Transformer fault diagnosis based on autoassociative neural networks

Adriana R. Garcez Castro, Vladimiro Miranda, *Fellow IEEE* and Shigeaki Lima

***Abstract* – This paper presents a new approach to incipient fault diagnosis in power transformers, based on the results of dissolved gas analysis. A set of autoassociative neural networks or autoencoders are trained, so that each becomes tuned with a particular fault mode. Then, a parallel model is built where the autoencoders compete with one another when a new input vector is entered and the closest recognition is taken as the diagnosis sought. A remarkable accuracy is achieved with this architecture, in a large data set used for result validation.**

***IndexTerms* – Auto-associative networks, transformer failure, failure diagnosis.**

I. Introduction

Transformer fault diagnosis via dissolved gas [1] analysis (DGA) is a problem that has deserved constant attention from many researchers, because of the economic importance of the early detection of evolving faults that may lead to equipment failure. Furthermore, it is a problem that was object of an IEC norm (IEC 60599 [2]) and a data base for diagnosed failures denoted IEC TC10 is available for researchers to study and develop new models [3].

Thus, it should not be surprising that a number of techniques have been proposed to deal with this problem. Knowledge may be extracted from the available data base and then applied to test cases to validate diagnosis models proposed. Among these, one may find expert systems [4], fuzzy set models [5], multi-layer feedforward artificial neural networks (ANN) [6][7], wavelet networks [8], hybrids fuzzy sets/ANN [9], radial basis function neural networks [10], Support Vector Machines (SVM) [11], Self-Organizing Maps (SOM) or Kohonen Neural Networks [12] are just some of the techniques that one finds in the literature, and this is just an enumeration of examples and not an exhaustive listing. Furthermore, benchmarking is always possible against the set of diagnosis rules in [2].

This paper describes a new approach to the problem which matches the highest accuracy with simplicity – a diagnosis system based on autoassociative neural networks, or simply autoencoders, which are special feedforward neural networks designed and trained in such a way that the output reproduces the input.

Adriana Castro is with UFPA – Federal University of Pará, Brasil (adcastro@ufpa.br).

V. Miranda is with INESC Porto – Institute of Systems and Computer Engineering of Porto, and also with FEUP, Faculty of Engineering of the University of Porto, Portugal (vmiranda@inescporto.pt).

Shigeaki Lima is with UFMA – Federal University of Maranhão, Brasil, and with INESC Porto, Portugal (shiilima@inescporto.pt)

The training has the specific purpose of making the autoencoder learn the non-linear manifold where data lies. Once learned, the autoencoder may be used as recognition machine – if a new data vector belongs to the manifold, the autoencoder will produce a small error; however, if this vector does not lie on the manifold (which should be the case if the new input vector is distinct from the global pattern of the data used for training), the autoencoder will return in the output a result not matching the input and the error will be high.

This property is used in the model proposed in this paper. First, for each fault type, a specific autoencoder is trained so that it learns this fault's characteristics. Then, when data for an unknown fault type is considered, each autoencoder for each type of fault will try to match output to input – but, hopefully, only one will stay tuned while all the other will display large errors. The fault is thus identified by recognizing which autoencoder presents minimum error.

II. Autoencoders

Auto-associative neural network encoders, or simply *autoencoders*, are feedforward neural networks that are trained to reproduce the input space S in the output. Because some inner layer will have a number of neurons n different to the set of m inputs/outputs, this layer will effectively be encoding variables from S into a different dimension space S'. There is no theoretical limitation on the architecture of autoencoders, either on the number of neurons or the number of layers. The simplest architecture keeps only one middle hidden layer. Fig. 1 illustrates this case, when dim(S) > dim(S'). Naturally, the input and output layers are of equal size.

An autoencoder can be split in two halves: the first half approximates a function f that maps the input space S onto S' while the second half approximates the inverse function f^{-1}.

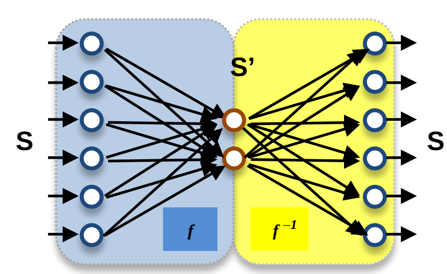

Fig. 1. An autoassociative neural network or autoencoder, with input and output layers of the same dimension and a different middle layer. If trained to reproduce the input variables in the output, one has in the middle layer a set of values that encode, in a different space S', the values in S.

Once properly trained, the autoencoder learns the data supporting manifold by memorizing, in its weights, its projection onto a different space and the inverse projection. When the hidden layer has a smaller dimension than the outer layers, one achieves an effective feature space reduction whenever dim(S) > dim(S'). This has been used to build data compression machines [13][14][15][16].

A point over the data manifold will be correctly projected back and forth by a trained autoencoder; however, a point not lying on the data manifold will not be correctly re-projected back and a large error ε will be detected between input and output. This property may be used in pattern recognition and classification, as well as in novelty detection [17].

The classical training of autoencoders uses the same tools as any other feedforward network, namely backpropagation. However, because of the special architecture of an autoencoder with a middle bottleneck, experience has shown that training may be difficult and the tuning of the weights in the first layers may require many epochs and extreme care in choosing the departing values. This is why, especially in architectures with several layers, specific training procedures have been developed, such as referred to in [13].

It has been shown that autoencoders with linear activation functions produce a mapping in the inner layer (space S') equivalent to Principal Component Analysis (PCA) [18]. This means that information is condensed along orthogonal axes such that variance is minimized. There is nevertheless loss of information. When the activation functions are non-linear (sigmoidal), it has been shown that the mapping is not equivalent to PCA and has better characteristics [19]. Of course, it is always assumed that the training is done in supervised mode, adopting a Minimum Square Error criterion. If **X** is the input vector and **Y** the output vector, then for N samples

$$MSE: \quad \min \varepsilon = \frac{1}{N}\sum_{k=1}^{N} \|\boldsymbol{X_k} - \boldsymbol{Y_k}\|^2 \qquad (1)$$

MSE is, of course, the minimization of the variance of the pdf (probability density function) of the error distribution – but is it known that this criterion is only optimal if this distribution is Gaussian. Therefore, the adoption of this criterion may be challenged whenever the error distribution cannot be assumed as Gaussian. A non-parametric method should be preferred.

III. Fault diagnosis data in power transformers

The process of detecting incipient failures in power systems usually departs from monitoring the evolution rate of dissolved gases in the oil. When this change is considered significant, possible fault is investigated through the DGA (dissolved gas analysis) technique and upon confirmation of the suspicion, other procedures are put in place namely to locate the fault.

One of the well-known diagnosis methods is the one described in the norm IEC 60599 [2] and summarized in Table 1. These rules, when applied to the transformer data set IEC TC10, lead to a number of mistaken classifications plus a number of non-classified patterns (non-identified failures).

Table 1 – IEC 60599 fault diagnosis rules

Case	Fault type	$\frac{C_2H_2}{C_2H_4}$	$\frac{CH_4}{H_2}$	$\frac{C_2H_4}{C_2H_6}$
PD	Partial discharge	NS	<0.1	<0.2
D1	Low energy discharge	>1	0.1-0.5	>1
D2	High energy discharge	0.6-2.5	0.1-1	>2
T1	Thermal fault – T<300°C	NS	>1 but NS	<1
T2	Thermal fault 300°C< T<700°C	<0.1	>1	1-4
T3	Thermal fault – T>300°C	<0.2	>1	>4

It is usual to lump together the cases T2 and T3 in many studies, because the number of cases in the data base is too small to an adequate training.

As far as quality or diagnosis accuracy is concerned, Table 2 presents some results of a few systems developed for transformer diagnosis based on DGA. One must bear in mind that a direct comparison of percentage of hits/misses should take in account that the data sets used were not the same in all works reported.

Table 2 – Results in different systems/publications

Model	No. samples in the database	% of correct diagnoses of the developed system
Zhang et al [6]	(?)	95% Classification of the 3 main faults
Huang et al [7]	711	90.30% -training set 93.81% - testing set
Wang [20]	210	99.72% - training set 95.34% - testing set
Liao et al [21]	711	96.2%
Guardado et al [22]	33	ANN training considering: Dornemburg method – 90.91% Modified Roger – 87.88% Roger – 90.91% IEC – 93.94%
Huang [23]	820	90.49% and 93.54%, depending on the number of inputs Classification of only 4 fault types
Castro et al [24]	431	100.0% -training set 97.84% - testing set
Miranda et al [9]	431	99,37% in IEC TC10 data

In the work reported in this paper, the data base from [3] was used, complemented with data from other origin, comprising 318 cases, from which 230 were selected to constitute a training set and the remainder 88 were used as validation or test set. Each sample in the data base includes information of dissolved gas concentration of H_2 (hydrogen), CH_4 (methane), C_2H_6 (ethane), C_2H_4 (ethylene) and C_2H_2 (acetylene) as well as the verified condition of the transformer. In order to have sets with minimally meaningful sizes, the types of faults were organized as in Table 3 (five fault types).

TABLE 3 – SAMPLES GROUPED FOR TRAINING AND IDENTIFICATION

	Type of Fault	No. of samples
T1	Thermal fault – T<300°C	77
T2	Thermal fault – T>300°C	71
PD	Partial discharge	30
DL	Low energy discharge	37
DH	High energy discharge	103

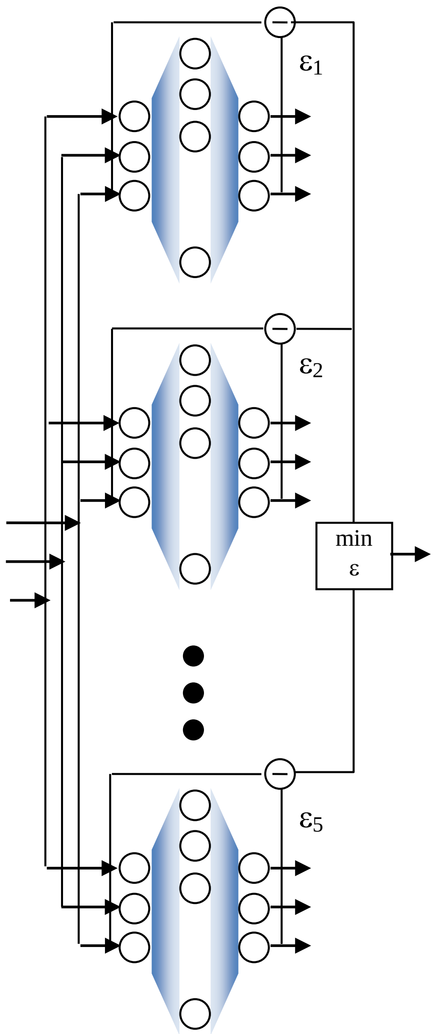

Fig. 2. General architecture of the new diagnosis system, based on a set of autoassociative neural networks in parallel, each tuned for a specific fault type, and generating competing outputs[25] .

IV. NEW DIAGNOSIS SYSTEM CONCEPT

Most automatic diagnosis systems based on neural networks and similar approaches rely on a single system that performs classification. When activated by a sample at its input, they produce an output signal indicating the proposed fault classification. In this paper, we describe a distinct and more successful approach.

The new idea behind the diagnosis system is to tune an independent autoassociative network for each cluster of data and then, for an unclassified sample, have the tuned autoassociative networks competing for the identification of the fault.

In line with Table 2, for this model one requires 5 distinct autoencoders, one for each fault. The input vectors were specified as being with exactly the same composition as used in IEC 60559 norm, meaning that we used the concentration ratios $(C_2H_2)/(C_2H_4)$, $(CH_4)/(H_2)$ and $(C_2H_4)/(C_2H_6)$.

Fig. 2 illustrates the competitive parallel architecture for the diagnosis system. Each autoassociative neural network is trained to learn the manifold where data for a specific fault lie, returning the same vector if a new case for the same fault is input and returning a vector with a large deviation from input if a different case is input.

Thus, when a gas concentration ratio vector is input, it is expected that only one autoencoder will display a small error, recognizing that the vector is close to a particular learned manifold, while the other autoencoders will not be able to display, at their output, a close reproduction of the input. Therefore, when the competing autoencoders present to the decision module their error value, the winner is the one with minimum error.

V. TRAINING AND TESTING

For the transformer diagnosis system, the error in each autoencoder was calculated as in Eq. (1), which is equivalent to a Euclidean distance between the two vectors (input and output). Each autoencoder was designed with 3 neurons in the input and output layers, corresponding to the 3 gas ratios, and a hidden layer with dimension 15. So, instead of a bottleneck on the middle layers, one generates a projection into a higher dimension space. With only 3 inputs, it would be senseless to compress the data into 2 or 1 dimension in the middle layer.

The activation functions used in the input and hidden layers were hyperbolic tangents, while in the output layer each neuron had a linear activation function. The training procedure adopted the Levenberg-Marquardt algorithm and was performed in Matlab.

In Table 4 one finds the results obtained with the new system, as opposed to the ones obtained when applying IEC 60599 to the same data. It is remarkable that no errors or misclassification were produced by the new system (318 hits in 318 cases!). The training and test sets were not especially doctored, except that there was care in having a training set covering as evenly as possible the domain. From the results of applying the IEC 60599 model, one may see that the validation set was not especially easy to diagnose.

TABLE 4 – RESULTS AND DIAGNOSIS ACCURACY COMPARISON

Model	% correctly identified faults in the training set	% correctly identified faults in the validation set	No. of non-identified faults	No. cases with wrong diagnosis
IEC 60599	95.65	89.77	14	5
Castro-Miranda Autoencoder	100 %	100 %	0	0

TABLE 5 – SAMPLES FROM THE IEC TC10 DATABASE AND PERFORMANCE COMPARISON BETWEEN IEC 60599 AND THE AUTOENCODER DIAGNOSIS SYSTEM

$\frac{C_2H_2}{C_2H_4}$	$\frac{CH_4}{H_2}$	$\frac{C_2H_4}{C_2H_6}$	Fault	Auto-encoder	IEC 60599
0.0417	1.1628	0.4444	T1	T1	T1
0.0164	1.0494	0.1452	T1	T1	T1
0.3333	4.0	0.1765	T1	T1	T1
0.0098	2.7497	4.0	T2	T2	T2
0.0412	1.8218	4.0	T2	T2	T2
0.0198	1.8438	4.0	T2	T2	T2
0.0001	0.1102	0.0001	PD	PD	NI
1.1667	0.1065	0.1000	PD	PD	NI
0.0001	0.0476	0.0001	PD	PD	PD
1.0	0.1667	1.0	DL	DL	NI
4.0	0.1607	4.0	DL	DL	DL
2.2233	0.125	1.0605	DL	DL	DL
0.6667	0.2250	4.0	DH	DH	DH
0.8800	0.1650	4.0	DH	DH	DH
0.6818	0.3950	3.1429	DH	DH	DH

To illustrate this result with a few examples, Table 5 presents some classification results, with the correct fault identification, the diagnosis produced by the autoencoder model and the result provided by IEC 60599.

The new system displays an absolute superiority over IEC 60599 and is better than any result reported and summarized in Table 4, where the best result is the one referenced in [9], resulting from 2 errors in the data from IEC TC10, but deriving from a larger database. In relation to IEC 60599, the autoencoder model was able to solve and correctly identify all undecided cases produced by this method.

The 100% hit is, indeed, a remarkable result. It can only be explained by the capacity of the autoencoders to really learn distinct manifolds for the distinct sets or clusters of data.

VI. CONCLUSIONS

The work reported in this paper presents an original model for incipient transformer fault diagnosis, which is based on a novel application of autoassociative neural networks.

The main idea behind the new system if to take advantage of the property of autoencoders that allows them to learn the manifold where data lie, by projecting inputs to a different space and reprojecting back to the input space. One can therefore tune an autoencoder to a particular fault mode. When activated by a new input vector, an autoencoder will reproduce it in its output with very small error if the input corresponds to the fault for which it was trained, otherwise the output will display a large dissimilarity with input.

The results presented show that, in the transformer fault diagnosis based on DGA, autoencoders do discriminate among the distinct fault modes and are able to pinpoint faults with a remarkable accuracy – so much so that the results obtained surpassed all published results referenced.

The architecture proposed for fault diagnosis is completely general and its application is not restricted to transformer fault diagnosis. It may be suitable for any problem with data in enough quantity to allow proper training of the autoassociative neural networks.

ACKNOWLEDGMENT

This work is partially integrated in project LASCA PTDC/EEA-EEL/104278/2008 and in project GEMS PTDC/EEA-EEL/105261/2008, both financed by FCT, Portugal.

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] M. Duval and A. Pablo, "Interpretation of Gas-in-oil Analysis using new IEC Publication 60599 and IEC TC10 Databases", IEEE Electrical Insulation Magazine, vol. 17, no. 2, pp. 31–41, March/April 2001.

[4] 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.

[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, no. 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, no. 4, pp 1836–1841, Oct. 1996.

[7] Y.-C. Huang. "Evolving Neural Nets for fault Diagnosis of Power Transformer", IEEE Transactions on Power Delivery, vol. 18, no. 3, pp 843–848, July 2003.

[8] L. Honglei, X. Dengming and C. Yazhu, "Wavelet ANN Based Transformer Fault Diagnosis Using Gas-in-Oil Analysis", in *Proc. of the 6th International Conference on Properties and Applications of Dielectric Materials,* 2000, pp. 147–150

[9] A. R. G. Castro and V. Miranda, "Improving the IEC Table for Transformer Failure Diagnosis with Knowledge Extraction from Neural Networks, IEEE Transactions on Power Delivery, vol. 20, pp. 2509–2516, Oct. 2005.

[10] J. P. Lee, D. J. LEE, P. S. Ji, J.Y. Lim and S. S. Kim, "Diagnosis of Power Transformer Using Fuzzy Clustering and Radial Basis Function Neural Network", 2006 International Joint Conference on Neural Networks, pp. 1398–1404, July 2006

[11] D-H Liu, J.-P. Bian and X.-Y. Sun, "The Study of Fault Diagnosis Model of DGA for Oil-Immersed Transformer Based on Fuzzy Means Kernel Clustering and SVM Multi-Class Object Simplified Structure", in Proc. of the Seventh International Conference on Machine Learning and Cybernetics, Kunming, July, 2008, pp. 12–15.

[12] K. F. Thang and R. K. Aggarwal, "Analysis of Power Transformer Dissolved Gas Data Using the Self-Organizing Map", IEEE Transactions on Power Delivery, vol. 18, no. 4, pp. 1241, October 2003.

[13] G. E. Hinton and R. R. Salakhutdinov, "Reducing the Dimensionality of Data with Neural Networks", *Science*, Vol. 313, no. 5786, pp. 504 – 507, July 2006

[14] G.W. Cottrell, P. Munro and D. Zipser**, "**Learning internal representations from gray-scale images: An example of extensional programming", *Proceedings of the Ninth Annual Conference of the Cognitive Science Society*, Seattle (WA), USA, 1987

[15] M.K. Fleming, G.W. Cottrell, "Categorization of faces using unsupervised feature extraction", *Proceedings of IJCNN - International Joint Conference on Neural Networks*, vol. 2, pp. 65-70, San Diego (CA), USA, 17-21 Jun 1990

[16] B. Golomb and T. Sejnowski, “Sex Recognition from Faces Using Neural Networks”, in A. Murray (ed.), in *Applications of Neural Networks*, pp. 71-92, Kluwer Academic Publishers, 1995
[17] B. B. Thompson, R.J. Marks II, J.J. Choi, M.A. El-Sharkawi, M.Y. Huang and C. Bunje , " Implicit learning in autoencoder novelty assessment", *Proceedings of the 2002 International Joint Conference on Neural Networks*, 2002, IEEE World Congress on Computational Intelligence, Honolulu (Hawaii), USA, pp. 2878-2883, May12-17, 2002.
[18] Sanger, T.D., "Optimal unsupervised learning in a single-layer linear feedforward neural network", *Neural Networks*, Vol. 2, pp. 459-473, 1989
[19] N. Japkowicz, S.J. Hanson, M.A. Gluck, "Nonlinear Autoassociation is not Equivalent to PCA", *Neural Computation*, Vol. 12 , Issue 3, pp. 531 - 545, March 2000
[20] 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.
[21] H-T Yang and C-C Liao, “Adaptive Fuzzy Diagnosis System for Dissolved Gas analysis of Power Transformers”, IEEE Transactions on Power Delivery, Vol 14, Nº. 4, pp 1342-1350, October 1999.
[22] J.L Guardado, J.L Naredo, P. Moreno, adn 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.
[23] Yann-Chang Huang. “Evolving Neural Nets for fault Diagnosis of Power Transformer”, IEEE Transactions on Power Delivery, Vol 18, Nº 3,pp 843-848, July, 2003.
[24] A. Castro and V. Miranda, " Knowledge Discovery in Neural Networks with Application to Transformer Failure Diagnosis", IEEE Transactions on Power Systems, Vol. 20, no.2, pag. 717-724, May 2005
[25] A. Castro and V. Miranda, "Sistema Inteligente para Diagnóstico de Faltas Incipientes em Transformadores baseado em Redes Neurais Auto-Associativas" (in Portuguese), Proceedings of SBSE 2010 – Simpósio Brasileiro de Sistemas Elétricos, Belém (PA), Brasil, May 2010

Vladimiro Miranda (M’90, SM’04, F’05) received his B.Sc. and Ph.D. degrees from the Faculty of Engineering of the University of Porto, Portugal (FEUP) in 1977 and 1982 in Electrical Engineering. In 1981 he joined FEUP and currently holds the position of Full Professor. He is also a researcher at INESC since 1985 and is currently Director at INESC Porto, an advanced research institute in Portugal. He has authored many papers and been responsible for many projects in areas related with the application of Computational Intelligence to Power Systems

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 Assistant Professor at UFPA. Her interests are in Power Systems and Control and the application of Computational Intelligence techniques.

Shigeaki Lima graduated and obtained his M.Sc. degree in Electrical Engineering from UFMA (Federal University of Maranhão, Brazil). He is a PhD student at UFMA and is developing his thesis work in 2010/2011 at INESC Porto, Portugal .