Detection and Classification of High Impedance Faults in Power Distribution Networks Using ART Neural Networks

I. Nikoofekr*, M. Sarlak**, and S.M. Shahrtash***
*Ferdowsi University of Mashhad, Iman.Nikoofekr@stu.um.ac.ir
**Jondi Shapour University, Sarlak@iust.ac.ir
***Center of Excellence for Power System Automation and Operation, Iran University of Science and Technology, Shahrtash@iust.ac.ir

Abstract: Adaptive Resonance Theory (ART) neural networks have several interesting properties that make them useful in the area of pattern recognition. Many different types of ART-networks have been developed to improve clustering capabilities. In this paper, five types of ART neural networks (ART1, ART2, ART2-A, Fuzzy ART and Fuzzy ARTMAP) are applied to detect and classify high impedance faults (HIF) in distribution networks. The features are extracted by applying TT-transform to one cycle of fault current signal. These features include energy, standard deviation and median absolute deviation. Then, they are applied to ART neural networks to detect and classify high impedance fault with broken conductor on gravel, asphalt and concrete, unbroken conductor on tree and also no fault condition. Finally, the results of these ART neural networks are compared with each other.

Keywords: Distribution Network Protection; High Impedance Fault; Pattern Recognition; TT-transform; ART Neural Network.

1. Introduction

High impedance faults (HIF) often occur when an overhead conductor breaks down and touches high impedance surfaces (broken conductor) such as asphalt, gravel or where the conductor become in contact with a high impedance object such as tree (unbroken conductor). These faults exhibit low current and for this reason are not detected by conventional protection in distribution networks. The main purpose of high impedance fault detection is protecting human lives and preventing fires.

The classification of high impedance faults becomes important at least as a matter of maintenance management and protection. It is obvious that the repair and maintenance procedures are different for a conductor which is broken and/or whenever trees are become in contact with it. In the case of unbroken conductor an alarm and indication to send the repairmen for cutting tree branches is enough, whereas for a broken conductor sending the trip signal to circuit breaker is necessary.

Several techniques for HIF detection have been reported so far. These are harmonics component methods based on the low-frequency energy components [1], Kalman filtering approach [2], low order current harmonics ratio [3], high-frequency harmonics of the feeder neutral line current [4], discrete wavelet transform [5], and continuous Morlet wavelet transform [6].

In this paper, a new pattern recognition algorithm (Fig. 1) based on adaptive resonance theory (ART) is presented to detect and classify HIFs in distribution networks. This algorithm is based on two steps: feature generation and classification. The TT-transform is used to extract the time–time distribution of the respective signals, i.e. the features are generated by applying the TT-transform to one cycle of the current signal, as shown in Fig. 1.

Fig. 1: Proposed algorithm for HIF classification
Afterwards, the features such as energy, standard deviation and median absolute deviation of the diagonal elements and TT-contour’s area are fed to ART neural networks for distinguishing HIF cases from non-fault conditions and classifying HIF cases, as well. Five categories for classification are selected, i.e. (1) non-fault conditions (such as insulator leakage current, harmonic loads, capacitor banks switching, no load transformer switching, load switching), (2) unbroken conductor, (3) broken conductor on asphalt, (4) broken conductor on concrete, and (5) broken conductor on gravel.

The structure of this paper is as follows. In Section 2, a short review of S-transform is introduced. Then, TT-transform and the extracted features upon this transform are described in Section 3. In Section 4, adaptive resonance theory is briefly explained. The test feeder and applied tests are introduced in Section 5. Finally, in Section 6 results and discussions are presented.

2. S-Transform

S-transform can be introduced as the “phase correction” of continuous wavelet transform (CWT)[6]. The S-transform is defined as follows:

\[ S(f,\tau) = \int_{-\infty}^{\infty} h(t) \frac{1}{\sqrt{2\pi}} e^{-\frac{(t-\tau)^2}{2}} e^{-i2\pi ft} dt \]

(1)

\[ \int_{-\infty}^{\infty} S(\tau,f)d\tau = H(i\omega) \]  

(2)

Equation (2) indicates that the S-transform is fully invertible. The S-transform localizes the phase spectrum as well as the amplitude spectrum. There is an ST-Amplitude Matrix (STA) where the rows are the frequencies and the columns are the time values, i.e. each row displays the S-transform amplitude in the same frequency, and each column displays the S-transform amplitude with all frequencies at the same time.

3. TT-transform

TT-transform [7] is a two-dimensional time–time representation of a one-dimensional time series based on S-transform. TT-transform provides the time-local view of the time series through the scaled windows. The TT-transform is obtained from the inverse Fourier transform of the S-transform:

\[ TT(\tau,t) = \int_{-\infty}^{\infty} S(\tau,f)e^{i2\pi ft} df \]

(3)

From Equation (2) and (3),

\[ \int_{-\infty}^{\infty} TT(\tau,t)d\tau = h(t) \]

(4)

So, similar to the S-transform, the TT-transform is invertible. Each column of TT-matrix is a time localized time series, referred to as the TT-series. The TT-transform (like the S-transform) concentrates higher frequencies more strongly around the midpoint of the Gaussian window (t=\tau) than lower frequencies. This is referred to as differential localization of frequencies. One consequence of differential localization of frequencies is an apparent increase in the amplitudes of higher frequencies, relative to the amplitudes of lower frequencies, on the TT-transform and TT-series.

The features used from TT-transform are energy, standard deviation and median absolute deviation of the diagonal elements and TT-contour’s area from TT-matrix. The diagonal elements and TT-contour’s area are given as follows:

\[ Diagonal = \text{diag}(TT \text{ matrix}) \]

(5)

\[ Area = \text{polyarea}(\text{contour}(TT \text{ matrix})) \]

(6)

The energy, standard deviation and median absolute deviation of diagonal elements and TT-contour’s area, used as input features, are as following:

\[ F_{Di}\text{TT} = \text{sum}(\text{Diagonal}^2) \]

(7)

\[ F_{Di}\text{TT} = \text{std}(\text{Diagonal}) \]

(8)

\[ F_{Di}\text{TT} = \text{mad}(\text{Diagonal}) \]

(9)

\[ F_{A}\text{TT} = \text{sum}(\text{Area}^2) \]

(10)

\[ F_{A}\text{TT} = \text{std}(\text{Area}) \]

(11)

\[ F_{A}\text{TT} = \text{mad}(\text{Area}) \]

(12)

4. Adaptive Resonance Theory (ART)

Adaptive Resonance Theory (ART) was developed by Carpenter and Grossberg[9]. There are a wide variety of ART neural networks implemented using both supervised and unsupervised learning schemes. Unsupervised learning means that the network is able to categorize input patterns based on compete mechanism and inherent similarities between patterns. In ART structure, the traditional clustering process is slightly modified to accommodate the concept of resonance. It means that after finding suitable cluster unit, the network signals “resonate” as the weights are updated. The clustering process in ART networks uses a minimum similarity degree to group patterns together into a cluster. A prominent property of ART neural networks is that if none of units can classify a given pattern, then a new cluster is created. Therefore, the resulting number of cluster units depends on the similarity between all the input patterns in the training set. This property is said to solve the “stability–plasticity dilemma,” that is, “how to learn new things without forgetting old ones” [10].

The first ART model, named by ART1, was essentially built to cluster only binary input vectors. To allow clustering real-valued (analog) input vectors, ART1 developers proposed an extended version, which accepts real-value vectors as input, named by ART2. Based on these two pioneer models, many modifications and improvements have been proposed. The ART2-A system
accurately reproduces the behaviour of ART2 in the fast-learn limit. Among the most applied members of the ART family are ARTMAP, which is a supervised network; Fuzzy ART, which combines ART1 features and fuzzy logic; Fuzzy ARTMAP, which is a supervised version of Fuzzy ART.

4.1 Basic architecture

The basic architecture of an adaptive resonance network involves three groups of neurons (as shown in Fig. 2): Input processing units (F1 layer), cluster units (F2 layer), and similarity check unit. The F1 and F2 layers are short-term memories (STM), with the F1 units used as a buffer for temporary storage of input patterns, and the F2 units used as a buffer to represent categories generated for all nodes. Layers F1 and F2 are fully connected through weights. The weights are long-term memories (LTM).

4.1.1 Vigilance parameter ($\rho$)

The criterion for an adequate match between an input pattern and a chosen neuron is determined by a vigilance parameter ($\rho$), which ranges between 0 and 1. The lower $\rho$ causes fewer categories and rougher the classification, while higher values of $\rho$ means more categories and finer classification. The choice of the vigilance parameter is critical to the performance of an ART network, but there are no guidelines for setting the value of vigilance. The choice of vigilance parameter is application-based, and is usually determined by experimentation. In some applications a variable vigilance parameter has also been used [11].

4.1.2 Learning

Learning updates top-down and bottom-up weights until equilibrium weights are obtained. Once a cluster has been selected for learning, the bottom-up and top-down signals are maintained for an extended period.

4.2 Algorithm

The basic ART training algorithm which is same for all ARTs is summarized here. The input pattern is presented to input layer and activate F2 neurons through bottom-up weights. As the F2 layer is a competitive layer, the neurons compete with each other and the largest neuron becomes winner to learn the input vector. The activations (outputs) of all other F2 units are set to zero. Then the top-down weights belonging to the winner neuron are sent back to the F1 layer. Similarity between top-down weights and input vector is the criterion for whether or not this cluster unit is allowed to learn the input vector. This decision is made by the similarity check unit, based on received signals from the input layer and top-down weights of the F2 layer and a vigilance parameter ($\rho$). If it does not pass the vigilance test (the input pattern is not similar to top-down expectation) the winner neuron is inhibited and a new neuron in F2 layer is selected as the winner. Otherwise, claim of the winner neuron for input pattern classification is accepted and bottom-up and top-down weights are updated and a new input pattern is applied to network. If none of the F2 neurons are able to classify input vector, then a new neuron in the F2 layer is created and the weights are updated [11]. This procedure is repeated until all the patterns are classified.

ART1 is an unsupervised neural network which is designed to cluster and recognize binary patterns only. ART2, same as ART1, is an unsupervised neural network for performing either analog vectors or binary vectors [12]. ART2-A is a simple computational system that models the essential dynamics of the ART2 analog pattern recognition neural network [13]. Its processing time is considered to be twice or more times faster than that ART2. The Fuzzy ART system incorporates computations from fuzzy set theory into ART1 [14]. For example, the intersection ($\cap$) operator used in ART1 learning is replaced by the MIN operator ($\wedge$) of fuzzy set theory. Fuzzy ART reduces to ART1 in response to binary input vectors, but can also learn stable categories in response to analog input vectors, while, the MIN operator reduces to the intersection operator. Fuzzy ARTMAP is a supervised neural network that the training is controlled by a base of examples, where each example is a connection of an input vector to a desired output vector [15].

5. Distribution Feeder

In order to test the proposed method, actual data from high impedance fault currents is gathered from Palash feeder in south west region of Tehran with the single line diagram shown in Fig. 3. This feeder is fed by 63/20 kV transformer with rated apparent power 30 MVA and the high voltage side has been grounded by a zigzag transformer and variable resistance adjusted in 29.5 $\Omega$. Also in this substation, two 2.4 MVAR capacitor bank are connected through high voltage circuit breakers.
Fig. 3: Single line diagram of the considered distribution network

HIF current signals and no-HIF current such as insulator leakage current (ILC) and harmonic load current are measured through the ION 7650 meter with sampling rate of 64 samples per cycle at the site. TABLE I has shown HIF condition tests on different surfaces (rows 1 to 5 have shown the broken conductor test conditions and the sixth row for unbroken conductor).

The physical model for HIF tests on concrete, gravel and asphalt, also HIF tests for unbroken cases are illustrated in Fig. 4 and Fig. 5, respectively.

Because of difficulties in field-testing for no-HIF conditions such as capacitor switching, load switching and no-load transformer switching, those conditions are simulated in ATP software. TABLE II has shown various phenomena of HIF and no-HIF for real network and various phenomena of no-HIF for simulated network.

<table>
<thead>
<tr>
<th>Surface</th>
<th>Conditions</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Concrete</td>
<td>Thickness 20 cm</td>
<td>Very short arc</td>
</tr>
<tr>
<td>2 Concrete</td>
<td>Thickness 10 cm</td>
<td>Arc with vary lengths</td>
</tr>
<tr>
<td>3 Asphalt</td>
<td>Very low thickness</td>
<td>Arc</td>
</tr>
<tr>
<td>4 Asphalt</td>
<td>Thickness 10 cm</td>
<td>No arc</td>
</tr>
<tr>
<td>5 Gravel</td>
<td>Thickness 10 cm</td>
<td>Arc</td>
</tr>
<tr>
<td>6 Tree</td>
<td>Dry</td>
<td>Very short arc</td>
</tr>
</tbody>
</table>

TABLE II: Number of HIF and no-HIF cases considered

<table>
<thead>
<tr>
<th>Different phenomena</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Real Network</strong></td>
<td></td>
</tr>
<tr>
<td>HIF</td>
<td></td>
</tr>
<tr>
<td>Broken</td>
<td>1148</td>
</tr>
<tr>
<td>Concrete</td>
<td>2492</td>
</tr>
<tr>
<td>Asphalt</td>
<td>532</td>
</tr>
<tr>
<td>Gravel</td>
<td>28</td>
</tr>
<tr>
<td><strong>No-HIF</strong></td>
<td></td>
</tr>
<tr>
<td>ILC</td>
<td>1512</td>
</tr>
<tr>
<td>Harmonic load</td>
<td>4200</td>
</tr>
<tr>
<td><strong>Simulated Network by ATP</strong></td>
<td></td>
</tr>
<tr>
<td>No-HIF</td>
<td></td>
</tr>
<tr>
<td>Capacitor switching</td>
<td>72</td>
</tr>
<tr>
<td>Load switching</td>
<td>288</td>
</tr>
<tr>
<td>No-load transformer switching</td>
<td>72</td>
</tr>
<tr>
<td>Three phase fault on adjacent feeder</td>
<td>144</td>
</tr>
<tr>
<td>Feeder energization</td>
<td>240</td>
</tr>
</tbody>
</table>

6. Results and Analysis

6.1 Feature Extraction

The extracted features have time-time information (TT-transform) of current signal that flows through the distribution feeder. The real HIF signal on concrete surface and its diagonal elements and TT-contours are shown in Fig. 6. As shown in Fig. 6-A, due to the different lengths in the formed arcs, the size of the current signal in different cycles and the maximum size of the current signal in two consecutive cycles are not the same. Also, these arcs initiate the presence of high frequency component in HIF current.
TT-transform increases amplitude of the higher frequency components and concentrates them, as it can be observed in zero crossings and peak points, as in Fig. 6-C. Fig. 7-A has shown the ILC current and Fig. 7-B and Fig. 7-C has illustrated its corresponding TT-contours and diagonal elements. As it can be seen in the Fig. 7-A, the ILC current has high frequency components which are bolded by TT-transform as illustrated in Fig. 7-B and Fig. 7-C.

Fig. 8 has shown extracted features in 3-D space of HIF and No-HIF current. As it can be seen in this figure each combination of features has different shape so that overlap between HIF and No-HIF vary with each combination. Also, because of non-linear properties of HIF, these two cases are not linearly separable.

### 6.2 Classification and Results

After extracting features of one cycle of the current signal, to evaluate the proposed method, data set is divided to training and test samples. In this paper, 60% of the total data is used for training stage and the remaining 40% is used for the test stage. In order to detect and classify HIF, test and training data collection is applied to the five types of ART including ART1, ART2, ART2-A, Fuzzy ART and Fuzzy ARTMAP.

TABLE III has indicated the data used for training and testing the ART neural networks. Moreover, the numbers of correct and incorrect cases in testing stage are shown in this table.
TABLE IX has revealed the results and accuracy related to each ART. As shown in TABLE III and TABLE IX, due to binary input ART1 has the worst results. Also, Fuzzy ART and Fuzzy ARTMAP have come with similar outcomes which have the best results among the compared neural networks.

### TABLE IX: Details of each class in the outputs of ART neural networks

<table>
<thead>
<tr>
<th></th>
<th>Number of samples</th>
<th>Correct</th>
<th>Incorrect</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ART1</td>
<td>4291</td>
<td>3933</td>
<td>360</td>
<td>91.61</td>
</tr>
<tr>
<td>ART2</td>
<td>4291</td>
<td>4233</td>
<td>58</td>
<td>98.65</td>
</tr>
<tr>
<td>ART2-A</td>
<td>4291</td>
<td>4248</td>
<td>43</td>
<td>99</td>
</tr>
<tr>
<td>Fuzzy ART</td>
<td>4291</td>
<td>4256</td>
<td>35</td>
<td>99.18</td>
</tr>
<tr>
<td>Fuzzy ARTMAP</td>
<td>4291</td>
<td>4256</td>
<td>35</td>
<td>99.18</td>
</tr>
</tbody>
</table>

#### TABLE IX: Results of ART neural networks

<table>
<thead>
<tr>
<th></th>
<th>Number of input data</th>
<th>Correct identification</th>
<th>Incorrect identification</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ART1</td>
<td>4291</td>
<td>3973</td>
<td>360</td>
<td>91.61</td>
</tr>
<tr>
<td>ART2</td>
<td>4291</td>
<td>4233</td>
<td>58</td>
<td>98.65</td>
</tr>
<tr>
<td>ART2-A</td>
<td>4291</td>
<td>4248</td>
<td>43</td>
<td>99</td>
</tr>
<tr>
<td>Fuzzy ART</td>
<td>4291</td>
<td>4256</td>
<td>35</td>
<td>99.18</td>
</tr>
<tr>
<td>Fuzzy ARTMAP</td>
<td>4291</td>
<td>4256</td>
<td>35</td>
<td>99.18</td>
</tr>
</tbody>
</table>

### 7. Conclusion

In this paper, a cooperating analysis based on adaptive resonance theory is employed in detecting and classifying HIFs in distribution networks, where five classes were defined for classification: (1) non-fault conditions (such as insulator leakage current, harmonic loads, capacitor banks switching, no load transformer switching, load switching), (2) unbroken conductor, (3) broken conductor on asphalt, (4) broken conductor on concrete, (5) broken conductor on gravel. The features are extracted using TT-transform from one cycle of the current signal. Then these features are applied to five types of ART neural networks including: ART1, ART2, ART2-A, Fuzzy ART and Fuzzy ARTMAP. The results have indicated that Fuzzy ART and Fuzzy ARTMAP, which have similar outcomes, came with best answers for classification. In addition, setting apart the ART1 network having poor results in the comparison, the other two networks, i.e. ART2 and ART2-A, have similar acceptable performances.

### References


