Wavelet Transform and Support Vector Machine Approach for Fault Location in Power Transmission Line

V. Malathi, N.S.Marimuthu

Abstract**—**This paper presents a wavelet transform and Support Vector Machine (SVM) based algorithm for estimating fault location on transmission lines. The Discrete wavelet transform (DWT) is used for data pre-processing and this data are used for training and testing SVM. Five types of mother wavelet are used for signal processing to identify a suitable wavelet family that is more appropriate for use in estimating fault location. The results demonstrated the ability of SVM to generalize the situation from the provided patterns and to accurately estimate the location of faults with varying fault resistance.

Keywords**—** Fault location, support vector machine, support vector regression, transmission lines, wavelet transform.

I. INTRODUCTION

ACCURATE fault location on power transmission line is important for both protection and maintenance purposes. Conventional fault location methods use the fault steady state components of voltage and current measured at one or more points along the transmission line. The fault distance can be estimated from the measured impedance of the transmission line at the power system frequency. The impedance is assumed to be proportional to the fault distance. The impedance measurement used in distance protection schemes is too inaccurate for precise fault location as the error in the estimated fault location can be as high as 10% of line length. Fault location based on reactance is a well known technique that has been used to improve the estimation of fault location [1]-[3]. The technique is based on linear relation between the reactance, estimated from the voltage and current of the fault, and the fault location. In most cases, the error in estimating the fault location using these techniques varies between 1% to 6%. The use of travelling waves to detect and locate faults on such line is another feasible alternatives [4]-[5]. The schemes are all based on determining the time needed for a wave to travel between the local end and the fault location. However, travelling wave schemes have problems with faults close to the bus and faults with close-to-zero incidence angle. Algorithms based only on local terminal current and voltage data need some simplifying hypothesis to allow the fault distance calculations [6] affecting the accuracy of the results. To overcome this fault location in transmission line using one terminal post voltage data is proposed in [7]. New tools and algorithms for better system control become a main interest. One of the new tool introduced is wavelet analysis [8]. The wavelet is used to estimate the time taken by the wave to travel between the fault and the local end. Using the travelling time and propagation speed, the fault location is estimated. Another tool that has been used in signal processing applications and introduced recently to the protection field is Prony method [9]. Recent reformation in the power industry such as open access and regulation may have an impact on the reliability and security of power systems. New methodologies for various protection and control schemes are a must to maintain system reliability and security within an acceptable level. Artificial intelligence techniques are among the top candidates to realize this new methodology. One of the AI technique used in fault location is Artificial Neural Network [10]-[14]. Although the neural network based approaches have been quite successful in determining the fault location, the main disadvantage of neural network algorithm is that it requires considerable amount of training effort for good performance.

Methods of locating power system faults introduced so far, can be broadly classified under two categories: one based on the power frequency components, and the other utilizing the higher frequency content of the transient fault signals. The latter method is adopted in this paper in order to locate the fault quickly [8].

In this paper, pre-processing module based on DWT in combination with SVM is used for fault location. Five types of mother wavelets, Daubechies (db5), Biorthogonal (bior5.5), Coiflets (coif5), Symlets (sym5), and reverse biorthogonal (rbio5.5) have been considered for signal processing. The SVM, which is based on statistical learning theory is a general classification method and its theoretical foundation is described in [15]-[16]. The great advantage of SVM approach is the formulation of its learning problem, leading to the quadratic optimization task. It greatly reduces the number of operations in the learning mode and hence SVM algorithm is usually much quicker for large datasets [17]. Because of these advantages of SVM, the proposed method is fast and easy implemental for practical large power systems.

V.Malathi is with Electrical and Electronics Engineering Department, Raja College of Engineering and Technology, Madurai, Tamiladu, India. Phone: +919894068686; e-mail: malathi_triven@yahoo.co.in

N.S.Marimuthu is with Electrical and Electronics Engineering Department, Government College of Engineering, Tirunelveli, Tamilnadu, India.

II. WAVELET TRANSFORM

Wavelet analysis is a relatively new signal processing tool and is applied recently by many researchers in power systems due to its strong capability of time and frequency domain analysis [18].

The definition of continuous wavelet transform (CWT) for a given signal x(t) with respect to mother wavelet Ψ(t) is

$$\text{CWT (a , b)} = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t)\Psi\left(\frac{t-b}{a}\right)dt \qquad (1)$$

where a is the scale factor and b is the translation factor. The Discrete wavelet transform (DWT) can be written as

$$\text{DWT (m,n)} = \frac{1}{\sqrt[m]{a}} \sum_k x(k)\Psi\left(\frac{k - nb_0\, a_0^m}{a_0^m}\right) \qquad (2)$$

where original a and b parameters (1) are changed to be the functions of integers m,n,k which is an integer variable and it refers to a sample number in an input signal. The wavelet transform is useful in analyzing the transient phenomena associated with transmission line faults and /or switching operations. This technique can be used effectively for realizing non-stationary signals comprising of high and low frequency components, through the use of a variable window length of a signal. The ability of the wavelet transform to focus on short time intervals for high frequency components and long time intervals for low frequency components improves the analysis of transient signals. For this reason, wavelet decomposition is ideal for studying transient signals and obtaining better current characterization and a more reliable discrimination.

The wavelet transform technique is first applied, in order to, decompose the different current signals into a series of wavelet components, each of which is a time domain signal, that covers a specific frequency band.

III. SUPPORT VECTOR MACHINE

SVM is a computational learning method based on the statistical learning theory. In SVM, the input vectors are non-linearly mapped into a high dimensional feature space. In this feature space optimal hyper plane is determined to maximize the generalization ability of the classifier.

The motivation for considering binary classifier SVM comes from the theoretical bounds on the generalization error [19]. The main features of SVM are:

(i)The upper bound on the generalization error does not depend on the dimension of the space.

(ii)The error bound is minimized by maximizing the margin γ

Support Vector Regression (SVR) is applied to locate the faults in transmission line.

A. ε - Support vector regression

Let the training data be $\{(x_1, y_1),\ldots(x_l, y_l)\} \subset X \times \Re$, where X denotes the space of input patterns. In ε - SVR , goal is to find a function f(x) that has at most ε deviation from actually obtained targets y_i for all the training data, and at the same time is as flat as possible. In other words, no need to father about errors as long as they are less than ε , but will not accept error larger than this [20].

$$f(x) = \langle w, x \rangle + b \text{ with } w \in x, b \in \Re \qquad (3)$$

where $\langle \cdot,\cdot \rangle$ denotes the dot product in x . A separating hyper-plane which generalises well can be found by solving the following quadratic programming problem

$$\text{Minimize } \frac{1}{2}\|w^2\| \qquad (4)$$

$$\text{Subject to } \begin{cases} y_i - \langle w, x_i \rangle - b & \le \varepsilon \\ \langle w, x_i \rangle + b - y_i & \le \varepsilon \end{cases} \qquad (5)$$

The tacit assumption in (5) was such a function *f* actually exists that approximates all pairs (x_i , y_i) with ε precision, or in other words, that the convex optimization problem is feasible. To cope with the infeasible constraints of the optimization problem one can introduce slack variables ξ_i , ξ_i^*

$$\text{Minimize } \frac{1}{2}\|w^2\| + C \sum_{i=1}^{l} (\xi_i + \xi_i^*) \qquad (6)$$

$$\text{subject to } \begin{cases} y_i - \langle w, x_i \rangle - b & \le & \varepsilon + \xi_i \\ \langle w, x_i \rangle + b - y_i & \le & \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* & \ge & 0 \end{cases}$$

The constant C > 0 determines the trade-off between the flatness of *f* and the amount up to which deviations larger than ε are tolerated. This corresponds to dealing with a so called ε - insensitive loss function $|\xi|_\varepsilon$ described by

$$|\xi|_\varepsilon := \begin{cases} 0 & if|\xi| \le \varepsilon \\ |\xi| - \varepsilon & otherwise \end{cases} \qquad (7)$$

The key idea is to construct a Lagrange function from the objective function and the corresponding constraints, by introducing a dual set of variables. The dual optimization problem is

$$\text{Maximize } \begin{cases} -\frac{1}{2}\sum_{i,j=1}^{l} (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)\langle x_i, x_j \rangle \\ -\varepsilon \sum_{i=1}^{l} (\alpha_i + \alpha_i^*) + \sum y_i(\alpha_i - \alpha_i^*) \end{cases} \qquad (8)$$

subject to $\sum_{i=1}^{l} (\alpha_i - \alpha_i^*) = 0$ and $\alpha_i, \alpha_i^* \in [0, C]$

The support vector expansion w is

$$w = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)x_i \text{ , thus } f(x) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)\langle x_i, x\rangle + b \quad (9)$$

where w can be completely described as a linear combination of the training patterns x_i. Hence the complexity of a function's representation support vectors is independent of the dimensionality of the input space X and depends only on number of support vectors.

IV. System Studied

The single line diagram of a sample system is shown in Fig.1. It consists of two areas connected by transmission line. The transmission line is modelled as a distributed parameter line, representing 200-Km, 240-kV line with positive sequence impedance of, $Z_L(1) = 8.05 + j\,110.66$ ohm and Zero sequence impedance of $Z_L(0)=79.19 + j302.77$ ohm. The Thevenin impedance of area A is $Z_A = 5 + j27.7$ ohm and the area B is $Z_B = 0.6 + j9.3$ ohm. The source voltages are $E_A = 240$KV and $E_B = 240$KV with a frequency of 50 Hz. The sample system is modelled and simulated using MATLAB7.

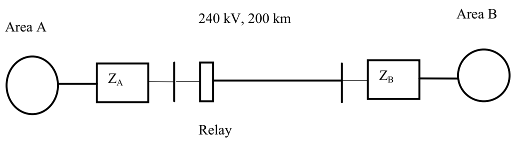


Fig.1 Sample transmission line model

V. Training and Testing for Fault Location Using SVR

The data set is obtained from simulation of a single line to ground fault on a transmission line. The sampling rate employed is 1.6 KHz (32 samples per cycle at 50HZ). The fault simulation has been conducted at a fault impedance of 10Ω and 100Ω. This also includes the fault at different locations of the transmission line as seen by a relay at one end of the transmission line. The location includes 10%, 20%, 30%, 40% and 50% of distance from one end of the line. The analysis is focussed on single line to ground faults, under the assumption that the pre-processing required for all kinds of faults would be similar, based on the preliminary analysis of data. The proposed system is trained with 400 data sets and tested with 640 data sets.

In this paper, the input pattern consists of higher frequency content of transient fault current samples. The data are sampled at 32 samples per cycle. The sampled signals perform as the input to DWT.

The pre-processed signals are used to train SVR. MATLAB SVM toolbox is used to train the network. In this case, the target value of each pattern is the distance from relay locations. The details of simulations carried for generating the training and test patterns are given in Table.1. The criterion for evaluating the performance of the fault locator is defined as

$$\%error = \frac{|SVM\ output - Fault\ location|}{length\ of\ the\ line} * 100 \quad (10)$$

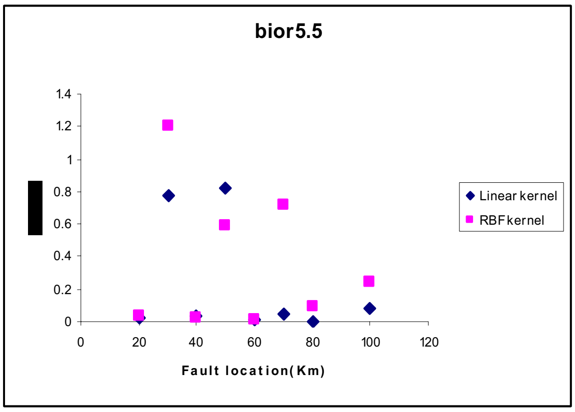
For single line to ground fault, %error for different mother wavelet with linear and RBF kernel are depicted in Tables.2,3. Figs.2-3 show the error in the estimation of the fault location for different mother wavelet with linear and RBF kernel.

TABLE I
TEST PATTERNS GENERATED FOR TRAINING AND TESTING

Sl.No	Patterns generated for training	Patterns generated for testing
1.	10 % , 20% , 30% , 40%, 50% of total transmission line length	15%, 20%, 25%, 30%, 35%, 40%, 50% of total transmission line length

VI. Conclusion

The results proved the ability of SVM to generalize the situation from the available patterns and to accurately locate the faults. The proposed trained networks are capable of providing fast and precise location for different types of mother wavelet. The percentage error for most of fault location is within 1% for all types of wavelet. In the proposed method using SVR, maximum error obtained is less than 1% for different fault location with Daubechies and reverse bi-orthogonal wavelet. Though, all five types of wavelet is giving better results for trained data, only db5 and rbio5.5 wavelet are giving better result for both trained and untrained data.



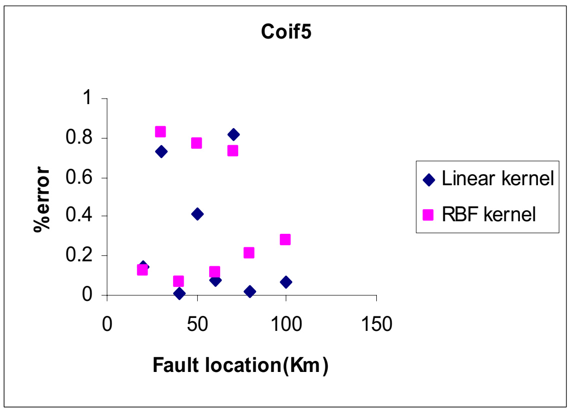


TABLE.II
% ERROR FOR DIFFERENT MOTHER WAVELET WITH L-G FAULT (R_F = 10 Ω)

Fault distance (km)	bior5.5		coif5		sym5		Rbio5.5		db5	
	Linear	RBF	Linear	RBF	Linear	RBF	Linear	RBF	Linear	RBF
20	0.0183	0.0353	0.1456	0.1235	0.0342	0.3324	0.0044	0.0593	0.0651	0.0294
30	0.7792	1.2021	0.7349	0.8234	0.5348	0.7119	0.2460	0.5712	0.4893	0.4956
40	0.0348	0.0254	0.0108	0.0634	0.0111	0.3756	0.0265	0.1087	1.5546e-004	0.1662
50	0.8259	0.5945	0.4105	0.7653	0.0497	0.6869	0.3870	0.0261	0.8005	0.7408
60	0.0126	0.0144	0.0779	0.1146	0.0105	0.1219	0.0288	0.0996	0.0733	0.0884
70	0.0412	0.7205	0.8153	0.7289	0.7468	1.0555	0.5090	0.1412	0.3356	0.8105
80	0.0056	0.0970	0.0235	0.2088	0.0344	0.1602	0.0146	0.0752	0.0234	0.1996
100	0.0784	0.2468	0.0679	0.2790	0.1423	0.2504	0.0158	0.2294	0.0365	0.2720

TABLE.III
% ERROR FOR DIFFERENT MOTHER WAVELET WITH L-G FAULT (R_F = 100 Ω)

Fault distance (km)	% error for bior5.5		% error for coif5		% error for sym5		% error for rbio5.5		% error for db5	
	Linear	RBF	Linear	RBF	Linear	RBF	Linear	RBF	Linear	RBF
20	0.0018	0.0283	0.1273	0.1283	0.0085	0.3242	0.0177	0.0468	0.0515	0.0414
30	0.7945	1.2457	0.7578	0.8847	0.5415	0.9325	0.22	0.6018	0.4993	0.5075
40	0.0497	0.0035	0.0785	0.0696	0.0521	0.3618	0.0357	0.1116	0.0681	0.2021
50	0.8598	0.6147	0.4406	0.6767	0.0311	0.8254	0.4148	0.0052	0.7947	0.7769
60	0.1787	0.0405	0.0708	0.1568	0.0495	0.0717	0.1389	0.0713	0.0736	0.079
70	0.2459	0.8066	0.8756	0.9364	0.628	0.6232	0.3456	0.0156	0.4862	0.7544
80	0.0063	0.1255	0.025	0.1985	0.0349	0.1621	0.0179	0.0989	0.0665	0.1978
100	0.2092	0.2457	0.1557	0.2454	0.2257	0.2179	0.1	0.222	0.1131	0.2597

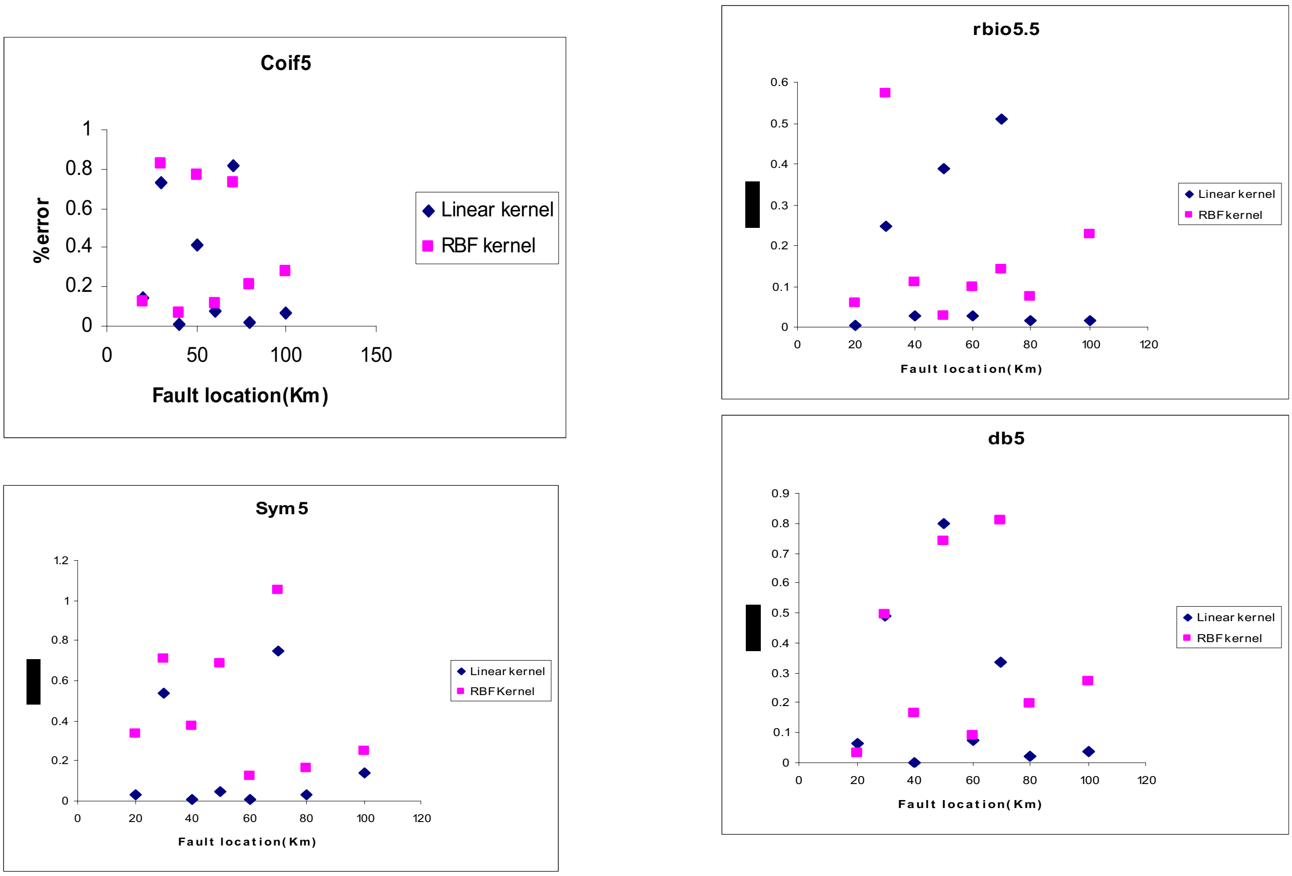
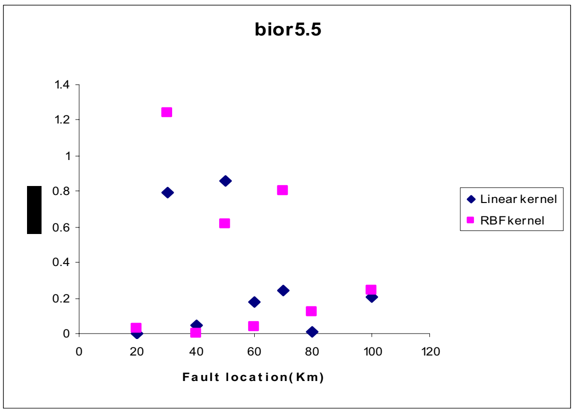
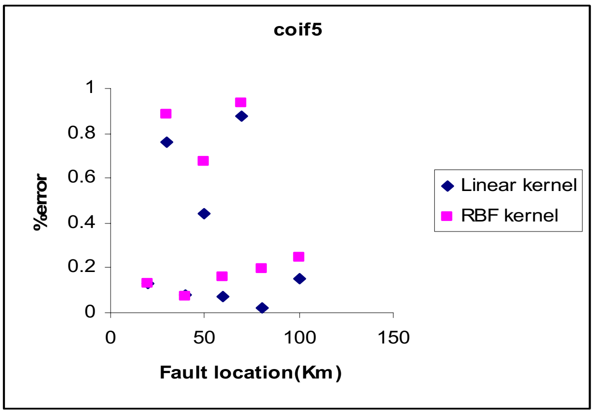
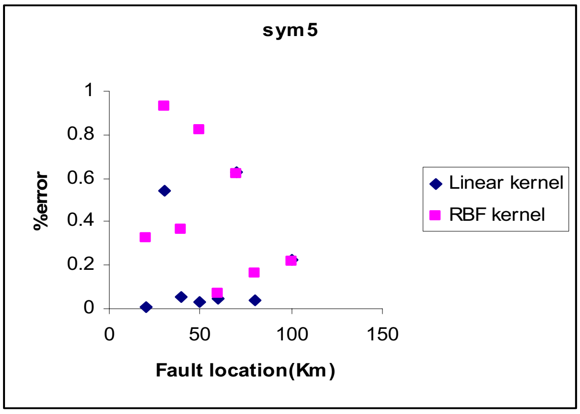
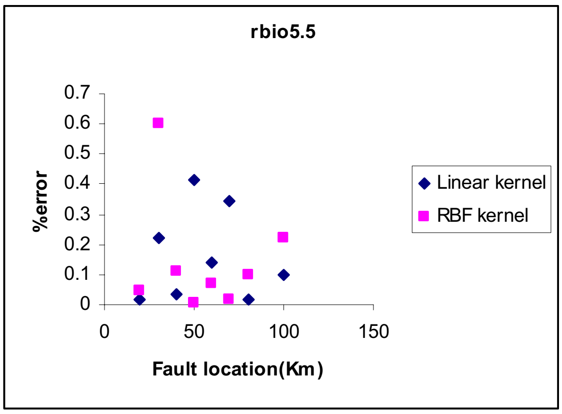


Fig. 2 Test results for different mother wavelet with R_f = 10 Ω









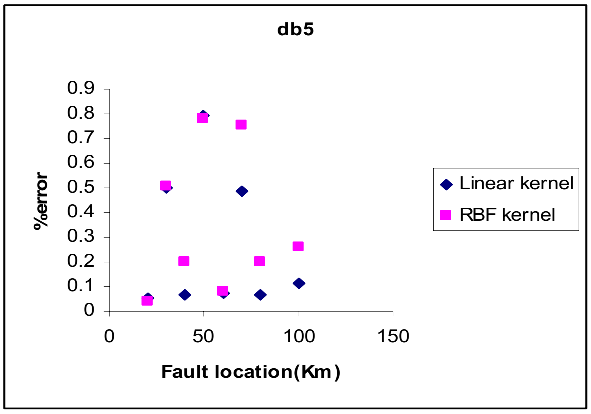


Fig.3 Test results for different mother wavelet with R_f= 100 Ω

REFERENCES

[1] Takagi, Y.Yamaura, R.Kondow, and T.Matsushima, "Development of a new type fault locator using the one terminal voltage and current data", *IEEE Trans. Power Apparatus and systems*, Vol.101, no.8, pp 2892-2898, Aug.1982.

[2] M.S.Sachdev and R.Agarval, "A technique for estimating transmission line fault locations from digital impedance relay measurements", *IEEE Trans. On Power Delivery*, Vol.3 No.1, pp121-129, Jan 1988.

[3] A.A. Girgis and C.M. Fallon, "Fault location techniques for radial and loop transmission system using digital fault recorded data", *IEEE Trans. On Power Delivery*, Vol.9, No.4, pp.1936-1945, oct.1992.

[4] E.Shehab-Eldin and P.Mclaren, "Travelling wave distance protection- problem areas and solutions, *IEEE Trans. on Power apparatus and Systems*", Vol.1, pp-894-802,July 1986.

[5] G.B.Ancell and N.C.Pahalawatha, "Maximum Likelihood estimation of fault location on transmission lines using travelling waves", *IEEE Trans. On Power Delivery*, vol.9 No.2 pp.680-688, April 1994.

[6] M.B.Djuric, Z.M.Redojevic, and V.V.Terzija, "Distance protection and fault location utilizing only phase current phasors", *IEEE Trans. On Power Delivery*, vol.13, pp 1020-1026, October 1998.

[7] Carlos Eduardo de, Morais Pereira and Luiz cera Zanetta, "Fault location in transmission lines using one terminal post fault voltage data", *IEEE Trans. On Power Delivery*, vol.19, No.2, pp 570-575, April 1998.

[8] F.H.Magnago and A.Abur, "Fault location using wavelets", *IEEE Trans. On Power Delivery*, vol.13, No.4, pp 1475-1480, 1998.

[9] O.Chaari, P.Bastard and M.Meunier, "Pronys method: An efficient tool for the analysis of earth fault currents in Peterson coil protected networks", *IEEE Trans. On Power Delivery*, vol.10, No.3, pp 1234-1241, 1995.

[10] Zhihong Chen and Jean- Claud Maun, "Artificial Neural Network Approach to single ended fault locator for transmission lines", *IEEE Trans. On Power Delivery*, vol.15, No.1, pp 370-375, 2000.

[11] M.M.TTawfik and M.M.Morcos, "ANN-Based Techniques for estimating fault location on transmission lines using Prony method", *IEEE Trans. On Power Delivery*, vol.16, No 2, pp 219-224, 2001.

[12] Tahar Bouthiba, "Fault location in EHV transmission lines using artificial Neural Networks", *International Journal of Applied mathematics and computer science*, 2004, Vol.14, No.1, 69-78.

[13] Julio Cesar Stacchini de Souza, A.P.Rodrigues, Marcus Theodor Schilling, and Milton Brown Do Coutto Filho, "Fault location in Electrical power systems using Intelligent system techniques", *IEEE Trans. On Power Delivery*, vol.16, No 1, pp 59-67, 2001.

[14] P.K.Dash, A.K.Pradhan, and G.Panda, "Application of minimal radial basis function neural network to distance protection", *IEEE Trans. on Power Delivery*, Vol.16.No.1, Jan 2001

[15] Vapnik V.N: '*The nature of statistical learning theory*', Springer Verlag,1995.

[16] Vapnik V.N: '*The Statistical Learning Theory*', Wiley,New York,1998.

[17] Ravikumar .B, Thukaram Dhadbanjan, H.P.Khincha, “Fault diagnosis in power transmission systems using Support vector machines”, *International Journal of Emerging Electric Power Systems*, vol.8, no.4, 2007, pp.1-32.
[18] I.Daubechies, “*Ten lectures on wavelets*”, SIAM, Philadelphia, Pennsylvania, 1992.
[19] Cortes.C, Vapnik V.N: “Support Vector Networks”, *Machine learning*, 1995,20(3) pp. 273-297.
[20] Alex.J.Smola and Benhard Scholkopf, “*A tutorial on support vector regression*”, 2003

V Malathi received her Bachelor of Engineering in Electrical and Electronics from College of Engineering Guindy, Anna University, Chennai, and Master of Engineering in Power Systems from Thiagarajar College of Engineering, Madurai, Tamil Nadu, India .

She is working as a Professor in Raja College of Engineering and Technology, Madurai, Tamil Nadu, India. Her area of research is intelligent technique to power system protection.

N.S Marimuthu obtained his B.E and M.Sc (Engg) from Annamalai University,Tamilnadu,India. He obtained his Ph.D. from IIT Kharagpur,India in 1988.

He is currently working as a Professor and Head in Government college of Engineering, Tirunelveli, Tamil Nadu, India. His area of interest is Power system dynamics.