High impedance fault detection in power distribution networks using time–frequency transform and probabilistic neural network

S.R. Samantaray, B.K. Panigrahi and P.K. Dash

Abstract: An intelligent approach for high impedance fault (HIF) detection in power distribution feeders using advanced signal-processing techniques such as time–time and time–frequency transforms combined with neural network is presented. As the detection of HIFs is generally difficult by the conventional over-current relays, both time and frequency information are required to be extracted to detect and classify HIF from no fault (NF). In the proposed approach, S- and TT-transforms are used to extract time–frequency and time–time distributions of the HIF and NF signals, respectively. The features extracted using S- and TT-transforms are used to train and test the probabilistic neural network (PNN) for an accurate classification of HIF from NF. A qualitative comparison is made between the HIF classification results obtained from feed forward neural network and PNN with same features as inputs. As the combined signal-processing techniques and PNN take one cycle for HIF identification from the fault inception, the proposed approach was found to be the most suitable for HIF classification in power distribution networks with wide variations in operating conditions.

1 Introduction

Faults on power distribution feeders are difficult to detect [1, 2] using the conventional over-current, ground fault relays and some versions of distance relaying schemes. Diversity, uncertainties, selectivity, suitability and operational constraints introduce malfunction, limitations and detection errors in case of high impedance faults (HIFs). This is notable when remote source loading, fault resistance nonlinearity, capacitive line currents, mutual coupling and back-feed effects are taken into consideration. HIF faults [3, 4] are usually characterised by the ripple-rich current harmonic content because of nonlinearity and are thus abnormal events that frequently occur in distribution feeders. There are two types of HIFs: the active faults and the passive ones. Active faults are followed by an electric arc and present currents below the threshold of the protection relays. Normally, these currents decay with time until the complete extinction of the arc [5]. The majority of the techniques used to detect active HIFs make use of signals generated by the electric arc (harmonic and non-harmonic components) [6–9]. However, the arc may vanish even before the detection system gathers enough information to confirm the fault. Passive faults do not present an electric arc. They are more hazardous to people since there is no indication of the energisation condition of the conductor. Because of the presence of low or no current in HIF, the conventional over-current protection system normally fails to detect the same. Thus, it is a challenging issue to detect the HIF and isolate the feeder.

Sedighi *et al.* [10] presented a combined wavelet transform and soft computing application to HIF classification. This work includes feature extraction using wavelet transform and then classification using soft computing methods. HIF detection using neo-fuzzy systems [11] uses an artificial neuron set, composed of 'neo-fuzzy' neurons, and is trained to recognise the standard responses. In another work, earth faults with high impedance earthing in electrical distribution networks are characterised [12]. In the occurrence of disturbances, the traces of phase currents, voltages, neutral currents and voltages were recorded at two feeders at two substations. The study dealt with the clearing of earth faults, relation between short circuits and earth faults, arc extinction, arcing fault characteristics, appearance of transients and magnitudes of fault resistances. The above works find limitations as wavelet transform is highly prone to noise and provides erroneous results even with an SNR of 30 dB [13]. Also the fuzzy neural networks are sensitive to system frequency changes and require large training sets and long training time.

In this paper, a hybrid pattern recognition technique using either S- or TT-transform and probabilistic neural network (PNN) is used to detect and classify HIF. The HIFs are created under linear and nonlinear loading conditions of the distribution network. HIFs are easily distinguished from no fault (NF) under linear loading condition as higher harmonic components are more pronounced in HIF compared with NF under linear loading condition. But HIF and NF current signals under nonlinear loading are mostly similar in nature, and it is difficult to distinguish HIF from NF under nonlinear loading condition as both contain similar harmonic components.

In the first method, the current signals for fault and NF conditions are processed through S-transform [14–19] to

doi:10.1049/iet-gtd:20070319

Paper first received 19th July and in revised form 3rd November 2007

S.R. Samantaray is with the Department of ECE, National Institute of Technology, Rourkela-769008, India

B.K. Panigrahi is with the Department of EE, Indian Institute of Technology, Delhi, Hauz Khas, New Delhi-110016, India

P.K. Dash is with the Centre for Research in Electrical, Electronics and Computer Engineering, Bhubaneswar-751023, India

E-mail: sbh_samant@yahoo.co.in

find out both time and frequency information with proper time and frequency resolution, respectively. The frequency information is extracted in better frequency resolution and poor time resolution to obtain the accurate frequency information, and the time information is extracted in better time resolution and poor frequency resolution to obtain the accurate time information. The features such as energy and standard deviation of the frequency and time information are computed for half-cycle HIF current signal from the fault inception and used to train and test the neural networks [(PNN and feedforward neural network (FNN)] for the classification of HIF from NF.

Similarly, in the second method, the TT-transform [20–22] is used to extract the time–time distribution of the respective signals. The current signals of the HIF and NF are processed through TT-transform to generate TT-contour and time index at a particular time. The features such as energy and standard deviation of the TT-contour and time index are found out and used to train and test the PNN and FNN to provide accurate classification. The PNN and FNN are tested with features for different loading conditions of the distribution feeder as well as under different HIF conditions on surfaces such as dry asphalt, wet asphalt, dry cement, wet cement, dry soil, wet soil and so on.

2 System studied

The systems studied in the proposed research are: (a) three-phase radial distribution feeder and (b) three-phase meshed network. The schematic diagrams are given in Figs 1*a* and 1*b*, respectively. The generator is of 15 kV and 10 MV A capacity and connected to the transformer with 15/25 kV and 10 MV A capacity. The distribution network operates at 25 kV voltage. The networks are simulated under linear and nonlinear loads with different loading conditions. The 6-pulse rectifier is used to represent the nonlinear load. The HIF model is developed using antiparallel diodes with nonlinear resistance and DC source connected together for each phase as shown in Fig. 1*c*. The simulation models are developed using Power System Blokset (SIMULINK) and the sampling rate chosen is 1.0 kHz. The typical HIF fault current under linear and nonlinear loads is shown in Figs. 2*a* and 2*b*, respectively. As seen from the figures, the HIFs are created after two cycles of normal condition. Thus, in case of HIF under linear loading condition (Fig. 2*a*), the post-fault current signal (HIF) contains higher harmonic components compared with the pre-fault current signal (NF under linear loading). Thus, extracted harmonic components can easily distinguish HIF from NF under linear loading condition. But in case of HIF under nonlinear loading condition (Fig. 2*b*), the post-fault current (HIF) and the pre-fault current (NF under nonlinear loading) contain higher harmonic components. Thus, it is very difficult to distinguish HIF from NF under nonlinear loading condition, which is a vital issue in case of power distribution network.

3 Generalised S- and TT-transforms

3.1 Generalised S-transform

The S-transform [14] is an extension to the idea of the Gabor transform and wavelet transform and is based on a moving and scalable localising Gaussian window. The S-transform falls within the broad range of multiresolution spectral analysis, where the standard deviation is an inverse function of the frequency, thus reducing the dimension of the transform. The localising Gaussian function g(t) is defined as

$$g(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp^{-[t^2 f^2/2\sigma^2]} \tag{1}$$

where σ is the standard deviation. The multiresolution ST is defined by

$$S(f, \tau, \sigma) = \int_{-\infty}^{\infty} h(t) g(\tau - t, \sigma) \mathrm{e}^{-i2\pi ft}\, \mathrm{d}t \tag{2}$$

This falls within the definition of the multiresolution Fourier transform. The Gabor transform $\Gamma(f, \tau)$ is a particular case of S (f, τ, σ) with σ held constant. The primary purpose of the dilation (or scaling) parameter is to increase the 'width' of the window function $g(t, \sigma)$ for lower frequency and vice versa and is controlled by selecting a specific functional dependency of σ with the frequency f. We have chosen the width of the window to be proportional to the period of the cosinusoid being localised

$$\sigma(f) = T = \frac{1}{|f|} \tag{3}$$

where T is the time period . The choice of unity for the constant in (5) makes the Gaussian window in (1) the narrowest in the time domain. The ST may be written as

$$\boldsymbol{S}(f, \tau) = \int_{-\infty}^{\infty} h(t) \times \frac{|f|}{\sqrt{2\pi}} \mathrm{e}^{-((\tau-t)^2 f^2/2)} \times \mathrm{e}^{-i2\pi ft}\, \mathrm{d}t \tag{4}$$

One can see here that the zero frequency of the S-transform is identically equal to zero for this definition of $\sigma(f)$. This adds no information. Therefore $\boldsymbol{S}(f,\tau)$ is defined as independent of time and is equal to the average of the function $h(t)$, that is

$$\boldsymbol{S}(o, \tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} h(t)\, \mathrm{d}t, \quad \text{for} \quad f = 0 \tag{5}$$

For the discrete S-transform, $h(t)$ can be written in discrete form as $h[pT]$, where p varies from 0 to $N-1$ and is known as discrete time series of the signal $h(t)$. Discrete Fourier transform of the time series $h[pT]$ can be expressed as

$$H\left[\frac{n}{NT}\right] = \frac{1}{N} \sum_{p=0}^{N-1} h[pT] \mathrm{e}^{-(i2\pi nk/N)} \tag{6}$$

where n = 0, 1, . . . , $N-1$ and the inverse discrete Fourier transform is

$$h[pT] = \sum_{n=0}^{N-1} H\left[\frac{n}{NT}\right] \mathrm{e}^{i2\pi nk/N} \tag{7}$$

The S-transform in discrete case is the projection of the vector defined by the time series $h[pT]$ onto a spanning set of vectors. Since spanning vectors are not orthogonal and the elements of $\boldsymbol{S}$-matrix are not dependent, each basis vector is divided into N localised vectors by an element-by-element product with N shifted Gaussians, such that sum of these N localised vectors is the original basis vector. The S-transform of the discrete time series $H[pT]$ is given by

$$\boldsymbol{S}\left[\frac{n}{NT}, jT\right] = \sum_{m=0}^{N-1} H\left[\frac{m+n}{NT}\right] \mathrm{e}^{-(2\pi^2 m^2/n^2)} \mathrm{e}^{i2\pi mj/N} \tag{8}$$

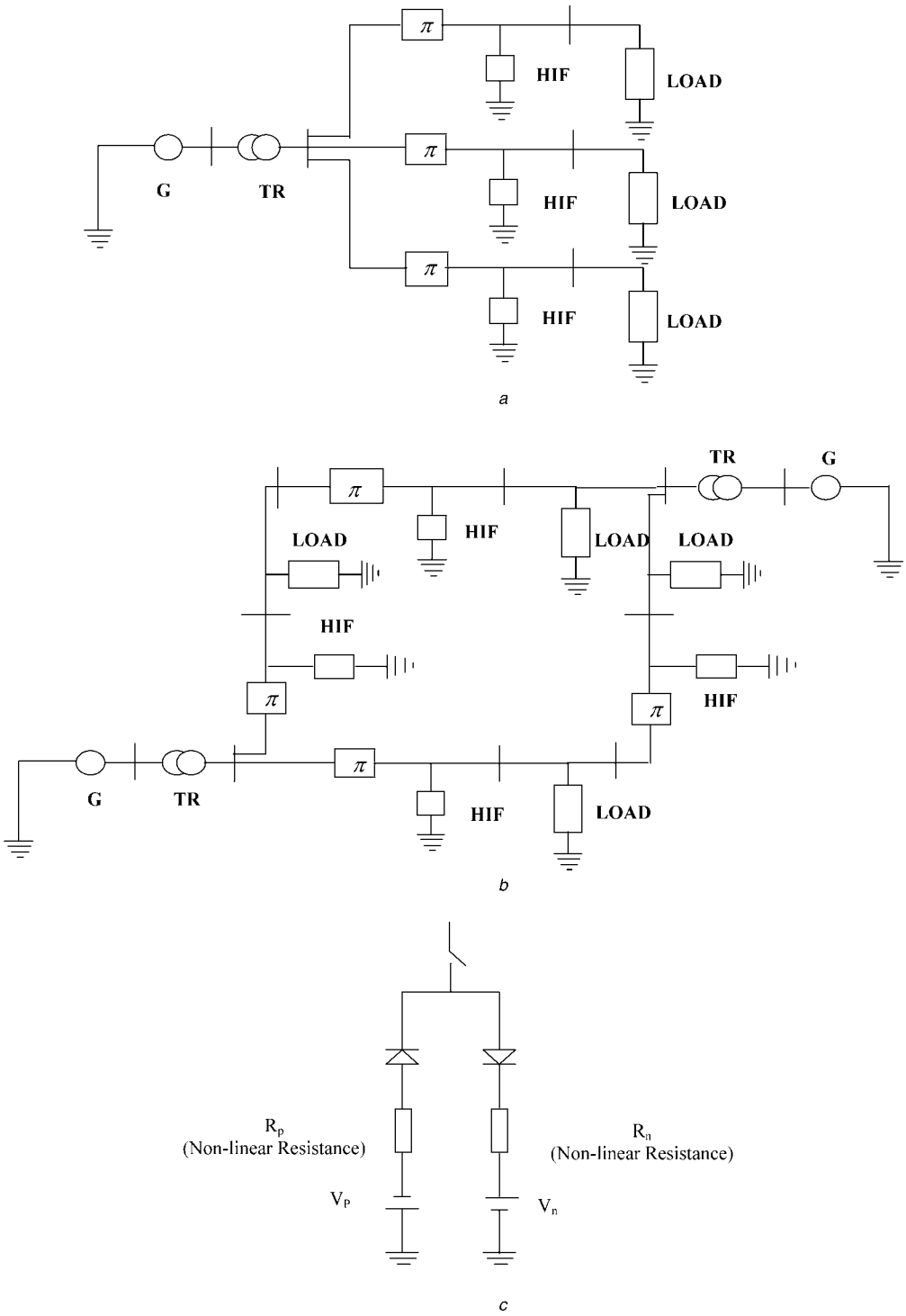

Fig. 1 *Schematic diagrams*

a Three-phase radial distribution feeder (single-line diagram)
b Three-phase meshed network (single-line diagram)
c HIF model

for $n = 0$

$$\boldsymbol{S}[0, jT] = \frac{1}{N}\sum_{m=0}^{N-1} h\left[\frac{m}{NT}\right] \tag{9}$$

where j, m and $n = 0, 1, \ldots, N-1$

3.2 TT-transform

TT-transform [22] is a two-dimensional time–time representation of a one-dimensional time series based on S-transform. TT-transform provides the time-local view of the time series through the scaled windows. It differs from the windowed time series of S-transform that the degree of localisation of the signal components is frequency-dependent rather than frequency invariant. Compared with the S-transform, the TT-transform provides better time-local properties of the time series and thus helps in localising the frequency components of the time series. The mathematical formulation of the TT-transform is given below.

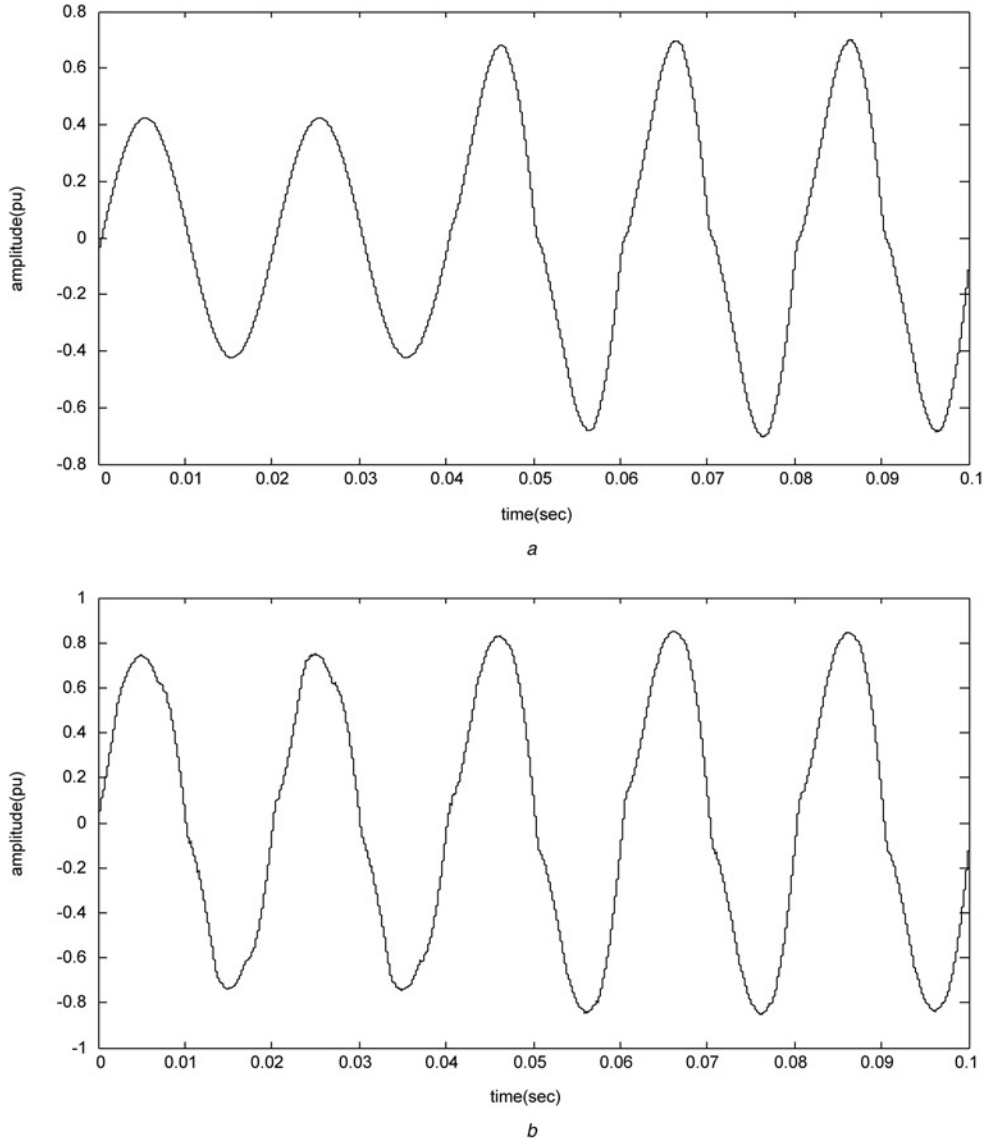

Fig. 2 *Typical HIF fault currents*

a HIF current under linear load for a-phase
b HIF current under nonlinear load for a-phase

The general expression for short-time Fourier transform (STFT) is given by

$$\mathrm{STFT}(t,f) = \int_{-\infty}^{\infty} h(\tau)w(t-\tau)\exp(-2\pi i f t)\,\mathrm{d}\tau \qquad (10)$$

Also it can be expressed through convolution sum as

$$+\mathrm{STFT}(t,f) = \int_{-\infty}^{\infty} H(\alpha+f)W(\alpha) \exp(+2\pi i\alpha t)\,\mathrm{d}\alpha \qquad (11)$$

where W and H are Fourier transform of w and h, respectively.

The inverse Fourier transform of the above equation leads to

$$[h(\tau)w(t-\tau] = \int_{-\infty}^{\infty} \mathrm{STFT}(t,f)\exp(+2\pi i f\tau)\,\mathrm{d}f \qquad (12)$$

for all values of 't' the windowed function becomes a two-dimensional function and is expressed as

$$\mathrm{STFT}_{\mathrm{TT}} = \int_{-\infty}^{\infty} \mathrm{STFT}(t,f)\exp(+2\pi i f\tau)\,\mathrm{d}f \qquad (13)$$

The $\mathrm{STFT}_{\mathrm{TT}}$ is the time–time distribution. Similarly, another time–time distribution can be obtained from the inverse Fourier transform of S-transform and given as follows

$$\mathrm{TT}(t,\tau) = \int_{-\infty}^{\infty} \boldsymbol{S}(t,f)\exp(+2\pi i f\tau)\,\mathrm{d}f \qquad (14)$$

where the S-transform (t, f) is defined as

$$\boldsymbol{S}(t,f) = \int_{-\infty}^{\infty} H(\alpha+f)\exp\left(\frac{2\pi^2\alpha^2}{f^2}\right) \exp(+2\pi i\alpha t)\,\mathrm{d}\alpha \qquad (15)$$

The expression is generalised using an windowing factor β and

$$\boldsymbol{S}(t,f) = \int_{-\infty}^{\infty} H(\alpha+f)\exp\left(\frac{2\pi^2\alpha^2\beta^2}{f^2}\right)\exp(+2\pi i\alpha t)\,\mathrm{d}\alpha \qquad (16)$$

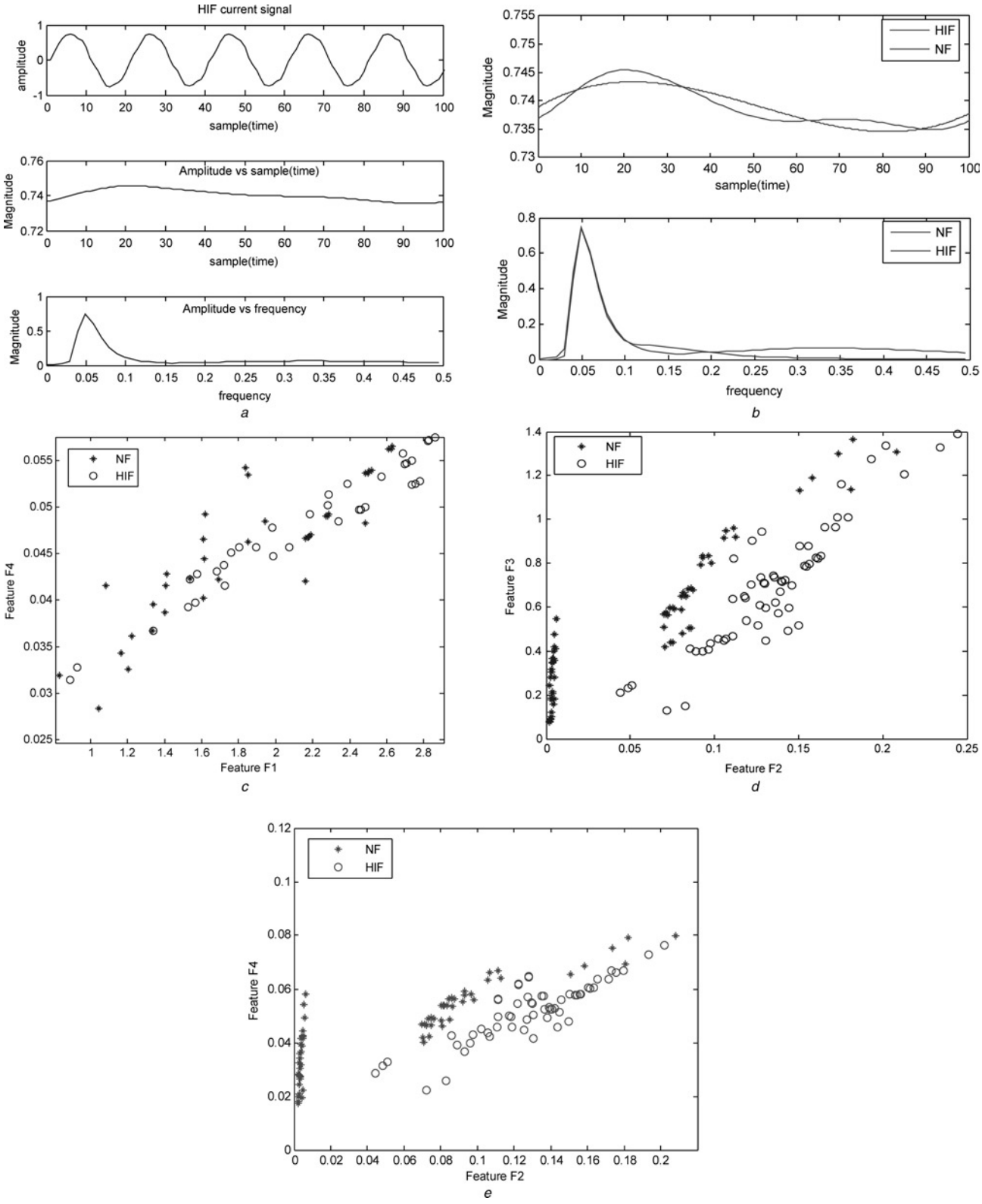

Fig. 3 *Feature plots*

a HIF signal, magnitude against sample (time) and magnitude against frequency (normalised) for HIF current signal of phase-a
b Comparisons in magnitude- sample (time) and magnitude–frequency (normalised) for HIF and NF signals, respectively
c Feature F1 against feature F4
d Feature F2 against feature F3
e Feature F2 against feature F4

If TT (t, τ) is considered at all τ but a specific t, the result is a time-local function, conceptually similar to a windowed function such as that in (12). Since a different window has been used to obtain S-transform at each value of 'f', this time-local function is different from the windowed functions of the STFT. The scaling properties of S-transform lead to higher amplitudes of high frequencies (as compared with low frequencies) around $\tau = t$.

The discrete S-transform is found out by sampling (15) in frequency and is given by

$$\boldsymbol{S}\left[jT, \frac{n}{NT}\right] \sum_{m=-N/2}^{N/2-1} H\left[\frac{m+n}{NT}\right] \exp\left[\frac{2\pi^2 m^2}{n^2}\right] \times \exp\left[\frac{2\pi imj}{NT}\right] \quad (17)$$

Similarly, the discrete TT-transform is obtained from the discrete form (14) of as

$$\mathrm{TT}(jT, kT) = \sum_{m=N/2}^{N/2-1} \boldsymbol{S}\left[jT, \frac{n}{NT}\right] \exp\left[\frac{2\pi ink}{N}\right] \quad (18)$$

4 Feature extraction using S- and TT-transforms

4.1 Feature extraction using S-transform

The time–frequency transform known as S-transform is used to extract features from HIF and NF current signals for different operating conditions of the power distribution networks. The features are extracted for half cycle fault current signal after fault inception. The HIF and NF current signals are generated from the designed power distribution models and processed through S-transform. The frequency and time information of the corresponding current signals is extracted from the generated S-matrix with proper frequency and time resolution.

The frequency information is extracted from the S-matrix as

$$a = \max\left(abs(S^{\mathrm{T}})\right) \quad (19)$$

where a represents the maximum of the absolute value of the transposed S-matrix generated from the S-transform which provides the amplitude–frequency information.

Similarly, the time information is extracted from the S-transform as

$$b = \max\left(abs(\boldsymbol{S})\right) \quad (20)$$

where b represents the maximum of the absolute value of the S-matrix generated from the S-transform which provides the amplitude–sample (time) information.

The energy and standard deviation of frequency and time information are given as follows

$$\text{Energy (frequency)} = \text{sum}\ (a^2) \quad (21)$$

$$\text{s tan dard deviation (frequency)} = \text{std}\ (a) \quad (22)$$

$$\text{Energy (time)} = \text{sum}\ (b^2) \quad (23)$$

$$\text{s tan dard deviation (time)} = \text{std}\ (b) \quad (24)$$

The frequency information is extracted with better frequency resolution and poor time resolution. This provides the information regarding the different frequency components present in the HIF and NF current signals. Similarly, the time information is extracted for both signals with better time resolution and poor frequency resolution. Fig. 3*a* shows the magnitude against sample (time) and magnitude against frequency (normalised) for HIF current signal. Here the magnitude–time characteristic is considered as time information and magnitude against frequency characteristic is considered as frequency information of the signal. Several numerical indices such as energy and standard deviation of both frequency and time information are computed to obtain the corresponding features. The energy and standard deviation of the time information are considered as features F1 and F2, respectively, and the energy and standard deviation of frequency information are regarded as features F3 and F4, respectively. Fig. 3*b* shows the difference in magnitude–sample (time) and magnitude–frequency (normalised) characteristics between HIF and NF signals. While comparing the amplitude–time and amplitude–frequency characteristics for HIF and NF signals, variations are observed between the two.

As seen from the feature plots shown in Figs. 3*c*–3*e*, some of the features are distinct, whereas some are

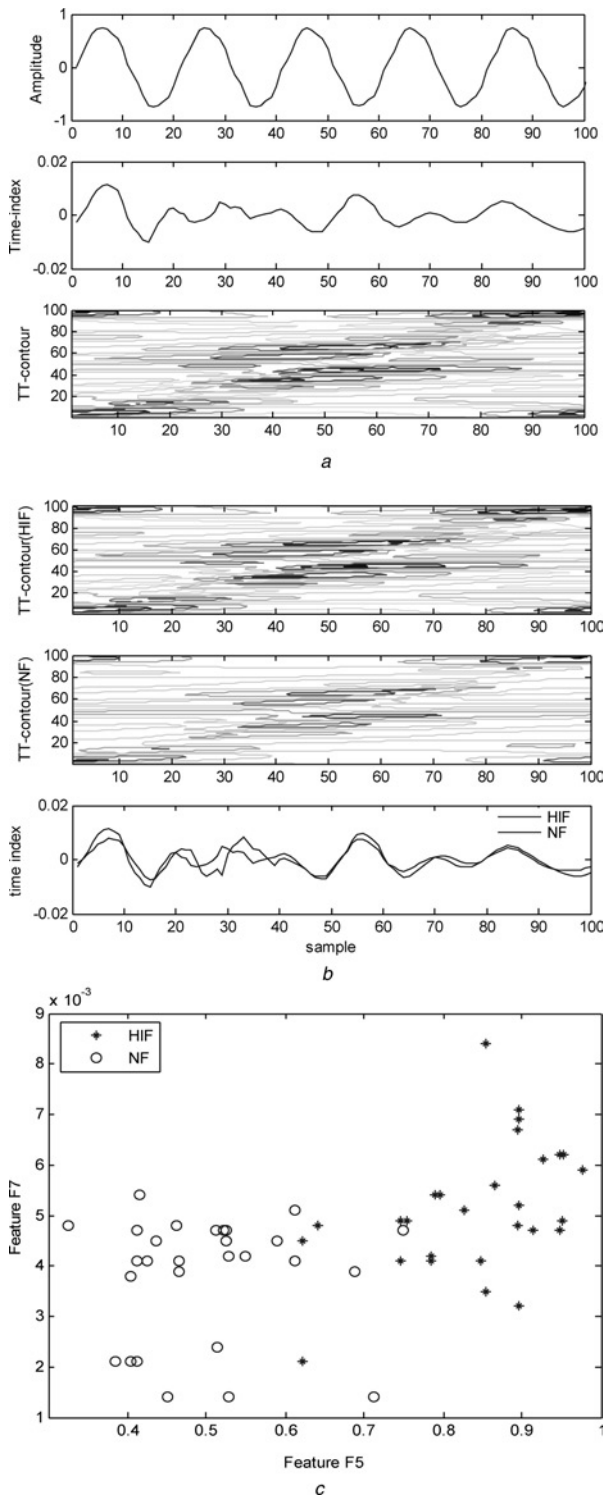

Fig. 4 *TT-contour and time-index for HIF and NF signals*

a Time index and TT-contour for HIF signal
b Comparison in TT-contour and time index for HIF and NF signals, respectively
c Feature F5 against feature F7

overlapping in nature. All the above plots are presented to provide information regarding the ability of the extracted features for classification in raw feature form so that the features can be used as inputs to the designed PNN and FNN for final classification between HIF and NF. Hence to classify the HIF and NF with nonlinear load from the above-mentioned features, intelligent classification techniques based on neural classifier are adopted.

4.2 Feature extraction from TT-transform

The features extracted from TT-transform are the energy and standard deviation of the TT-contour and time index, respectively, for half-cycle of the fault current after fault inception. TT-contours provide the frequency localisation of the time-series in time–time distribution as TT-transform provides better time local properties of the time series and thus helps in localising the frequency components of the time series. Thus, frequency information can be extracted from the TT-contour. Time index is the time localisation of a time series at a particular time instant. Fig. 4*a* shows the TT-contour and time index at a particular time for HIF signal. TT-contour and time index for HIF and NF signals are compared in Fig. 4*b*. While comparing the TT-contour and time index for HIF and NF signals, variations are observed in both cases. As seen in Fig. 4*b*, the TT-contours for both cases appeared to be the same, but if judged properly it can be observed that the TT-contours are highly pronounced in HIF signal (first window, Fig. 4*b*) when compared with that of the NF signal (second window, Fig. 4*b*), which reflects the frequency localisation of the time series in time–time distribution. Several numerical indices such as energy and standard deviation of both TT-contour and time index are computed to obtain the corresponding features. Thus F5 and F6 are the features for energy of the TT-contour and time index, respectively. Similarly, F7 and F8 are the features for standard deviation for the TT-contour and time index, respectively. Fig. 4*c* shows the relationship between feature F5 and F7 for HIF and NF. All the above plots are presented to provide information regarding the ability of the extracted features for classification in raw feature form so that the features can be used as inputs to the designed PNN and FNN for final classification between HIF and NF. It is observed that the corresponding features are separable, but overlapping to some extent. Thus these features are used to train and test the FNN and PNN to classify the HIF from NF.

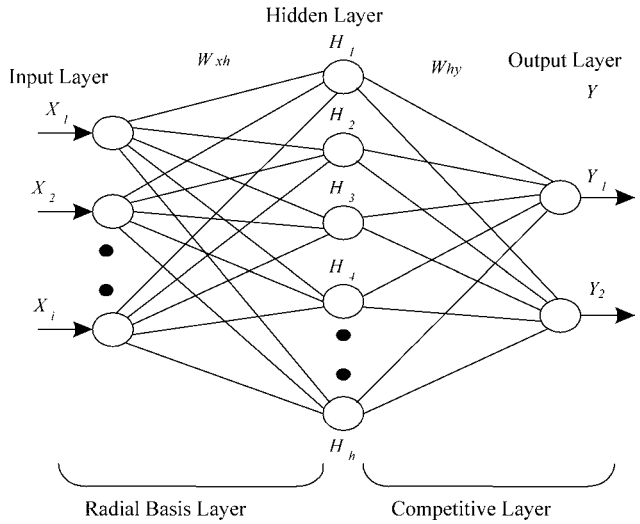

Fig. 5 *Architecture of the proposed PNN*

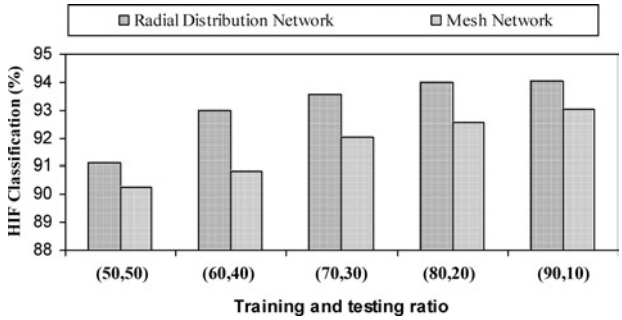

Fig. 6 *Classification rate of FNN with F1–F4*

The following section deals with the designed PNN and FNN for classification purposes.

5 Classification of HIF and NF using FNN and PNN

5.1 FNN and computational results

FNNs are the most popular and most widely used models in many practical applications. They are known by many different names, such as 'multilayer perceptrons'. The proposed FNN has one hidden layer with inputs $x_1, x_2, \ldots, x_n$ (inputs), and outputs y_1 and y_2 (two outputs). The input layer consists of features as the inputs to the network and the hidden layer has five hidden neurons. Each neuron performs a weighted summation of the inputs, which then passes through a nonlinear activation function. FNN provides an output corresponding to HIF and NF classes, respectively.

There are 500 that cases are simulated and used to train and test the designed FNN. The FNN is trained and tested with different combinations of inputs to test the impact on classification rate. The classification rates are calculated on the testing data sets. Figs. 6–8 show the HIF classification rate of FNN using features from S- and TT-transforms for both radial distribution and mesh networks. The *y*-axis of the figures shows the percentage HIF classification, and *x*-axis shows the training and testing ratio of data sets. FNN with F1–F4 (S-transform features) provides up to 94.04% and 93.02% classification rate (Fig. 6) for radial distribution network and mesh network, respectively. Similarly, FNN with features F2 and F3 only (S-transform features) provides up to 93.04% and 92.86% classification rate (Fig. 7) for radial distribution and mesh networks, respectively. FNN is trained and tested with features from TT-transform provides maximum 94.16% and 93.55% classification rate (Fig. 8) for radial distribution and mesh networks, respectively. It is found that FNN

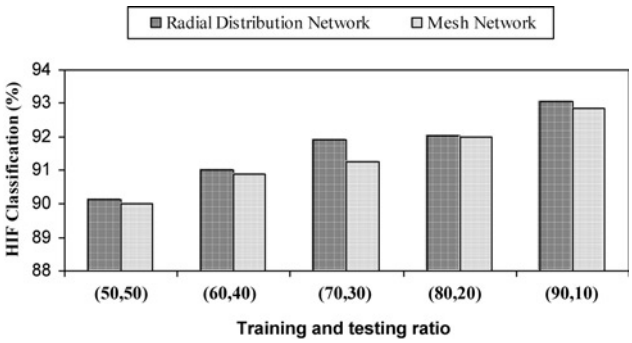

Fig. 7 *Classification rate of FNN with F2 and F3*

Table 1: Training and testing time of FNN for different data sets

Training, %	Testing, %	FNN with F1–F4		FNN with F2 and F3		FNN with F5–F8	
		Training time, s	Testing time, s	Training time, s	Testing time, s	Training time, s	Testing time, s
50	50	9.30	0.3	7.14	0.28	7.14	0.28
60	40	8.47	0.25	5.96	0.19	5.96	0.19
70	30	8.91	0.25	5.27	0.16	5.27	0.16
80	20	9.36	0.25	5.44	0.16	5.44	0.16
90	10	9.63	0.25	5.56	0.19	5.56	0.19

provides up to 94% classification rate for HIF identification with variations in operating conditions of the power distribution network. Table 1 depicts the training and testing time of FNN for different training and testing data sets.

5.2 PNN and computational results

The PNN model is one among the supervised learning networks and has the following features different from those of other networks in the learning processes.

- It is implemented using the probabilistic model, such as Bayesian classifiers.
- A PNN is guaranteed to converge to a Bayesian classifier provided that it is given enough training data.
- No learning processes are required.
- No need to set the initial weights of the network.
- No relationship between learning and recalling processes.
- The difference between the inference and the target vectors is not used to modify the weights of the network. The learning speed of the PNN model is very fast, making it suitable for fault diagnosis and signal classification problems in real time. Fig. 5 shows the architecture of a PNN model that is composed of the radial basis layer and the competitive layer.

In the signal classification application, the training examples are classified according to their distribution values of probabilistic density function (PDF), which is the basic principle of the PNN. A simple PDF is as follows

$$f_k(X) = \frac{1}{N_k}\sum_{j=1}^{N_k}\exp\left(-\frac{\|\boldsymbol{X}-\boldsymbol{X}_{kj}\|}{2\sigma^2}\right) \tag{25}$$

Modifying the output vector $\boldsymbol{H}$ of the hidden layer in the PNN is

$$\boldsymbol{H}_{\mathrm{h}} = \exp\left(\frac{-\sum_i (X_i - W_{i\mathrm{h}}^{x\mathrm{h}})^2}{2\sigma^2}\right) \tag{26}$$

$$\mathrm{net}_j = \frac{1}{N_j}\sum_{\mathrm{h}} W_{\mathrm{h}j}^{\mathrm{h}y}\,\boldsymbol{H}_{\mathrm{h}} \text{ and } N_j = \sum_{\mathrm{h}} W_{\mathrm{h}j}^{\mathrm{h}y},\ \mathrm{net}_j = \max_k(\mathrm{net}_k) \text{ then } y_j = 1,\ \text{else } y_j = 0 \tag{27}$$

where i is the number of inputs, h the number of hidden units, j the number of output (two in this case), k the number of training examples, N the number of classifications (clusters), σ the smoothing parameter (standard deviation), $\boldsymbol{X}$ the input vector, $\|\boldsymbol{X}-\boldsymbol{X}_{kj}\|$ the Euclidean distance between the vectors $\boldsymbol{X}$ and $\boldsymbol{X}_{kj}$, that is $\|\boldsymbol{X}-\boldsymbol{X}_{kj}\| = \sum_i (\boldsymbol{X}-\boldsymbol{X}_{kj})^2$, $W_{i\mathrm{h}}^{x\mathrm{h}}$ the connection weight between the input layer X and the hidden layer H and $W_{\mathrm{h}j}^{\mathrm{h}y}$ the connection weight between the hidden layer H and the output layer Y.

Figs. 9–11 provide the computational results from PNN. There are 500 cases that are simulated and used to train and test the designed PNN. The PNN is trained and tested with different combinations of inputs to test the impact on classification rate. The classification rates are calculated on the testing data sets. The maximum classification rate for radial distribution and mesh networks are 98.02% and 97.89%, respectively (Fig. 9) with F1, F2, F3 and F4 (S-transform features). PNN with features F2 and F3 only (S-transform features) provides classification rate up to 98.06% and 97.85% (Fig. 10) for radial distribution and mesh networks, respectively. Similarly, PNN with features from TT-transform provides classification rate up to 98.05% and 97.09% (Fig. 11) for radial distribution and mesh networks, respectively. Table 2 depicts the training

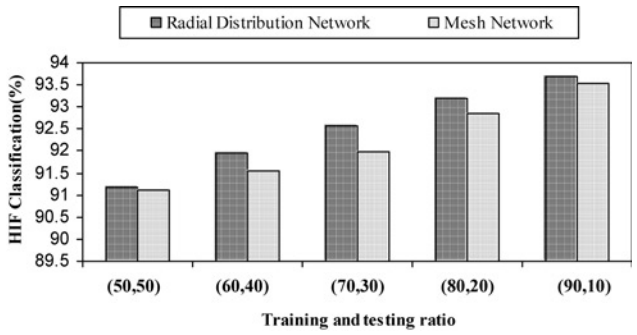

Fig. 8 *Classification rate of FNN with F5–F8*

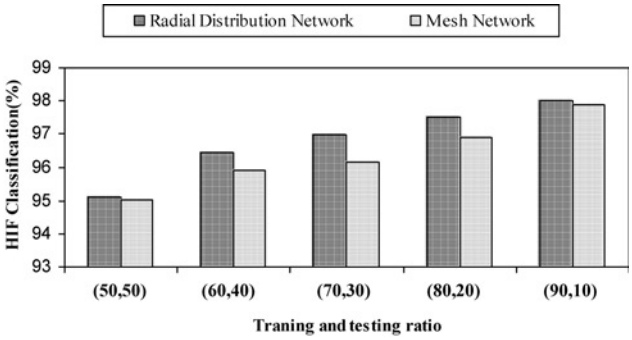

Fig. 9 *Classification rate of PNN with F1–F4*

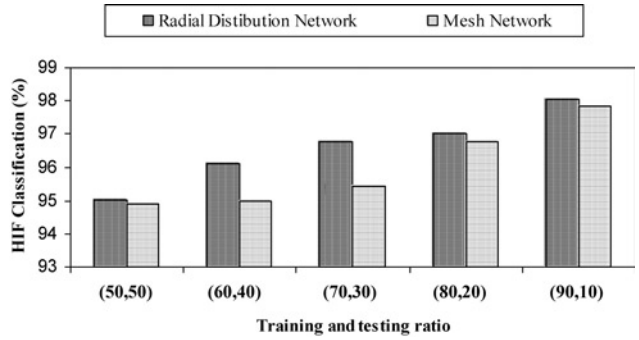

Fig. 10 *Classification rate of PNN with F2 and F3*

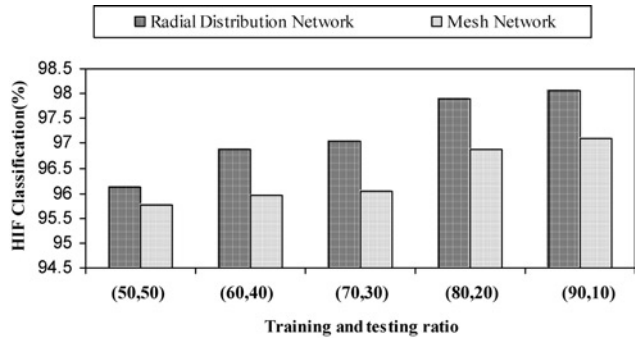

Fig. 11 *Classification rate of PNN with F5–F8*

and testing time of PNN for different training and testing data sets.

Fig. 12 shows the performance comparison between RBFNN, support vector machine (SVM) and PNN for same features as inputs. SVM with Gaussian kernel is used for designing the classifier. SVM provides 94.82% classification accuracy compared with 94.58% provided by RBFNN with 90% training and 10% testing data sets. As seen in Fig. 12, for other training and testing data sets,

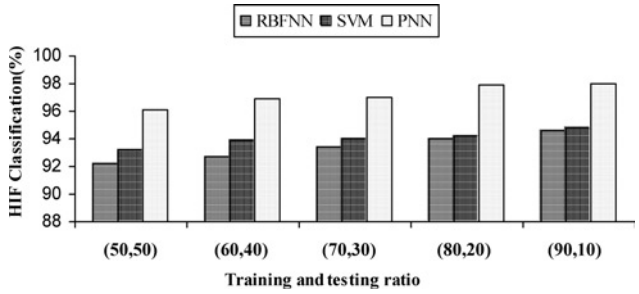

Fig. 12 *Performance comparison between RBFNN, SVM and PNN with different training and testing data sets for radial distribution network*

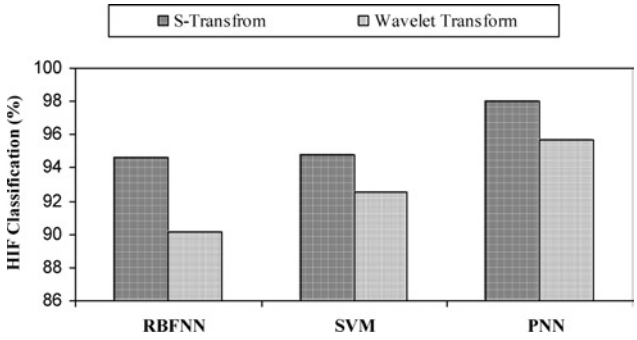

Fig. 13 *Performance comparison between RBFNN, SVM and PNN with features from S-transform and wavelet-transform with 90% training and 10% testing data sets for radial distribution network*

the classification accuracy resulting from SVM is almost the same compared with RBFNN. With same training and testing data sets, PNN provides an HIF classification accuracy of 98.05%, which is substantially high compared with SVM and RBFNN. Similar observations are made for different training and testing data sets. It is also observed that RBFNN suffers from slowness because of high training and testing time compared with PNN, which is not desired for fault diagnosis and signal classification problems in real time. HIF classification accuracy is also observed with features from wavelet transform. Db-4 is selected for feature extraction from the HIF and NF signals. Features such as energy and standard deviation of the detailed coefficients at level-1 (D-1) and level-2 (D-2) are extracted which keep information regarding higher harmonic components. The performance comparison for HIF classification rate between S- and wavelet transforms is shown in Fig. 13. It is found that the classification rate is substantially higher with features from S-transform compared with the wavelet transform. Also wavelet transform is highly prone to noise and provides erroneous results in noisy environment [13], whereas S-transform is highly immune to noise. Thus the proposed technique is found to be highly effective for classifying HIF from NF compared with the existing techniques.

6 Discussion

In the proposed technique, a qualitative comparison is made between PNN and FNN for HIF detection in power distribution network. From the results obtained, it is found that PNN provides better results compared with FNN with same input features to the network. The classification rate for radial distribution network is 98.02% from PNN compared with 94.04% from FNN with features F1–F4. Similarly, the classification rate is 97.89% from PNN

Table 2: Training and testing time of PNN for different data sets

Training, %	Testing, %	PNN with F–F4		PNN with F2 and F3		PNN with F5–F8	
		Training time, s	Testing time, s	Training time, s	Testing time, s	Training time, s	Testing time, s
50	50	0.72	0.01	0.45	0.01	0.45	0.01
60	40	0.72	0.01	0.45	0.01	0.45	0.01
70	30	0.75	0.01	0.44	0.01	0.44	0.01
80	20	0.69	0.01	0.47	0.01	0.47	0.01
90	10	0.70	0.01	0.45	0.01	0.45	0.01

compared with 93.02% from FNN with same features from S-transform for mesh networks. Similar observations are made with features from TT-transform. PNN provides a classification rate of 98.05% compared with 94.16% from FNN with features from TT-transform. It is found that the training and testing time for PNN is drastically reduced compared with the training and testing time of FNN. As seen from Table 1, the training and testing times are 9.30 and 0.3 s, respectively, from FNN compared with the training and testing time of 0.72 and 0.01 s, respectively, from PNN as depicted in Table 2, with 50% training and 50% testing data sets. Similar observations are made with other features from TT-transform. As the PNN testing takes 0.01 s (half-cycle on 50 Hz cycle) and feature are extracted for half cycle post-fault current signal, thus the combined S- or TT-transform with PNN will take one cycle time for HIF identification from the inception of fault.

7 Conclusions

Intelligent techniques for HIF detection and classification are presented in the proposed study. An attempt is made to classify the HIF from NF under nonlinear loading. In this study, the time–frequency and time–time distributions of the HIF and NF current signals are extracted using S- and TT-transforms, respectively, and different features like energy, standard deviation are computed and used to train and test the PNN for HIF classification. As futures are extracted for half cycle post-fault HIF signal and PNN testing takes half-cycle time (0.01 s), thus the combined approach takes one cycle for HIF classification from the fault inception. Also HIF classification rate is more than 98%, obtained from PNN. Thus, the proposed approach is fast and accurate for HIF identification and can be extended for protection of large power distribution network.

8 References

1 Johns, A.T.: 'Correspondence on microprocessor based algorithm for high-resistance earth-fault distance protection', *Proc. Inst. Electr. Eng. C*, 1985, **132**, pp. 94–95

2 Yang, Q., and Morrison, I.: 'Microprocessor-based algorithm for high-resistance earth-fault distance protection', *Proc. Inst. Electr. Eng. C*, 1983, **130**, pp. 306–310

3 Sharaf, A.M., E1-Sharkawy, R.M., Talaat, H.E.A., *et al.*: 'Novel alpha-transform distance relaying scheme'. Proc. IEEE-CECE Conf., Calgary, Canada, 1996

4 Sharaf, A.M., and El-Sharkawy, R.M.: 'High impedance ripple based detection scheme on radial systems'. Proc. GCC CIGRE Conf., Muscat, Sultanate of Oman, October 1996

5 Aucoin, M., and Russell, B.D.: 'Detection of distribution high impedance faults using burst noise signals near 60 Hz', *IEEE Trans. Power Deliv.*, 1987, **2**, (2), pp. 342–348

6 Sultan, A.F., Swift, G.W., and Fedirchuk, D.J.: 'Detection of high impedance arcing faults using a multilayer perceptron', *IEEE Trans. Power Deliv.*, 1992, **7**, (4), pp. 1871–1877

7 Sultan, A.F., Swift, G.W., and Fedirchuk, D.J.: 'Detection arcing downed wires using fault current flicker and half cycle asymmetry', *IEEE Trans. Power Deliv.*, 1994, **9**, (1), pp. 461–470

8 Russell, B.D., and Chinchali, R.P.: 'Signal processing algorithm for detecting arcing faults on power distribution feeders', *IEEE Trans. Power Deliv.*, 1989, **4**, (1), pp. 132–140

9 Russell, B.D., Chinchali, R.P., and Kim, C.J.: 'Behaviour of low frequency spectra during arcing fault and switching events', *IEEE Trans. Power Deliv.*, 1988, **3**, (4), pp. 1485–1492

10 Sedighi, A.-R., Haghifam, M.-R., and Malik, O.P.: 'Soft computing applications in high impedance fault detection in distribution systems', *Electr. Power Syst. Res.*, 2006

11 Jota, P.R.S., and Jota, F.G.: 'Fuzzy detection of high impedance faults in radial distribution feeders', *Electr. Power Syst. Res.*, 1999, **49**, pp. 169–174

12 Hänninen, S., and Lehtonen, M.: 'Characteristics of earth faults in electrical distribution networks with high impedance earthing', *Electr. Power Syst. Res.*, 1998, **44**, pp. 155–161

13 Yang, H.-T., and Liao, C.-C.: 'A de-noising scheme for enhancing wavelet-based power quality monitoring system', *IEEE. Trans. Power Deliv.*, 2001, **16**, (3), pp. 353–360

14 Stockwell, R.G., Mansinha, L., and Lowe, R.P.: 'Localization of the complex spectrum: the S-transform', *IEEE Trans. Signal Process.*, 1996, **44**, (4), pp. 998–1001

15 Pinnegar, C.R., and Mansinha, L.: 'The S-transform with windows of arbitrary and varying shape', *Geophysics*, 2003, **68**, (1), pp. 381–385

16 Livanos, G., Ranganathan, N., and Jiang, J.: 'Heart sound analysis using S-transform', *Comp. Cardiol.*, 2000, **27**, pp. 587–590

17 McFadden, P.D., Cook, J.G., and Forster, L.M.: 'Decomposition of gear vibration signals by the generalized S-transform', *Mech. Syst. Signal Process.*, 1999, **13**, pp. 691–707

18 Dash, P.K., Panigrahi, B.K., and Panda, G.: 'Power quality analysis using S-transform', *IEEE Trans. Power Deliv.*, 2003, **18**, (2), pp. 406–411

19 Pinnegar, C.R., and Mansinha, L.: 'Time-local spectral analysis for non-stationary time series: the S-transform for noisy signals', *Fluct. Noise Lett.*, 2003, **3**, (3), L357–L364

20 Rioul, O., and Flandrin, P.: 'Time-scale energy distributions: a general class extending wavelet transforms', *IEEE Trans. Signal Process.*, 1992, **40**, pp. 1746–1757

21 Papandreou-Suppappola, A., Hlawatsch, F., and Boudreaux-Bartels, G.F.: 'Quadratic time–frequency representations with scale covariance and generalized time-shift covariance: a unified framework for the affine, hyperbolic, and power classes', *Digit. Signal Process.*, 1998, **8**, pp. 3–48

22 Pinnegar, C.R., and Mansinha, L.: 'A method of time–time analysis: the TT-transform', *Digit. Signal Process.*, 2003, **13**, pp. 588–603