Neurocomputing 73 (2010) 2160–2167

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Intelligent approaches using support vector machine and extreme learning machine for transmission line protection

V. Malathi [a,*], N.S. Marimuthu [b], S. Baskar [c]

[a] Raja College of Engineering and Technology, Madurai 625 020, Tamilnadu, India
[b] National Engineering College, Kovilpatti, Tamilnadu, India
[c] Centre for Advanced Research in Evolutionary Computation (CAREC), Thiagarajar College of Engineering, Madurai, Tamilnadu, India

ARTICLE INFO

Article history:
Received 24 October 2008
Received in revised form
18 December 2009
Accepted 7 February 2010
Communicated by A. Zobaa
Available online 25 February 2010

Keywords:
Extreme learning machine
Fault classification
Fault location
Support vector machine
Transmission line
Wavelet transform

ABSTRACT

This paper proposes two approaches based on wavelet transform-support vector machine (WT-SVM) and wavelet transform-extreme learning machine (WT-ELM) for transmission line protection. These methods uses fault current samples for half cycle from the inception of fault. The features of the line currents are extracted by first level decomposition of the current samples using discrete wavelet transform (DWT) and extracted features are applied as inputs to SVM and ELM for faulted phase detection, fault classification, location and discrimination between fault and switching transient condition. The feasibility of the proposed methods have been tested on a 240-kV, 225-km transmission line for all the 10 types of fault using MATLAB Simulink. Upon testing on 9600 fault cases with varying fault resistance, fault inception angle, fault distance, pre-fault power level, and source impedances, the performance of the proposed methods are quite promising. The performance of the proposed methods is compared in terms of classification accuracy and fault location error. The results indicate that SVM based approach is accurate compared to ELM based approach for fault classification. For fault location, the maximum error is less with SVM than ELM and the mean error of SVM is slightly higher than ELM.

1. Introduction

Transmission line protection is the most elaborate and challenging function in power system protection. Very often, fault classification is an overall protection scheme. This is particularly so in techniques based on a modular approach whereby correct fault discrimination is very much dependent upon accurate fault type classification.

For transmission line protection, it is desirable to have a system of relays to generate a trip signal whenever a fault is detected within its protection zone. The relay system will be modular where fault area estimation is very much dependent upon accurate fault type classification, in which fault type is verified before verifying fault location [1,2].

Fault location estimation is another desirable feature in any protection scheme. Locating the fault on the transmission line accelerates line restoration and reduces the power disruption to consumers. Fault location based on the line reactance calculation is a well known technique that has been used to estimate the fault location [3–5]. The technique is based on linear relation between the reactance, estimated from the voltage and current measured at the relay point during the fault, and the fault location. In most cases, the error in estimating the fault location using these techniques varies between 1% and 6%.

Application of pattern recognition techniques could be employed to discriminate between the healthy and faulty states of the power system. It can also be used for faulty phase detection, fault classification and location. Recently, different attempts have been made using pattern recognition techniques for fault classification and location. Some of the recent papers have used fuzzy logic [6,7] artificial neural network (ANN) [8–11], and support vector machine [12,13] for this purpose. In fuzzy based approach [6], only the nature of the fault has been identified, but the phases involved in the fault have not been determined. In recent years, neural networks have been trained to recognize fault patterns based on the voltage and current waveforms measured at the relaying point. The success of this technique is due to neural network's superior ability to learn and generalize from training patterns. Although the neural-network based approaches have been quite successful, the main disadvantage of ANN is that it requires a considerable amount of training effort for good performance, especially under a wide variation of operating conditions [7]. In [12], section identification and classification of TCSC compensated transmission line is discussed. In [13], section identification of series compensated transmission line is discussed. Support vector machine is a recently used popular

* Corresponding author.
E-mail address: malathi_triven@yahoo.co.in (V. Malathi).

0925-2312/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2010.02.001

method for classification and regression problems [14,15] because of its generalization capability [16]. Extreme learning machine is another recent popular method for classification and regression problems [17,18] because of its universal approximation capability [19].

In this paper, an attempt is made to protect transmission line using WT-SVM and WT-ELM with half cycle data. The proposed system is tested on a 240-kV, 225-km transmission line under variety of fault conditions. The proposed approaches classifies all 10 types of short circuit faults (e.g., a–g, b–g, c–g, a–b, b–c, c–a, a–b–g, b–c–g, c–a–g, a–b–c/a–b–c–g) and locates the fault accurately. These methods are capable of protecting transmission line under wide variations in operating conditions (i.e. fault resistance, fault inception angle, fault distance, source impedance and pre-fault power level) in about half-a-cycle period of fundamental frequency. The performance of the WT-SVM and WT-ELM has been tested over a large data set (9600 test cases) considering wide variation in system operating conditions. The results of WT-SVM based approach is compared in terms of classification accuracy and fault location error with WT-ELM.

2. 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 [20].

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

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

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

$$\mathrm{DWT}(m,n)=\frac{1}{\sqrt{a_0^m}}\sum_k x(k)\Psi\left(\frac{n-k_{b0}a_0^m}{a_0^m}\right) \tag{2}$$

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 choice of mother wavelet plays a major role in the characterization of the signal under study. The mother wavelet, whose characteristics matches closely with the signal under consideration, would be the best choice. For studying power system fault signals, it has been reported in the literature that Daubechies wavelet is the most suitable one [13].

3. Support vector machine

SVM is a computational learning method based on the statistical learning theory. In SVM, the input vectors are nonlinearly 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 [21]. The main features of the 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.

3.1. Support vector classification

The set of training samples

$$X=\{x_1,x_2,x_3,\ldots,x_n\},\quad x_i\in R^M \tag{3}$$

where each training samples x_i has M features describing a particular signature and belongs to one of one of two classes

$$Y=\{y_1,y_2,\ldots,y_n\},\quad y_i\in\{+1,-1\} \tag{4}$$

When data is linearly separable there exists a vector $w\in R^N$ and a scalar $b\in R$ such that $y_i(w\cdot x_i+b)\geq 1$ for all patterns in the training set (i=1,2,…,l). Thus, canonical hyper plane is such that $w\cdot x+b=1$ for closest points on one side and $w\cdot x+b=-1$ for closest points, on other side. The optimal hyper-plane separates points lying on opposite classes yielding the maximum margin of separation. A separating hyper-plane which generalizes well can be found by solving the following quadratic programming problem.

$$\text{Minimize}\quad \frac{1}{2}\|w^2\|+C(\sum_{i=1}^{l}\varepsilon_i)\text{subject to}\quad y_i(w\cdot x_i+b)\geq 1-\varepsilon_i,\varepsilon_{i\geq 0}\ \forall i \tag{5}$$

The constrained optimization problem is solved by constructing a Lagrangian

$$\lambda(w,b,\alpha)=\frac{1}{2}\|w^2\|-\sum_{i=1}^{l}\alpha_i(y_i(w\cdot x_i+b)-1) \tag{6}$$

The Lagrangian has to be minimized with respect to the primal variables w and b and maximized with respect to the dual variable α_i. The Karush–Kuhn–Tucker conditions lead to find the solution vector in terms of the training pattern, $w=\sum_{i=1}^{l}\alpha_i y_i x_i$ for some $\alpha_i\geq 0$. Notice that $\alpha_i\neq 0$ only for a subset of the training patterns, precisely those few vectors that lie on the margin, called the support vectors (SVs). In the case where a linear decision boundary is inappropriate the SVM can map the input vector, x_i, to higher dimensional feature space. Under this conditions, a kernel function $K(.,.)$ can be introduced such that $k(x_i,x_j)=x_i\cdot x_j$. An SVM uses then the convolution of the scalar product to build, in input space, the nonlinear decision function

$$f(x)=\mathrm{sgn}\Big(\sum_{i=1}^{l}\alpha_i y_i k(x,x_i)+b\Big) \tag{7}$$

where x is test vector, b is found from the primal constraints and is computed by $\alpha(y_i(w\cdot x_i+b)-1)=0$, i=1,…,l, such that α_i is not zero and sgn is the signal function.

3.2. Support vector regression (SVR)

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. In ε-SVR, goal is to find a function f(x) that has at most deviation from actually obtained targets y_i for all the training data. An overview of - SVR algorithm