

# Improving the IEC Table for Transformer Failure Diagnosis with Knowledge Extraction from Neural Networks

Vladimiro Miranda, *Senior Member, IEEE*, and Adriana R. Garcez Castro, *Student Member, IEEE*

**Abstract** – The paper describes how mapping a neural network into a rule-based fuzzy inference system leads to knowledge extraction. This mapping makes explicit the knowledge implicitly captured by the neural network during the learning stage, by transforming it into a set of rules. By applying the method to transformer fault diagnosis using dissolved gas-in-oil analysis (DGA), one could not only develop intelligent diagnosis systems, providing better results than the application of IEC 60599 Table, but also generate a new rule table whose application also leads to better diagnosis results.

*Index Terms*— Fault diagnosis, fuzzy logic, neural networks

## I. INTRODUCTION

This paper is meant to present a practical Table for incipient transformer fault diagnosis, improving the existing IEC 60599 Table based on DGA - Dissolved Gas Analysis. This presentation serves as an example of a new technique of knowledge extraction from case analysis, which we have named TFRENN - Transparent Fuzzy Rule Extraction from Neural Networks. Its basic idea is the following: given a data base of cases, if a certain type of Artificial Neural Network (ANN) can be trained to recognize features in it, then we can apply a specific new mathematical transform and convert the ANN in a Fuzzy Inference System (FIS) of the Takagi-Sugeno type. Associated with the FIS there is a rule base formed by a set of fuzzy rules, which may be conveniently manipulated so that the knowledge implicitly captured by the ANN becomes explicit. This allows knowledge recognition and, hopefully, knowledge discovery - we call this "opening the black box".

In fact, ANN are often referred to as black boxes, because it is argued that they do not have explaining capability - although they perform well, no logic explanation may be offered on why an ANN, facing a given input, generates its output. Mathematically, the ANN performs interpolation on data, but there are no explicit or individualized rules providing an explanation that is significant in human terms.

The problem of the diagnosis of incipient transformer faults has been dealt with DGA methods such as Dornenburg Ratios or Rogers Ratios, which are commonly used by utilities and

manufacturers[1]. The concentration, relative proportion, and generation rate of certain byproduct gases dissolved in the insulating oil have been extensively used for the estimation of the condition of a transformer. One particular technique of interest is translated in IEC 60599 document, presenting the so-called IEC Ratios [2]. However, the characterization achieved so far still has some margin for improvement.

Both ANN [3] [4] and FIS [5] have been tried in developing fault diagnosis systems, but a connection has never been established between them. In particular, ANN lack explanation capability and FIS efficiency depended on the completeness of the knowledge of human specialists. The TRFENN approach, however, allows the extraction of a set of rules exhibiting the property of transparency, which is a requirement for human understanding of a set of rules with linguistic correspondence.

## II. PREVIOUS WORKS WITH ANN

A number of works on incipient transformer fault diagnosis have been published, based on neural networks. In general, the systems developed presented promising results and this success is mainly due to the neural network ability to learn the hidden relations among fault types and dissolved gas concentrations. Besides learning abilities, neural networks also have the advantage of being able to acquire new information by incremental training from newly obtained samples. This is not so natural in systems based on fuzzy rules, with the exception of those implementations that also rely on a Back propagation procedure to evolve parameters, such as the Takagi-Sugeno FIS.

In [6] we found an artificial neural network approach to diagnose and detect faults in transformers based on DGA analysis. A two step ANN method is used to detect faults with or without cellulose involved. The first ANN provided the classification of the three major fault types: overheating, corona and arcing while the second ANN was constructed to determine if the cellulose was involved. The authors claimed that the results of the two-ANN approach were promising even with limited sample data. However, if more complex relationships had to be learned by the ANN, such as the classification of more specific faults, more training data were needed.

A combined ANN and expert system tool (ANNEPS) for transformer fault diagnosis is presented in [8]. With the aim of including as many known diagnosis rules as possible, the knowledge base of the expert system was derived from the

---

Vladimiro Miranda is with INESC Porto, Portugal and also with FEUP, Faculty of Engineering of the University of Porto, Portugal (email: vmiranda@inescporto.pt).

Adriana R. Garcez Castro is with INESC Porto, Portugal and also with NESCFUPPA – Federal University of Pará, Brasil. She is sponsored by CAPES - Brazilian Government (email: acastro@inescporto.pt).

IEEE and IEC DGA standards as well as from expert experience. The ANNEPS provided a weighted final diagnostic result based on the combination of the expert system and ANN output. The system also provided maintenance action recommendations. The ANNEPS could determine thermal faults (distinguishing between four thermal stages and overheating of the oil), low energy discharge (corona), high energy discharge and arcing, and cellulose degradation from normal condition.

In [10] a comparative study of ANN efficiency for the detection of faults in power transformers is presented. Five conventional DGA methods were used to train the ANN. The ANN diagnosis results were compared with those obtained by inspection and a specialist. The studies showed that the ANN rate of successful diagnosis were dependent on the criterion under consideration. The percentage of correct diagnosis was in the range of 87-100%.

In [11] an evolving neural net (ENNs) for fault diagnosis was proposed. An evolutionary algorithm was used to tune the ANN parameters to achieve the best model. The results showed that the use of the concentration of the gases as input of the ANN provided better results than the use of the relation between the gases. Another result of this work was that the ENNs improved the diagnosis accuracy and learning speed of the conventional ANN approaches.

Table I presents some results of some systems developed for transformer diagnosis based on DGA analysis.

TABLE I – RESULTS OF SOME SYSTEMS FOR TRANSFORMER INCIPENT FAULT DIAGNOSIS

Reference	No. samples in the database	Percentage of correct diagnoses of the developed system
Zhang et al [6]	-	95% Classification of the three main faults
Huang et al [7]	711	90.30% - training database 93.81% - testing database
Wang [8]	210	99.72% - training database 95.34% - testing database
Liao et al [9]	711	96.2%
Guardado et al [10]	33	ANN training Considering: Doemberg method – 90.91% Modified Roger – 87.88% Roger – 90.91% IEC – 93.94%
Huang [11]	820	90.49% and 93.54%, depending on the number of inputs Classification of 4 fault types

### III. NEURAL NETWORKS

A Multilayer Feedforward Neural Network, also known as Multilayer Perceptron (MLP), basically consists of a finite number of successive layers, each having a finite number of processor units called neurons (Figure 1). Each neuron of every layer is connected to every neuron of a following layer through synaptic weights.

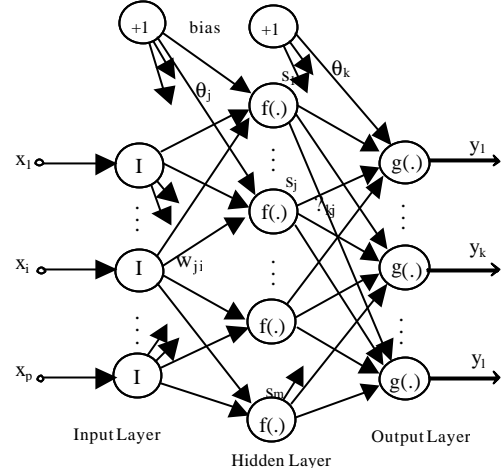


Figure 1. Multilayer Feedforward Neural Network

Every neuron in a hidden layer calculates:

$$s_j = f\left(\sum_{i=1}^n x_i w_{ij} + q_j\right) \quad (1)$$

where  $x_i$  is the  $i$ -th input to the net,  $w_{ij}$  is the weight of the connection from input neuron  $i$  to hidden neuron  $j$ ,  $q_j$  is the bias of the  $j$ -th hidden neuron and  $f(\cdot)$  is the activation function of the neuron.

For the output layer, each neuron calculates:

$$y_k = g\left(\sum_{j=1}^m b_{jk} s_j + q_k\right) \quad (2)$$

where

$b_{jk}$  is the weight of the connection from hidden neuron  $j$  to output neuron  $k$ ,  
 $y_k$  is the  $k$ -th output of the net  
 $q_k$  is the bias of the  $k$ -th output neuron and  
 $g(\cdot)$  is the activation function of the neuron.

ANNs are universal approximators. It has been extensively demonstrated that a MLP working with arbitrary squashing functions in hidden neurons can approximate virtually any function of interest to any desired degree of accuracy [12].

### IV. FUZZY INFERENCE SYSTEMS

In fuzzy inference systems of the Takagi-Sugeno type (TS-FIS), the relationships among variables of the system are represented by fuzzy IF-THEN rules in the form:

$$\text{Rule } R_l: \text{ If } x_1 \text{ is } C_1^l \text{ and } \dots \text{ and } x_n \text{ is } C_n^l \\ \text{ then } y^l = f(x_1, \dots, x_n) \quad (3)$$

where  $C_i^l$  are fuzzy sets,  $x_i$  is the input of the system. When  $y^l$  is a constant, the fuzzy inference system is called a zero-order TS fuzzy model.

The firing strength of each rule is calculated by:

$$v_j = \bigcap_{i=1}^n m_{ij}(x_i) \quad (4)$$

where  $\mathbf{m}_j(x_i)$  is the membership function associated to the fuzzy set  $C_i^j$  and  $\cap$  represents the product operator (AND operator). The consequent of a rule is an affine linear or non-linear function of the input variables.

The output of the system is computed as the weighted average of the  $y^j$ , that is:

$$y = \frac{\sum_{j=1}^N y_j v_j}{\sum_{j=1}^N v_j} \quad (5)$$

where  $N$  is the number of rules of the system.

## V. MAPPING ANN INTO A FIS MODEL

### A. Definition of the topology of the ANN

This paper deals with a particular subset of ANNs, having one hidden layer and with only one output neuron with a linear activation function.. The hidden neurons have a particular activation sigmoid function, which we will call *positive sigmoid function* and whose graphic is shown in Figure2:

$$f(x) = \begin{cases} 1 - e^{-x} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (6)$$

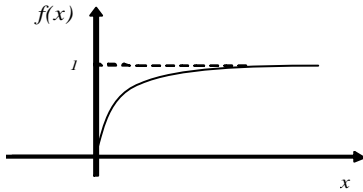


Figure 2. Positive Sigmoid function

### B. Discovering fuzzy logic operators in the ANN

The concept of *f-duality*, originally introduced in [13], allows us to define a new transform providing an equivalent mathematical operation to the operation performed by the hidden neurons in the ANN – equation (1). We will not describe the details of the concept. Additional Lemmas established and proven by us provide the adequate framework for transforming an ANN into a FIS [14] [15].

Applying the concept of *f-duality* in an ANN without bias  $\mathbf{q}_j$ , where the activation function  $f$  is a positive sigmoid such as in (6), and considering that  $\sum_{i=1}^n x_i w_{ij} \geq 0$ , we have proven that the output signal of the hidden neurons can be calculated by:

$$\begin{aligned} s_j &= f\left(\sum_{i=1}^n x_i w_{ij}\right) = f(x_1 w_{1j}) * \dots * f(x_n w_{nj}) \\ &= 1 - (1 - f(x_1 w_{1j})) \dots (1 - f(x_n w_{nj})) \\ &\text{if } \sum_{i=1}^n x_i w_{ij} \geq 0 \text{ and } x_i w_{ij} \geq 0 \end{aligned} \quad (7)$$

This transformation is crucial because we now recognize in (7) a logic operator, well known in Fuzzy Logic as Algebraic Sum, which is an S-norm (OR-operator).

Now we can interpret  $f(x_i w_{ij})$  as a membership function whose linguistic expression may be “ $x_i$  is greater than  $2.3/w_{ij}$ ”, by fixing  $f(2.3/w_{ij}) = 0.9$ . The reason for this is that  $f(x)$  can reach 1 only asymptotically; thus we have set the  $\alpha$ -cut for  $\mathbf{a} = 0.9$  as a significant linguistic threshold.

### C. Extracting Rules from ANN

In an ANN as shown in Figure 1 (considering the ANN with one output neuron), having the hidden neurons without bias,  $\sum_{i=1}^n x_i w_{ij} \geq 0$  and  $x_i w_{ij} \geq 0$ , for each neuron in the hidden layer, one rule can be extracted as:

$$\text{Rule } R_j: \text{ If } \sum_{i=1}^n x_i w_{ij} \text{ is } A \text{ then } y_j = \mathbf{b}_j \quad (8)$$

where  $A$  is a fuzzy set whose membership function is the positive-sigmoid function.

According to (7), rules as in (8) can be written as:

$$\begin{aligned} \text{Rule } R_j: &\text{ If } (x_1 w_{1j} \text{ is } A) * \dots * (x_i w_{ij} \text{ is } A) * \dots * (x_n w_{nj} \text{ is } A) \\ &\text{then } y_j = \mathbf{b}_j \end{aligned} \quad (9)$$

Expression “ $x_i w_{ij}$  is  $A$ ” may be interpreted as “ $x_i$  is  $A_{ij}$ ” if one defines the weight  $w_{ij}$  as a scaling factor of the slope of  $f(\cdot)$ . Then, the fuzzy set  $A_{ij}$  becomes defined by a membership function  $\mathbf{m}(A_{ij}) = f(x_i w_{ij})$ . Recognizing the operation  $*$  as the Algebraic Sum operator (OR), we may rewrite (9) as:

$$\begin{aligned} \text{Rule } R_j: &\text{ If } (x_1 \text{ is } A_{1j}) \text{ or } \dots \text{ or } (x_i \text{ is } A_{ij}) \text{ or } \dots \text{ or } (x_n \text{ is } A_{nj}) \\ &\text{then } y_j = \mathbf{b}_j \end{aligned} \quad (10)$$

where its firing strength is calculated by the algebraic sum operator:

$$v_j = \mathbf{m}(A_{1j}) * \dots * \mathbf{m}(A_{nj}) = 1 - ((1 - \mathbf{m}(A_{1j})) \dots (1 - \mathbf{m}(A_{nj}))) \quad (11)$$

Finally, the output of the fuzzy system can be extracted as:

$$y = \sum_{j=1}^m \mathbf{b}_j s_j \quad (12)$$

Since  $s_j = v_j$  and  $\mathbf{b}_j = y_j$ , equation (12) can be rewritten:

$$y = \sum_{j=1}^m y_j v_j \quad (13)$$

We have finally obtained a mathematical description of the operation performed by the ANN that has the form of an inference system that resembles to a zero-order Takagi-Sugeno model. However, the fuzzy logic operator used to calculate the firing strength of each rule is an S-norm (OR) and not a T-norm (AND) as in the canonical form of TS-FIS.

T-norms and S-norms (representing AND and OR operations) may be associated in couples such that for each S-norm there is a T-norm such that the two together satisfy DeMorgan’s Law [16].

The T-norm associated with the Algebraic Sum operator  $S(a,b) = 1 - (1-a)(1-b)$  is the Algebraic Product operator  $T(a,b) = ab$ . Therefore, each rule extracted in (10) can be transformed into:

*Rule  $R_j$* : If ( $x_i$  is Not  $A_{ij}$ ) and... ..and ( $x_n$  is Not  $A_{nj}$ )  
then  $y_j = \mathbf{b}_j$  (14)

where the rule firing strength is calculated by the algebraic product operator (AND operator); the system output becomes

$$y = \sum_{j=1}^m \mathbf{b}_j (1 - v_j) \quad (15)$$

Rearranging (15),

$$y = \sum_{j=1}^m \mathbf{b}_j (1 - v_j) = \sum_{j=1}^m \mathbf{b}_j - \sum_{j=1}^m y_j v_j \quad (16)$$

where  $\sum_{j=1}^m \mathbf{b}_j$  is the default value of the fuzzy system output.

If a bias is used in the hidden neuron, then the consequent of rule  $R_j$  will change from  $y_j = \mathbf{b}_j$  to  $y_j = \mathbf{b}'_j = \mathbf{b}_j (1 - f_p(\mathbf{q}_j))$ . If a bias ( $\mathbf{q}_{out}$ ) conditions the output neuron, the system output will change to:

$$y = \sum_{j=1}^m \mathbf{b}'_j - \sum_{j=1}^m y_j v_j + \mathbf{q}_{out} \quad (17)$$

where  $\sum_{j=1}^m \mathbf{b}'_j + \mathbf{q}_{out}$  is the new default value of the rule [15].

#### D. Enforcing a correct training

The basic idea to produce a mapping of ANNs into FIS is condensed in the previous section. However, there is a further objective: that the rule antecedents extracted from the ANN may be meaningful and subject to human interpretation. To guarantee this, some constraints have to be enforced.

If the negation (NOT) is applied to the extracted membership  $\mathbf{m}(A_i) = f(x_i w_{ij})$ , we will have a new membership defined as:

$$f(x_i w_{ij}) = \begin{cases} e^{-x_i w_{ij}} & , x_i w_{ij} \geq 0 \\ 1 & , x_i w_{ij} < 0 \end{cases} \quad (18)$$

Weight  $w_{ij}$  acts as a scaling factor of the  $f(\cdot)$ . Taking the  $\alpha$ -cut for  $\mathbf{a} = 0.999$ , we can approximate (18) to:

$$f(x_i) = \begin{cases} e^{-x_i w_{ij}} & , x_i \geq 0.001 / w_{ij} \\ 1 & , x_i < 0.001 / w_{ij} \end{cases} \quad (19)$$

where  $f(0.001 / w_{ij}) = 0.999 \approx 1$  (Figure 3).

This new set fuzzy may be described by “smaller than  $0.001 / w_{ij}$ ”. It is meaningful if  $0 \leq 0.001 / w_{ij} \leq 1$ , which requires that  $w_{ij} \geq 0.001$ . With  $w_{ij} \geq 0.001$ ,  $0 \leq x_i \leq 1$  (normalized inputs) and  $\mathbf{q}_j \geq 0$ , the correct use of (7) is guaranteed since we will always have  $\sum_{i=1}^n x_i w_{ij} \geq 0$  and  $x_i w_{ij} \geq 0$ . However, during the training of the neural net the bias values can assume any value in  $[-\infty + \infty]$ . To overcome this problem, and considering the constraints  $w_{ij} \geq 0.001$  and  $\mathbf{q}_j \geq 0$ , the weights and bias of equation (1) are transformed using the exponential function:

$$s_j = f\left(\sum_{i=1}^n x_i (0.001 + e^{w_{ij}}) + e^{q_j}\right) \quad (20)$$

The new weight  $w'_{ij} = 0.001 + e^{w_{ij}}$  will always be greater than 0.001 and the new bias  $\mathbf{q}'_j = e^{q_j}$  will be greater than zero.

The backpropagation algorithm commonly used for training a neural network can be used taking in account these conditions, simply adjusting  $w_{ij}$  and  $\mathbf{q}_j$  in  $[-\infty + \infty]$ .

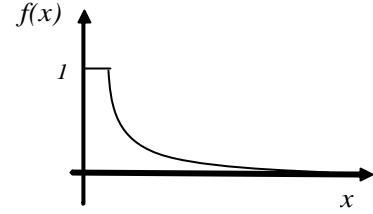


Figure 3. New Membership Function Extracted

#### E. Extraction of a Transparent Fuzzy System

Transparency is a property of fuzzy rule systems that allows the interpretation of rules by humans. A fuzzy system is transparent only if all rules of the rule base are transparent [17]. A rule is considered transparent if, fired by itself, produces an output that is meaningful and according to the expected behavior of the physical system represented. This means that at firing strength

$$v_j = \bigcap_{i=1}^n \mathbf{m}_j(x_i) = 1 \quad (21)$$

the system output is

$$y = y_j \quad (22)$$

where  $y_j$  is the center of the output membership associated with the rule. For the case of zero-order Takagi-Sugeno models,  $y_j$  is equal to the constant consequent. This means that the effect of a single rule may be isolated. In non transparent fuzzy rule systems, the output is meaningful only with interpolation of rules.

To provide the desired transparency to the rule system extracted from the ANN, we need to develop a further approximation process for all membership functions. In our work, this process is performed by using a combination of 5 membership functions (Figure 4 and Figure 5):

$$\mathbf{m}(x) = a_1 \mathbf{m}_{small}(x) + a_2 \mathbf{m}_{medium}(x) + a_3 \mathbf{m}_{high}(x) + a_4 \mathbf{m}_{verysmall}(x) + a_5 \mathbf{m}_{quitesmall}(x)$$

where  $x \in [0,1]$  and  $[a_1, a_2, \dots, a_5]$  are parameters that have to be identified. In this work the recursive least squares algorithm is used for this task. Based on experience, TFRENN is applied in the following manner:

1. When the learning of the constrained ANN, for a specific application, can provide the desired result with the value of  $w_{ij}$  restrict to  $[0.001, 13]$ , then the extracted membership functions will be approximated by using the 5 membership functions in Figure 4 (case 1). The restriction of  $w_{ij}$  to  $[0.001, 13]$  can be

guaranteed during the learning of the ANN. The restriction to  $w_{ij} > 0.001$  is already guaranteed using  $w'_{ij}$  whereas the restriction to  $w_{ij} \leq 13$  will be guaranteed by avoiding variations in the weights if the training leads them to values greater than 13. The restrictions imposed on the weights will reduce the search space during the ANN learning phase, and luckily this can in fact lead to a better generalization.

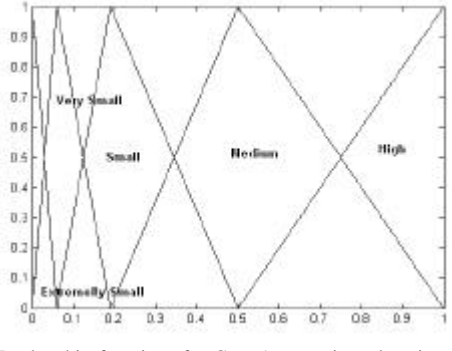


Figure 4. Membership functions for Case 1 – x axis: values in a normalized domain; y axis: membership value.

- When the learning of the constrained ANN, for a specific application, cannot provide the desired result with the value of  $w_{ij}$  restrict to  $[0.001, 13]$ , then the constrained ANN has to be trained normally with the restriction  $w_{ij} > 0.001$ . In this case, the extracted membership functions for values of  $w_{ij}$  in  $[0.001, 13]$  will be approximated by using the 5 membership functions in Figure 5 (Case 2), whereas the extracted membership functions for  $w_{ij} > 13$  will not be approximated; they will be used as hedges of the membership function represented by the linguistic variable “small”.

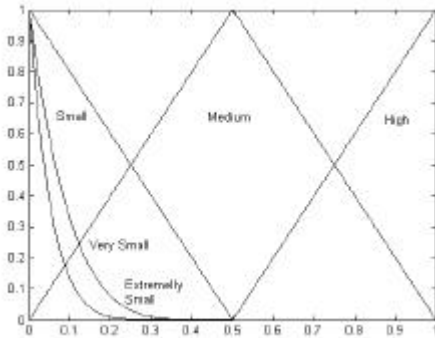


Figure 5. Membership functions for Case 2 – x axis: values in a normalized domain; y axis: membership value.

For each rule, the approximated membership functions calculated in (23) for each input are then combined. As result, a new rule-based system with a total number of transparent rules equal to  $5^n$  is formed, where  $n$  is the number of inputs of the system.

## VI. TRANSFORMER FAULT DIAGNOSIS

### A. The transformer fault diagnosis system proposed

The detection of incipient faults on transformers using DGA begins with the observation of the evolution of rates of combustible gases that exceed “normal” quantities. If the evolution rate per day is greater than a determined level then the transformer is suspected to have an active internal fault.

TABLE II  
IEC 60599 CRITERIA FOR THE INTERPRETATION OF DGA

Case	Characteristic fault	$\frac{C_2H_2}{C_2H_4}$	$\frac{CH_4}{H_2}$	$\frac{C_2H_4}{C_2H_6}$
PD	Partial discharge	NS	<0.1	<0.2
D1	Discharges of low energy	>1	0.1-0.5	>1
D2	Discharges of high energy	0.6-2.5	0.1-1	>2
T1	Thermal fault $T < 300^\circ\text{C}$	NS	>1 but NS	<1
T2	Thermal fault $300^\circ\text{C} < T < 700^\circ\text{C}$	<0.1	>1	1-4
T3	Thermal fault $T > 700^\circ\text{C}$	<0.2	>1	>4

NS = Non-significant whatever the value

After a first diagnosis of the possible fault, in order to obtain confirmation and more detailed information, such as the location of the fault, other tests are needed.

Many techniques for the detection of possible faults of transformer using the measurement of gases have been established. Table II presents the IEC 60599 criteria, which is widely used to interpret the DGA [2].

In our work, we trained a neural network of the type described above, with 60 neurons in the hidden layer, to receive as input data the ratios  $\frac{C_2H_2}{C_2H_4}$ ,  $\frac{CH_4}{H_2}$  and  $\frac{C_2H_4}{C_2H_6}$ , and

then classify the transformer faults in the five classes existing in our database, according to Table III. This database was composed of data of faulty equipment inspected in service, used in Publication IEC 60599 [2][18]. Additionally, we had available also a database derived from the literature and a database obtained from CELPA (Power Stations of Pará, SA - Brazil).

TABLE III – FAULT TYPES IN THE DATABASE

	Fault Type	Number of samples
T1	Thermal fault - $T < 300^\circ\text{C}$	77
T2	Thermal fault - $T > 300^\circ\text{C}$	71
PD	Partial Discharges (corona)	30
DL	Discharges of Low Energy	37
DH	Discharges of High Energy	103

For the classification of the faults the ANN works as a discriminating function. If six discriminatory lines were created in -0.5, 0.5, 1.5, 2.5, 3.5 and 4.5, the classification of the faults would be according to:

If ANN output is between  $[-0.5 \ 0.5]$  then T1  
 If ANN output is between  $[0.5 \ 1.5]$  then T2  
 If ANN output is between  $[1.5 \ 2.5]$  then PD  
 If ANN output is between  $[2.5 \ 3.5]$  then DL  
 If ANN is output between  $[3.5 \ 4.5]$  then DH

However, to take in account possible cases with not clear symptoms (remember that the IEC table leaves regions with the label “not identified”) we have decided to fuzzify the ANN output using the membership functions of Figure 6. It is important to point out that this fuzzification is reasonable since the neighboring codes of the ANN output represent neighboring faults.

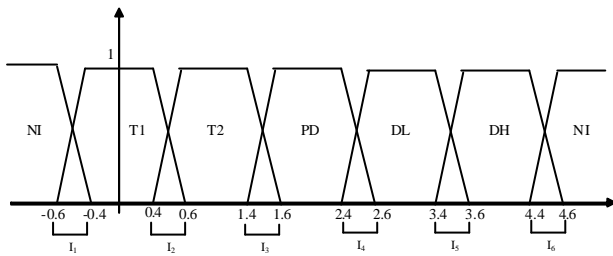


Figure 6. Fuzzy classification of the output of the ANN or the non-transparent FIS as well as the transparent FIS (NI means non-identified).

This fuzzification of the output means that a case may be classified as follows:

*Diagnosis:* Discharge of low energy (possibility = 0.65) or Partial Discharge (possibility = 0.35).

Or, in another case,

*Diagnosis:* Thermal Fault  $T < 300 \text{ }^\circ\text{C}$  (possibility = 1)

Because the membership functions are trapezoidal, there are intervals where the diagnosis only formulates one hypothesis, and intervals of indecision where two hypotheses are kept. This means that the hypothesis with a smaller degree should not be discarded as a possibility.

After training, we have proceeded to the knowledge extraction phase, by using our transform to convert the ANN to a FIS with 60 rules (recalling that the ANN has 60 hidden neurons). These rules formed a non-transparent set and we have further applied the transformation described in the previous section - as we had three inputs, we have been led to a transparent rule set of 125 rules.

From the results provided for all training and testing data, we could see that among the 125 rules there were 30 rules that were never activated and could be excluded from the rule base. Thus, the final transparent fuzzy system has 95 fuzzy rules.

Finally, these fuzzy rules may be transformed in crisp rules, by taking the maximum membership of rule output in each case. Examining the crisp rules, we realized that they could be grouped and we thus formed the final Table IV.

TABLE IV - IMPROVED IEC TABLE

Case	$C_2H_2/C_2H_4$	$CH_4/H_2$	$C_2H_4/C_2H_6$
<b>PD Partial discharge</b>			
	<0.1	<0.1	>2.6 or <0.36
	<0.1	0.1-0.36	<0.1 or >1.06
	0.1-0.36	<0.1	<0.1
	0.1-0.36	0.36-2.6	>2.6
	0.36-1.06	<0.36	<0.36
	0.36-1.06	1.06-2.6	1.06-2.6
	1.06-2.6	<0.36	<0.36
	1.06-2.6	0.36-1.06	0.36-1.06
	>2.6	<0.1	<0.1
	>2.6	0.1-0.36	<0.36
	>2.6	0.36-1.06	0.36-1.06
<b>D1 Discharges of low energy</b>			
	<0.1	<0.1	1.06-2.6
	0.1-0.36	<0.36	>1.06
	0.1-0.36	<0.1	0.36-1.06
	0.36-1.06	0.36-1.06	1.06-2.6
	0.36-1.06	1.06-2.6	>2.6
	1.06-2.6	<0.1	>0.36
	1.06-2.6	0.1-0.36	0.36-2.6
	1.06-2.6	0.36-2.6	1.06-2.6
	>2.6	<0.1	>0.1
	>2.6	0.1-0.36	>0.36
	>2.6	0.36-1.06	>1.06
	>2.6	1.06-2.6	<0.1 or 1.06-2.06
<b>D2 Discharges of high energy</b>			
	0.36-1.06	<0.36	>1.06
	1.06-2.6	0.36-1.06 or 1.06-2.6	>2.6
	0.36-1.06	0.36-1.06	>2.6
	>2.6	1.06-2.6	>2.6
<b>T1 Thermal fault <math>T &lt; 300 \text{ }^\circ\text{C}</math></b>			
	<0.1	0.36-1.06	0.1-1.06
	<0.1	>1.06	<1.06
	0.1-0.36	0.36-1.06	0.1-1.06
	<0.36	>1.06	<1.06
	0.36-1.06	1.06-2.6	0.1-0.36
<b>T2 Thermal fault <math>T &gt; 300 \text{ }^\circ\text{C}</math></b>			
	<0.1	0.1-0.36	0.1-0.36
	<0.1	0.36-1.06	<0.1 or >1.06
	<0.1	>1.06	>1.06
	0.1-0.36	0.36-1.06	<0.1 or 0.36-1.06
	0.1-0.36	>2.6	>1.06
	0.36-1.06	>2.6	<1.06

This Table deserves some comments. First of all, it should be interpreted as a proposal to be carefully examined. It derives from the database analyzed and from the rule system extracted from a ANN. In this respect, it performs well and, in fact, better than the IEC Table, as shown in the following section.

Some adjustments may perhaps still be done in order to present a more compact table. The lines in the Table are, in fact, rules, but these rules are not evenly represented in the database of transformer failures used – there are some rules that are fired by many cases while other are only activated by a few cases.

Another point that deserves further consideration is the interpretation of the table by experts in dissolved gas analysis, in order to confirm the coherence between the knowledge extracted and the scientific knowledge about the phenomena occurring in a transformer. For instance, it is believed that at 300 °C acetylene is not formed; yet, the Table includes rules for high levels of  $C_2H_2/C_2H_4$  ratios.

One must keep in mind that the Table refers to ratios and not to absolute values of gas concentrations. When examining the database, we have confirmed that the cases justifying the existence of such rules corresponded to cases of extremely low concentrations of  $C_2H_2$  and  $C_2H_4$ . Therefore high ratios may appear and the rule is not in contradiction with known physics and chemistry phenomena. It could be argued, however and in these cases, that an option such as in the IEC 60599 Table, of indicating some ratios as NS (non significant) would be preferable. We have decided to present Table IV without these references to preserve the information existing in the database.

### B. Comparing several diagnosis systems

As a result of our work, we have six diagnosis systems to compare: the constrained ANN with positive sigmoid activation functions (which is equivalent to the non-transparent FIS) with fuzzified output (ANN/FIS) and crisp output (ANN/FIS crisp), the TFRENN transparent FIS with fuzzified output (TFRENN) and crisp output (TFRENN crisp), the improved table of rules (IEC improved - Table IV) and IEC 60599 criteria (Table II). This comparison is in Table V, made over a base of 318 cases.

It is obvious that the Intelligent Diagnosis systems all perform better than the crisp sets of rules, and that the IEC 60599 rule set has the worst performance. It is also interesting to notice that the strategy of having a fuzzified output has advantages over a crisp classification. In fact, in all cases where the diagnosis fell on the indecision band, where two hypotheses were produced, always the correct diagnosis has been kept as a possibility. However, just taking the highest possibility as the final diagnosis may lead to error, as the systems with crisp output prove.

Still, the analysis of the transparent rules extracted by TFRENN has allowed knowledge discovery - which led to producing a new table for incipient transformer fault diagnosis that represents an improvement over the Table included in IEC 60599. It is important to underline that the same data were used in the generation of IEC 60599 and of the new Table included in this paper.

TABLE V - COMPARISON OF METHODS

	% Corr	Und	NI	Errors
ANN/FIS	100	5	0	0
ANN/FIS crisp	99.05	0	0	3
TFRENN	100	7	0	0
TFRENN crisp	99.37	0	0	2
IEC improved	95.91	0	8	5
IEC 60599	94.02	0	14	5

The columns in this Table correspond to the following:

- % Corr - Percentage of cases where the output included the correct answer
- Und - No. of cases when the fuzzy output indicated two possibilities, but one of them still was the correct answer
- NI - No. of non-identified cases
- Errors - No. of errors in classification

## VII. CONCLUSION

From the results obtained with TFRENN we could observe the percentage of correct diagnoses was greater than the percentage obtained with IEC 60599. The problem of “no decision”, characteristic of the IEC method and other ratio methods, has been overcome. All the non-identified faults by IEC 60599 were correctly classified, and no errors committed.

With TFRENN, we were able not only to produce a diagnosis system, but a system that may be translated into a set of rules with linguistic expression. This allowed knowledge discovery and a new table, improving IEC 60599, could be built. Although the Intelligent Systems perform better, a rule table may allow better understanding of the diagnosis and the phenomena involved.

The advantage of Fuzzy Systems has also been put to evidence. In fact, the practical examples have shown that when fuzziness is recognized in the linguistic expression of knowledge, the diagnosis of faults in transformers has always pointed out to, or at least (in marginal cases) not excluded, the correct cause of failure, while a crisp representation of answers led, in some cases, to error in the diagnosis. A Fuzzy answer is, therefore, more informative and more correct.

Finally, we produced evidence that the separation of clusters found in Dissolved Gas Analysis is non-linear or, at least, not representable by an interval partition of the domain. This evidence results from the fact that both the Fuzzy Inference Systems and the ANN based systems produced results with accuracy superior to the IEC code. This non-interval-separable data structure makes it more difficult for humans to apprehend knowledge and reinforces the usefulness of having an approach involving intelligent systems.

## ACKNOWLEDGMENT

The authors are grateful to GRICES (Portugal) and CAPES (Brazil) for the support brought to the NESC/INESC cooperation protocol, which allowed Mrs. Castro to develop the research reported.

## REFERENCES

- [1] ANSI/IEEE Std C57.104.1991, IEEE Guide of gases generated in oil-immersed Transformer, IEEE Power Engineering Society, 1992.
- [2] IEC Publication 60599, "Interpretation of the analysis of gases in transformers and other oil-filled electrical equipment in service", March 1999.
- [3] Yann-Chang Huang. "Evolving Neural Nets for fault Diagnosis of Power Transformer", IEEE Transactions on Power Delivery, Vol 18, N° 3, pp 843-848, July, 2003.

- [4] Y. Zhang, X. Ding, Y. Liu and P. J. Griffin, "An Artificial Neural Approach to Transformer Fault Diagnosis", IEEE Transactions on Power Delivery, Vol 11, N° 4, pp 1836-1841, October, 1996.
- [5] K. Tomsovic, M. Tapper and T. Ingvarsson, "A Fuzzy Information Approach to Integrating different Transformer Diagnostic Methods", IEEE Transactions on Power Delivery, Vol. 8, N° 3, pp 1638-1644, July, 1993.
- [6] Y. Zhang, X. Ding, Y. Liu and P. J. Griffin, "An Artificial Neural Approach to Transformer Fault Diagnosis", IEEE Transactions on Power Delivery, Vol 11, N° 4, pp 1836-1841, October, 1996.
- [7] Y-C Huang, H-T Yang and C-L Huang, "Developing a New Transformer Fault Diagnosis System through Evolutionary Fuzzy Logic", IEEE Transactions on Power Delivery, Vol. 12, N° 2, pp 761-767, April 1997.
- [8] Z. Wang, Y. Liu, and P. J. Griffin, "A combined ANN and expert system tool for transformer fault diagnosis", IEEE Transactions on Power delivery, vol.13, pp 1224-1229, October 1998.
- [9] H-T Yang and CC Liao, "Adaptive Fuzzy Diagnosis System for Dissolved Gas analysis of Power Transformers", IEEE Transactions on Power Delivery, Vol 14, N°. 4, pp 1342-1350, October 1999.
- [10] J.L Guardado, J.L Naredo, P. Moreno, and C.R. Fuerte, "A Comparative study of neural networks efficiency in power transformers diagnosis using dissolved gas analysis", IEEE Transactions on Power delivery, Vol. 16, pp 643-647, October 2001.
- [11] Yann-Chang Huang, "Evolving Neural Nets for fault Diagnosis of Power Transformer", IEEE Transactions on Power Delivery, Vol 18, N° 3, pp 843-848, July, 2003.
- [12] K. Hornik, M. Stinckcombe, and H. White, "Multilayer feedforward networks are Universal Approximators", Neural Network, Vol. 2, pp:359-366, 1989.
- [13] J. M. Benitez, J. L. Castro and I. Requena, "Are Artificial Neural Networks Black Boxes?", IEEE Transactions on Neural Networks, Vol 8, N° 5, pp 1156-1164, September, 1997.
- [14] Adriana R.G. Castro, Vladimiro Miranda, "Mapping Neural Networks Into Rule Sets and Making Their Hidden Knowledge Explicit – Application to Spatial Load Forecasting", *Proceedings of PSCC02 - 14th Power Systems Computation Conference*, Sevilla, Spain, June 2002
- [15] Adriana R.G. Castro, "Knowledge Extraction from Artificial Neural Networks - an Application to Transformer Fault Diagnosis", Ph.D. Thesis, Ed. FEUP - Faculty of Engineering, University of Porto, Portugal, July 2004
- [16] Li-Xin Wang, A Course in Fuzzy Systems and Control. PrenticeHall, 1997.
- [17] A. Riid and E. Rustern, "Transparent Fuzzy Systems and Modeling with Transparency Protection", In Proc. IFAC Symposium on Artificial Intelligence in real Time Control Three Control", pp 229-235, October, 2000.
- [18] M. Duval and A. Pablo, " Interpretation of Gas-in-oil Analysis using new IEC Publication 60599 and IEC TC10 Databases", IEEE Electrical Insulation Magazine, March/April, Vol17, N°2, pp 31-41.

**Adriana Rosa Garcez Castro** graduated and obtained her M.Sc. degree in Electrical Engineering from UFPA (Federal University of Pará, Brazil) in 1992 and 1995. She finished her Ph.D. degree at INESC Porto and FEUP (Faculty of Engineering of the University of Porto, Portugal) in 2004. She is presently lecturer at UFPA. Her interests are in Power Systems and Control and the application of Computational Intelligence techniques.

**Vladimiro Miranda** received his graduation, Ph.D. and Agregado degrees from the Faculty of Engineering of the University of Porto, Portugal (FEUP) in 1977, 1982 and 1991, all in Electrical Engineering. In 1981 he joined FEUP and currently holds the position of Professor Catedrático. He is also currently Director of INESC Porto. He has authored many papers and been responsible for many projects in areas related with the application of Computational Intelligence to Power Systems.