non-sinusoidal systems, and power quality. Dr. Ginn's current research activities are focused on the control and coordination of power electronic converters in special case distribution systems such as self-contained vehicular systems and micro-grids.

Yong-June Shin (S’98-SM’04) received his B.S. degree from the Department of Electrical Engineering, Yonsei University, Seoul, Korea, in 1996 with early completion honors and the M.S. degree from The University of Michigan, Ann Arbor, in 1997. He received the Ph.D. degree from the Department of Electrical and Computer Engineering, The University of Texas at Austin, in 2004. Upon his graduation, he joined the Department of Electrical Engineering, The University of South Carolina as an Assistant Professor. He was promoted to Associate Professor with tenure in 2011. He joined School of Electrical and Electronic Engineering, at Yonsei University, Seoul, Korea as Associate Professor in 2012. He is a recipient of the United States National Science Foundation CAREER award in year 2008, and General Electric Korean-American Education Commission Scholarship. Dr. Shins current research interests are characterized by the application of novel digital signal processing techniques to a wide variety of important transient and nonlinear problems in smart electric power grid.