CHAPTER 5

Transformer Switching Detection
5. TRANSFORMER SWITCHING DETECTION

Introduction

Power system fault is defined as an abnormal event which needs to be detected and cleared quickly to maintain stability and continuity of power supply. The detection of the fault should be accurate enough to distinguish between the healthy and abnormal condition; failure to which may cause unnecessary tripping. Conventional methods of fault detection usually employ electromagnetic relays which are generally of overcurrent type. These relays are sluggish and often prone to false tripping. Also in most of the conventional techniques the fault detection is based on measurement of power frequency voltages and currents; and the high frequency signals are discarded, which contains more information of fault. To overcome the difficulties of conventional techniques some attempts based on digital calculations and high frequency signals are reported in [42-45]. Unfortunately, these approaches do not have ability to adapt dynamically to changes in the system operating condition.

Conventional power frequency methods sometimes give inaccurate decision; transformer switching is one such event where chances of inaccurate decisions are more. When a transformer is first connected to a source of voltage, there is transient inrush of magnetizing current, which may be as great as ten times of full load current and may decay with a time constant as long as 2 seconds [46]. The magnitude of inrush current depends on the instant of switching. This inrush current may be seen as some fault situation by relay. Also sometimes, the sudden switching of load at some buses may cause difficulty to relay in discriminating between normal and abnormal situation. Existing solution to these problems are use of time delay relay or harmonic
restraint to block the operation of relay. These methods lead to delayed
decision and no protection during the blocked period.

The ANNs are best suited for pattern recognition task and they have
inherent adaptivity. Distinction of transformer switching from fault situation
can be achieved with ANN, if the voltage and current signals are presented in
the form of patterns. This chapter presents the use of feedforward ANN with
feedback in the output layer to distinguish among three events viz. transformer
switching, single line to ground fault in phase-a and load switching. The
method to be discussed in this chapter uses high frequency signals with some
preprocessing. Distinction among the three events is also attempted with
Elman Neural Network using low as well as high frequency signals without
any preprocessing. Both of these networks have been trained using
backpropagation learning algorithm.

5.1 Switching Detection with Preprocessing of Input

5.1.1 System Model Used for Simulation: A typical system with one
generator bus (A) and one load bus (B) (see Fig. 5.1) has been used for
simulating three switching events mentioned earlier with MATLAB-
SIMULINK software. System data used are as under:

**Generator:** Modeled as constant inductive voltage source.
Capacity = 6 MVA. Line Voltage = 25 kV. Resistance = 0.0625 Ohms.
Inductance = 1.675 mH.

**Three Phase Line:** Modeled as Pi section of 100 km
Positive Sequence Resistance = 0.01273 Ohm/km. Positive Sequence
Inductance = 0.9337 mH/km. Positive Sequence Capacitance = 0.01274
µF/km. Zero Sequence Resistance = 0.3864 Ohm/km. Zero Sequence
Inductance = 4.126 mH/km. Zero Sequence Capacitance = 0.007751 µF/km.

**Transformer:** Modeled as Y-Y connected and two winding type.
Capacity = 6 MVA.
Magnetizing Characteristic (B-H) = (0.0-0.0, 1.0-0.0012, 1.52-1.0) pu
Residual Flux: a-phase = 0.8 pu, b-phase = -0.4 pu, c-phase = -0.4 pu.

Shunt Branch Resistance = 500 pu

Winding-1, Solidly Grounded Star:
Line Voltage = 25 kV, Resistance = 0.002 pu, Leakage Reactance = 0.08 pu.

Winding-2, Solidly Grounded Star:
Line Voltage = 0.4 kV, Resistance = 0.002 pu, Leakage Reactance = 0.30 pu.

**Three Phase Load**: Modeled as resistive elements

Load Switched Without Transformer:
Capacity = 6 MVA, Line Voltage = 25 kV.

Load Connected to Secondary of the Transformer:
Capacity = 3 MVA, Line Voltage = 0.4 kV.

**Fig. 5.1**, Single Line Diagram of Power System Used for Simulation

**Fig. 5.2**, Current for SLG Fault
The current and voltage waveforms generated by software simulation for the three events under consideration at positive excursion zero crossing are shown in Fig. 5.2, Fig. 5.3 and Fig. 5.4. Similar waveforms were generated for different switching instants spaced at 0.6 millisecond up to quarter cycle of power frequency for making training and testing data sets. Observation of these waveforms indicates that the power frequency current magnitude ranges from 1.0 pu to 4.0 pu. Overcurrent relaying depends on current magnitude, which is often dependent on system condition, fault location and fault inception time. ANN based technique uses the features of these waveforms to arrive at decision and hence is more reliable. Learning method used for detection of switching event is discussed in next section.
5.1.2 Backpropagation Learning Algorithm

This learning rule is an extension of the delta rule for multilayer feedforward network and is known as the generalized delta rule. The objective of this rule is to calculate the optimum set of weights with gradient descent along the error surface. The error is defined as the squared difference between the desired output and the actual output that is produced at the output layer on corresponding input presentation. The output is calculated using the initial setting of weights in all the layers. The optimum weights may be obtained if the weights are updated in such a way that the gradient descent is made along the total error surface.

Backpropagation learning rule is derived in this section for one hidden layer which can be extended for any number of layers. Let \((a_p, d_p), p = 1, 2, ..., P\) is the set of training pattern pairs to determine weight update for each presentation of an input-output pair. The given data may be used several times therefore let us use index \(m\) to indicate presentation step.

![Diagram of a single hidden layer feedforward neural network](image)

Fig. 5.5, A Single Hidden Layer Feedforward Neural Network
For training a multilayer feedforward neural network following estimate of gradient descent along the error surface is used.

\[ \Delta w_{ji} (m) = -\eta \frac{\partial E(m)}{\partial w_{ji}}, \quad (5.1) \]

Where \( \eta > 0 \) is a learning rate parameter and \( \Delta w_{ji} \) is the change in weight connecting the unit \( j \) and \( i \). The weight update equation is given by:

\[ w_{ji} (m + 1) = w_{ji} (m) + \Delta w_{ji} (m) \quad (5.2) \]

Hence the backpropagation rule consists of finding the expression for \( \Delta w_{ji} \) for the connections at different layers. Let a multilayer feedforward network with one hidden layer is as shown in Fig. 5.5. This network has \( I \) linear input units indexed by \( i \), \( J \) nonlinear units in hidden layer indexed by \( j \) and \( K \) nonlinear units in the output layer indexed by \( k \).

Let \([a (m), d (m)]\) be the current samples of the function mapping the input space to output space. Let \( y (m) \) is the actual output of the network for the input \( a (m) \) at the step \( m \). The mean squared error at the \( m^{th} \) step is given by:

\[ E(m) = \frac{1}{2} \sum_{k=1}^{K} [d_k (m) - y_k (m)]^2 \quad (5.3) \]

\[ E(m) = \frac{1}{2} \sum_{k=1}^{K} [d_k (m) - s_k^o]^2 \quad (5.4) \]

Where,

\[ x_k^o = \sum_{j=1}^{J} w_{kj} s_j^h \quad (5.5) \]

\[ s_j^h = f_j^h (x_j^h) \quad (5.6) \]
\[ x_j^h = \sum_{i=1}^{1} w_{ji}^h s_i \quad (5.7) \]

\[ s_i = x_i = a_i \text{ (m)} \quad (5.8) \]

The superscript 'o' refers to the output unit quantities, the superscript 'h' refers to the hidden unit quantities and \( a_i, x_i \) and \( s_i \) refer to the input, activation and output values for the unit \( i \), respectively. For the weights leading to the units in the output layer,

\[ \Delta w_{kj} (m) = - \eta \frac{\partial E (m)}{\partial w_{kj}} \quad (5.9) \]

\[ \frac{\partial E (m)}{\partial w_{kj}} = \frac{1}{2} \sum_{j=1}^{J} \left[ d_k - f_k^o \left( \sum w_{kj} s_j^h \right) \right]^2 / \partial w_{kj} \quad (5.10) \]

\[ \frac{\partial E (m)}{\partial w_{kj}} = - \left( d_k - f_k^o \right) \left( \frac{df_k^o}{dx_k^o} \right) s_j^h \quad (5.11) \]

\[ \frac{\partial E (m)}{\partial w_{kj}} = - \delta_k^o s_j^h \quad (5.12) \]

Where \( \delta_k^o = (d_k - f_k^o) \left( \frac{df_k^o}{dx_k^o} \right) \). In above equations iteration index \( m \) is omitted for convenience.

Therefore

\[ \Delta w_{kj} (m) = \eta \delta_k^o s_j^h \quad (5.13) \]

and

\[ w_{kj} (m + 1) = w_{kj} (m) + \Delta w_{kj} (m) \quad (5.14) \]
\[ w_{kj} (m + 1) = w_{kj} (m) + \eta \delta_k s_j \]  (5.15)

For the weights leading to the units in the hidden layer.

\[ \Delta w_{j}^h (m) = - \eta \frac{\partial E (m)}{\partial w_j^h} \]  (5.16)

\[ \frac{\partial E (m)}{\partial w_j^h} = - \sum_{k=1}^{K} (d_k - f_k^o) \left( \frac{\partial f_k^o}{\partial w_{kj}^h} \right) \]  (5.17)

\[ \frac{\partial E (m)}{\partial w_j^h} = - \sum_{k=1}^{K} (d_k - f_k^o) \left( \frac{\partial f_k^o}{\partial w_{kj}^h} \right) \]  (5.18)

Since \( s_j^h = f_j^h (x_j^h) \), we get

\[ \frac{\partial s_j^h}{\partial w_j^h} = (\frac{\partial f_j^h}{\partial x_j^h}) (\frac{\partial x_j^h}{\partial w_j^h}) \]  (5.19)

Since \( x_j^h = \sum_{i=1}^{I} w_{ji}^h s_i \), we get

\[ \frac{\partial x_j^h}{\partial w_j^h} = s_i \]  (5.20)

Therefore.

\[ \frac{\partial E (m)}{\partial w_j^h} = - \sum_{k=1}^{K} (d_k - f_k^o) \left( \frac{\partial f_k^o}{\partial w_{kj}^h} \right) w_{kj} (\frac{\partial f_j^h}{\partial x_j^h}) s_i \]  (5.21)

\[ \frac{\partial E (m)}{\partial w_j^h} = - \delta_j^h s_i \]  (5.22)

Where,

\[ \delta_j^h = (\frac{\partial f_j^h}{\partial x_j^h}) \sum_{k=1}^{K} w_{kj} \delta_k^o \]  (5.23)
Fig. 5.6: Flow of Signal and Error in Backpropagation Learning
Hence,

\[ \Delta w_{ji}^h (m) = \eta \delta_j^h s_i = \eta \delta_j^h a_i (m) \]  \hspace{1cm} (5.24)

Since \( s_i = x_i = a_i (m) \). Therefore,

\[ w_{ji}^h (m + 1) = w_{ji}^h (m) + \Delta w_{ji}^h (m) \]  \hspace{1cm} (5.25)

\[ w_{ji}^h (m + 1) = w_{ji}^h (m) + \eta \delta_j^h a_i (m) \]  \hspace{1cm} (5.26)

where \( \delta_j^h \) represents the error propagated back to the output of hidden units from the next layer that is why this method of weight update is known as backpropagation learning algorithm. Fig. 5.6 shows the details of signal and error flow diagram for a simple case of three linear units in the input layer, three nonlinear units in the hidden layer and two nonlinear units in the output layer. Dot over activation function in the figure indicates its derivative [1, 32-34].

5.1.3 Pattern Formation

Switching currents of Fig. 5.2, Fig. 5.3 and Fig. 5.4 are not convenient as input to ANN therefore preprocessing is needed. These currents have been passed through A/D converter with sampling frequency of 10 kHz. DC components and low frequency currents are rejected with FIR filter of pass band 750 Hz-2000Hz. The filtered waveforms are shown in Fig. 5.7, Fig. 5.8 and Fig. 5.9. These waveforms have different feature from each other, which forms the basis for their classification. The power spectral densities have been evaluated to extract the features of three events. Spectral energy distribution is shown in Fig. 5.10, Fig. 5.11 and Fig. 5.12. Power spectrum reveals that whole of the information is within 2000 Hz, as higher frequencies have already been rejected. Spectral density for fault is distinct from the other two. Spectral
energy for load switching and transformer switching are almost similar, except some energy near 1000 Hz for later event, this forms the basis of differentiation between them. For feature extraction frequencies up to 2325 Hz have been considered and this is subdivided in four parts to ease the training burden of ANN [12]. For spectral energy calculation, 128 samples of filtered current have been included and such data windows at time step of 1 ms are presented to ANN after decomposing and normalizing to neuronal range. With this data the ANN relay can be shown as in Fig. 5.13. Three prefault and three postfault data window of ten switching instants for each event forms the 180 patterns for training. Target for these patterns have been taken as ‘−1’ to indicate absence and ‘+1’ for occurrence of the event in the respective output element.

Fig. 5.7, Filtered Current for SLG Fault

Fig. 5.8, Filtered Current for Transformer Switching
Fig. 5.9, Filtered Current for Load Switching

Fig. 5.10, Power Spectrum for SLG Fault

Fig. 5.11, Power Spectrum for Transformer Switching

Fig. 5.12, Power Spectrum for Load Switching
5.1.4 Training and Testing of ANN Relay

(a) Training

For classification of three switching events, ANN with 4 inputs and 12 neurons in hidden layer and 3 outputs has been chosen. Four inputs correspond to selected frequency decomposition and three outputs indicate occurrence of three events. This configuration with five delayed self-feedback in output neurons was found suitable to accommodate noisy situations. As ‘+1’ and ‘−1’ has been chosen as the threshold limits of target, hence tangent hyperbolic transfer function will be suitable in hidden layer and in output layer. The ANN was fed with preprocessed current samples as shown in Fig. 5.13 (where, ‘f’ is fault, ‘x’ is transformer and ‘l’ is load switching outputs). On the basis of errors between the calculated output and target, training has been carried out using already discussed backpropagation algorithm. For mean square error criterion of 0.00001, convergence occurred in 275-300 epochs. Fig. 5.14 shows all the component of the ANN relay thus obtained.

(b) Testing

ANN relay trained with above patterns has been tested for samples, which were not included in the training phase and results indicate correct classification. Table 5-1 shows the comparison between expected output and
actually obtained outputs. If the criterion for correct classification is taken as 80% of threshold value then the results are well within this range. The correct classifications were obtained with change in fault /switching inception time. On presentation of data window with combination of normal condition and switching event, the ANN relay gives indication of switching merely by inclusion of 20 samples of fault wave in normal condition data. This indicates that it is able to identify the spectral energy of 20 samples of fault. Same results were found for transformer switching. For load switching there is a requirement of 70 samples, which indicates poor feature learning for this event as this event is analogous to normal condition. However it is evident that it will give correct classification for all the cases when sufficient samples are included. This shows the ability of ANN relays to function in noisy situation also. As there are 128 samples in each data window and there is separation of 1 ms between them the operating time for the relay will be 13.8 ms (excluding the simulation time) which corresponds to 0.69 cycle of power frequency. The operating time can be reduced by increasing the sampling frequency, reducing the width of data window and improvement in preprocessing.

![ANN Relay Diagram](Fig. 5.14)
<table>
<thead>
<tr>
<th>Switching Event</th>
<th>Desired Output</th>
<th>Actual Output</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>f</td>
<td>x</td>
</tr>
<tr>
<td>LG Fault in Phase-a</td>
<td>+1</td>
<td>-1</td>
</tr>
<tr>
<td>Transformer Switching</td>
<td>-1</td>
<td>+1</td>
</tr>
<tr>
<td>Load Switching</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>Normal Condition</td>
<td>-1</td>
<td>-1</td>
</tr>
</tbody>
</table>

5.2 Switching Detection without Preprocessing of Input

In section 5.1 there is a preprocessing of the input data for feature extraction which calls for extra computational burden. Also because of preprocessing some extra time is needed which can not be afforded in many situations. In this section alternate method which uses Elman neural network is proposed. Elman network is a feedback type of network and is very good in learning features in temporal sequences. Details of this network are presented in Chapter-6. The Elman neural network gives the classification among three discussed switching events without the need of preprocessing because it learns the relationship of a pattern with neighbouring patterns and also learns features within the patterns itself.

Similar current and voltage waveforms were generated for this case also for the same power system model. For making the training and testing set these waveforms were generated for various instants spaced at 1 ms up to half cycle of power frequency. Here the voltages and currents of phase-a are considered for input.

5.2.1 Training and Testing of ANN Relay

(a) Training

ANN based relay with two input, three output and thirty units in hidden layer is proposed for three switching events (see Fig. 5.15). As all the
information regarding different switching events is contained in voltage and current signals, therefore patterns with these signals have been formed for training purpose [12]. To prepare training set, voltage and current signals of phase-a for switching at 40, 42, 44, 46, 48 and 50 ms have been used with sampling frequency of 10 kHz. Hence these signals contain information up to 5 kHz and information higher than this frequency is discarded. Sequence of data window with 20 samples in each window in the time step of 1 ms has been formed for voltage and current to constitute two input of the ANN relay. Targets for the three outputs f (ground fault in phase-a), x (transformer switching) and l (load switching) have been set to +1.0 to indicate the presence of switching and −1.0 to indicate the absence.

![Fig. 5.15, ANN Used for Switching Detection](image)

Input and target sequences thus prepared for six instants over half cycle of power frequency were presented to the relay for training, to learn the pattern features. To meet the mean square error criterion of 0.001 the network required 14235 epochs.

**(b) Testing**

ANN relay trained with input output sets as discussed in section-5.2.1(a) was tested for the three switching events. Testing data window was prepared with the voltage and current signals for switching at 43, 45 and 47ms. These switching instants were not included during training. Fig. 5.16, Fig.5.17 and
Fig. 5.18 show the three outputs of ANN relay vs. data window for different switching events. Results show that the relay gives correct output almost within 2 to 3 data window presentation i.e. 4 to 5 ms, this shows the fastness of detection. Also correct outputs are obtained for patterns which are not included in training, which shows the generalization capability of the relay.

Fig. 5.16, Testing Results for Switching at 43 ms

Fig. 5.17, Testing Results for Switching at 45 ms
5.3 Conclusion

This chapter has presented ANN relays using feedforward network with feedback in the output layer and Elman Network to solve some conflicting problems of power system, which are not solved satisfactorily by existing methods. Almost all the existing methods are based on power frequency signals only; on the other hand the method suggested here is based on power frequency as well as high frequency which contain more information of faults. Also with the use of Elman network there is no need of preprocessing of input signal for feature extraction which saves processing time and facilitates the fast detection of switching events. Unlike conventional technique ANN method has ability to classify the healthy and unhealthy situations of power system even for the noisy or partial signals. This chapter has shown successful classification among LG fault in phase-a, transformer switching and load switching. It can be extended to include all the phases and all possible faults with increased training data and complex ANN architecture.

**********