Power Transformer Condition Monitoring and Fault Diagnosis using Support Vector Machine

A. Y. Hatata  A. El-Saeed
Electrical Engineering Dept., Faculty of Eng., Mansoura University, Egypt
a_hatata@yahoo.com

M. S. kandil
Electrical Engineering Dept., Faculty of Eng., Mansoura University, Egypt

Abstract: Differential protection is the best protection technique for power transformers. However, differential relays cannot provide discrimination between internal fault and inrush currents; they always detect the inrush current as an internal fault state. This paper presents a proposed algorithm based on Support Vector Machines (SVM) for monitoring the three phase two windings transformer. The algorithm diagnosis different transformer conditions: normal, over excitation, inrush current, internal and external faults. Both primary and secondary currents and their second order harmonics are used to distinguish between these conditions. The transformer conditions are simulated using PSCAD/EMTDC in order to obtain the primary and secondary current signals. These signals are then used to train and test the SVM using MATLAB/Simulink software. The results are compared with corresponding results obtained by applying Recurrent Neural Network (RNN)-based method. The results prove that the proposed algorithm is accurate and very fast in detecting and classifying different transformer conditions.

Keywords: Support Vector Machine (SVM), transformer monitoring, magnetizing inrush current, internal fault, transformer differential protection.

1. Introduction

Power transformer is one of the vital components of electric power systems. The continuity of the transformer operation is of vital importance in maintaining the reliability of power system. In order to detect faults, high speed, highly sensitive and reliable protective relays are required. For this purpose, differential protection has been employed as the primary protection for most of the power transformers. However, the magnetizing inrush current always exists during energization of power transformers. This inrush current always flows on one side of the transformer, it looks like an internal fault to the differential relay and ends up as spill current and the relay mal-operates. [1]. Distinguishing between inrush current and internal fault current is still a challenging power transformer protection problem. The inrush current typically contains a large second-harmonic component whereas the internal fault consists of the fundamental and small amount of second harmonics. Earlier protection schemes use second harmonic component as the discriminator factor between an inrush and internal fault current [2, 3]. On other hand, inter-turn (turn-to-turn) fault is one of the most important causes of failures occurring in power transformers [4]. Turn-to-turn faults will give rise to a heavy fault current in the short-circuited turns, but changes in the transformer terminal currents will be very small. For that reason, the traditional transformer differential protection is not sensitive enough to detect such winding turn-to-turn faults [5].

It is very difficult to detect these faults since they induce negligible changes in the currents at the transformer terminals, although they give rise to a heavy fault current in the short-circuited turns [5]. In addition, when the transformer tap changer is moved up and down with respect to the middle point at which the relay is adjusted to, the differential relay might initiate a trip signal without the presence of any fault. Such mal-operation of differential relays can affect both the reliability and stability of the whole power system.

To enhance the reliability of differential protection, signals other than current have also been utilized. To enhance the security and dependability of differential protection, signals other than current have also been utilized [6]. A method based on an equivalent circuit composed of inverse inductance has been proposed to recognize internal fault and
inrush current conditions [7]. Another method discriminates internal fault from inrush current by the sum of active power flowing into transformer from each terminal [8]. These methods need to use voltage transformers and require accurate measurements which may increase the protection system cost. Some other methods use second harmonic component as the discriminator factor between an inrush and internal fault current [9-10]. The main drawback of these schemes is the possibility of generation of second harmonic component during faults due to current transformer (CT) saturation. Other techniques detect faults based on waveform fluctuations of differential current. Regards to the fact that the time interval between two respective peaks in inrush current is smaller than the time intervals in the internal fault current [10], delayed fault detection is the main disadvantage of such techniques.

Wave shaped recognition technique was presented in [11]; this technique depends on fixed threshold index and requires large computational burden. In [12], a wavelet-based method has been presented. The method is based on measuring both voltages and currents which may increase the cost of hardware implementation. Alternative improved digital protective methods for accurate and effective discrimination between inrush and internal fault currents have thus to be found.

Recently, Artificial Neural Networks (ANNs) techniques have been applied to power transformer protection to distinguish internal faults from magnetizing inrush currents [13-16]. These classifiers were associated with feature extraction techniques based on either time or frequency domain signals. However, no one of these methods seems to reach a practical level yet, and the second harmonic component filtering technique still to be widely used regardless of its shortcoming.

Support vector machine (SVM) is a computational learning method based on the statistical learning theory. It is originally developed by Cortes and Vapnik [17]; and it has emerged as a potent tool for data analysis. It is a powerful tool for the nonlinear and high dimension problems. It provides a unique solution and is a strongly regularized method appropriate for most ill-posed problems. It can successfully solve the problems of ‘over-fitting’, local optimal solution and low-convergence rate existing in the ANN [18].

SVMs have been widely used for many applications in power system [19-25]. The SVM classified the transformer faults using dissolved gas analysis [26-28] and it is used to identify internal faults as well as inrush current [29-31], however; no one of these papers has applied the harmonic restraint principle.

This paper presents an SVM based algorithm to discriminate between different transformer conditions (normal, over excitation, inrush current, internal and external faults). Various transformer conditions are simulated to obtain the training and testing data. The accuracy and reliability of the proposed algorithm is validated for all case studies.

2. Harmonics restrain

Energization of unloaded power transformers results in magnetizing inrush current very often with high amplitude. The inrush current exists whenever the residual flux does not match the instantaneous value of the steady-state flux which would normally be required for the particular point on the voltage waveform at which the circuit is closed [32]. There are many factors affecting the inrush currents such as the residual flux, saturation flux and the energization voltage angle.

Harmonics restrain is based on the fact that the inrush current has a large second-harmonic component of the differential current which is much larger in the case of inrush than for a fault. Figure 1 shows the simulation of internal fault and inrush currents occurred at 0.1 sec. These harmonics can be used to restrain the relay from tripping during inrush current condition, so it is important to use them to obtain better discrimination between inrush and internal fault currents.

![Figure 1 Second harmonic component of magnetizing inrush and internal fault currents](image)

3. Support Vector Machine

SVM is a relatively new computational learning method based on the statistical learning theory. In SVM, the original input space is mapped into a high-dimensional dot product space called a feature space. In the feature space, the optimal hyperplane is determined to maximize the generalization ability of the classifier. SVMs have
the potential to handle very large feature spaces, because the training of SVM is carried out so that the dimension of classified vectors does not have as a distinct influence on the performance of SVM as it has in the conventional classifier. Also, SVM-based classifiers are claimed to have good generalization properties compared to conventional classifiers, because in training SVM classifier, the so-called structural missclassification risk is to be minimized, whereas traditional classifiers are usually trained so that the empirical risk is minimized. SVMs may have problems with large data sets, but in the development of fault classification routines, these are usually not even available [33].

Let n-dimensional input $X_i (i = 1, 2, \ldots, M)$, $M$ is the number of samples belong to Class-I or Class-II, and associated labels are $Y_i = 1$ for Class-I and $Y_i = -1$ for Class-II respectively. For linearly separable data a hyperplane $f(x) = 0$ which separates the data can be determined as:

$$f(x) = w^T x + b = \sum_{i=1}^{n} w_i x_i + b = 0 \quad (1)$$

Where $w$ is a dimensional vector and $b$ is a scalar. The vector $w$ and the scalar $b$ determine the position of the separating hyperplane. This separating hyperplane satisfies the constraints:

$f(x_i) \geq 1$ if $Y_i = 1$, and $f(x_i) \geq -1$ if $Y_i = -1$ and this results in

$$y_i f(x_i) = y_i (w^T x_i + b) \geq 1 \quad (2)$$

The separating hyperplane that creates the maximum distance between the plane and the nearest data is called the optimal separating hyperplane as shown in the Figure 2. The geometrical margin is found to be $\|w^2\|$ [33].

**Figure 2 Optimal separating planes**

Taking into account the noise with slack variables $\xi_i$ and the error penalty $C$, the optimal hyperplane separating the data can be found by solving the following convex quadratic optimization problem:

**Minimize**

$$\frac{1}{2} \|w\|^2 + C \sum_{i=1}^{M} \xi_i \quad (3)$$

**Subject to**

$$y_i (w^T x_i + b) \geq 1 - \xi_i \quad \text{for} \quad i = 1, 2, \ldots, M \quad (4)$$

$$\xi_i \geq 0 \quad \text{for all values of} \quad i$$

Where $\xi_i$ is the measured distance between the margin and the examples $x_i$ that lying on the wrong side of the margin. The calculations can be simplified by converting the problem with Kuhn–Tucker condition into the equivalent Lagrange dual problem, which will be:

**Maximize**

$$w(\alpha) = \sum_{i=1}^{M} \alpha_i - \frac{1}{2} \sum_{i=1}^{M} \sum_{k=0}^{M} \alpha_i \alpha_k y_i y_k x_i^T x_k \quad (5)$$

**Subject to**

$$\sum_{i=1}^{M} y_i \alpha_i = 0, \quad C \geq \alpha_i \geq 0 \quad \text{for} \quad i = 1, 2, \ldots, M \quad (6)$$

The number of variables of the dual problem is the number of training data. Let us denote the optimal solution of the dual problem with $w^*$ and $\alpha^*$, the equality conditions in (2) holds for the training input-output pair $(x_i, y_i)$ only if the associated $\alpha_i^* \neq 0$. In this case, the training example $x_i$ is a support vector (SV). The number of SVs is considerably lower than the number of training samples, making SVM computationally very efficient.

The value of the optimal bias $b^*$ is found from the geometry:

$$b^* = -\frac{1}{2} \sum_{sv} y_i \alpha_i (S_i^T x_i + S_i^T x_i) \quad (7)$$

Where $S_1$ and $S_2$ are arbitrary SVs for Class-I and Class-II, respectively. Only the samples associated with SVs are summed, because the other elements of the optimal Lagrange multiplier $\alpha^*$ are equal to zero.

The final decision function is given by:

$$f(x) = \sum_{sv} y_i \alpha_i x_i^T X + b^* \quad (8)$$

The unknown data sample $x$ is then classified as:

$$x \in \{ \text{class } - I, \quad \text{if} \ f(x) \geq 0 \}
\{ \text{class } - II, \quad \text{otherwise} \} \quad (9)$$

The classified data is mapped into a high-dimensional feature space where the linear classification is possible. Using nonlinear vector function:

$$\Phi(x) = \Phi_1(x), \Phi_2(x), \ldots, \Phi_m(x), \quad m \gg n \quad (10)$$

To map the n-dimensional input vector $x$ into the m-dimensional feature space, the linear decision function in dual form is given by:

$$f(x) = \sum_{sv} y_i \alpha_i \Phi^T(x_i) \Phi(x) \quad (11)$$

Notice that in (11) as well as in (4), the inner products are used. The nonlinear classification problems can also be solved by using SVM applying...
a kernel function. This function returns with a dot product of the feature space mapping of the original data points is called a kernel function,

\[ K(x, z) = \phi^T(x)\phi(z) \].

The learning in the feature space does not require the inner products where a kernel function is applied. Using kernel function, the decision function can be written as:

\[ f(x) = \sum_{i=1}^{N} y_i \alpha_i k(x_i, x) \] (12)

There are different kernel functions used in the literature. Mercer’s theorem states that any symmetric positive-definite matrix can be regard as a kernel matrix. In this paper, Gaussian radial basis kernel function is selected. The radial basis kernel function is defined as:

\[ K(x, z) = e^{-\frac{||x-z||^2}{\gamma}} \] (13)

\[ \gamma = \frac{1}{2\sigma^2} \]

Where \( \sigma \) is the width of the Gaussian function [34, 35]

4. Applying SVM for Discrimination between Transformer Conditions

In this section, SVMs classification technique is generalized to discriminate between different transformer conditions.

4.1 SVM classifier structure

As shown in Figure 3, the diagnostic model includes four SVM classifiers. These classifiers are used to identify the normal and faulty states (internal, over excitation, external and inrush).

The first SVM1 is trained to identify the normal state. When the input pattern of SVM1 represents a normal state, the output is 0; otherwise it will be +1. The second SVM2 is trained to identify the internal fault. When the input pattern of SVM2 represents internal fault, the output is +1; otherwise it will be +2. The third SVM3 is trained to identify the saturation state. When the input pattern of SVM3 represents the saturation fault, the output is +2; otherwise it will be +3. The fourth SVM4 is trained to identify external fault. The output is +3 for external fault; otherwise it will be +4. All the four SVMs adopt Gaussian as their kernel function. In the studied SVM, the parameters \( \sigma \) and C of SVMs model are optimized by the cross validation method. The adjusted parameters with maximal classification accuracy are selected as the most appropriate parameters. Then, the optimal parameters are utilized to train the SVM model. The output codification is presented in Table 1.

<table>
<thead>
<tr>
<th>Transformer conditions</th>
<th>SVM1</th>
<th>SVM2</th>
<th>SVM3</th>
<th>SVM4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal operation</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Internal fault</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Over excitation</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>External fault</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Inrush current</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 1 The output of SVM.

Figure 3 Diagnostic model of power transformer based on SVM classifiers.

4.2 Training and Testing Data

Consecutive samples of phase currents are usually chosen as the input to SVM classifier. In modern power transformers, the harmonic components in inrush currents are very low. In order to increase the security and dependability of the
SVM classifier for these transformers, harmonic sharing method is used to determine the harmonic input to the SVM. In this method the second harmonic is obtained by dividing the summation of the second harmonics of the three phases by the summation of the fundamental currents of these phases.

\[ I_{2nd} (pu) = \frac{I_{2ndA} + I_{2ndB} + I_{2ndC}}{I_{fundA} + I_{fundB} + I_{fundC}} \]  

(14)

The appropriate input data window length is a major factor which should be considered. Long data window of inputs enables protective algorithms to get more information and in turn resulting in stable performance. Through a series studies, it was found that a short window of length of 5 samples of each current at sample rate 1 kHz for a 50 Hz power frequency have a good results. So, each of these currents is represented by 5 samples and second harmonic current, resulting in 36 inputs for all patterns.

5. Simulation and Results

The studied power system consists of a three phase source connected to a load through a three phase power transformer 110/10.5 kV, 100 MVA, as shown in Figure 4. The transformer has a star-star-ground connection. The data required for training and testing the SVMs are developed by simulating the power system using the PSCAD/EMTDC software package. Power transformer internal faults are implemented by PSCAD/EMTDC program as shown in Figure 4. The necessary data required to generalize the problem are obtained from this simulation.

Different types of internal winding faults are simulated at different percentage of windings, and inception time. The CT ratios are chosen as 1257:1 and 120:1 for secondary and primary sides of the power transformer respectively. The prescribed five operating conditions are studied.

![Figure 4 Single line diagram](image)

Many cases have been implemented on the studied system at the prescribed transformer conditions. Total number of patterns used for SVMs training and testing is equal to 1152 patterns. These data are divided into two data sets: the training data set (576 samples) and the testing data set (576 samples). The training data sets have been provided as: normal operation (54 samples), internal fault (378 samples), overexcitation (54 samples), external fault (32 samples) and inrush current overheating (58 samples) as illustrated in Table 2.

<table>
<thead>
<tr>
<th>Transformer conditions</th>
<th>No of pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal operation</td>
<td>108</td>
</tr>
<tr>
<td>Internal fault</td>
<td>378</td>
</tr>
<tr>
<td>overexcitation</td>
<td>108</td>
</tr>
<tr>
<td>External fault</td>
<td>64</td>
</tr>
<tr>
<td>Inrush current</td>
<td>116</td>
</tr>
</tbody>
</table>

Table 2. Input patterns of SVM.

The simulations of the proposed SVMs classifier have been performed through MATLAB environment using Bioinformatics toolbox.

The performances of the four SVMs are assessed for each of these values by calculating the classification accuracy which is defined by [34]:

\[
\%\text{Classification accuracy of SVM} = \frac{\text{Total patterns} - \text{Incorrect patterns}}{\text{Total patterns}} \times 100
\]  

(15)

The highest training efficiency obtained with the kernel function is 99.83% for SVM. The best accuracy for SVM has been achieved with the values of $C = 10$, and $\gamma = 0.1$. These parameters are used for “learning” the SVMs. Once the training is completed, the trained SVMs are used for testing the new patterns. Table 3 illustrates the classification accuracy of the designed SVMs for testing data.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Training</th>
<th>Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of pattern</td>
<td>576</td>
<td>576</td>
</tr>
<tr>
<td>Correct pattern</td>
<td>576</td>
<td>575</td>
</tr>
<tr>
<td>Incorrect pattern</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Classification accuracy</td>
<td>100</td>
<td>99.83</td>
</tr>
</tbody>
</table>

Table 3. Over all test results

To help judging the value and the accuracy of the proposed system, the SVMs results have compared with those obtained using recurrent neural network (RNN) based system. In that study the RNN-based method was applied to detect and classify different conditions of power transformer. The same training and testing data described above have been used to train the RNN-based system. A comparison between the two methods is depicted in Tables 5 These results...
show that the proposed SVMs is faster and more accurate (both for training and testing patterns) than RNN-based method as it is clear from Table 5.

Table 5: Comparison between accuracy and training time for RNN and SVMs methods

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy for training patterns %</th>
<th>Accuracy for testing patterns %</th>
<th>Training time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>100</td>
<td>99.83</td>
<td>5</td>
</tr>
<tr>
<td>RNN</td>
<td>100</td>
<td>98</td>
<td>1800</td>
</tr>
</tbody>
</table>

6. Conclusion

This paper presents a novel technique to classify different power transformer conditions such as inrush, saturation, internal and external fault conditions based on SVM technique. Both primary and secondary currents and their second order harmonics are used to distinguish between different transformer conditions. The use of second harmonic current as an addition input helps to improve the ability of the SVM to discriminate between transformer conditions and to increase the security and dependability of the technique.

The proposed SVM classification technique has been proven to be highly reliable and very fast in detecting and classifying transformer conditions with accuracy of 99.83% average for all the test cases.

Comparing the proposed SVM with the RNN-based method has proven that the proposed SVM-based method is faster and more accurate. Applying the proposed method to several case studies has shown that the SVM based classifier has consistently accurate detection and discrimination in all operating conditions.

7. Reference