High Impedance Fault Detection and Isolation in Power Distribution Networks using Support Vector Machines

Muhammad Sarwar∗, Faisal Mehmood, Muhammad Abid, Abdul Qayyum Khan, Sufi Tabassum Gul, Adil Sarwar Khan

Department of Electrical Engineering,
Pakistan Institute of Engineering and Applied Sciences
Islamabad, Pakistan

Abstract

This paper proposes an accurate High Impedance Fault (HIF) detection and isolation scheme in a power distribution network. The proposed schemes utilize the data available from voltage and current sensors. The technique employs multiple algorithms consisting of Principal Component Analysis, Fisher Discriminant Analysis, Binary and Multiclass Support Vector Machine for detection and identification of the high impedance fault. These data driven techniques have been tested on IEEE 13-node distribution network for detection and identification of high impedance faults with broken and unbroken conductor. Further, the robustness of machine learning techniques has also been analysed by examining their performance with variation in loads for different faults. Simulation results for different faults at various locations have shown that proposed methods are fast and accurate in diagnosing high impedance faults. Multiclass Support Vector Machine gives the best result to detect and locate High Impedance Fault accurately. It ensures reliability, security and dependability of the distribution network.

Keywords: Fisher Discriminant Analysis, High Impedance Fault, Principal Component Analysis, Support Vector Machines

1. Introduction

Detection of high impedance faults poses a highly challenging problem because of the random, asymmetric and non-linear nature of high impedance fault (HIF) current. Most of the time, these faults cannot be detected and isolated by conventional over-current schemes, because magnitude of fault current is considerably lower than nominal load current (Jones, 1996). High impedance faults typically occur when an energised conductor comes in contact with ground through any high impedance object such as dry asphalt, wet sand, dry grass and sod etc. which limits the flow of current towards ground (Jones, 1996). Timely detection of high impedance faults is necessary for efficient, reliable and safe operation of power systems. Probability of occurrence of high impedance fault in distribution networks is more than in transmission network because distribution feeders are more likely to come in contact with high impedance objects like trees etc. However, in underground cables high impedance faults are caused by insulation degradation that exposes the energised conductor to high impedance objects (Bakar et al., 2014). High impedance faults occur at voltage level of 15KV or below in most of the cases. Magnitude of HIF current is independent of the conventional short circuit fault current level (Bishop, 1985).

High impedance faults are extremely difficult to detect and isolate by conventional protection schemes, because fault current magnitude is much lower than nominal current. According to report by Power System Relaying Committee (PSRC), only 17% of HIFs can be detected by conventional relaying schemes (Jones, 1996). Detection of HIFs helps in prognostic maintenance in power distribution system. High impedance faults involve arcing which makes fault current asymmetric and non-linear. As a result of arcing, HIFs involve high frequency components similar to load and capacitor switching which makes detection much more difficult (Sahoo and Baran, 2014).

Previous research on diagnosis of HIFs was focused on lab-based staged fault studies. However, with the advancement in technology and better understanding of features of HIFs, the focus has shifted towards simulations and software studies (Brahma, 2013).

Due to its critical nature, researchers from both industry and academia have proposed various techniques to detect HIFs in distribution networks. Majority of the studies were reported as early as 1980s and 1990s, but the simulation methods and advanced detection techniques are still being developed and proposed. HIF detection methods can be broadly classified into time domain algorithms, frequency domain algorithms (Lima et al., 2018), hybrid algorithms (Samantaray et al., 2008) and knowledge-based systems (Etemadi and Sanaye-Pasand, 2008).

K. Zoric and M. B. Djuric presented a method to detect high impedance fault based on harmonic analysis of voltage signals (Zoric et al., 1997). James Stoupis introduced a new relaying scheme manufactured by ABB in the area of artificial neural networks (Stoupis et al., 2004). Sedighi proposed two methods based on soft computing for detection of HIF (Sedighi et al., 2005). Mark Adamiak proposed signature based high impedance fault diagnoses which involves expert system pat-
tern recognition on harmonic energy levels in arcing current (Adamiak et al., 2006).

S.R. Samantaray presented an intelligent approach to detect high impedance faults in distribution systems (Samantaray, 2012). In (Zanjani et al., 2012), authors have proposed a new approach to detect HIF based on PMU (Phasor Measurement Unit). F. V. Lopes presented a method to diagnose HIF in smart distribution systems (Lopes et al., 2013). Authors in (Brahma, 2013; Hou, 2007) presented a method to detect HIF based on mathematical morphology. A new model for high impedance faults has been present in (Torres et al., 2014). Results of this model are quite closer to what is observed in staged faults. This research activity also detected HIF using harmonic analysis of current waveform.

Recently, Kavi presented a method to detect HIF in Single Wire Earth Return (SEWR) system (Kavi et al., 2016). Sekar and Mohanty proposed fuzzy rule base approach for high impedance fault detection in distribution systems (Sekar and Mohanty, 2018). They present a filter-based morphology gradient (MG) to differentiate non-HIF events from HIF events. W. C. Santos presented a transient based approach to identify HIF in power distribution systems (Santos et al., 2017). In (Ferdowsi et al., 2017), real time complexity measurement (RCM) based approach is used to detect HIF. In (Ghaderi et al., 2016) HIF detection techniques are evaluated and compared with each. With the advancement in technology, the trend has been shifted towards smart grids and smart distribution systems. Smart distribution systems include measurements (such as voltage and current) at each node that helped to discover and develop digital signal processing based fault detection techniques (Lai et al., 2005; Elkalashy et al., 2007; Sheelvant, 2015).

Prior studies have helped to reveal many of the hidden characteristics of High Impedance Faults. But the major drawback of aforementioned techniques is that they are not capable of detecting all types of high impedance faults. Furthermore, active methods of HIF detection use signal injection which deteriorates the power quality. Some methods employ data gathered by PMUs, which are quite expensive and also many distribution systems currently don’t deploy PMUs. Another disadvantage is that most of the proposed methods use a lot of computing power and thus can not be implemented on an embedded system as a portable numerical relay.

The motivation for this research work lies in manifold shortcomings of the prior research. The proposed method is computationally less rigorous, so it can be implemented on an embedded system as a numerical relay. The method ensures power quality as it does not inject any signal into power system for HIF detection. The technique developed can detect all types of HIF i.e., broken and unbroken conductor HIFs and can locate the faulty section of network. Furthermore, input data is gathered using CTs and PTs which are already deployed in all distribution networks, so no additional hardware installation is needed for data acquisition. The proposed method is accurate and highly reliable as it can distinguish load switching from faults, and can also detect and isolate multiple high impedance faults in the power network. The training model used for detecting faults is of low order and can easily be implemented on any embedded hardware for real time prototyping and HIF detection.

In proposed method, data obtained from voltage and current sensors is fed to three fault detection algorithms i.e., Principal Components Analysis (PCA), Fisher Discriminant Analysis (FDA) and Support Vector Machines (SVM). PCA extracts principal components of data and use them for fault detection. FDA reduces the data to a lower dimension to maximise distance among various classes for increased accuracy. Binary class SVM is used only for detection of HIFs. To determine type and location of fault multiclass SVM is deployed to signal the presence of an incipient or sudden high impedance fault. Once the fault is detected, the faulty system can be isolated from the network by issuing a trip signal to traditional over current relays in substation. The speed and accuracy of proposed method is comparable to conventional fault detection of over-current faults.

The rest of research paper is organised as follows. Section II details characteristics and features of high impedance faults. Theoretical foundation for data driven techniques is laid down in Section III. Proposed techniques have been tested on IEEE 13-node test system and results are discussed in Section IV. Section V concludes the research.

2. High Impedance Fault Characteristics

In this section, prominent features of high impedance faults are described. To obtain training data from simulation, a fault model is simulated in Simulink and training data is obtained from voltage and current sensors installed in the network.

2.1. Properties of High Impedance Faults

Arcing is a prominent phenomena in HIFs. The arc is formed due to air gap between energised conductor and high impedance object. Arc ignition occurs when magnitude of voltage is higher than air gap breakdown voltage. Consequently, arc extinction occurs when voltage is lower than breakdown voltage (Ghaderi et al., 2016). The value of break down voltage changes during each cycle. Thus, in every cycle of voltage, the HIF current includes two arc re-ignitions and two arc extinctions. Therefore, current conducting path changes during each cycle which changes the magnitude of HIF current making it non-linear, also HIF current is intermittent in nature (Ghaderi et al., 2016).

Some of the typical features of high impedance faults are as follows:

- **Non linearity:** Voltage-current characteristics are highly nonlinear due to change in current conducting path (Chen et al., 2013).
- **Asymmetric nature of HIF current:** Peak values of current are different in positive and negative half cycle due to the presence of varying break down voltage (Ghaderi et al., 2016; Biswal, 2017).
- **Intermittent nature:** HIF current is not steady due to intermittent nature of arc (Mai et al., 2012).
• **Build up:** Current magnitude progressively increases till it reaches its maximum value (Biswal, 2017)

• **Randomness:** Magnitude of HIF current and its shape changes with time due to change in impedance of conducting path (Sykulski, 2006).

• **Low and high frequency components:** HIF current includes low frequency components due to non-linearity of HIF. Additionally, HIF current also contains high frequency components due to intermittent nature of arc (Ghaderi et al., 2016).

2.2. Simulation of High Impedance Fault

To obtain training data from simulation, an appropriate model of HIF is required which would show the real behaviour of an HIF. This paper utilises an HIF model shown in Fig. 1, connected between any of phase and ground (Brahma, 2013). The model is simulated in MATLAB, the parameters of model are tuned according to test feeder.

In this HIF model, two diodes $D_p$ and $D_n$ are connected to two DC voltage sources $V_p$ and $V_n$, respectively. The DC sources have different magnitude and their magnitude randomly changes around $V_p$ and $V_n$ every 0.11 ms. This models the asymmetric nature of arc current and intermediate arc extinction.

To model the randomness in duration of arc extinction in high impedance fault, voltage polarity also changes at every sampling instant (Brahma, 2013). When the instantaneous value of phase voltage is greater than $V_p$, current flows towards ground; when the instantaneous value of phase voltage is less than $V_n$, current reserves its direction, when the instantaneous value of phase voltage is between $V_p$ and $V_n$, no current flows. In order to incorporate varying arc resistance, the model of HIF also includes two variable resistances, $R_p$ and $R_n$, such that values of these resistances vary randomly every 0.11 ms.

The parameters used for HIF model with IEEE 13-node test feeder in Simulink are:

$V_p = 1.0 \text{ kV}, \text{ with } \pm 10\% \text{ variation}$  
$V_n = 0.5 \text{ kV}, \text{ with } \pm 10\% \text{ variation}$  
$R_p, R_n = 1000 \Omega - 1500 \Omega, \text{ with random variation}$

The above model is simulated in MATLAB. There are two steps involved in modelling HIF; in the first step, variable DC voltage sources are modelled using controlled voltages source; in the second step, variable resistances are modelled using controlled current sources. First part is implemented using only random number generator, constant block, and controlled voltage source to obtain varying DC voltages. Second step involves a build-up series R-L circuit and a sinusoidal signal of 60 Hz. Both of these generate an exponentially growing sine wave. The sine wave is multiplied with a random number of amplitude 1 and variance of 0.12 to obtain a randomly varying resistance.

Fig. 2 shows arc current and voltage waveforms obtained as a result of modelling HIF in Simulink, using HIF model of Fig. 1. It is clear that the arc current is small, random, asymmetric, and nonlinear in nature. The voltage waveform in Fig. 2
3. Theoretical Foundation

Data driven techniques are the perfect candidate for fault diagnoses in large systems where enough amount of data is available. Principal Component Analysis, Fischer Discriminant Analysis, and Support Vector Machines are widely used for addressing diagnosis problems due to their simplicity and efficiency in processing large amount of data. Here the theoretical basis for applied algorithms is given.

3.1. Principal Component Analysis

Principal Component analysis is linear dimensional reduction technique, it projects higher dimensional data into lower dimensions while keeping significant features. PCA has ability to retain maximum variation that is possible in lower dimensions such that transformed features are linear combination of primary features. In reduced dimensions, different statistical plots such as T2 or Q-charts are utilised for visualisation of different trends. PCA is known as powerful tool for feature extraction and data reduction in fault detection techniques because of its simplicity and ability to process large amount of data (Jamil et al., 2015).

Application of PCA for fault diagnoses consists of three steps; first of all, loading vectors (transformation vectors) are calculated by performing offline computations on training data; in second step, the loading vectors are utilised to transform online data (higher dimensional data) into lower dimensions; in third step, test statistics such as T2 are used to detect fault (Jamil et al., 2015).

Let us assume that a training set of m process variables, with set of n observations, is normalised to unit variance and zero mean by subtracting each process variable by its mean and dividing by standard deviation of data, and is shown in the form of input matrix X∈R^m×n. With the help of singular value decomposition of input data matrix X, loading vectors or transformation vectors are calculated.

\[
X = U\Sigma V^T
\]  

(1)

In equation (1), U and V are unitary matrices, and \(\Sigma\) is called diagonal matrix and its singular values are in decreasing order. The transformation vectors are orthonormal vectors of matrix \(V \in R^m\times m\). Training set’s variance projected along the \(i^{th}\) column of V is equal to \(\sigma_i^2\). In PCA, loading vectors or transformation vectors related to a largest singular value are kept to capture large data variation in lower dimensions. Let us assume that P \(\in R^m\times n\) is the matrix with first a column of V \(\in R^m\times m\), and projection of observed data X into reduced dimensions are incorporated in the score matrix T is given as,

\[
T = XP
\]  

(2)

Once the data is projected in lower dimensions, Hotelling’s \(T^2\) statistics is used for fault detection. Hotelling’s \(T^2\)-statistics can be calculated as (Jamil et al., 2015; Chiang et al., 2000),

\[
T^2 = x^T\Sigma_a^{-1}P^Tx
\]  

(3)

where \(\Sigma\) is the diagonal matrix of first a singular values, P is the loading vector matrix corresponding to first a singular values. The Hotelling’s \(T^2\)-statistics (3) is scaled squared 2-norm of observation space X, measures systematic variations of the process, and if there is violations, it will indicate that systematic variations are out of control. If \(\alpha\) is the level of significance, the threshold of \(T^2\) statistics can be calculated as (Chiang et al., 2000),

\[
T^2_a = \frac{m\alpha}{n}F_{a}(m,n-m)
\]  

(4)

Where \(F_{a}(m,n-m)\) is known as F-distribution with m and (n-m) degree of freedom (Chiang et al., 2000). Essential condition for fault detection occurs if Hotelling’s \(T^2\)-statistics exceeds its threshold value, that is,

\[
T^2 \leq T^2_a \quad \text{Fault free case}
\]

\[
T^2 > T^2_a \quad \text{sFault case}
\]

A complete flowchart for offline and online computation of the PCA algorithm for fault detection is shown in Fig. 4 and 5, respectively.

3.2. Fisher discriminant analysis

Fisher discriminant analysis is one of the most powerful methods for dimensionality reduction. In case of fault detection, PCA gives very good results. However, it has poor properties of fault classification because it does not consider information (variance) among different classes of data during computation of loading vectors (Jamil et al., 2015). FDA considers information among different classes of data, so it is more favourable for fault classification. It determines a set of transformation vectors, known as FDA vectors. FDA vectors maximise the information (distance) among different classes of data,
while minimising information within each class in projected space. FDA tries to centralise different data classes and feature recognition rates of FDA is better than PCA. According to (Adil et al., 2016), performance of FDA for fault detection and classification is quite better than that of PCA.

The procedure to implement FDA is similar to PCA. First of all, FDA vectors are computed using training data, then these FDA vectors are utilised to transform online data into lower dimensional space. Finally, a discriminant function isolates the fault. In FDA training data, both normal and faulty data is used for computation of FDA vectors, however, in PCA only normal data is used for computation of loading vectors (Adil et al., 2016; Chiang et al., 2000). In order to detect fault with the help of FDA, Hotteling’s $T^2$- statistics is used.

Let us assume that a training set of $m$ process variables, with set of $n$ observations, is shown in the form of input matrix $X \in R^{n \times m}$. Consider $q$ as number of classes in different faults and $n_k$ is number of observations in $k$-th class, let $x_i$ be the transpose of $i$-th row of matrix $X$. The transformation vector $v$ is computed using training data such that following optimisation is solved:

$$ J_{FDA}(v) = \max_{v \neq 0} \frac{v^T S_w v}{v^T S_b v} $$

(5)

Where $S_w$ shows within class scatter matrix given by

$$ S_w = \sum_{k=1}^{q} S_k $$

(6)

With

$$ S_k = \sum_{i \in c_k} (x_i - \bar{x}_k)(x_i - \bar{x}_k)^T $$

(7)

and the mean of $k$th class $\bar{x}_k = \frac{1}{n_k} \sum_{i \in c_k} x_i$ similarly $S_b$ is between class scatter matrix given by

$$ S_b = \sum_{i} (x_i - \bar{x}_k)(x_i - \bar{x}_h)^T $$

(8)

With $\bar{x}$ shows the combined (total) mean vector given by $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ it is stated that solution to above optimisation problem is identical to eigenvalue decomposition problem (Ding, 2014),

$$ S_b v_h = \lambda_h S_w v_h $$

(9)

Where $\lambda_h$ is generalised eigenvalue representing the extent of separability between classes and $\lambda_h$ are respective eigenvectors. Equation (5) shows optimisation problem that ensures minimum scatter within class and maximum scatter between different data classes. This feature helps to classify faults. In order to project online data into lower dimensional space, a matrix $V_q \in R^{p \times q}$ with $q$-1 FDA vectors is defined as, such data projected data $z_i \in R^{q-1}$ is given by

$$ z_i = V_q^T x_i $$

(10)

For fault detection Hotteling’s $T^2$- statistics is used (Yin et al., 2012), given by

$$ T^2_k = x^T V_q (V_q^T S_k V_q)^{-1} V_q^T x $$

(11)

Where $k$ shows the number of non-zero eigenvalues. For a given level of significance $\alpha$, threshold for Hotteling’s $T^2$- statistics is given by:

$$ T^2_k = \frac{a(n-1)(n+1)}{n(n-1)} F_{n-k, n-k} $$

(12)

$$ T^2_k \leq T^2_k \quad \text{Fault free case} $$

$$ T^2_k > T^2_k \quad \text{Fault case} $$

For fault classification, the discriminant function is used as given below:

$$ g_k(x) = -\frac{1}{2}(x - \bar{x}_k)^T V_q (\frac{1}{n_k - 1} V_q^T S_k V_q)^{-1} V_q^T (x - \bar{x}_k) $$

$$ + \ln(q_i) - \frac{1}{2} \ln[det(\frac{1}{n_k - 1} V_q^T S_k V_q)] $$

(13)

In above equation, $g_k(x)$ is the discriminant function associated with class $k$, provided a data vector $x \in R^m$, online data is associated with class $i$ provided that the discriminant function belonging to $i$-th class is maximum for a fault in class $i$, can be expressed as,

$$ g_i(x) > g_i(x) $$

(14)

A complete flowchart for offline training of the FDA algorithm is shown in Fig. 6.

3.3. Support Vector Machines (SVM)

SVM is a well-known data driven technique used for detection and classification of faults due to its generalisation ability and being less susceptible to the curse of dimensionality (Burges, 1997). For the first time, Support vector machines were used by Vapnik (Zhang, 2010). It is one of the new machine learning tools for classification of linear and nonlinear data. SVM is a binary classifier that maximises the margin between two data classes through a hyper-plane as shown in Fig. 7. SVMs maximise the margin near separating hyperplane. The decision of separation is fully identified by the support vectors. Solution of SVM is obtained through solution of quadratic programming.

In SVM, a discriminant function is used to differentiate different classes of data given by:

$$ f(x) = w^T x + b $$

(15)
Where \( b \), the bias, \( x \), the data points, and \( w \), the weighting vector, are obtained through training data. In two-dimensional space, the discriminant function is a line, in three-dimensional space, the discriminant function is a plane, and in \( n \)-dimensional space, the discriminant function is a hyperplane. SVM generates the optimal separating hyperplane by calculating the value of bias, weighting vector in such way that maximum margin is achieved. The points in training set with least perpendicular distance to the hyperplane are known as support vectors. The margin of the optimal separator can be defined as width of separation between support vectors.

\[
\rho = 2 \frac{f(x^0)}{w} = 2r
\]  

### 3.3.1. The Kernel Trick (Feature Space)

The cases in which training data is not linearly separable in the original space using above methods, then, this kind of data can be mapped to a higher-dimensional space which makes the data separable (Nayak, 1998), as shown in Fig. 8.

A kernel function is a type of function that corresponds to an inner product in the higher dimensional space. For example, if data is mapped to feature space through a transformation \( \Phi : x \rightarrow \varphi(x) \), then, the inner product results:

\[
K(x_i, x_j) = \varphi(x_i)^T \Phi(x_j)
\]  

There are different types of kernels, such as, polynomial, linear, Radial Basis Function (RBF) etc. The discriminant function of SVM, can be written as:

\[
f(x) = w^T x + b
\]  

According to Representer theorem, \( w \) can be written as linear combination of input vectors.

\[
w = \sum_{j=1}^{N} \alpha_j x_j
\]  

Thus

\[
f(x) = w^T x + b = b + \sum_{i=1}^{N} \alpha_i x_i^T x
\]  

All the dot products can be replaced with

\[
k(c, d) = c^T d
\]  

Optimisation problem:

\[
\min a, b \quad \frac{1}{2} \sum_{j=1}^{N} (\alpha_j) k(x_i, x_j) + C \sum_{j=1}^{N} \xi_j
\]  

Where \( \xi_j > 0 \)

\[
y_j \sum_{i=1}^{N} \alpha_i k(x_i, x_j) + b \geq 1 - \xi_j
\]  

In order to test the pattern, we use:

\[
f(x) = b + \sum_{i=1}^{N} \alpha_i k(x_i, x)
\]  

Euclidean dot product can be substituted with dot product in feature space “\( \Phi \)”, which will permit nonlinear classification.

\[
k(c, d) = \Phi(c)^T \Phi(d)
\]  

\( k(c, d) \) is known as kernel function and corresponding SVM is called kernelized SVM. This type of SVM can solve the issue of classification of not linear separable data. Steps involved in implementation of kernelized SVM are:
Step 1: Input data is normalised.
Step 2: Training of SVM.
  Step 2.1: Selection of kernel function.
  Step 2.2: Selection of kernel parameter.
  Step 2.3: Optimisation of penalty factor (C).
  Step 2.4: Cross validation.
Step 3: Classification of SVM test data.

3.3.2. Multiclass SVM

Binary class SVM can be used for fault detection, but it cannot be used for fault classification. However in practical cases, discrimination of more than two classes is required, hence, multiclass pattern recognition is often required in real world problems (Xue, 2014). In majority of cases, multiclass pattern recognition problems are decomposed into series of binary problems such that binary pattern recognition techniques can easily be applied in practical cases. multiclass SVM algorithms such as one-versus-one, one-versus-all, can be applied to classify more than two faults.

4. Application of Data Driven Techniques to Diagnose HIF

HIF is introduced at different positions and different phase conductors of IEEE 13-node test feeder as shown in Fig. 9. In data structure, data is generated from Simulink model of test feeder. There are 29 variables of singles phase, two phase and three phase voltages of 13-node test feeder. Data has been placed in input matrix in such a way that each column of input matrix represents voltage and each row of input matrix represents number of observations. There were 400 observations recorded for bus voltages, first 100 observations correspond to normal data, while other 300 observations correspond to three HIF locations at different positions of test feeder.

4.1. Detection of HIF using PCA and Hotteling’s $T^2$ statistics

For diagnoses of HIF using PCA, training data consisting of 60 samples of normal condition (without fault) has been selected while testing data is consist of 100 samples of non-faulty data and 100 samples of faulty data. PCA algorithm has been applied on training data and 29 principal components are obtained. Out of 29 principal components only 5 principal components have been retained, the decision is made on the basis of total variance captured by these 5 principal components. The value of $\alpha$, as mentioned in (4), is taken 0.001. As we have retained 5 principal components so $(1.6983/1.72)=98\%$ of total variance has been captured by first five principal components.

PCA can successfully detect high impedance fault as shown in Fig. 11. In some cases, it is required to classify different types of HIFs such as broken conductor and unbroken conductor HIFs at different locations of feeder. For this purpose, High impedance faults at three different locations are analysed. Fig. 12 shows plot of Hotteling’s $T^2$ statistics to detect HIFs at three different locations. Results show that PCA can successfully detect these three HIFs. Only 5 principal components are retained such that $(1.6983/1.72)=98\%$ of total variance has been captured. Fig. 13 shows projection of training data and test data in two dimensional space. It can be observed that first two components capture most of variation in higher dimensional data. Fig. 11 shows the results of Hotteling’s $T^2$ statistics to detect HIF after applying (3.3) on test data. It can be seen that normal data (first 100 samples) lies below threshold value of $T^2$ statistics, where threshold value is 22.0108. This threshold value was found using significance level of 0.1% and confidence region of 99%.

PCA can successfully detect high impedance fault as shown in Fig. 11. In some cases, it is required to classify different types of HIFs such as broken conductor and unbroken conductor HIFs at different locations of feeder. For this purpose, High impedance faults at three different locations are analysed. Fig. 12 shows plot of Hotteling’s $T^2$ statistics to detect HIFs at three different locations. Results show that PCA can successfully detect these three HIFs. Only 5 principal components are retained such that $(1.6983/1.72)=98\%$ of total variance has been captured. Fig. 13 shows projection of training data and test data in two dimensional space. It can be observed that first two components capture most of variation in higher dimensional data. Fig. 11 shows the results of Hotteling’s $T^2$ statistics to detect HIF after applying (3.3) on test data. It can be seen that normal data (first 100 samples) lies below threshold value of $T^2$ statistics, where threshold value is 22.0108. This threshold value was found using significance level of 0.1% and confidence region of 99%.

PCA can successfully detect high impedance fault as shown in Fig. 11. In some cases, it is required to classify different types of HIFs such as broken conductor and unbroken conductor HIFs at different locations of feeder. For this purpose, High impedance faults at three different locations are analysed. Fig. 12 shows plot of Hotteling’s $T^2$ statistics to detect HIFs at three different locations. Results show that PCA can successfully detect these three HIFs. Only 5 principal components are retained such that $(1.6983/1.72)=98\%$ of total variance has been captured. Fig. 13 shows projection of training data and test data in two dimensional space. It can be observed that first two components capture most of variation in higher dimensional data. Fig. 11 shows the results of Hotteling’s $T^2$ statistics to detect HIF after applying (3.3) on test data. It can be seen that normal data (first 100 samples) lies below threshold value of $T^2$ statistics, where threshold value is 22.0108. This threshold value was found using significance level of 0.1% and confidence region of 99%.
data in two dimensional space, it can be seen that PCA cannot
discriminate between different types of HIFs, this is due to the
reason that PCA do not consider information among different
classes of data. We can conclude that PCA is suitable for HIF
detection but it cannot classify different types of HIFs.

4.2. Detection of HIF using FDA

FDA is applied for detection and isolation of high impedance
faults in power distribution systems. In order to compute FDA
vectors, both normal and faulty data is used, in this work, 60
samples in training data and 40 samples in test data correspond-
ing to each scenario, that is, faulty and non-faulty case. Fig. 14
shows projection of training data in two-dimensional space by
FDA. In second step, after computation of transformation vec-
tors (FDA vectors), the discriminant function is used to test on-
line data. Fig. 15 shows plot of discriminant function in each
category. It can be observed that up to first 30 samples, value of
discriminant function corresponding to normal case has max-
imum magnitude, which shows that there is no fault in test
feeder. Similarly, after 30 samples, value of discriminant func-
tion corresponding to fault at position A has maximum magni-
tude, which shows that fault at position has occurred. Same is
the case with fault at position B and C. Zoomed view of plot is
shown in Fig. 16.

The above results have shown that FDA can successfully
isolate/locate HIF. This technique is very well suited for moni-
toring of power distribution systems.

4.3. Detection of HIF using SVM

Support vector machine algorithm has been applied for 29-
dimensional data without any dimensional reduction technique.
Selection of optimal value of penalty factor is important, this
is done by performing nested 3-fold cross validation in original
data. With the help of cross validation, average area under the
curve was computed for 1000 values of penalty factor between
0.1 and 100. After selection of optimal value of C, SVM classi-
fiers were trained with optimal penalty factor and validated on
training data so that generalisation would be checked. Test data
of HIF was classified by validated SVM classifier. The pre-
dicted labels of test data fairly detects the occurrence of fault,
that is -1 for non-faulty data and +1 for faulty data as shown in

Fig. 14: Projection of training data in two dimensional space by FDA

Fig. 15: Plot of discriminant function for multiple HIF detection using FDA

Fig. 16: Zoomed view of discriminant function for multiple HIF detection using FDA
4.4. Detection and Classification of HIF using M-SVM

Binary class SVM can be used for fault detection, but it cannot be used for fault classification. However, in Power distribution systems, discrimination of more than two classes is required, hence, multiclass pattern recognition is often required in monitoring Power distribution systems. Multiclass SVM (M-SVM) classifier is obtained using training of non-fault cases with class label 4, fault at position A with class label 3, fault at position B with class label 2, and fault at position C with class 1. In each classifier, during training, a Gaussian Radial Basis Function kernel with a scaling factor, sigma ($\sigma$), of 0.5 and a penalty factor of 10 is used. The tolerance value for Karush-Kuhn-Tucker (KKT) condition for the training of data is taken as 0.001. The value of regularization parameter, lambda ($\lambda$) is 1. Test data is classified using the trained classifiers for 50 observations of each data class and predicted labels were differentiated with known data labels.

Fig. 18 shows that up to 50 samples, the predicted labels belong to normal class data, indicating that there is no fault. After first 50 samples, predicted labels belong to class label 3, indicating that the fault is occurred at position A. After first 100 samples, predicted labels belong to class label 2, indicating that the fault is occurred at position B. Similarly, after first 150 samples, predicted labels belong to class label 1, indicating that the fault is occurred at position C. Similarly, Fig. 19 show the score plot of test data for each class of data.

A comparison is presented to evaluate the results of the proposed technique with those from literature and observations have been recorded in Table 1. After testing the technique on 400 test cases, it is found that proposed method is extremely quick and efficient in detecting HIFs. It is demonstrated that PCA successfully detects HIF but it cannot classify HIFs. Compared to that, FDA can also successfully classify/locate the fault. Further superior results are achieved by M-SVM, fault classification rate of SVM is better than FDA. M-SVM algorithm can detect all types of HIF and is also robust against capacitor and load switching transients in distribution network.

5. Conclusion

In this research paper, high impedance fault detection and classification in power distribution systems has been studied using data driven techniques. Source-diode-resistance model consisting of two diodes with opposite polarity connected to DC sources is utilised to simulate the high impedance fault. Data driven techniques including PCA, FDA, and SVM are applied to detect/classify HIFs. PCA along with Hotteling’s $T^2$ statistics to detect HIFs, it is demonstrated that PCA successfully detects HIF but it cannot classify HIFs. Compared to that, FDA can also successfully classify/locate the fault. Further superior results are achieved by M-SVM, fault classification rate of SVM is better than FDA. M-SVM algorithm can detect all types of HIF and is also robust against capacitor and load switching transients in distribution network.

Conflict of Interest

The authors declare no conflict of interest.
Table 1: Comparison of performance indices of the proposed M-SVM method with previous techniques

| Method                                      | Security (%) | Dependability (%) |
|---------------------------------------------|--------------|-------------------|
| Wavelet transform (Chen et al., 2016)       | 68.5         | 72                |
| Time frequency transform Samantaray et al. (2008) | 81.5         | 98.3              |
| Morphological gradient (Sarlak and Shahrtafs, 2011) | 96.3         | 98.3              |
| Mathematical Morphology (Gautam and Brahama, 2012) | 100          | 100               |
| The proposed method (M-SVM)                 | 100          | 100               |

References

Adamiak, M., Wester, C., Thakur, M., Jensen, C., 2006. High impedance fault detection on distribution feeders. GE Industrial solutions.

Adil, M., Abid, M., Khan, A.Q., 2016. Comparison of PCA and FDA for monitoring of coupled liquid tank system. Pakistan.

Bakar, A.H.A., Ali, M.S., Tan, C., Mokhlas, H., Arof, H., Illias, H.A., 2014. High impedance fault location in 11 kV underground distribution systems using wavelet transforms. Electrical Power and Energy Systems 55, 723–730.

Bishop, R.E.L.M.T., 1985. A comparison of measured high impedance fault data to digital computer modeling results. IEEE Transactions on Power Apparatus and Systems 104, 2754–2758.

Biswal, A.P.K.T., 2017. Detection of high impedance fault in distribution systems considering distributed generation, Bangalore, India.

Brahma, S.G.S.M., 2013. Detection of high impedance fault in power distribution systems using mathematical morphology. IEEE TRANSACTIONS ON POWER SYSTEMS 28, 1226–1234.

Burgess, C.J.C., 1997. A Tutorial on Support Vector Machines for Pattern Recognition. Kluwer Academic Boston.

Chen, J., Phung, T., Blackburn, T., Ambikairajah, E., Zhang, D., 2016. Detection of high impedance faults using current transformers for sensing and identification based on features extracted using wavelet transform. IET generation, transmission & distribution 10, 2990–2998.

Chen, J.C., Phung, B.T., Zhang, D.M., Blackburn, T., 2013. Study on High Impedance Fault Arcing Current Characteristics. TAS, Australia/A, Hobart.

Chiang, L., Russel, E., Braatz, R., 2000. Advanced Textbooks in Control and Signal Processing. Springer New York.

Ding, S.X., 2014. Data-driven Design of Fault Diagnosis and Fault-tolerant Control Systems. Duisburg, Germany: Springer.

Elkalasy, N.I., Lehtonen, M., Darwish, H.A., 2007. A novel selectivity technique for high impedance arcing fault detection in compensated mv networks. EUROPEAN TRANSACTIONS ON ELECTRICAL POWER 18, 344–363.

Etemadi, A., Sanaye-Pasand, M., 2008. High-impedance fault detection using multi-resolution signal decomposition and adaptive neural fuzzy inference system. IET generation, transmission & distribution 2, 110–118.

Ferdowsi, F., Vahedi, H., Edrington, C.S., 2017. High impedance fault detection utilizing real-time complexity measurement. TX, USA.

Gautam, S., Brahma, S.M., 2012. Detection of high impedance fault in power distribution systems using mathematical morphology. IEEE Transactions on Power Systems 28, 1226–1234.

Ghaderi, A., Ginn, H.L., Mohammadpour, H.A., 2016. High impedance fault detection: A review. Electric Power Research Systems 143, 376–388.

Hou, D., 2007. Detection of High-Impedance Faults in Power Distribution Systems. Duisburg, Germany: Springer.

Janul, F., Abid, M., Haq, I., Khan, Q.A., Iqbal, M., 2015. Fault diagnosis of pakistan research reactor-2 with data-driven techniques. Annals of Nuclear Energy 90, 433–440.

Jones, B.M.A.R.H., 1996. High impedance fault detection implementation issues. IEEE 11, 139.

Kavi, M., Mishra, Y., Vilathgamuwa, D.M., 2016. Detection and identification of high impedance faults in single wire earth return distribution networks. QLD, Australia, Brisbane.

Lai, T.M., Snider, L.A., Lo, E., Sutanto, D., 2005. High-impedance fault detection using discrete wavelet transform and frequency range and rms conversion. IEEE TRANSACTIONS ON POWER DELIVERY 20, 397–407.

Lima, E.M., dos Santos Junqueira, C.M., Brito, N.S.D., de Souza, B.A., de Almeida Coelho, R., de Medeiros, H.G.M.S., 2018. High impedance fault detection method based on the short-time fourier transform. IEEE Generation, Transmission & Distribution 12, 2577–2584.

Lopes, F.V., Santos, W.C., Fernandes, D., Neves, W.L., 2013. A transient based approach to diagnose high impedance faults on smart distribution networks. Mai, W., Phung, B.T., Ambikairajah, E., 2012. Detection of high impedance faults in medium voltage distribution networks. Vietnam.

Nayak, C.M.P., 1998. Support vector machines and machine learning on documents. In Introduction to information Retrieval.

Sahoo, S., Baran, M.E., 2014. A method to detect high impedance faults in distribution feeders, in: 2014 IEEE PES T&D Conference and Exposition. IEEE. pp. 1–6.

Samantaray, S., Panigrahi, B., Dush, P., 2008. High impedance fault detection in power distribution networks using time–frequency transform and probabilistic neural network. IET generation, transmission & distribution 2, 261–270.

Samantaray, S.R., 2012. Ensemble decision trees for high impedance fault detection in power distribution network. Electrical Power and Energy Systems 43, 1048–1055.

Santos, W.C., Lopes, F.V., Brito, N.S.D., 2017. High-impedance fault identification on distribution networks. IEEE TRANSACTIONS ON POWER DELIVERY 32, 23–32.

Sarlak, M., Shahtash, S., 2011. High impedance fault detection using combination of multi-layer perceptron neural networks based on multi-resolution morphological gradient features of current waveform. IET generation, transmission & distribution 5, 588–595.

Sedighi, A.R., Haghifam, M.R., Malik, O.P., 2005. Soft computing applications in high impedance fault detection in distribution systems. Electric Power Systems Research 76, 136–144.

Sekar, K., Mohanty, N.K., 2018. A fuzzy rule base approach for high impedance fault detection in distribution system using morphology gradient filter. Journal of King Saud University-Engineering Sciences.

Sheelvant, B.P.P.V.R., 2015. High-impedance fault detection using wavelet transform. International Journal of Engineering Research and General Science 3, 166–172.

Stoupis, J., Maharsi, M., Nuqui, R., 2004. Ground alert: Reliable detection of high-impedance fault caused by downed conductor. ABB Review, Jan.

Sukatsuki, N.Z.I.K., 2006. Modelling arcing high impedances faults in relation to the physical processes in the electric arc. Lisbon, Portugal.

Torres, V., Guardado, J.L., Ruiz, H.F., Maximov, S., 2014. Modeling and detection of high impedance faults. Electrical Power and Energy Systems 61, 163–172.

Xue, Z.W.K., 2014. Multi-Class Support Vector Machine in Support Vector Machines Applications. Springer.

Yin, S., Ding, S.X., Hahami, A., Hao, H., Zhang, P., 2012. A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark tennessee eastman process. Journal of Process Control 22, 1567–1581.

Zanjani, M.G.M., Kargar, K.H., Zanjani, M.G.M., 2012. High impedance fault detection of distribution network by phasor measurement units. Israel, Iran, pp. 2–3.

Zhang, Z.Y.Z., 2010. A New Multi-class Classification Algorithm of Support Vector Machine. Guangzhou, China.

Zoric, K.J., Djuric, M.B., V.V., 1997. Arcing faults detection on overhead lines from the voltage signals. International Journal of Electrical Power & Energy Systems 19, 299–303.