Application of Wavelet Transform and MLP Neural Network for Ferroresonance Identification

G. Mokryani and M. R. Haghifam

***Abstract*—A novel method for ferroresonance detection is presented in this paper. Using this method ferroresonance can be discriminate from other transients such as capacitor switching, load switching, transformer switching. Wavelet transform is used for decomposition of signals and Multi Layer Perceptron (MLP) neural network used for classification. Ferroresonance data and other transients are obtained by simulation using EMTP program. Results show that the MLPNN trained with the Levenberg–Marquardt algorithm is effective for discriminating Ferroresonance from other transients.**

***Index Terms*—MLP neural network, ferroresonance, EMTP program, wavelet transform, Levenberg–Marquardt algorithm.**

I. Introduction

DISTURBANCE due to ferroresonance is a common phenomenon in electric power distribution system operation. Depending on circuit conditions, its effect may be a random over voltage that could be either a short transient for few cycles, a continuous over voltage or even a jump resonance. It causes both phase-to-phase and phase-to-ground high sustained oscillating over voltages and over currents with sustained levels of distortion to the current and voltage waveforms, leading to transformer heating together with excessively loud noise due to magnetostriction, electrical equipment damage, thermal or insulation breakdown and mal-operation of the protective devices. Detection of ferroresonance presents still important and unsolved protection problem, especially in distribution networks. In the area of power quality several studies have been carried out to detect and locate disturbances, for example, Kalman Filter System [1], Short time-filter Fourier Transform [2], Fuzzy Expert Systems for classification of Power Quality disturbances [3], automatic classification based on extraction of characteristics of wavelets [4], Bayesian classification [5], Neural Networks [6], and Hidden Markov Model [7]. The wavelet transform has been used to detect and locate various types of power quality disturbances, decomposing a disturbance into its wavelet coefficients using a Multi-Resolution Analysis (MRA) technique. Santoso *et al.* set up an investigation line on this area with the work presented in [8], then the authors in [9] make the proposal that, based on uniqueness of squared WT coefficients at each scale of the power quality disturbance, a classification tool such as neural network may be employed for the classification of disturbances [10], [11]. In high frequencies, wavelet transform has a good time resolution and a weak frequency resolution.

On the contrary, in low frequencies, it has a good frequency resolution and a weak time resolution. In this paper a new ferroresonance detection method that uses wavelet transform and MLP neural network is presented. In high frequencies, wavelet transform has a good time resolution and a weak frequency resolution. Ferroresonance data was gathered from a 20 kV radial distribution feeder in a real network. Transient state data was produced by simulation using EMTP program. The ferroresonance phenomenon and theory are introduced in Sections II and III, a case study and data collection are explained in Section IV. Wavelet transform and MLP neural network are introduced in Sections V and VI, respectively. Simulation results are shown in Section VII.

II. Ferroresonance Phenomenon

Ferroresonance is a particular type of oscillation, which can occur when a non-linear inductance is connected in series or parallel with a capacitance. Power networks are made up of a large number of saturable inductances (power transformers, voltage measurement inductive transformers (VT), shunt reactors), as well as capacitors (cables, long lines, capacitor voltage transformers, series or shunt capacitor banks, voltage grading capacitors in circuit-breakers). Therefore, ferroresonance may occur in any power system. The main feature of this phenomenon is that more than one stable steady state response is possible for the same set of the network parameters [12]. Inductive voltage transformers (VTs) are widely used to measure power system voltages. These equipments are highly favorable to ferroresonance in vicinity of capacitive sources. VTs are connected between phases or between phase and ground. One form of ferroresonance involving VTs can occur in a three-phase system as the result of an unbalanced switching operation when the VT is connected between two phases. In this condition, the capacitance of the open phase is energized through the magnetizing inductance of the VT. Another example is energization of a VT through grading capacitors of an open circuit breaker while the VT is connected between phase and ground. Since in these cases the inductance and capacitance are in series, this type of ferroresonance is sometimes referred to

Manuscript received November 8, 2007; revised June 8, 2008.

G. Mokryani is with Islamic Azad University, Soofian Branch, Soofian, I. R. Iran (e-mail: gmokryani@gmail.com).

M. R. Haghifam is with Tarbiat Modares University, Tehran, I. R. Iran (e-mail: haghifam@modares.ac.ir).

Publisher Item Identifier S 1682-0053(09)1663

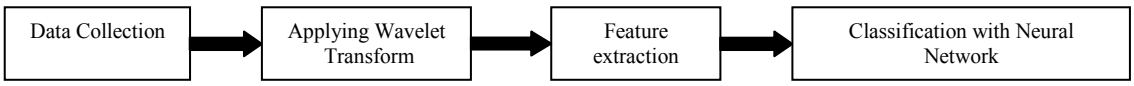

Fig. 1. Proposed algorithm.

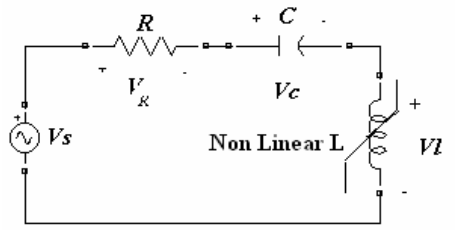

Fig. 2. Ferroresonance circuit.

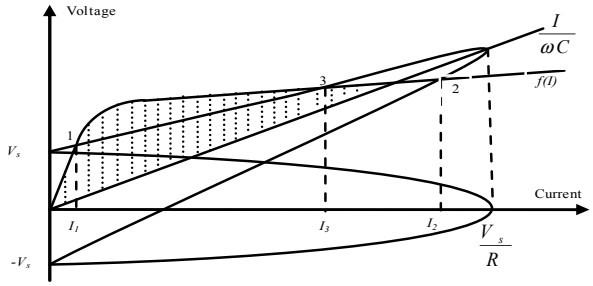

Fig. 3. Graphical solution of ferroresonance circuit.

as series ferroresonance. When a VT is connected between phase and ground, the magnetizing inductance of the VT is in parallel with the zero sequence capacitance of the system. With such VT, a second type of ferroresonance can occur during a temporary overvoltage due to an islanding condition with insufficient voltage regulation or a single phase to ground fault in an isolated neutral system. In these conditions, as the VT saturates, there is an exchange of energy between the zero sequence capacitance of the system and the highly nonlinear magnetizing inductance of the VT. The rapid changes in VT core flux during this period can produce high overvoltages. Since in this case the inductance and capacitance are in parallel, this second type of ferroresonance is sometimes referred to as parallel ferroresonance. The probability of both types of ferroresonance occurring is somewhat unpredictable as both depend on such factors as the cable lengths, the amount of system capacitance, the connection and saturation characteristics of the transformers, the amount of load or burden, etc. Numerous articles have been written on various aspects of ferroresonance phenomenon, from recording actual cases of ferroresonance or qualitative analysis of ferroresonance cases to computer aided modeling [13]. Most solutions use a simple mathematical expression to represent the non-linear B-H characteristic of the transformer core. Using the proposed algorithm that shown in Fig. 1, we can predict some possibilities in happening ferroresonance and so we can face it with making some relays.

III. FERRORESONANCE THEORY

Consider the circuit of Fig. 2, in which the linear inductance has been replaced with a nonlinear inductor.

When the series resistance is ignored, the sum of the voltages around the only mesh of the circuit can be rewritten as

$$V_s(t) - V_c(t) - V_l(t) = 0 \qquad (1)$$

The value of V_c can be replaced by its time-integral expression and V_l as the total derivative of $L(i)i(t)$. Then (1) can be written as

$$V_s(t) - \frac{1}{C}\int i(t)dt - \frac{d}{dt}[L(i)i(t)] = 0 \qquad (2)$$

Evaluating (2) and substituting $q(t) = \int i(t)dt$ will result in

$$V_s(t) - \frac{1}{C}q(t) - L(i)\frac{d^2q(t)}{dt^2} - \frac{dq(t)}{dt}\frac{dL(i)}{di(t)} = 0 \ . \qquad (3)$$

From (3) it is evident that finding a closed-form solution for this nonlinear circuit will be quite difficult. This would be made more evident by adding a source impedance and by proving a complete equivalent of the transformer. Historically, methods of graphical solution for the circuit of Fig. 2, including the series resistance, can be obtained from two independent relationship for the voltage across the inductance and the capacitance. The voltage across the inductor is proportional to the frequency, and the voltage across the capacitance is proportional to the current and inversely proportional to the frequency and capacitance.

$$V_l = \omega f(I)$$
$$V_c = -\frac{I}{\omega C} \qquad (4)$$

The total magnitude of voltage for the circuit is

$$V_s = \sqrt{(V_l + V_c)^2 + (RI)^2} \qquad (5)$$

From (4) and (5), the voltage across the nonlinear inductor can be written as

$$V_l = \sqrt{V_s^{\,2} - (RI)^2} + \frac{I}{\omega C} \qquad (6)$$

The first term in the right-hand side of (6) ($\sqrt{V_s^{\,2} - (RI)^2}$) represents an ellipse whose main axes have the magnitude of V_s and V_s / R, and the second term is a straight line having slope of $I / \omega C$. Adding these two quantities represents an oblique ellipse, whose intersection with the characteristic of V_l presents the three possible states of the oscillation of the circuit. Fig. 3 shows the graphical solution for the ferroresonance circuit of Fig. 2. Points 1 and 2 in Fig. 3. represent the stable solution, whereas point 3 represents an unstable solution.

To show this, rewrite (6) as

$$(RI)^2 = V_s^{\,2} - [f(I) - \frac{I}{\omega C}]^2 \qquad (7)$$

If the quantity $[f(I) - I/\omega C]$ increases in magnitude with an increase in current, then according to (7), $(RI)^2$ tends to decrease, and this suppresses any further increase in current. Thus stability is achieved. However, if the quantity $[f(I) - I/\omega C]$ decreases in magnitude, with an increase of current, the magnitude of $(RI)^2$ tends to

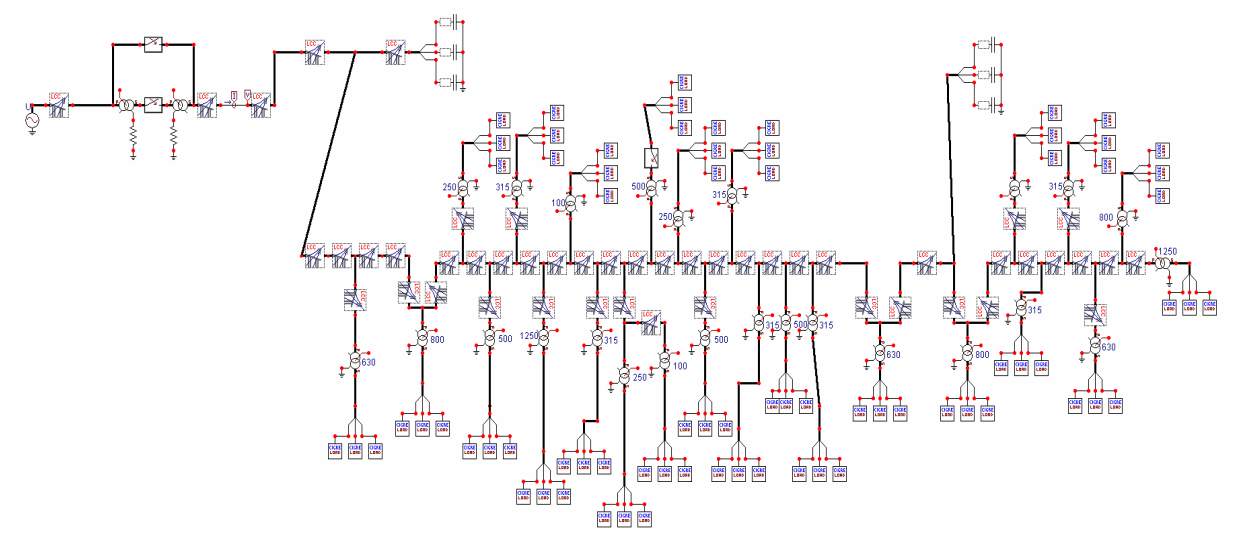

Fig. 4. 20 kV feeder.

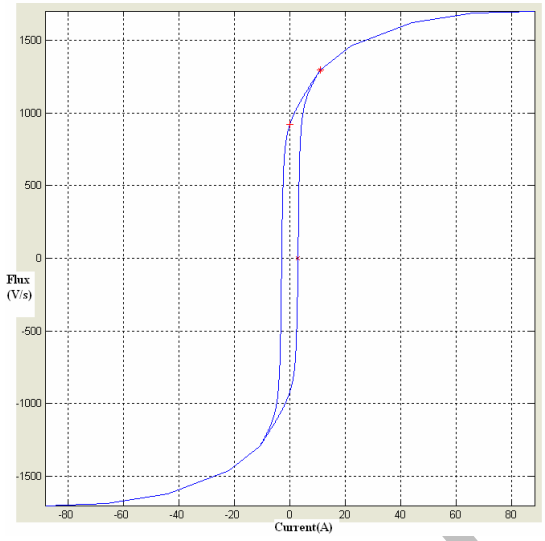

Fig. 5. Magnetizing curve.

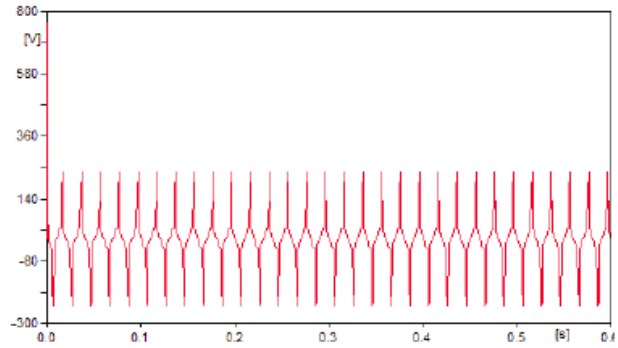

Fig. 6. An example of ferroresonance.

increase and under this condition, the current continues to increase and the solution is unstable. The dashed area in the Fig. 3 shows the variation of the magnitude of $[f(I)-I/\omega C]$. Point 3 in Fig. 3 corresponds to an unstable solution since $[f(I)-I/\omega C]$ decrease with an increase of current. In the same figure points 1 and 2 represent the stable solution.

IV. Data Collection

In order to obtain the signals, a part of a 20kV feeder has been selected in Qeshm Island which is illustrated in Fig. 4

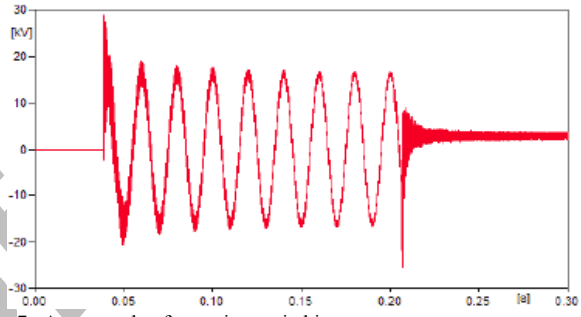

Fig. 7. An example of capacitor switching.

[14]. These signals include: ferroresonance, capacitor switching, load switching, and transformer switching signals. The models determined to be simulated by the EMTP software are, π and load frequency model (CIGRE), for line and load respectively, saturable model is used for all transformers. The inductor with hystersis loop of TYPE 96 was used for modeling hystersis loop in EMTP, which was connected to the outlet magnetizing branch of the transformer. The magnetization curve of the transformers is illustrated in Fig. 5. Feeder information is provided in the appendix. Load data and transformers data is provided in Tables I and II, respectively. All kind of ferroresonance that different parameters such as switching types, transformer connection type, hystersis phenomenon, line capacitance feature, line length and load impact which can be influential in the occurrence of this phenomenon have been simulated. Fig. 6 illustrates a type of ferroresonance which has been simulated by the EMTP. Different types of capacitor switching have been obtained through the switching of the two capacitor banks of the feeder in various forms. For simulating different types of load switching, we switch the loads in different arrangements. For example, we firstly switch them one at a time, then two at a time, and other arrangements can be achieved by switching one or two of the loads with a part of the feeder. Thus, different signals are obtained. An example of which is provided in Fig. 8. For simulating the transformer switching signals, we switch the transformers in different orders. For example, we switch the transformers one at a time, then two at a time, and different

TABLE I
CONSTANT PARAMETERS OF THE CIGRE LOAD MODEL USUALLY CONSIDERED IN THE EMTP PROGRAM

No.	I_a (A)	I_b (A)	I_c (A)	I_n (A)	Capacity of connected transformers (kVA)
1	115	78	110	90	630
2	295	200	220	165	800
3	40	60	55	0	500
4	200	250	220	0	1250
5	40	40	40	8	315
6	20	25	25	10	250
7	80	50	40	0	100
8	85	40	70	40	500
9	145	130	120	40	315
10	205	180	205	65	500
11	125	100	105	25	630
12	30	60	50	20	800
13	65	55	55	25	315
14	155	140	105	99	630
15	60	55	55	17	250
16	33	57	45	32	315
17	5	20	20	15	100
18	60	65	75	25	500
19	25	65	60	35	250
20	80	85	75	28	315
21	15	15	15	5	100
22	175	130	145	45	315
23	165	175	150	55	800
24	125	150	150	45	1250

types can be achieved by switching one or two of the transformers with a part of the feeder. Thus, different signals are obtained. An example of which is provided in Fig. 9. This way, for each group of signals, 100 types can be obtained. Then we normalize (scale) them in the max-min range (0 to 1). This is very influential in the exact determination of the features and every pattern.

V. WAVELET TRANSFORM

Wavelet Transform (WT) was introduced by Morlet at the beginning of 1985 and has attracted much interest in the fields of speech and image processing. Applications of DWT in power systems are reported for:

- Power system transients [15].
- Power quality assessment [16].
- Modeling of system component in wavelet domain [17].

In this section an introduction to wavelet transform is presented. More details can be found in [18], [19]. The WT was developed as an alternative to the short time Fourier Transform (STFT) to overcome problems related to its frequency and time resolution properties. More specifically, unlike the STFT that provides uniform time resolution for all frequencies, the DWT provides high time resolution and low frequency resolution for high frequencies and high frequency resolution and low time resolution for low frequencies. The DWT is a special case of the WT that provides a compact representation of a signal in time and frequency that can be computed efficiently. The DWT is defined by the following equation

$$W(j,K)=\sum_{j}\sum_{k}x(k)2^{-j/2}\varphi(2^{-j}n-k) \qquad (8)$$

where $\varphi(t)$ is a time function with finite energy and fast decay called the mother wavelet. The DWT analysis can be performed using a fast, pyramidal algorithm related to multi-rate filter banks. As a multi-rate filter bank DWT can

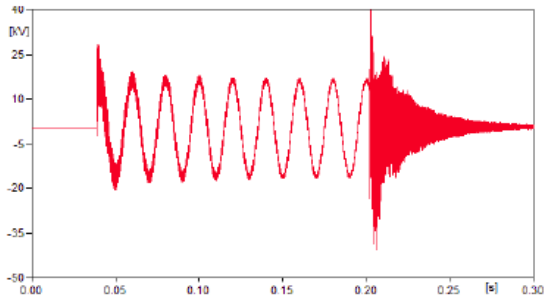

Fig. 8. An example of Load switching.

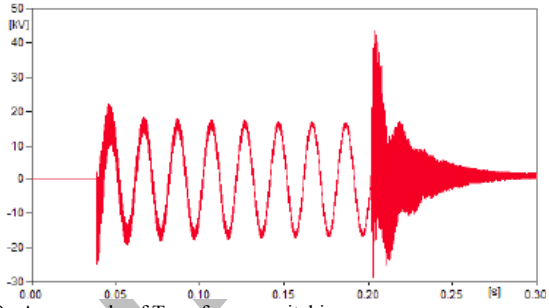

Fig. 9. An example of Transformer switching.

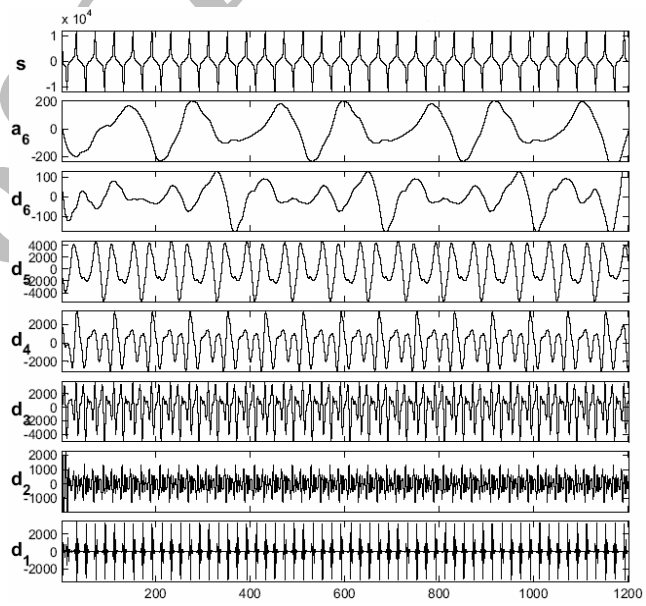

Fig. 10. Decomposition of ferrosonance by Daubechies mother wavelet.

be viewed as a constant Q filter bank with octave spacing between the centers of the filters. Each sub band contains half the samples of the neighboring higher frequency sub band. In the pyramidal algorithm the signal is analyzed at different frequency bands with different resolution by decomposing the signal into a coarse approximation and detail information. The coarse approximation is then further decomposed using the same wavelet decomposition step. This is achieved by successive high pass and low pass filtering of the time domain signal and is defined by the following equations:

$$y_{high}[k]=\sum_{k}x[n]g[2k-n] \qquad (9)$$

$$y_{low}[k]=\sum_{k}x[n]h[2k-n] \qquad (10)$$

where $y[k]$ high and $y[k]$ low are the outputs of the high pass (g) and low pass (h) filters, respectively after

TABLE II
TRANSFORMER DATA

No.	S (kVA)	Connection	N_1/N_2	UK%	P_{oc} (W)	$In1$ %	P_{sc} (W)
1	3000	Yd1	63/20 kV	14	22410	2.83	151247
2	1250	Dy5	20/0.4 kV	6	2100	1.4	16400
3	1000	Dy5	20/0.4 kV	6	1750	1.4	13500
4	800	Dy5	20/0.4 kV	6	1450	1.5	11000
5	630	Dy5	20/0.4 kV	6	1200	1.6	9300
6	500	Dy5	20/0.4 kV	6	1000	1.7	7800
7	400	Dy5	20/0.4 kV	6	850	1.8	6450
8	315	Dy5	20/0.4 kV	6	720	2	5400
9	250	Dy5	20/0.4 kV	6	650	2.3	4450
10	100	Dy5	20/0.4 kV	6	340	2.6	2150
11	50	Dy5	20/0.4 kV	6	210	2.8	1250

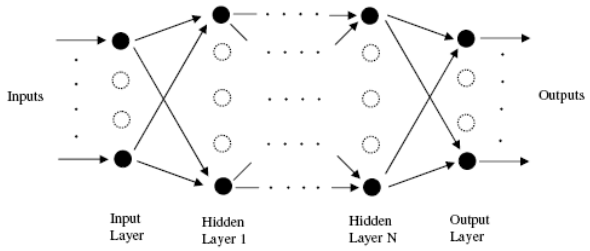

Fig. 11. MLP neural network topology.

sub sampling by 2. Down sampling the number of resulting wavelet coefficients becomes exactly the same as the number of input points. A variety of different wavelet families have been proposed in the literature. The choice of mother wavelet plays a significant role in time frequency analysis. It also depends on a particular application. In this work all wavelets available in the Wavelet Toolbox of MATLAB program [20] were used for the decomposition of the signals and the best answer was obtained with Daubechies mother wavelet. It was found to have the most correlation with the decomposed signals and was selected for this procedure.

A. Applying Wavelet Transform and Feature Extraction

The decomposition is done by modifying the wavelet transform through passing the signal via a digital half band low pass filter. This digital half band low pass filter excludes all the signals which are higher than the half of the value of the largest signal frequency. If a signal having Nyquist rate(which is twice the largest frequency in the signal) was taken as a sample, the largest frequency present in the signal would be π radian. That is, Nyquist frequency in the range of discrete frequency corresponds π (rad/s). After a signal passes through a digital half band low pass filter, according to the theory of nyquist, half of the signals can be excluded, for now the signal has the maximum frequency of $\pi/2$ (rad/s). Thus the obtained signal has a length half of that of the original one. This procedure is repeated for 6 times and the signals omitted by the low pass filter at each time, are considered as detail signals. The energies of these detail signals are the features extracted from the patterns to feed into the neural network. In Fig. 10 a pattern of 4 signals with 6 detail signals and an approximation signal obtained by applying the Db wavelet transform up to six levels is illustrated. According to the definition, the energy of every discreet signal such as $x(n)$ is defined as follows: (N equals the length of the signal)

$$E(x) = \sum_{n=<N>} |x(n)|^2 . \quad (11)$$

VI. MLP NEURAL NETWORK

ANNs may be defined as structures comprised of densely interconnected adaptive simple processing elements (neurons) that are capable of performing massively parallel computations for data processing and knowledge representation. ANNs can be trained to recognize patterns and the non-linear models developed during training allow neural networks to generalize their conclusions and to make applications to patterns not previously encountered. The MLPNNs, which have features such as the ability to learn and generalize, smaller training set requirements, fast operation, ease of implementation and are, therefore, the most commonly used neural network architectures, have been adapted for describing the alertness level of an arbitrary subject. Presently the most widely used ANN type is a MLP neural network which has been playing a central role in applications of neural networks. The MLP neural network shown in Fig. 11 is a nonparametric technique for performing a wide variety of detection and estimation tasks [21]-[23]. In the MLPNN, each neuron j in the hidden layer sums its input signals x_i after multiplying them by the strengths of the respective connection weights w_{ji} and computes its output y_j as a function of the sum

$$y_j = f(\sum w_{ji} x_i) \quad (12)$$

where f is activation function that is necessary to transform the weighted sum of all signals impinging onto a neuron. The activation function (f) can be a simple threshold function, or a sigmoidal, hyperbolic tangent, or radial basis function. The sum of squared differences between the desired and actual values of the output neurons E is defined as

$$E = \frac{1}{2}\sum_i (y_{dj} - y_j)^2 \quad (13)$$

where is y_{dj} is the desired value of output neuron j and y_j is the actual output of that neuron. Each weight w_{ji} is adjusted to reduce E as rapidly as possible. How w_{ji} is adjusted depends on the training algorithm adopted. Training algorithms are an integral part of ANN model development. An appropriate topology may still fail to give a better model, unless trained by a suitable training algorithm. A good training algorithm will shorten the

TABLE III
IDENTIFICATION PERCENTAGE OF MLP NN

Signal	WT	Percentage of NN identification
First phase current	Db1	85%
First phase current	Db2	85%
First phase current	Db3	92%
Third phase current	Db2	90%
Second phase current	Db1	85.66%
Second phase current	Db2	80.33%
Second phase current	Db3	90.33%
Second phase voltage	Db6	95.33%
Third phase voltage	Db4	95.65%
Third phase voltage	Db3	97%
Third phase voltage	Db5	95%
Overall	**90.11%**	

training time, while achieving a better accuracy. Therefore, training process is an important characteristic of the ANNs, whereby representative examples of the knowledge are iteratively presented to the network, so that it can integrate this knowledge within its structure. There are a number of training algorithms used to train a MLP neural network and a frequently used one is called the backpropagation training algorithm [21]-[23]. The backpropagation algorithm, which is based on searching an error surface using gradient descent for points with minimum error, is relatively easy to implement. However, backpropagation has some problems for many applications. The algorithm is not guaranteed to find the global minimum of the error function since gradient descent may get stuck in local minima, where it may remain indefinitely. In addition to this, long training sessions are often required in order to find an acceptable weight solution because of the well known difficulties inherent in gradient descent optimization. Therefore, a lot of variations to improve the convergence of the backpropagation were proposed. Optimization methods such as second-order methods (conjugate gradient, quasi-Newton, Levenberg–Marquardt) have also been used for ANN training in recent years. The Levenberg–Marquardt algorithm combines the best features of the Gauss–Newton technique and the steepest-descent algorithm, but avoids many of their limitations. In particular, it generally does not suffer from the problem of slow convergence [24], [25]. Therefore, in this study the MLP neural network was trained with the Levenberg–Marquardt algorithm.

A. *Levenberg–Marquardt algorithm*

ANN training is usually formulated as a nonlinear least squares problem. Essentially, the Levenberg–Marquardt algorithm is a least-squares estimation algorithm based on the maximum neighborhood idea. Let $E(w)$ be an objective error function made up of m individual error terms $e_i^2(w)$ as follows

$$E(w) = \sum_{i=1}^{m} e_i^2(w) = \left\| f(w) \right\|^2 \tag{14}$$

where $e_i^2(w) = (y_{di} - y_i)^2$ and y_{di} is the desired value of output neuron i, y_i is the actual output of that neuron. It is assumed that function $f(\cdot)$ and its Jacobian J are known at point w. The aim of the Levenberg–Marquardt algorithm is to compute the weight vector w such that $E(w)$ is minimum. Using the Levenberg–Marquardt algorithm, a new weight vector w_{k+1} can be obtained from the previous weight vector w_k as follows

$$w_{k+1} = w_k + \delta w_k , \tag{15}$$

where δw_k is defined as

$$\delta w_k = -(J_k^T f(w_k))(J_k^T J_k + \lambda) I^{-1} . \tag{16}$$

In (14), J_k is the Jacobian of f evaluated at w_k, λ is the Marquardt parameter, I is the identity matrix.

VII. SIMULATION RESULTS

The obtained signals were analyzed by the Daubechies mother wavelet and the energies of the detail signals obtained through the applying wavelet transform up to six levels have been used as the features fed into the neural network. For the MLP neural network, 16 neurons are determined in the hidden layer, four of which are allocated to ferroresonance signals and the rest to capacitor switching, load switching, and transformer switching signals. For training the network all four types of signals are used; 95 signals for learning and 90 for testing. Also the learning rate of the neural network is 0.0001 and the number of epochs is selected 500. The Daubechies wavelet transform is applied in all the three phases of current and voltage of signals. The results are provided in Table III. It should be noted that the currents and the voltages are the primary currents and voltages of the feeder shown in Fig. 4. Using all wavelets available in the MATLAB Wavelet Toolbox program, many simulations were performed. After many trial and error the best answer was received with Daubechies mother wavelet. So, Daubechies, which has the most correlation with the signals, was chosen as the suitable mother wavelet in this procedure. By applying the Db2 in the second phase current of the signals, the neural network has the least precisian of 80.33% and by applying the Db3 in the third phase voltage of the signals, neural network shows the most precisian of 97%. The above results can be justified using Fig. 12. This figure compares the average of the components correspondent to the feature vectors extracted by applying Db2 and Db3 in the second phase current and the third phase voltage of signals, respectively (the rectangles corresponding the ferroresonance signals are darker). According to the figure, the features extracted by applying Db1 in the second phase current are much similar. Thus the precision of algorithm is less in this case. But the features exacted by the applying Db2 in the third phase voltage are least similar. Thus the precision of algorithm is more in this case.

VIII. CONCLUSIONS

In this paper, the MLP neural network and the wavelet transform have been used to distinguish ferroresonance from other transients. The presented algorithm has the highest precision on third phase voltage and lowest precision on the second phase current of the signals. One of the main advantages of this algorithm is capability of changing the number of extracted features by changing the number of wavelet transform levels. Also, the chosen network has the ability to classifying the nonlinear feature vectors in multi-dimension space. By increasing complexity, only the number of hidden layer neurons should be increased. The applied network has an acceptable

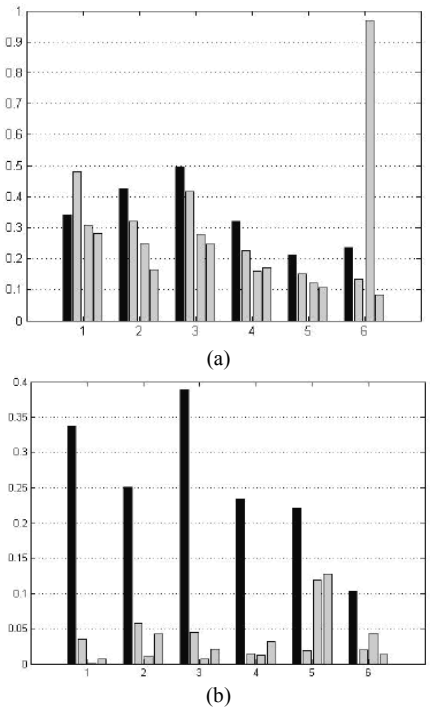

Fig. 12. Average of the components correspondent to the feature-vectors extracted in second phase current and third phase voltage, (a) Second phase current and (b) Third phase voltage.

precision in the recognition of unused patterns for learning. This fact highlights the practical importance of algorithm.

IX. Appendix

The data of feeder are: $R = 0.509\ \Omega/\text{km}$, $X = 0.3561\ \Omega/\text{km}$, Outside Radius of conductor=0.549 cm. Configuration of phases and mechanical data are:

Height of pole = 12 m, Sag in mid span = 2.32 m. Constant parameters of the CIGRE load model usually considered in the EMTP program are as follows:

$$A = 0.073,\ B = 6.7,\ C = 0.74$$

References

[1] H. Ma and A. Girgis , "Identification and tracking of Harmonics Source in a Power System using Kalman filter," *IEEE Trans. on Power Delivery*, vol. 11, no. 3, pp. 1659-1665, Jul. 1996.

[2] I. Gu, M. Bollen, "Time frequency and Timescale domain analysis of voltage Disturbances," *IEEE Trans. on Power Delivery*, vol. 15, no. 4, pp. 1279-1284, Oct. 2000.

[3] S. Santoso, J. Lamoree, M. Grady, E. powers, and S. Bhatt, "A Scalable PQ event identification system," *IEEE Trans. on Power Delivery*, vol. 15, no. 2, pp. 738-743, Apr. 2000.

[4] A. Gaouda M. salama, M. Sultan, and A. Chikhani, "Power quality detection and classification using wavelet multuresolution signal decomposition," *IEEE Trans. on Power Delivery,* vol. 14, no. 4, pp. 1469-1476, Oct. 1999.

[5] M. karami, H. Mokhtari, and R. Irvani, "Wavelet based on line disturbance detection for power quality application," *IEEE Trans. on Power Delivery*, vol. 15, no. 4, pp. 1212-1220, Oct. 2000.

[6] J. Chung, E. J. Powers, W. M. Grady, and S. C. Bhatt, "Electric power transient disturbance classification using wavelet based hidden markov models," in *Proc. of 2000 IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, vol. 6, 2000.

[7] B. Perunicic, M. Mallini, Z. Wang, Y. Liu, and G. T. Heydt, "Power quality disturbance detection and classification using wavelets and artificial neural networks," in *Proc. 8th Int. Conf. on Harmonics and Quality of Power*, pp. 77-82, 1998.

[8] C. L. Huang, H. Y. Chu, and M. T. Chen, "Algorithm comparison for high impedance fault based on staged fault test," *IEEE Trans. on Power Delivery,* vol. 3, no. 3, pp. 1427-1435, Oct. 1988.

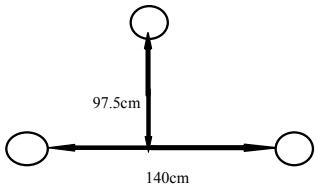

Fig. 13. Configuration of phases and mechanical data.

[9] H. Calhoun, M. T. Bishop, C. h. Eiceler, and R. E. Lee, "Development and testing of an electro-mechanical relay to detect fallen distribution conductors," *IEEE Trans. on Power apparatus and systems*, vol. 101, no. 6, pp. 1643-1650, Jun. 1998.

[10] Y. Sheng and S. M. Rovnyak, "Decision tree- based methodology for high impedance fault detection," *IEEE Trans. on Power Delivery,* vol. 19, no. 3, pp. 533-536, Apr. 2004.

[11] B. D. Russell and R. P. Chinchali, "A digital signal processing algorithm for detecting arcing fault on power distribution feeders," *IEEE Trans. on Power Delivery,* vol. 4, no. 11, pp. 132-140, Jan. 1989

[12] P. Ferracci, "Ferroresonance," *Group Schneider: Cahier*, vol. 3, no. 190, pp. 1-28, Mar. 1998.

[13] M. R. Iravani, *et. al*., "Modeling and analysis guidelines for slow transients-part III: the study of ferroresonance," *IEEE Trans. on Power Delivery*, vol. 15, no. 1, pp. 255-265, Jan. 2000.

[14] A. R. Sedighi, M. R. Haghifam, O. P. Malik, M. H. Ghassemian, "High impedance fault detection based on wavelet transform and statistical pattern recognition," *IEEE Trans. on Power Delivery,* vol. 20, no. 4, pp. 2414-2421, Oct. 2005.

[15] D. C. Robertson, O. I. Camps, J. S. Mayer, and W. B. Gish, "Wavelets and electromagnetic power system transients," *IEEE Trans. on Power Delivery*, vol. 11, no. 11, pp. 1050-1058, Apr. 1996.

[16] S. Santoso, E. J. Powers, W. M. Grady, and P. Hofman, "Power Quality Assessment Via Wavelet Transform Analysis," *IEEE Trans. on Power Delivery,* vol. 11, no. 11, pp. 924-930, Apr. 1996.

[17] Z. Tongxin, E. B. Makram, and A. A. Girgis, " Power system transient and harmonic studies using wavelet transform," *IEEE Trans. on Power Delivery,* vol. 14, no. 14, pp. 1461–1468, Oct. 1999.

[18] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1998.

[19] J. C. Goswami, *Fundamentals of Wavelets*, John Wiely&Sons, 1999.

[20] *Wavelet toolbox for Matlab, User manual*, MathWorks, Natick, US, 1997.

[21] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Macmillan, New York, 1994.

[22] I. A. Basheer and M. Hajmeer, "Artificial neural networks: fundamentals, computing, design, and application," *Journal of Microbiological Methods*, vol. 43, no. 1, pp. 3-31, 2000.

[23] B. B. Chaudhuri and U. Bhattacharya, "Efficient training and improved performance of multilayer perceptron in pattern classification," *Neurocomputing*, vol. 34, no. 1, pp. 11-27, Sep. 2000.

[24] M. T. Hagan and M. B. Menhaj, "Training feedforward networks with the Marquardt algorithm," *IEEE Trans on. Neural Networks*, vol. 2, no. 3, pp. 989-993, Jun. 1994.

[25] R. Battiti, "First- and second-order methods for learning: between steepest descent and Newton's method," *Neural Computation*, vol. 4, no. 2, pp. 141-166, 1992.

Geev Mokryani was born in Mahabad, Iran on September 18, 1981. He received the B.Sc. and M.Sc. degrees in Control and Power engineering in 2004, 2006, respectively. Currently, he is a faculty member of Electrical Engineering Department at Islamic Azad University, Soofian Branch, Iran. His field of interests and experience includes transients in power systems, power quality, protection of power systems and control systems.

M. R. Haghifam received the B.Sc., M.Sc., and Ph.D. degrees in 1988, 1990, 1995, respectively, all in power engineering. Currently, he is a Professor of Electrical Engineering at Tarbiat Modares University, Tehran, Iran. His main research interests are electric distribution systems and power system reliability and soft computing applications in power systems. He has published more than 200 technical papers in these areas. He is a senior member of IEEE and a research fellow of Alexandervonhumboldt.