Fault detection in distribution networks in presence of distributed generations using a data mining–driven wavelet transform

Youness Mohammadnian1, Turaj Amraee1, Alireza Soroudi2
1Faculty of Electrical Engineering, K. N. Toosi University of Technology, Tehran, Iran
2University College Dublin, Dublin, Ireland

Abstract: Here, a data mining–driven scheme based on discrete wavelet transform (DWT) is proposed for high impedance fault (HIF) detection in active distribution networks. Correlation between the phase current signal and the related details of the current wavelet transform is presented as a new index for HIF detection. The proposed HIF detection method is implemented in two subsequent stages. In the first stage, the most important features for HIF detection are extracted using support vector machine (SVM) and decision tree (DT). The parameters of SVM are optimised using the genetic algorithm (GA) over the input scenarios. In second stage, SVM is utilised to classify the input data. The efficiency of the utilised SVM-based classifier is compared with a probabilistic neural network (PNN). A comprehensive list of scenarios including load switching, inrush current, solid short-circuit faults, HIF faults in the presence of harmonic loads is generated. The performance of the proposed algorithm is investigated for two active distribution networks including IEEE 13-Bus and IEEE 34-Bus systems.

1 Introduction

High impedance faults (HIFs) could not be detected by insensitive overcurrent protective relays due to low fault current in distribution networks. HIFs often occur when the conductor is broken and comes into contact with high impedance surfaces such as branches of a tree. As of low fault current, the main objective of HIF detection is not to protect the system, but to protect the human and animals lives and prevent fire hazards [1].

The HIF detection methods are categorised into time domain methods, frequency domain techniques, time-frequency domain methods, and intelligent methods. Current or voltage waveforms of HIF faults have some unique characteristics due to non-linearity and randomness of these faults. Proportional relaying [2], ratio ground relaying [3], fault current flicker and half-cycle asymmetry [4], and fractal techniques [5] are examples of time domain methods for HIF detection. As of the arcing phenomenon during this fault, current waveform of HIF contains low- and high-frequency components. Frequency domain algorithms use third harmonics [6], second-fourth-sixth harmonics [7], and high-frequency components (2–10 kHz) of fault current to detect HIF faults [8]. Another category of HIF detection methods is the time-frequency domain technique, in which some features of HIF is detected based on discrete wavelet transform (DWT) [9–12], S transforms [13, 14], and Kalman filtering [15]. Intelligent methods act based on training over a given predetermined scenario such as decision tree (DT) [16], neural networks [1, 17, 18], ANFIS algorithms [19], and support vector machine (SVM) [13, 20–22].

Some recent works have proposed new method for HIF detection. The proposed method in [23] detects HIF faults using the even harmonics of the voltage waveforms measured by smart meters. In [24], the proposed algorithm utilises mathematical morphology (MM) techniques for HIF detection based on filtering functions. In [25], the authors utilise one-cycle sum of superimposed components of residual voltage, that is the maximum value of one-cycle sum of superimposed components of negative-sequence current for HIF detection. In [26], a transient-based algorithm for HIF detection in distribution systems has been developed using the DWT.

To increase reliability and security of HIF detection method, it is necessary to present more effective features with high information gain. For this purpose here, a two-stage scheme is presented using DWT. In the first stage, the feature selection is carried out using SVM technique. Based on DWT, the correlation between phase current signal and its details is presented as a new HIF predictor. In the second stage, some classifier methods are utilised to detect HIF conditions. The performance of the proposed index is compared through a general proposed structure to other classification methods via the dependability and security criterion. The two major advantages of the current work with respect to the method discussed in literature review are: (a) The correlation between phase current signal and its details is presented as a new HIF predictor. This feature needs just the current signal and (b) Optimising the SVM parameters using GA algorithm.

This paper is organised as follows. In Section 2, the general structure of proposed method is explained. Section 3 describes the simulation results of the proposed algorithm based on the security and dependability criterion. Finally, the conclusion is given in Section 5.

2 HIF detection scheme

Generally, data mining techniques for HIF detection have two major parts. In the first part, a classifier is trained using a large number of input-output training scenarios. These scenarios include all credible HIF and non-HIF conditions. In other word, the first part is an offline procedure to construct the HIF classifier. The second part is devoted to the application of trained classifier for the online detection of HIF fault. Indeed, the second part has a very low computational burden and could identify the HIF in a fraction of second.

The flowchart of the proposed HIF detection method is depicted in Fig. 1. Based on this structure, the proposed method has three major parts including generating input scenarios, feature extraction, and HIF detection. Each part is described as follows:

2.1 Feature extraction

Here, feature extraction is carried out by wavelet transform. The most important features are extracted using DWT. DWT is utilised to determine the approximation and details coefficients of fault current signal. The correlation of fault current signal with its details is developed as a HIF detection feature.
2.1.1 Discrete wavelet transform: DWT can provide time and frequency domain characteristics simultaneously. In DWT, the scaling function is given by (1), and wavelet function as given in (2) is used to decompose the signal to different levels.

\[
\varphi_{j,k}(t) = 2^{j/2} \varphi(2^j t - k), \quad -\infty < j, k < +\infty
\]

(1)

\[
\varphi_{j,k}(t) = 2^{j/2} \phi(2^j t - k), \quad -\infty < j, k < +\infty
\]

(2)

According to [27], \(\varphi(t)\) and \(\varphi(t)\) can be expressed in terms of a weighted sum of shifted \(\varphi(2t)\) and \(\phi(2t)\) as:

\[
\varphi(t) = \sum_{n} h(n) \varphi(2t - n) \quad n \in Z
\]

(3)

\[
\phi(t) = \sum_{n} g(n) \phi(2t - n) \quad n \in Z
\]

(4)

where \(h(n)\) are scaling function coefficients and \(g(n)\) can be obtained by (5).

\[
g = (-1)^n h(1 - n)
\]

(5)

After the first level of decomposition, DWT divides input signals into high-frequency and low-frequency sub-bands by high-pass and low-pass filters. Therefore, the input signal can be rebuilt using its approximation and detail coefficients as given by (6).

\[
x(t) = \sum_{k=-\infty}^{\infty} a_j(k) \varphi_{j,k}(t) + \sum_{k=-\infty}^{\infty} \sum_{L=1}^{L} d_L(k) \varphi_{j-L,k}(t)
\]

(6)

where

\[
a_j(t, \varphi_{j-1,k}(t)) = \int x(t) 2^{(j-1)j} \varphi(2^{j-1} t - k)
\]

(7)

By using (3) and (7), (8) can be written as

\[
\varphi(2^{j-1} t - k) = \sum_{n} h(n) 2^{j/2} \varphi(2^{j-1} t - n - k)
\]

(8)

By assuming that \(m = 2k + n\), (8) and (7) are written as follows.

\[
\varphi(2^{j-1} t - k) = \sum_{n} h(m - 2k) 2^{j/2} \varphi(2^{j-1} t - m)
\]

(9)

\[
a_j = \sum_{m} h(m - 2k) \int x(t) 2^{j/2} \varphi(2^{j-1} t - m) dt
\]

(10)

Finally, the approximations and details of the input signals are determined as given by (11) and (12) respectively.

\[
a_{j}(k) = \sum_{m} h(m - 2k) a_{j-1}(m)
\]

(11)

\[
d_{j}(k) = \sum_{m} g(m - 2k) a_{j-1}(m)
\]

(12)

After the first level of signal decomposition, a high-pass and low-pass signals are obtained. This process is then applied on low-pass signal for several levels. Fig. 2 shows signal decomposition using high- and low-pass filters into three levels. The output of low-pass decomposition (i.e. \(a_j\)) and high-pass decomposition (i.e. \(d_j\)) gives the approximation and detail coefficients, respectively.

2.1.2 Correlation coefficient: Covariance is an index that shows the dependence between two variables. For two vectors \(X\) and \(Y\), the covariance will be defined as (13).

\[
COV(X, Y) = \frac{1}{n} \sum_{i=1}^{n} (X_i - \mu_X)(Y_i - \mu_Y)
\]

(13)

where \(\mu_X\) and \(\mu_Y\) are means of \(X\) and \(Y\), respectively. If there is no correlation between two vectors, their covariance will be equal to zero and they have no linear dependency. The amount of
correlation coefficient for continuous and discrete variables is calculated using (14) and (15), respectively.

\[
\text{Correlation}(X, Y) = \frac{\text{COV}(X, Y)}{\sqrt{\text{Var}(X) \cdot \text{Var}(Y)}} \tag{14}
\]

\[
\text{Correlation}(X, Y) = \frac{\sum_{i=1}^{n} (X_i - \mu_X)(Y_i - \mu_Y)}{\sum_{i=1}^{n} (X_i - \mu_X)^2 \sum_{i=1}^{n} (Y_i - \mu_Y)^2} \tag{15}
\]

The value of correlation coefficient given by (14) varies between -1 and 1. The correlation of detail coefficients of phase current signal in DWT is developed as a new feature for HIF detection. For calculating the new index, phase current signal and related detail coefficients are needed. The correlation between these two signals as a new feature is explained in (16).

\[
\text{Correlation}(X_d, Y_d) = \frac{\sum_{i=1}^{n} (X_i - \mu_X)(Y_i - \mu_Y)}{\sqrt{\sum_{i=1}^{n} (X_i - \mu_X)^2 \sum_{i=1}^{n} (Y_i - \mu_Y)^2}}, \quad k = 1, 2, \ldots, 8 \tag{16}
\]

This feature can provide remarkable information gain for distinguishing HIFs. DWT is applied to all simulated scenarios, and the extracted details are recorded as vectors. The value of correlation between every detail coefficients of the wavelet transform in each level and phase current signal is computed using (16). The developed correlation coefficients along with the other indices are used as input features for HIF detection.

### 2.2 Dimension reduction methods

A critical goal in pattern recognition problem or data classification is finding the optimal combination of indices that classify data with high accuracy. The existence of redundant features besides the essential features will cause computational burden and complexity in classification systems. Therefore, the classification systems may result in misclassification. By dimension reduction and feature selection, the computational burden of classification algorithm will be reduced, and extracted patterns can be easily implemented in hardware [28]. To reach this goal, the accuracy error in classification algorithm is considered as the fitness function in optimising algorithm. The dataset is divided into training and test data. The error of classification for each subset will be examined by K-Fold cross-validation. The subset with minimum classification error will be chosen as the best subset of final features. Here, SVM is utilised to extract the optimal input features. The error of SVM is used as fitness function of genetic algorithm (GA), and the best subset is then selected.

#### 2.2.1 Support vector machine

Due to the non-linearity of HIF phenomena, the non-linear SVM algorithm is developed for HIF detection. The non-linear SVM has two stages. In the first stage, input data are mapped to a higher dimension space. After that, in the second stage, the algorithm searches for linear separating hyper planes in new space. The task of this stage is formulated as a quadratic optimisation problem. In SVM, the original finite-dimensional space may not be linearly separable. Therefore, the original finite-dimensional space is mapped into a high-dimensional feature space. However, working with such high-dimensional space increases the computational burden of classification. Therefore, kernel functions are utilised. The kernel function is applied on initial data, and the dimension of these new data is lower than the initial data. Thus using the kernel function, the direct calculation of mapping function could be ignored [29].

\[
\text{Min} \frac{1}{2} \sum_{i,j} a_i a_j y_i y_j K(X_i, X_j)
\]

\[
= \frac{1}{2} \sum_{i,j} a_i a_j y_i y_j K(X_i, X_j)
\]

where

\[
K(X, X_j) = (X)^{T} F(X_j)
\]

Here, radial basis function (RBF) is utilised as the kernel.

\[
K(X, X_j) = \exp(- \frac{1}{2 \delta^2} \| X_j - X_j \|^2)
\]

where \(\delta\) is the spread parameter of the kernel function. The objective function is formed for a specified number of inputs and the support vectors (SV) \(a_i\) are extracted by solving (9). For example, the minimisation of \(n\) input data like \((X, Y)\) is performed as follows:

\[
\text{Min} \frac{1}{2} a^T \alpha^T + f^T \alpha
\]

s.t.

\[
H = \sum_{i} \sum_{j} y_i y_j K(X_i, X_j)
\]

\[f = -[1]_{i=1}^n\]

\[Y, \alpha = 0 \quad 0 \leq \alpha \leq C\]

SV set includes Support Vectors which is expressed as given in (10):

\[
\text{Support Vectors}: SV : \{i \mid 0 \leq \alpha_i \leq C\} \quad i = 1, 2, \ldots, n
\]

In the next stage, the value of \(b\) can be obtained using SV as given in (22).

\[
b = \frac{1}{|\alpha|} \sum_{i=1}^{n} \left(y_i - \sum_{j} a_j K(X_i, X_j)\right)
\]

Finally, the class label of test data is predicted as follows:

\[
\text{Outputs}: Y = \text{Sign} \left(\sum_{j} a_j K(X_i, X_j) + b\right)
\]

#### 2.2.2 Dimension reduction

The GA [30] is used for selecting the best features and determines the optimal parameters of SVM. Here, the classification error and the number of features are integrated to design the fitness function of GA algorithm. Therefore, the fitness function value is minimum for any member of population which has the high accuracy and fewer features. The utilised fitness function is expressed as follows:

\[
\text{Fitness} = W_A \times (1 - \text{SVM}_{\text{accuracy}}) + W_F \times \sum_{i=1}^{n} F_i
\]

where \(\text{SVM}_{\text{accuracy}}\) is the accuracy of classification based on 5-fold cross-validation and \(n\) is the number of features. If features \(i\) is selected, the value of \(F_i\) is equal to 1 and vice versa. \(W_A\) and \(W_F\) are weighting factors for error of classification and number of features. Here, the values of \(W_A\) and \(W_F\) are assumed as 18 and 0.55, respectively.

Indeed, a high priority is considered for the classification accuracy. Based on the required levels of accuracy and the maximum desired number of input features, these weighting factors can be adjusted properly. Without considering the second term of
the fitness function \( W_F = 0 \), the proposed classifier tends to select a high number of input features which is not practically preferred. Using these weighting factors, the proposed algorithm seeks for the minimum accuracy with minimum possible number of input features.

It is noted that the premature convergence is a common problem in GAs, as it leads to a suboptimal solution due to rapid convergence. However, there are some strategies to prevent this phenomenon such as increasing the population size or utilising a uniform crossover.

3 Simulation results and discussion

Here, the performance of the proposed method is verified. The proposed algorithm is implemented in IEEE 34-Bus and IEEE 13-Bus. The results of simulations are presented in six parts including modelling of test systems, HIF model, scenario generation, features extraction, dimension reduction, and classification.

Here, db4 as a most commonly used mother wavelet in Daubechies family wavelets has been selected. This type of mother wavelet is asymmetric, orthogonal, and biorthogonal. The details of utilised mother wavelet are similar to the settings utilised in [10]. The sampling frequency is 10 kHz. The scale of frequency conversion (i.e. the frequency bound) is assumed as (2.5–5 kHz), (1.25–2.55 kHz), (0.626–1.25 kHz), (0.313–0.626 kHz), and (0.157–0.313 kHz), for level 1 to level 5 of decomposition, respectively.

3.1 Test systems

HIF and other similar events are simulated in Matlab/Simulink. Different scenarios are discussed in test systems with and without photovoltaic. Single line diagram IEEE 34-Bus and IEEE 13-Bus test systems are shown in Fig. 3. Data of IEEE 13-Bus test system and IEEE 34-Bus test grid could be found in [31]. It has been assumed that a photovoltaic generating unit is installed at bus 680 and bus 840 in IEEE 13-bus and IEEE 34-Bus test systems, respectively. Data of this generating unit have been given in Table 1. Also the parameters of GA have been reported in Table 2.

3.2 Model of HIF

This paper uses the modified Emanuel model for HIF, which has been obtained based on actual tests [32]. The HIF current in the steady state has half-cycle asymmetry and non-linear variations. According to Fig. 4, two resistances, \( R_p \) and \( R_n \), represent the fault resistance with different values that build asymmetric behaviour of fault currents. In order to model the type of contact surfaces of HIF model, two anti-parallel DC sources are used. The amplitudes of these sources depend on the density and moisture of the soil. By decreasing the soil density and moisture, the value of these DC sources is increased.

When a broken conductor falls on a high impedance surface, some arcs coincide. To model such non-linearity, this paper uses several HIF models in parallel as illustrated in Fig. 5. The fulfilment of the utilised HIF model is investigated by measuring some current harmonics such as third, fifth, and seventh.

The current waveform for a given HIF condition has been depicted in Fig. 6. It has all steady state (i.e. asymmetric and non-linear behaviour) and transient (i.e. build-up and shoulder) characteristics.
3.3 Simulated scenarios

To evaluate the proposed method, a comprehensive list of input scenarios including load switching, capacitor bank switching, transformer inrush current, non-linear and harmonic loads, and different short circuit faults are simulated. Also, different HIFs are simulated in variant states.

Table 3 describes the type of events and number of input scenarios. The total number of non-HIF and HIF scenarios in each of IEEE 13-Bus and IEEE 34-Bus are 494 and 582 scenarios, respectively.

![Image](52x472 to 288x626)

**Table 7** Frequency components of simulated HIF current

According to Fig. 7, the modified Emanuel HIF model produces characteristics similar to the actual HIF current.

3.4 Dimension reduction

In Section 2, the procedure of dimension reduction using GA and SVM algorithms was presented. The dimension of each sample in the prepared dataset for HIF detection is 56. It is noted that just some effective and important features are selected by using SVM and GA. Results of the proposed method are compared with DT feature selection results [33, 34]. Table 4 introduces the input features. The results of two methods imply that correlation index is one of the most important features.

3.5 Classification

Here, the selected features are used as input features for HIF classification. The performance of classification methods is verified using the proposed indices. Table 4 and Table 5 give the results of classification for both test system using the input features determined by the GA and DT algorithms, respectively. Values of security and dependability indices and accuracy of classification algorithm are reported in Table 5 and Table 6. According to the obtained results, it can be seen that the proposed correlation coefficient has increased the accuracy of the classification significantly. For IEEE 13-Bus system, the features including \( F_{24}, F_{20}, F_{15}, F_1 \) and \( F_4 \) are selected. Also, for IEEE 34-Bus system, the features \( F_{24}, F_{20}, F_{15}, F_1 \) are selected. In both selected subsets of features, the presented feature (correlation coefficients) is an efficient feature.

According to Table 4, the initial number of input features is equal to 56. According to Table 5, by the proposed weighting factors, the best accuracy is obtained just using four input features. Table 6 gives the information subsets of selected features by DT for both test systems. To reduce obtained by DT includes too many features, just the six features that have the highest information gain values would be selected as the selected features for classifying algorithms input. In order to prove the efficacy of the correlation

![Image](52x653 to 288x793)

**Fig. 6** Fault current using modified Emanuel HIF model

**Fig. 7** Frequency components of simulated HIF current

**Table 3** Description of simulated events

<table>
<thead>
<tr>
<th>Event type</th>
<th>Number of scenarios</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IEEE 13-bus</td>
</tr>
<tr>
<td>harmonic loads</td>
<td>40</td>
</tr>
<tr>
<td>load switching</td>
<td>44</td>
</tr>
<tr>
<td>inrush current</td>
<td>42</td>
</tr>
<tr>
<td>short circuit</td>
<td>192</td>
</tr>
<tr>
<td>capacitor switching</td>
<td>32</td>
</tr>
<tr>
<td>high impedance faults</td>
<td>144</td>
</tr>
<tr>
<td>total</td>
<td>494</td>
</tr>
</tbody>
</table>

**Table 4** Input features description for eight decomposition

<table>
<thead>
<tr>
<th>Selected feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( F_{v} \ldots F_{8} )</td>
<td>energy for detail coefficients</td>
</tr>
<tr>
<td>( F_{v} \ldots F_{14} )</td>
<td>power of detail coefficients</td>
</tr>
<tr>
<td>( F_{25} \ldots F_{24} )</td>
<td>RMS for detail coefficients</td>
</tr>
<tr>
<td>( F_{4} \ldots F_{9} )</td>
<td>mean for detail coefficients</td>
</tr>
<tr>
<td>( F_{1} \ldots F_{3} )</td>
<td>entropy for detail coefficients</td>
</tr>
<tr>
<td>( F_{53} \ldots F_{56} )</td>
<td>correlation between phase current and details</td>
</tr>
</tbody>
</table>

**Table 5** Results for different HIF detection methods using selected features by GA

<table>
<thead>
<tr>
<th>Method</th>
<th>Security</th>
<th>Dependability</th>
<th>Train accuracy</th>
<th>Test accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>98.50</td>
<td>100</td>
<td>99.49</td>
<td>98.99</td>
</tr>
<tr>
<td>PNN</td>
<td>100</td>
<td>96.87</td>
<td>98.99</td>
<td>98.99</td>
</tr>
</tbody>
</table>

\[ \text{Dependability} = 100 \times \frac{\text{number of HIFs detected as HIF}}{\text{number of all HIF cases}} \]  
\[ \text{Security} = 100 \times \frac{\text{number of NonHIFs detected as NonHIF}}{\text{number of all NonHIF cases}} \]
The classification accuracy, security, and dependability indices without considering the correlation index has been reported in Table 6. It can be seen that without considering the correlation index, the overall classification is reduced significantly. Comparison between Table 6 and 7 implies that the dimensions of selected features are increased by eliminating the correlation index. This issue incurs more computational burden and complexity in extracted patterns and hence the classifier efficacy is deteriorated accordingly.

In other words, the proposed correlation index between the original fault current signal and its details will promote the overall accuracy of the HIF detection. Fig. 8 shows the best three-dimensional view of the three selected features obtained by GA-SVM in both simulated test cases.

Figs. 9 and 10 display the performance of the proposed method for a given HIF fault and a load switching event, respectively. In order to verify the effect of each event on detail coefficients, the noise is removed from the waveform of fault current signal. As shown in Fig. 9c, the HIF detection signal in HIF condition is activated two cycles after the fault inception. However, as shown in Fig. 10c, the HIF detection signal under the load switching event is not activated. According to the simulation results obtained for IEEE 13-Bus and IEEE 34-Bus test systems, a remarkable improvement in HIF detection accuracy is obtained by utilising the correlation coefficients index beside the other extracted statistical indices. As reported in Table 3, the highest accuracy is achieved by GA-SVM method for both test systems. The highest accuracy is 98.99 and 98.27% for IEEE 13-Bus and IEEE 34-Bus test systems.

In [20], the accuracies of the SVM algorithm, Bayes, Parzen, and Nearest Neighbour (NN) algorithms have been reported. As given in [20], the best accuracy for Bayes, NN, Parzen, SVM-Linear, SVM-Polynomial, SVM_RBF has been obtained as 80, 97.5, 92.5, 77.5, 97.5, and 97.5%, respectively. However, according to Table 4, the accuracy of the proposed method of this paper is equal to 98.27 and 98.29% for IEEE 13-Bus and IEEE 34-Bus test systems, respectively.

### 4 Conclusion

Here, a reliable HIF detection method based on pattern recognition is proposed. To extract features in time-frequency domain, the DWT is used. This process is performed by decomposing two cycles of fault current signal and extracting statistical features from detail coefficients of DWT in each level. After fault signal decomposition, correlation coefficients are introduced as a new index. Only some features from 56 features are chosen by DT and

---

**Table 6** Results for different HIF detection methods using selected features by DT

<table>
<thead>
<tr>
<th>Method</th>
<th>Dependability</th>
<th>Security</th>
<th>Train accuracy</th>
<th>Test accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>98.50</td>
<td>81.25</td>
<td>99.24</td>
<td>92.93</td>
</tr>
<tr>
<td>PNN</td>
<td>95.45</td>
<td>90.625</td>
<td>96.71</td>
<td>93.88</td>
</tr>
<tr>
<td>feature</td>
<td>$F_{24} - F_{35} - F_{54} - F_{7} - F_{11} - F_{1}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SVM</td>
<td>97.14</td>
<td>97.83</td>
<td>99.14</td>
<td>97.41</td>
</tr>
<tr>
<td>PNN</td>
<td>91.43</td>
<td>100</td>
<td>99.14</td>
<td>94.83</td>
</tr>
<tr>
<td>feature</td>
<td>$F_{26} - F_{35} - F_{2} - F_{11} - F_{15} - F_{1}$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>IEEE 13-bus</th>
<th>IEEE 34-bus</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>99.23</td>
<td>98.23</td>
</tr>
<tr>
<td>PNN</td>
<td>90.66</td>
<td>89.80</td>
</tr>
<tr>
<td>feature</td>
<td>$F_{31} - F_{51} - F_{3} - F_{4} - F_{2}$</td>
<td>$F_{31} - F_{51} - F_{3} - F_{4} - F_{2}$</td>
</tr>
</tbody>
</table>

**Table 7** Results of HIF detection methods using selected features by GA without correlation index

<table>
<thead>
<tr>
<th>Method</th>
<th>Dependability</th>
<th>Security</th>
<th>Train accuracy</th>
<th>Test accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>96.97</td>
<td>96.87</td>
<td>98.23</td>
<td>96.97</td>
</tr>
<tr>
<td>PNN</td>
<td>98.48</td>
<td>71.87</td>
<td>90.66</td>
<td>89.80</td>
</tr>
<tr>
<td>feature</td>
<td>$F_{23} - F_{17} - F_{8} - F_{6} - F_{2}$</td>
<td>$F_{23} - F_{17} - F_{8} - F_{6} - F_{2}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SVM</td>
<td>94.29</td>
<td>100</td>
<td>98.28</td>
<td>96.55</td>
</tr>
<tr>
<td>PNN</td>
<td>91.43</td>
<td>97.83</td>
<td>96.35</td>
<td>94</td>
</tr>
<tr>
<td>feature</td>
<td>$F_{45} - F_{32} - F_{28} - F_{7} - F_{4} - F_{2}$</td>
<td>$F_{45} - F_{32} - F_{28} - F_{7} - F_{4} - F_{2}$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Fig. 8** Best 3D representation of HIF and Non-HIF feature in

(a) IEEE 13-bus,

(b) IEEE 34-bus

---

This is an open access article published by the IET under the Creative Commons Attribution-NoDerivs License (http://creativecommons.org/licenses/by-nd/3.0/)
GA as input features for classification methods. The results of dimension reduction proved that correlation coefficients have high information gain and priority. Also, the results of simulation in IEEE standard test systems have illustrated that the proposed method is more accurate in comparison with other implemented methods.

**Fig. 9** Performance of the algorithm for HIF fault

(a) Three phase current waveform,  
(b) Magnitude of detail coefficients at level 3, and  
(c) Detection signal
5 References


---

Fig. 10 Performance of algorithm for load switching event
(a) Three phase current waveform,
(b) Magnitude of detail coefficients at level 3, and
(c) Detection signal