Fault Diagnosis of Power Transformer using Duval Triangle Based Artificial Intelligence Techniques

R.Geetha
Coimbatore Institute of Engineering and Technology

Abstract- The fault diagnosis of the dissolved gas analysis (DGA) of the power transformer is to be enhanced than previous adopted techniques; this paper proposes a novel adaptive neuro Fuzzy inference system for the incipient fault recognition through enhanced approach. Complying with the practical DGA records and associated fault causes as much as possible, an ANFIS algorithm is presented to establish fault diagnosis system. After 584 groups of training and testing gas-in oil samples, this work compares and analyzes the network training process and diagnostics results of ANFIS with the previous adopted ANN system. The Diagnostics accuracy of the ANFIS is much more accurate than ANN system. The faults are identified for the records using Duval triangle analysis which is more accurate than other Rogers ratio method, IEC codes etc.

IndexTerms—Dissolved gas analysis (DGA), fault diagnosis, ANFIS algorithm, power transformers.

I. INTRODUCTION

Dissolved gas analysis (DGA) techniques [1] are simple, inexpensive, and widely used to interpret gases dissolved due to the deterioration of the insulating oil of power transformers. Various diagnostic criteria based on gas analysis have been developed. A successful diagnosis using these criteria largely depends on the user’s skill in interpreting the gas datum. Therefore, the search for more reliable methods for incipient fault diagnosis in power transformers using some new DGA’s is still a topic of interest in many utilities.

In the past decades, there has been extensive research on the use of artificial intelligence (AI) techniques to assist the DGA. These investigations include a self adaptive neuro inference system based on emotional learning method which reduces structure complexity, learning time of the networks[1] and the expert system approach [2]–[4], the fuzzy system approach [5]–[8], and the artificial neural-Network (ANN) approach [9]–[12], according to the gas contents in the insulation oil of transformers. The expert system is a decision support system that provides fault diagnosis and maintenance advice. DGA information, such as transformer type, voltage level, gassing trend, and maintenance history is incorporated to build diagnostic rules. We can summarize three functions to be included in this fault diagnostic algorithm: 1) the “central diagnostic engine” incorporated several DGA methods; 2) the “expertise handler” considers special rules; and 3) the “maintenance adviser” proposes the date of next analysis and maintenance actions to be taken. However, many limitations exist in this algorithm. For example, the effectiveness of an expert system depends on the precision and completeness of the knowledge base, which is usually very complicated and must be constructed manually.

The major problem is that it cannot adjust its diagnostic rules automatically and, thus, cannot acquire knowledge from new data samples through a self-learning process. The fuzzy information theory was used to diagnose transformer faults, and the major issue is to tune the fuzzy membership function based on existing DGA methods and experiences [5]. Conventionally, this is done manually [5], [6], and later adaptively using sophisticated mathematical methods such as evolutionary computation, genetic algorithm, and adaptive pruning [7], [8]. An advantage of a fuzzy diagnosis system is that it is
insensitive to errors in the oil sampling, storage, and testing processes. One drawback is that it is bound with conventional DGA methods, and cannot learn directly from data samples.

Among these AI methods, the feed forward ANN is widely designed to diagnose transformer faults based on quite a few advantages. An important advantage is that it can learn directly from the training samples, and update its knowledge when necessary. However, with theoretical analysis and practical application of ANNs certain shortcomings exist [13]. On the other hand, the recently introduced ANFIS emerges as a new powerful tool for approximation. The ANFIS theory has found many applications in function approximation, numerical analysis, and signal processing.

II. FAULT DIAGNOSIS OF POWER TRANSFORMERS

A. Input Vector

The key gases used to diagnose the power transformer faults including six kinds of characteristic gases: H\textsubscript{2}, CH\textsubscript{4}, C\textsubscript{2}H\textsubscript{2}, C\textsubscript{4}H\textsubscript{2}, C\textsubscript{6}H\textsubscript{2}, CO\textsubscript{2}, which are set as the corresponding input vector: [x\textsubscript{1}, x\textsubscript{2}, x\textsubscript{3}, x\textsubscript{4}, x\textsubscript{5}, x\textsubscript{6}]. In order to avoid the input vector being too big and to distill the data information effectively, we use fuzzy technology to preprocess these gas data. According to expert experience, Table I shows the attention values of these gases and the preprocessing is done by

\[ y_i = (1 + e^{-x_i/x_{in}})^{-1} \]  

Where, \( x_i \) and \( x_{in} \) indicate the \( i^{th} \) gas measurement value and attention value, respectively, and \( y_i \) is the input vector to the ANN and ANFIS vector.

<table>
<thead>
<tr>
<th>Gas</th>
<th>H\textsubscript{2}</th>
<th>CH\textsubscript{4}</th>
<th>C\textsubscript{2}H\textsubscript{2}</th>
<th>C\textsubscript{4}H\textsubscript{2}</th>
<th>C\textsubscript{6}H\textsubscript{2}</th>
<th>CO\textsubscript{2}</th>
<th>C\textsubscript{2}H\textsubscript{4}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attention values (in ppm)</td>
<td>100</td>
<td>375</td>
<td>35</td>
<td>50</td>
<td>2500</td>
<td>780</td>
<td></td>
</tr>
</tbody>
</table>

B. Output Pattern

In many cases, there is most likely a single fault in the power transformer, occasionally with multiple incipient faults. Common single incipient faults are as follows: the thermal fault, arching, and partial discharge. The multiple faults are combined cases of the faults, i.e. partial discharge and thermal fault, arching and thermal fault etc. Along with the other faults no fault cases also exists.

C. Samples

About 584 DGA records and samples are tested for the fault condition (either single fault or multiple fault etc) using Duval triangle and provided to train and test in the ANFIS and compared with the ANN adopted system.

III. DUVAL TRIANGLE DIAGNOSTICS

Michel Duval of Hydro Quebec developed this method in the 1960s using a database of thousands of DGAs and transformer problem diagnosis. More recently, this method was incorporated in the Transformer Oil Analyst Software version 4 (TOA 4), developed by Delta X Research and used by many in the utility industry to diagnose transformer problems. This method has proven to be accurate and dependable over many years and is now gaining in popularity. In the above analysis, three gases are
taken for the testing (CH₄, C₂H₂, C₂H₄). The fault types considered are Partial discharge, thermal fault localized overheating, and overloading.

Legend
PD = Partial Discharge, T1 = Thermal Fault Less than 300°C, T2 = Thermal Fault between 300°C and 700°C, T3 = Thermal Fault Greater than 700°C, D1 = Low Energy Discharge (Sparking), D2 = High Energy Discharge (Arcing), DT = Mix of Thermal and Electrical Faults.

Practical gas data for analysis (in ppm):

**TABLE-II**

<table>
<thead>
<tr>
<th>H₂</th>
<th>CH₄</th>
<th>CO₂</th>
<th>C₂H₄</th>
<th>C₂H₆</th>
<th>C₂H₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>102</td>
<td>527</td>
<td>3630</td>
<td>773</td>
<td>28</td>
<td>35</td>
</tr>
<tr>
<td>104</td>
<td>531</td>
<td>3621</td>
<td>785</td>
<td>26</td>
<td>36</td>
</tr>
<tr>
<td>106</td>
<td>523</td>
<td>3633</td>
<td>779</td>
<td>29</td>
<td>37</td>
</tr>
<tr>
<td>103</td>
<td>556</td>
<td>3605</td>
<td>776</td>
<td>27</td>
<td>38</td>
</tr>
<tr>
<td>95</td>
<td>548</td>
<td>2410</td>
<td>770</td>
<td>23</td>
<td>36</td>
</tr>
<tr>
<td>101</td>
<td>553</td>
<td>3600</td>
<td>775</td>
<td>20</td>
<td>37</td>
</tr>
<tr>
<td>102</td>
<td>560</td>
<td>3589</td>
<td>783</td>
<td>22</td>
<td>45</td>
</tr>
<tr>
<td>100</td>
<td>552</td>
<td>3608</td>
<td>774</td>
<td>23</td>
<td>44</td>
</tr>
<tr>
<td>107</td>
<td>607</td>
<td>3530</td>
<td>848</td>
<td>23</td>
<td>46</td>
</tr>
</tbody>
</table>
Testing procedure:

Once a problem has been determined to exist, use the total accumulated amount of the three Duval Triangle gases and plot the percentages of the total on the triangle to arrive at a diagnosis. An example is shown below. Also, calculate the amount of the three gases used in the Duval Triangle, generated since the sudden increase in gas began. Subtracting out the amount of gas generated prior to the sudden increase will give the amount of gases generated since the fault began. Detailed instructions and an example are shown below.

a. Take the amount (ppm) of methane (CH₄) in the DGA and subtract the amount of CH₄ from an earlier DGA, before the sudden increase in gas. This will give the amount of methane generated since the problem started.
b. Repeat this process for the remaining two gases, ethylene (C₂H₄) and acetylene (C₂H₂).

1. Add the three numbers (differences) obtained by the process of step a above. This gives 100 percent (%) of the three key gases generated since the fault, used in the Duval Triangle.
2. Divide each individual gas difference by the total difference of gas obtained in step 1 above. This gives the percentage increase of each gas of the total increase.
3. Plot the percentage of each gas on the Duval Triangle, beginning on the side indicated for that particular gas. Draw lines across the triangle for each gas parallel to the hash marks shown on each side of the triangle.

Steps to Obtain the Diagnosis on the Duval Triangle:

Let us illustrate with a gas data

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CH₄</td>
<td>192</td>
</tr>
<tr>
<td>C₂H₄</td>
<td>170</td>
</tr>
<tr>
<td>C₂H₂</td>
<td>7</td>
</tr>
<tr>
<td>TOTAL</td>
<td>369</td>
</tr>
</tbody>
</table>

1. Use the total accumulated gas from DGA = 369
2. Divide each gas by the total to find the percentage of each gas of the total.
   \[
   \% \text{ CH}_4 = \frac{192}{369} = 52\%, \quad \% \text{ C}_2\text{H}_4 = \frac{170}{369} = 46\%, \quad \% \text{ C}_2\text{H}_2 = \frac{7}{369} = 2\%
   \]
3. Draw three lines across the Duval Triangle starting at the percentages obtained in step 2. These lines must be drawn parallel to the hash mark on each respective side.
4. Point 1 is obtained where the lines intersect within the T2 diagnostic area of the triangle, which indicates a thermal fault between 300 and 700 °C

By using the above procedure we can determine the fault that has occurred in the sample data. Once the fault is determined, the samples are implemented in the ANN and ANFIS system. The fault is probably a bad connection on a bushing bottom, a bad contact or connection in the tap changer, or a problem with a core ground. These problems are probably all reparable in the field. Any of these problems can cause the results revealed by the Duval Triangle diagnosis above.

The only disadvantage of the Duval triangle is that it can’t be used to determine whether or not a transformer has a problem. Notice, there is no area on the triangle for a transformer that does not have a
problem. The triangle will show a fault for every transformer whether it has a fault or not. The Duval Triangle is used only to diagnose what is the problem aroused.

Fig.2. Duval Triangle diagnostic of the power transformer.

**NUMERICAL TESTS AND DISCUSSIONS**

These proposed mathematic models are developed in MATLAB programming language based on the 584 testing and training datas.

**IV. TRAINING IN ANN**;

Under the same training conditions, including the same network input, output pattern, and input data-preprocessing methods, we use a conventional ANN based on the BP algorithm. The results of training in ANN is shown. The sample datas are trained and the results are shown.

Fig.3. Mean error curve of the ANN.
Testing in ANN

An Artificial Neural Network included selection of inputs, outputs, network topology and weighted connection of node. Input feature-selection constitutes an essential first step. This is chosen very carefully so that the input features will correctly reflects the characteristics of the problem. Another major task of the ANN design is to choose network topology. This is done experimentally through a repeated process to optimize the number of hidden layers and nodes according to training performance and prediction accuracy. In the current study, six key gases H₂, CH₄, C₂H₆, C₂H₄, and C₂H₂ and CO₂ are chosen as input features. Since overheating, partial discharge and arcing are the three major fault types in power transformers; hence there will be four output patterns to be identified including the normal condition.

For many years there was no theoretically sound algorithm for training multi-layer ANN, and therefore, the applications of ANN were severely limited. The invention of Back propagation algorithm has played a vital role in the resurgence of interest in ANN. Back-propagation is a systematic method for training multi-layer ANN. It has a strong mathematical set of inputs is applied from outside or from previous layer. Each of these is multiplied by a corresponding weight w. The sum of the weighted inputs and the bias b forms the input n to the transfer/activation function. Neurons may use any differentiable, monotonic increasing transfer functions to generate their outputs. Back-propagation networks often use the log-sigmoid and tan-sigmoid transfer functions.

Structure of ANN;

The ANN structure has adopted with the input, hidden, and output layer. The input layer has 6 neuron and 150 in the hidden layer and with 1 neuron in the output layer. The learning rate obtained in the ANN is 0.04. The feed forward network in this is resilient propagation used for the algorithm. The testing is done in the power transformer of capacity 150 MVA transformer, 22/66KV of voltage rating which is in the Bangalore substation, India.

V. ANFIS ALGORITHM

The neuro-adaptive learning method works similarly to that of neural networks. Neuro-adaptive learning techniques provide a method for the fuzzy modeling procedure to learn information about a data set. Fuzzy Logic Toolbox software computes the membership function parameters that best allow the associated fuzzy inference system to track the given input/output data. The Fuzzy Logic Toolbox function that accomplishes this membership function parameter adjustment is called anfis. The anfis function can be accessed either from the command line or through the ANFIS Editor GUI. Because the functionality of the command line function anfis and the ANFIS Editor GUI is similar, they are used somewhat interchangeably in this discussion, except when specifically describing the GUI.

The acronym ANFIS derives its name from adaptive neuro-fuzzy inference system. Using a given input/output data set, the toolbox function anfis constructs a fuzzy inference system (FIS) whose membership function parameters are tuned (adjusted) using either a back propagation algorithm alone or in combination with a least squares type of method. This adjustment allows your fuzzy systems to learn from the data they are modeling.

In a fuzzy inference system, basically there are three types of input space partitioning: grid, tree, and scattering partitioning. GENFIS1 uses the grid partitioning and it generates rules by enumerating all possible combinations of membership functions of all inputs; this leads to an exponential explosion even when the number of inputs is moderately large. For instance, for a fuzzy inference system with 10 inputs, each with two membership functions, the grid partitioning leads to 1024 (=2^10) rules, which is
inhibitively large for any practical learning methods. The "curse of dimensionality" refers to such situation where the number of fuzzy rules, when the grid partitioning is used, increases exponentially with the number of input variables.

FIS structure:

A network-type structure similar to that of a neural network, which maps inputs through input membership functions and associated parameters, and then through output membership functions and associated parameters to outputs, can be used to interpret the input/output map. The parameters associated with the membership functions changes through the learning process. The computation of these parameters (or their adjustment) is facilitated by a gradient vector. This gradient vector provides a measure of how well the fuzzy inference system is modeling the input/output data for a given set of parameters.

When the gradient vector is obtained, any of several optimization routines can be applied in order to adjust the parameters to reduce some error measure. This error measure is usually defined by the sum of the squared difference between actual and desired outputs. Anfis uses either back propagation or a combination of least squares estimation and back propagation for membership function parameter estimation.

Structure of ANFIS:
ANFIS information;

- Number of nodes : 1503
- Number of linear parameters : 729
- Number of nonlinear parameters : 54
- Total number of parameters : 783
- Number of training data pairs : 584
- Number of checking data pairs : 0
- Number of fuzzy rules : 729

VI. TRAINING IN ANFIS

To train a FIS, you must begin by loading a Training data set that contains the desired input/output data of the system to be modeled. Any data set you load must be an array with the data arranged as column vectors, and the output data in the last column. The DGA samples are given as input to the ANFIS model and Trained by Grid partitioning to gain a zero error tolerance. With 20 iterations ANFIS is trained to get the required mean error curve with minimum error. The error Tolerance is reduce in the ANFIS with a value of 0.456. The trained data is then tested in ANFIS for the required fault output. The Practical data is implied in the testing structure for the output and the output fault is detected with minimum error. The anfis model is trained and tested with the actual and practical gas data of the 150 MVA 22/66KV transformer.

![Fig: 6 Mean Error curve in ANFIS](image)

**ANFIS error measure;**

Error measure $E_k$ is calculated by the following formula. (for the $k^{th}$ entry of the training data)

$$E_k = \sum_{i=1}^{N(L)} (d_i - x_{L,i})$$

Where:

- $N(L)$ = number nodes in layer $L$
- $d_i$ = $i^{th}$ component of desired output vector
- $x_{L,i}$ = $i^{th}$ component of actual output vector
ANFIS Creates a fuzzy decision tree to classify the data into one of $2^n$ (or $p^n$) linear regression models to minimize the sum of squared errors (SSE):

$$\text{SSE} = \sum_{j} e_j^2$$

Where:
- $e_j$ is the error between the desired and the actual output
- $p$ is the number of fuzzy partitions of each variable
- $n$ is the number of input variables

The output got is with minimum error in the ANFIS is compared to other conventional methods adopted previously. The DGA samples are given as input to the ANFIS model and Trained by Grid partitioning to gain a zero error tolerance. With 20 iterations ANFIS is trained to get the required mean error curve with minimum error.

**Fig: 7 ANFIS output for Thermal fault.**

**Fig: 8 ANFIS output for Arcing & Partial discharge**
The ANFIS model is tested for all the practical gas data and the output for each data got in the ANFIS gives the error for that data. The blue dot specifies the actual encoded value for the fault and the red dot specifies the error obtained for the fault in this algorithm. The error got in this approach is least compared to other adopted methods. In this work, the main intention is to improve the fault diagnostic accuracy of the transformer than the ANN and previous adopted ANFIS algorithm with emotional learning. The ANFIS structure varies for the sub-clustering and grid partitioning with only the variation of the rules and the membership functions. The diagnostics accuracy of the various methods is compared below.


### Diagnostic Accuracy of the Networks:

**Table-III**

<table>
<thead>
<tr>
<th>Fault type</th>
<th>ANN NETWORK</th>
<th>ANFIS NETWORK</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Successful diagnosis</td>
<td>Diagnostic accuracy</td>
</tr>
<tr>
<td>Thermal fault</td>
<td>109/120</td>
<td>91%</td>
</tr>
<tr>
<td>Overloading</td>
<td>93/100</td>
<td>93.43%</td>
</tr>
<tr>
<td>OL &amp; Partial Discharge</td>
<td>81/90</td>
<td>90.56%</td>
</tr>
<tr>
<td>Arching &amp;PD</td>
<td>71/75</td>
<td>94.75%</td>
</tr>
<tr>
<td>No fault</td>
<td>74/81</td>
<td>90.96%</td>
</tr>
</tbody>
</table>

### VII. CONCLUSION

With the above diagnostic survey, the ANFIS network is much more accurate than the other adopted system analysis where the ANN is only compared here with the ANFIS. According to the mathematics foundation, the ANFIS, which are suitable for fault diagnosis, were taken for the analysis first and then compared with the other efficient networks. In addition, as a complex nonlinear system composed of gas inputs and fault type outputs, the power transformer state is very uncertain and unpredictable, while some problems exist in the BP of ANN, such as slow convergence, searching space tapping in local minima, and oscillation. To resolve these problems, in this paper, the ANFIS model is taken for the diagnosis to reduce error and accelerate convergence speed; this paper employs the ANFIS for achieving the satisfied fault diagnostic model.

### REFERENCES