Discrimination between Inrush and Fault in Transformer: ANN Approach

SRParaskar, M.A.Beg, G.M.Dhole

Department of Electrical Engineering, S.S.G.M.College of Engineering
Shegaon.(M.S.),44203,India

Corresponding Author Email: ***srparaskar@gmail.com***

Abstract

Transformer protection is critical issue in power system as the issue lies in the accurate and rapid discrimination of magnetizing inrush current from internal fault current. Artificial neural network has been proposed and has demonstrated the capability of solving the transformer monitoring and fault detection problem using an inexpensive, reliable, and noninvasive procedure. This paper gives algorithm where statistical parameters of detailed d1 level wavelet coefficients of signal are used as an input to the artificial neural network (ANN), which develops in to a novel approach for online detection method to discriminate the magnetizing inrush current and inter-turn fault, and even the location of fault i.e. whether the interturn fault lies in primary winding or secondary winding through the use of discrete wavelet transform and artificial neural-nets (ANNs). A custom-built single-phase transformer was used in the laboratory to collect the data from controlled experiments. After the feature extraction using discrete wavelet transform (DWT), a neural network models MLP has been designed and trained rigorously. The proposed on line detection scheme is also discussed.

Keywords: *Neural networks, transformer, fault detection, discrete wavelet transform (DWT), inrush current*

1. Introduction

Power transformers are important elements of power system. So it is very important to avoid any maloperation of required protective system. For many years, differential protection has been used as the primary protection of power systems. It contains the differential relay, which operates for all internal fault types of power transformer and block due to inrush current. The major drawback of the differential protection relays stem from its potential for mal-operation caused by the transient inrush current, which flow when the transformer is energized. The inrush current contains a large second harmonic component. Most of the methods for digital differential protection of transformers are based on harmonic content of differential current. These methods are based on this fact that the ratio of the second harmonics to the fundamental component of differential current in inrush current condition is greater than the ratio in the fault condition.

However, the second harmonic may also be generated during faults on the transformers. It might be due to saturation of CTs, parallel capacitances or disconnected transformers. The second harmonic in these situations might be greater than the second harmonic in inrush currents.

Thus, the commonly employed conventional differential protection based on second harmonic restraint will face difficulty in distinguishing inrush current and internal faults. Thus, an improved technique of protection is required to discriminate between inrush current and internal faults [1].

To overcome this difficulty and prevent the mal-function of differential relay, many methods have been presented to analyze and recognize inrush current and internal fault currents. As both inrush current and internal faults are non-stationary signals, wavelet based signal processing technique is an effective tool for power system analyze and feature extraction [2-6]. However the wavelet-based methods have better ability of time-frequency analysis but they usually require long data windows and are also sensitive to noise. The method presented in [6] uses WT and ANFIS to discriminate internal faults from inrush current. Since the values of wavelet coefficients at detail 5 (D5) are used for pattern recognition process, the algorithm is very sensitive to noise.

In [5], a new algorithm was presented which discriminate between the inter-turn fault and magnetizing inrush current. The algorithm used wavelet coefficients as a discriminating function. Two peak values corresponding the |d5| level following the fault instant are used to discriminate the cases studied. As criterion compare the two peak values, hence no threshold settings are necessary in this algorithm, but it is observed that in noisy environment it is difficult to identify correct switching instant and there the strategy fails.

Moreover, feed forward neural network (FFNN) [7-10] has found wide application for detection of inrush current from internal faults but they have two major drawbacks: First, the learning process is usually time consuming. Second, there is no exact rule for setting the number of neurons to avoid over-fitting or underfitting.To avoid these problems, a Radial Basis Function Network (RBFN) has been developed [11]. RBFs are well suited for these problems due to their simple topological structure and their ability to reveal how learning proceeds in an explicit manner. In some methods differential current harmonics are used as inputs to fuzzy logic [6], [12].

The problem associated with these methods is the need to design neural networks and fuzzy laws, which require a huge number of training patterns produced by simulation of various cases. In [14] an energy index is defined by calculation of 9-level frequency contours using S-transform to distinguish inrush current from internal fault currents. But the disadvantage of this method is determining the threshold value which can be different in transformers with different capacity and may change in noisy environment. Support Vector Machine (SVM) [14], Hidden Markov Model (HMM) [15] and Gaussian Mixture Models (GMM) [16] are used as new classifiers for detection of internal fault and inrush currents. In [14] the extracted features are chosen from differential currents which due to large data window could not be effective than those methods use less features based on preprocessing step like WT, but the performance and detection capability of SVM is better than HMM and GMM.

In this paper, Artificial Neural Network (ANN), however, have been proposed and have demonstrated to be an effective alternative for performing transformer fault detection even location of fault, while avoiding the need for a mathematical model. In addition, the ANN can perform this function on-line through the use of inexpensive monitoring devices. These devices obtain the necessary measurements in a noninvasive manner. Different advantages of using ANN'S instead of other fault detection techniques are discussed in more detail in [7].

The main problems facing the use of ANN are the selection of the best inputs and how to choose the ANN parameters making the structure compact, and creating highly accurate networks. For the proposed system, the feature selection is also an important process since there are many features after feature extraction. Many input features require a significant computational effort to calculate, and may result in a low success rate.

2. Wavelet Transform

Wavelet analysis is about analyzing the signal with short duration finite energy functions. They transform the considered signal into another useful form. This transformation is called Wavelet Transform (WT).

Let us consider a signal f(t), which can be expressed as-

$$f(t) = \sum_{l} a_l \; \varphi_l(t) \tag{1}$$

Where, l is an integer index for the finite or infinite sum. Symbol a_l are the real valued expansion coefficients, while $\varphi_l(t)$ are the expansion set.

If the expansion (1) is unique, the set is called a basis for the class of functions that can be so expressed. The bases are orthogonal if-

$$\langle \psi_l(t)\psi_k(t) \rangle = \int \psi_l(t)\psi_k(t)dt = 0 \quad k \neq l \tag{2}$$

Then coefficients can be calculated by the inner product as-

$$\left\langle f(t), \varphi_k(t) \right\rangle = \int f(t) \; \varphi_k(t) \, dt \tag{3}$$

If the basis set is not orthogonal, then a dual basis set $\varphi_k(t)$ exists such that using (3) with the dual basis gives the desired coefficients.

For wavelet expansion, equation (1) becomes-

$$f(t) = \sum_{k} \sum_{j} a_{j,k} \; \varphi_{j,k}(t) \tag{4}$$

In (4), j and k are both integer indices and $\varphi_{jk}(t)$ are the wavelet expansion function that usually form an orthogonal basis. The set of expansion coefficients a_{jk} are called Discrete Wavelet Transform (DWT).

There are varieties of wavelet expansion functions (or also called as a Mother Wavelet) available for useful analysis of signals. Choice of particular wavelet depends upon the type of applications. If the wavelet matches the shape of signal well at specific scale and location, then large transform value is obtained, vice versa happens if they do not correlate. This ability to modify the frequency resolution can make it possible to detect signal features which may be useful in characterizing the source of transient or state of post disturbance system. In particular, capability of wavelets to spotlight on short time intervals for high frequency components improves the analysis of signals with localized impulses and oscillations particularly in the presence of fundamental and low order harmonics of transient signals. Hence, Wavelet is a powerful time frequency method to analyze a signal within different frequency ranges by means of dilating and translating of a single function called Mother wavelet.

Formulation of DWT is related to filter bank theory in many of the good references. It divides the frequency band of input signal into high and low frequency components by using high pass h(k) and low pass g(k) filters. This operation may be repeated recursively, feeding the down sampled low pass filter output into another identical filter pair, decomposing the signal into approximation c(k) and detail coefficients d(k) for various resolution scales. In this way, DWT may be computed through a filter bank framework, in each scale, h(k) and g(k) filter the input signal of this scale, giving new approximation and detailed coefficients respectively. The filter bank framework is shown in Fig 1. The down pointing arrow denotes decimation by two and boxes denote convolution by h(k) or g(k).

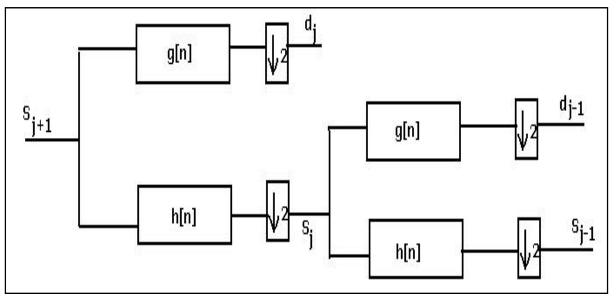

Fig. 1: Two band Multi-resolution analysis of Signal

The coefficients of filter pair are associated with the selected mother wavelet. The sampling frequency in this paper is taken to be 10 kHz and Table I shows the frequency levels of the wavelet function coefficients.

Table I: Frequency levels of Wavelet Functions Coefficients

Decomposition Level	Frequency Components (Hz)
d1	5000-2500
d2	2500-1250
d3	1250-625
d4	625-312.5
d5	312.5-156.25
a5	0-156.25

3. Experimentation & Data Collection

The setup for experiments has a custom built 220V/220V, 2KVA, 50Hz single-phase transformer with externally accessible taps on both primary and secondary to introduce faults.

The primary winding and secondary winding has 272 turns respectively. The load on the secondary comprises of static and rotating elements. The current and voltages on both primary and secondary were acquired using a real time Data acquisition system containing data acquisition card by National Instruments and appropriate signal conditioning devices. These signals were recorded at a sample rate of 10,000 samples/sec. Different cases of inter turn short circuit are staged, considering the effect of number of turns shorted on primary and secondary and load condition. Experimental set up is as shown in Fig 2.

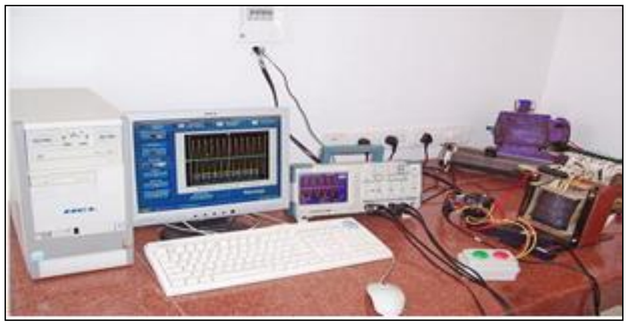

Fig.2: Experimental set up

Primary (Ip) and secondary (Is) currents were captured using the experimental setup. The Tektronix current probes of rating 100 mV/A, input range of 0 to 70 Amps AC RMS, 100A peak and frequency range DC to 100KHz are used. The captured current signals for inrush and faulted condition simulated on mains feed custom built transformer were decomposed up to second level using Haar wavelet. Various statistical parameters of differential current of level D1 for one cycle were calculated and are used as an input to the neural network.

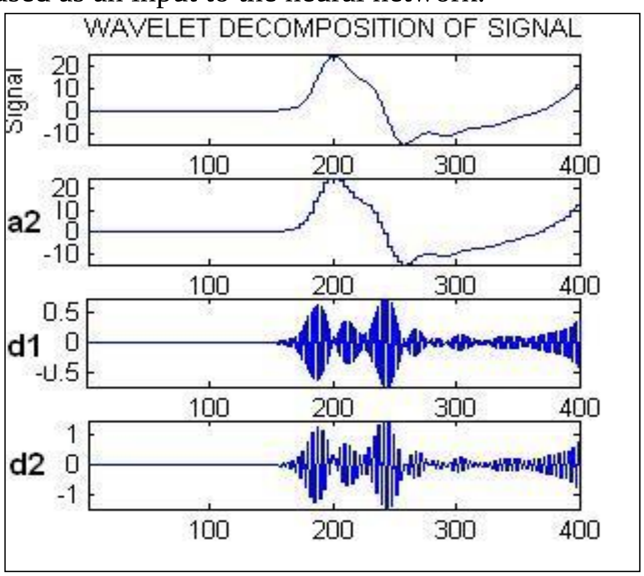

Fig. 3(a): Wavelet Decomposition of differential Inrush Current

Fig.3 (a) Fig.3 (b) and Fig.3 (c) shows decomposition of differential current signal for inrush, fault in primary and fault in secondary respectively. The decomposition is carried out using the Haar wavelet.

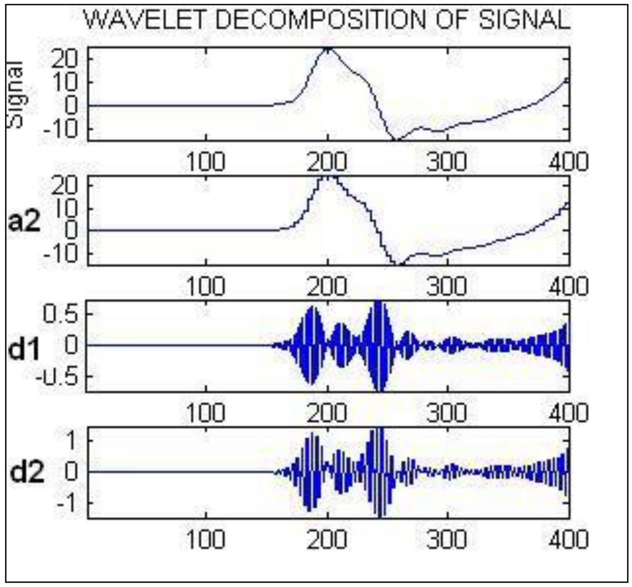

Fig. 3(b): Wavelet Decomposition of differential Current for primary inter-turn fault

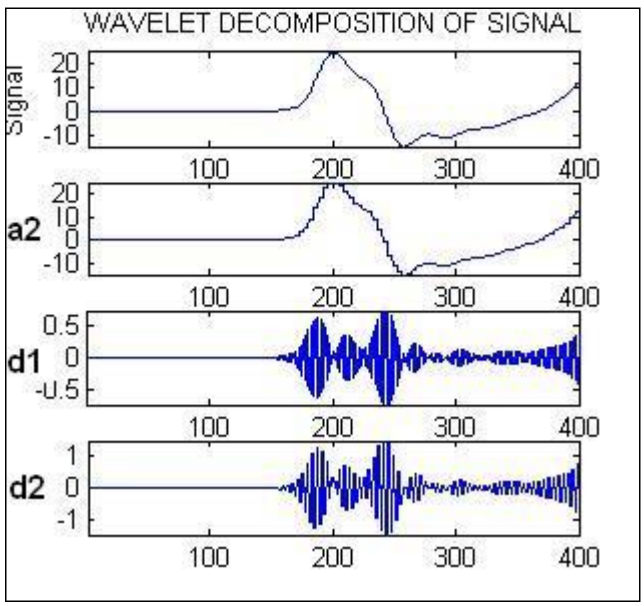

Fig. 3(c): Wavelet Decomposition of differential Current for secondary inter-turn fault

In Fig.3 (a), (b) and (c) level ‘d1’ to ‘d2’ indicates the detailed coefficients of wavelet transform, while ‘a5’ represents approximation coefficients of wavelet transform. From these figures there is no classification between inrush and fault current by visual inspection, and from changes in the detailed and approximation coefficients for both cases it is very difficult to draw any inference, therefore some suitable artificial intelligence technique should be used to classify these events in transformer.

4. Neural Network as A Classifier: Artificial Neural Network

The application of artificial neural networks to discriminate the fault has given a lot of attention recently. The simplest definition of a neural network, more properly referred to as an 'artificial' neural network (ANN), is provided by the inventor of one of the first neuro computers, Dr. Robert Hecht-Nielsen. He defines a neural network as**:** ***"...a computing system made up of a number of simple, highly interconnected processing elements, which process information by their dynamic state response to external inputs."*** An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process.

4.1 Architecture of neural networks

Neural networks are typically organized in layers. Layers are made up of a number of interconnected 'nodes' which contain an 'activation function'. Patterns are presented to the network via the 'input layer', which communicates to one or more 'hidden layers' where the actual processing is done via a system of weighted 'connections'. The hidden layers then link to an 'output layer' where the answer is output as shown in Fig.4.

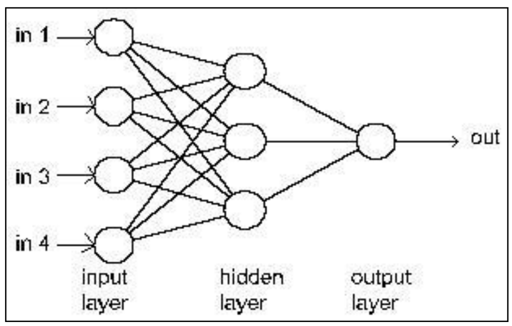

Fig .4 Architecture of ANN

4.2 Feed-forward networks

Feed-forward ANNs allow signals to travel one way only; from input to output. There is no feedback (loops) i.e. the output of any layer does not affect that same layer. Feed-forward ANNs tend to be straight forward networks that associate inputs with outputs as shown in Fig.5. They are extensively used in pattern recognition. This type of organization is also referred to as bottom-up or top-down.

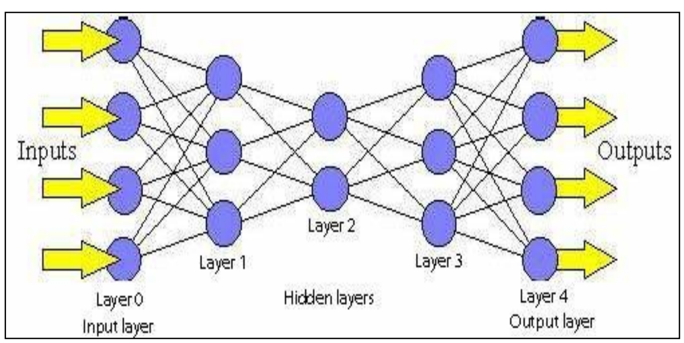

Fig.5. Feed-forward network

4.3 Feedback networks

Feedback networks can have signals traveling in both directions by introducing loops in the network. Feedback networks are very powerful and can get extremely complicated. Feedback networks are dynamic; their 'state' is changing continuously until they reach an equilibrium point. They remain at the equilibrium point until the input changes and a new equilibrium needs to be found. Feedback architectures are also referred to as interactive or recurrent, although the latter term is often used to denote feedback connections in single-layer organizations.

4.4 The Learning Process

All learning methods used for adaptive neural networks can be classified into two major categories:

- Supervised learning
- Unsupervised learning

Supervised learning which incorporates an external teacher, so that each output unit is told what its desired response to input signals ought to be. During the learning process global information may be required. Paradigms of supervised learning include error-correction learning, reinforcement learning and stochastic learning.

Important issue concerning supervised learning is the problem of error convergence, i.e. the minimization of error between the desired and computed unit values. The aim is to determine a set of weights, which minimizes the error. One well-known method, which is common to many learning paradigms, is the least mean square (LMS) convergence.

Unsupervised learning uses no external teacher and is based upon only local information. It is also referred to as self-organization, in the sense that it self-organizes data presented to the network and detects their emergent collective properties. Paradigms of unsupervised learning are Hebbian learning and competitive learning.

From Human Neurons to Artificial Neurons other aspect of learning concerns the distinction or not of a separate phase, during which the network is trained, and a subsequent operation phase. We say that a neural network learns off-line if the learning phase and the operation phase are distinct. A neural network learns on-line if it learns and operates at the same time. Usually, supervised learning is performed off-line, whereas unsupervised learning is performed on-line.

In this paper, the fully-connected multilayer feed-forward neural network (FFNN) was used and trained with a supervised learning algorithm called back-propagation. The FFNN consists of an input layer representing the input data to the network, some hidden layers and an output layer representing the response of the network. Each layer consists of a certain number of neurons; each neuron is connected to other neurons of the previous layer through adaptable synaptic weights ***w*** and biases ***b*,** cf. Fig.6.

If the inputs of neuron j are the variables x1, x2, . . , xi, . . . , xN, the output uj of neuron j is obtained as follows:

$$u_j = \varphi\left(\sum_{i=1}^{N} \omega_{ij}\, x_i + b_j\right) \tag{5}$$

where ωij represents the weight of the connection between neuron j and the i-th input, bj represents the bias of neuron j and ' is the transfer function (activation function) of neuron j.

An FFNN of three layers (one hidden layer) is considered with N, M and Q neurons for the input, hidden and output layers, respectively. The input patterns of the ANN represented by a vector of variables x = x1, x2, . . . , xi, . . . , xN) submitted to the ANN by the input layer are transferred to the hidden layer. Using the weight of the connection between the input and the hidden layer, and

the bias of the hidden layer, the output vector u = (u1, u2, . . . , uj , . .. , uM) of the hidden layer is then determined.
The output uj of neuron j is obtained as follows:

$$u_j = \varphi hid\left(\sum_{i=1}^{N} w_{ij}^{hid}\, x_i + b_j^{hid}\right) \qquad (6)$$

where w_{ij}^{hid} represents the weight of connection between neuron j in the hidden layer and the i-th neuron of the input layer, b_j^{hid} represents the bias of neuron j and φhid is the activation function of the hidden layer.
The values of the vector u of the hidden layer are transferred to the output layer. Using the weight of the connection between the hidden and output layers and the bias of the output layer, the output vector y = (y1, y2, . . . , yk, . . . , yQ) of the output layer is determined.
The output yk of neuron k (of the output layer) is obtained as follows:

$$y_k = \varphi out\left(\sum_{j=1}^{M} w_{jk}^{out}\, u_j + b_k^{out}\right) \qquad (7)$$

where w_{jk}^{out}represents the weight of the connection between neuron k in the output layer and the j-th neuron of the hidden layer, bout k represents the bias of neuron k and 'out is the activation function of the output layer.
The output yk (corresponding to the given input vector x) is compared with the desired output (target value) Yd k. The error in the output layer between yk and yd k (ydk − yk) is minimized using the mean square error at the output layer (which is composed of Q output neurons), defined by

$$\mathbf{E}= \frac{1}{2}\sum_{k=1}^{Q}\left(y_k^d - y_k\right)^2 \qquad (8)$$

Training is the process of adjusting connection weights w and biases b. In the first step, the network outputs and the difference between the actual (obtained) output and the desired (target) output (i.e., the error) is calculated for the initialized weights and biases (arbitrary values). During the second stage, the initialized weights in all links and biases in all neurons are adjusted to minimize the error by propagating the error backwards (theback-propagation algorithm). The network outputs and the error are calculated again with the adapted weights and biases, and the process (the training of the ANN) is repeated at each epoch until a satisfied output y_k (corresponding to the values of the input variables x) is obtained and the error is acceptably small The adjustment by the back-propagation algorithm, which is required in the weights and biases to minimize the total mean square error, is computed as

$$\Delta w = w^{new} - w^{old} = -\eta \frac{\partial E}{\partial w} \qquad (9a)$$

$$\Delta b = b^{new} - b^{old} = -\eta \frac{\partial E}{\partial b} \qquad (9b)$$

where η is the learning rate. The computation in (5) reflects the generic rule used by the backpropagation algorithm. Equations (6) and (7) illustrate this generic rule of adjusting the weights and biases. For the output layer, we have

$$\Delta w_{jk}^{new} = \propto \Delta w_{jk}^{old} + \eta \delta_k y_k \,, \quad (10a)$$

$$\Delta b_k^{new} = \propto \Delta b_k^{old} + \eta \delta_k, \quad (10b)$$

Where $\propto$ is the momentum factor (a constant between 0 and 1) and $\delta_k = y_k^d - y_k$. For the hidden layer, we get :

$$\Delta w_{ij}^{new} = \propto \Delta w_{ij}^{old} + \eta \delta_j y_j \quad (11a)$$

$$\Delta b_j^{new} = \propto \Delta b_j^{old} + \eta \delta_j \quad (11b)$$

Where, $\delta_j = \sum_k^Q \delta_k w_{jk}$ And $\delta_k = y_k^d - y_k$

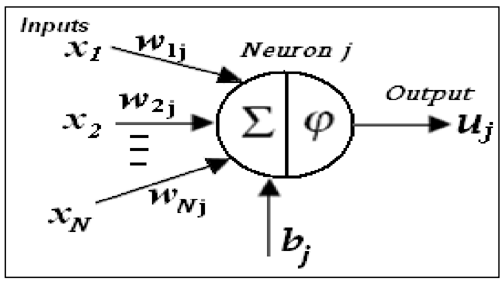

Fig.6 Information processing in a neural network unit

5. Experimental Results

The following types of conditions / events were staged on the custom built transformer, at different time instants of the supply voltage waveform and at different loading conditions.

a. *Magnetizing Inrush current*
b. *Short circuit of 04% primary turns only.*
c. *Short circuit of 04% secondary turns only.*

Various types of wavelets were tried; however application of Haar wavelet gave encouraging results for extracting relevant features and subsequently classifying the type of event, using ANN.

One hundred and four cases of each event, i.e. inrush, primary winding interturn short circuit and secondary winding interturn short circuit were staged on the custom built transformer. One cycle was taken for the purpose of analysis.

A visual inspection of the D1 level coefficient for all the three events mentioned above suggest a specific pattern for the given type of event. However for on line implementation of the strategy the statistical parameters like mean, maximum, standard deviation, variance, kurtosis, RMS value, energy, absolute sum, shape factor, and crest factor of the D1 level coefficients are given as an input to the ANN for classification the events.

Here, Multilayer Perceptron (MLP) Neural Network is used as a classifier network. This input dataset obtained through experiments was found to be non-linear non-separable mixed data. ANN possesses ability to classify such mixed datasets and can be used effectively in obtaining the correct classifications of the events in transformer. For generalization the randomized data is fed to the network and is trained for different hidden layers. The numbers of Processing Elements (PEs) in the hidden layer are varied. The network is trained and minimum MSE and percent accuracy for both Inrush and faults with respect to number of processing elements in the hidden layer is obtained.

For training method 'trainlm' of Levenberg- Marquardt with Learning Rate LR = 0.8, Momentum MM = 0.6, training data percentage TR = 50 %, cross validation data percentage CV = 20% and Testing data percentage TS = 30% the variation of percent accuracy with respect to number of processing elements in the hidden layer is plotted in Fig.7.

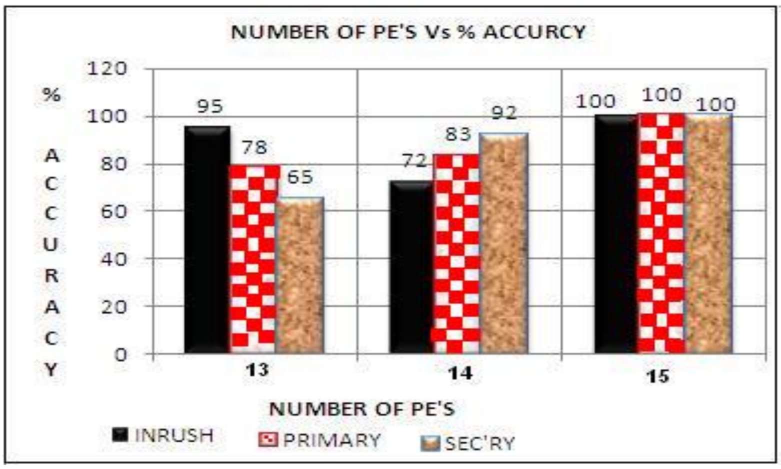

Figure 7: Variation of % Accuracy with number of processing elements in the hidden layer.

From Fig.7. it is found that for fifteen processing elements in the hidden layer the output result is hundred percent, that means there is a clear discrimination between Inrush and Fault. Hence, this network of training method 'trainlm' with fifteen numbers of processing elements in the hidden layer is the best suited network.

6. Proposed Algorithm

MLP with 15 hidden neurons has shown the capability to discriminate the inrush current from internal fault current. Long term memory weights were then used at the processor level to take decisions regarding the classification of inrush and faults. The on-line discrimination process of inrush and faults is illustrated in Fig.8. The step for carrying out the on-line detection scheme is presented as under-

1.Captured one cycle of primary and secondary current by data acquisition system.
2. Obtain differential current Id= Ip- Is
3. If the rms value of differential current value is less than threshold value, go to step 1.
4. Calculate DWT of differential current whose value is above threshold value.

5. Obtain statistical parameters of decomposed level d1.
6. These obtained parameters are given to ANN as input data to discriminate the faults and inrush that is healthy condition.
7. If ANN output is discriminated as fault, then issue trip signal otherwise proceed further i.e. monitor the differential current.

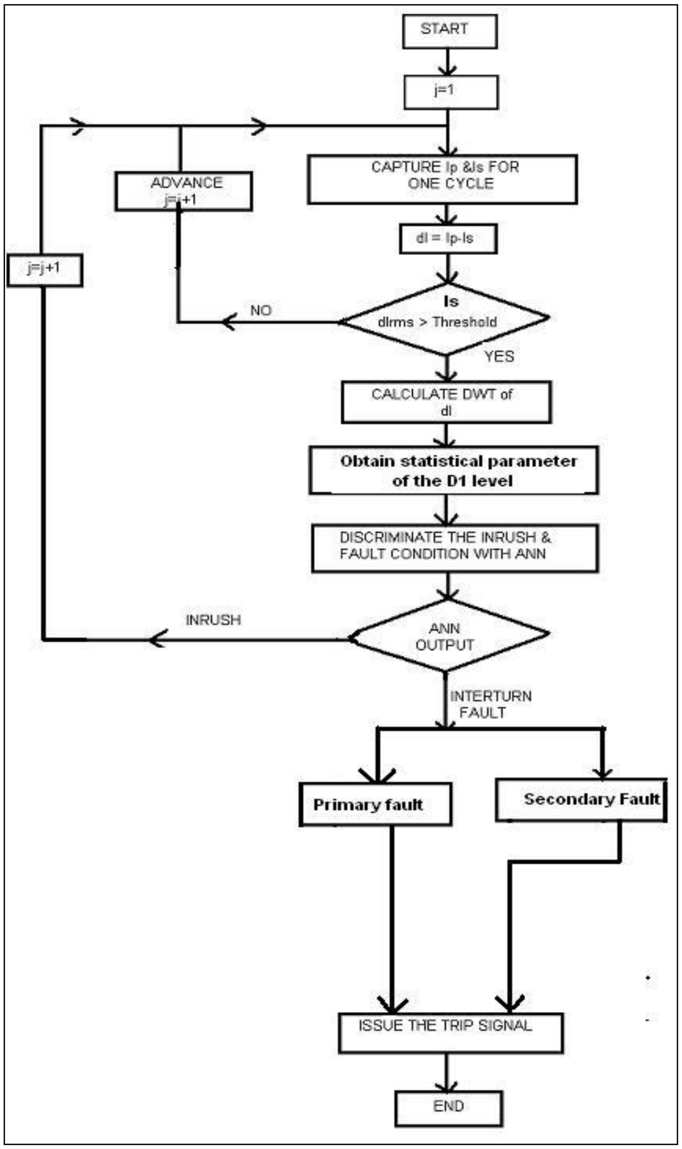

Figure 8: Flow chart for on-line discriminating scheme

7. Conclusion

A new method of discriminating magnetizing inrush current from interturn faults in a transformer is presented in this paper. Wavelet transform with its inherent time frequency localization property is employed to extract discriminating features from the differential current.

The ANN was successes in classifying the type of event from the extracted features given as input.

The algorithm has been tested successfully online, by staging these events on the custom built transformer. These events are identified in less than one cycle after their inception. This classification may occur for situations in which inception angle, fault resistance and other parameters are very different from those used during the ANN's learning. If this is the case, it is necessary to add the misclassified fault record, to the learning database and retrain the ANN.

References

[1] Moravej, Z., Vishvakarma D. N., and Singh S. P. "Digital filtering algorithms for the differential relaying of power transformer: an over view," Electric Machines and Power Systems, vol. 28, pp.485-500, 2000.

[2] Omar, A., and Youssef S. "A Wavelet-Based Technique for Discrimination between Faults and Magnetizing Inrush currents in Transformers," IEEE Transaction on Power Delivery, vol. 18, no. 1, January. 2003.

[3] Hamedani Golshan, M. E., Saghaian-Nejad, Saha, M., and Samet, A. H. "A new method for recognizing internal faults from inrush current conditions in digital differential protection of power transformers," Electric Power System Research, vol. 71, pp. 61-71, 2004.

[4] Eissa, M. M. "A Novel Digital Directional Transformer Protection Technique Based on Wavelet Packet," IEEE Transaction on Power Delivery, vol. 20, no.3, July 2005.

[5] J. Faiz and S. Lotfifard. "A Novel Wavelet Based Algorithm for Discrimination of Internal Faults from Magnetizing Inrush Current in Power Transforms," IEEE Transaction on Power Delivery, vol. 21, no. 4, October2006.

[6] Monsef, H., and Lotfifard, S. "Internal fault current identification based on wavelet transform in power transformers," Electric Power System Research, vol. 77 pp. 1637-1645, 2007.

[7] Moravej, Z., Vishwakarma, D. N., and Singh, S. P. "ANN based protection scheme for power transformer,"Electric Machines and Power Systems, vol.28, No.9, pp.875-884, 2000.

[8] Bastard, P., Meunier, M., and Regal, H. "Neural network-based algorithm for power transformer differential relays," Proceeding Industrial Electric Engineering, vol. 142, no. 4, pp. 386-392, 1995.

[9] Mao P. L., and Aggarwal, R. K. "A novel approach to the classification of the transient phenomena in power transformers using combined wavelet transform and neural network," IEEE Transaction on Power Delivery, vol.16, pp.654–660, 2001.

[10] Segatto, E. C., and Coury, D. V. "A differential Relay for Power Transformer Using Intelligent Tools," IEEE Transaction on Power Delivery, vol. 21, no. 3, August 2006.

[11] Moravej, Z., Vishwakarma, D. N., and Singh, S. P. "Radial basis function (RBF) neural network for protection of a power transformer", Electric power components and systems, vol. 29, pp. 307-320, 2001.

[12] Shin, M. C., Park C. W., and Kim J. H. "Fuzzy logic-based for large power transformer protection," IEEE Transaction on Power Delivery, vol. 18, no, 3, pp. 718-724, July 2003.

[13] Moravej, Z.Power transformer protection using support vector machine network," PES (2008) U.S.A, pp.29-33.

[14] Jazebi, S., Vahidi, B., and Hosseinian, S. H. "A new stochastic method based on hidden Markov models to transformer differential protection," IEEE Proceeding OPTIM'08, vol. 1, pp. 179–184, 2008.