INTRODUCTION

National Institute of Technology, Kurseong, India

Transformer is one of the most important but complex component of electricity transmission and distribution system. The trend toward a deregulated global electricity market has put the electric utilities under severe stress to reduce operating costs, enhance the availability of the generation, transmission and distribution equipment and improve the supply of power and service to customers. Much attention is needed on maintenance of transformers in order to have fault free electric supply and to maximize the life and efficiency of a transformer. Thus, it is important to be aware of possible faults that may occur. It is equally important to know how to detect them early.

1.1 Formation of Gases in Transformer Oil: The faults that occur within the transformer protection zone are internal faults. Transformer internal faults can be divided into classification: internal short circuit faults and internal incipient faults. Incipient fault detection in power transformer can provide information to predict failures ahead of time so that the necessary corrective actions are taken to prevent outages and reduce downtime. Incipient faults can produce hydrocarbon molecules and carbon oxides due to the thermal decomposition of oil, cellulose, and other solid insulation. Because the insulating oil used in power transformer is organic (i.e., composed primarily of hydrocarbons), certain fingerprint gases are generated at specific temperature ranges therefore, allowing the traditional methods to identify a possible fault temperature range and therefore the possible fault type. In the normal operation of the transformers, the released gases: Hydrogen (H₂), methane (CH₄), ethane (C₂H₆), ethylene (C₂H₄), acetylene (C₂H₂), carbon monoxide (CO), carbon dioxide (CO₂) and so on are small in quantities as ageing effects. When there is an abnormal situation, as occurring a fault, some specific gases are produced more than in normal operation and the amount of them in transformer oil increases. This decreases the insulation properties of the transformer oil.

1.2 Dissolved Gas Analysis: The faults in power transformers can be detected and monitored for abnormal conditions with dissolved gas analysis (DGA). Throughout the world, different countries/utilities are using different fault interpretation techniques/tools to diagnose the faults; According to the pattern of the gases composition, their types and quantities, the conventional interpretation approaches below for dissolved gases are extensively followed as:

- Key Gas Ratio
- IEEE Gas Guide C57.104™- 2008
- Dorenenberg Gas Ratio
- Roger’s Gas Ratio
IEC Standards 60599
IS 10593: 1992 Standards

At first a sample of transformer oil is taken. Then the dissolved gases is extracted, separated and measured by means of chromatography. In order to interpret the results of experiment a data in suitable form to diagnose the faults is produced. The forming of the data is based on different standards.

1.3 Computer Applications for Fault Diagnosis: Technology has become very dynamic over the past few years. Computers are the prime components for this whole revolution. Computers have found wide spread applications 'human-like abilities', capabilities to make judgments, guesses, change of opinions in fault diagnosis of power transformers in last two decades. Various computer based artificial intelligent (AI) techniques have been proposed for transformer fault diagnosis; however they have some limitations such as accuracy of diagnosis, requirement of inference rules, and determination of the detection architecture. Above stated conventional interpretation techniques/standards are applied to diagnose internal faults includes; thermal faults, electrical faults, and faults with cellulosic insulation degrading through computer applications. Broadly, computer applications as artificial intelligence are divided in two groups:

A. Individual Artificial Intelligence Techniques
B. Hybrid Artificial Intelligence Techniques

2. INDIVIDUAL ARTIFICIAL INTELLIGENCE TECHNIQUES

2.1 Fuzzy Logic/Fuzzy Inference System: [1]-[15]

The fuzzy analysis consists of three parts: fuzzification, fuzzy inference and defuzzification. Fuzzification is the process of transforming crisp input values into grades of membership for linguistic terms of fuzzy sets. The membership function is used to associate a grade to each linguistic term. A chosen fuzzy inference system (FIS) is responsible for drawing conclusions from the knowledge-based fuzzy rule set of if-then linguistic statements. Fault types are first listed and then form the fuzzy rule set for the diagnosis system. Defuzzification then converts the fuzzy output values back into crisp output actions.

For example, the IEC codes for three gases C\textsubscript{2}H\textsubscript{4}/C\textsubscript{2}H\textsubscript{6}, CH\textsubscript{4}/H\textsubscript{2} and C\textsubscript{2}H\textsubscript{4}/C\textsubscript{2}H\textsubscript{4} are labeled by codes of 0, 1 and 2, respectively. To make clear the relationship between the range of each gas ratio and its corresponding codes rearranged in Table 1 and Table 2 show the new arrangement of the code.

| Table 1 The Rearrangement of IEC Codes |
| Gas Ratio | Code 0 | Code 1 | Code 2 |
| X=C\textsubscript{2}H\textsubscript{2}/C\textsubscript{2}H\textsubscript{6} | X=CH\textsubscript{4}/H\textsubscript{2} | Z=C\textsubscript{2}H\textsubscript{4}/C\textsubscript{2}H\textsubscript{6} |
| X<0.1 | 0.1<X<3 | Z<1 |
| X>3 | Y<0.1 | 1<Z<3 |
| Y>1 |

One of the problems in fault diagnosis of transformer based on dissolved gas is lack of matching the result of fault diagnosis of different standards with the real world. The result of the different standards is analyzed using fuzzy and the result is compared with the empirical test. The comparison between the suggested method and existing methods indicate the capability of the suggested method in on-line fault diagnosis of the transformers. In addition, in some cases the existing standards are not able to diagnose the fault. However, there are situations of errors and misleading results occurring due to borderline and multiple faults. Fuzzy logic is implemented here as an improved DGA interpretation method that provides higher reliability and precision of fault diagnosis.

2.2 Artificial neural network (ANN) method [16]-[26]

Artificial Neural Networks is a massively parallel distributed processor, having a natural tendency to acquire sufficient experimental knowledge and making it available for use.

DESIGN OF ANN
Among these AI methods, the ANN is widely designed to diagnose transformer faults. An important advantage of fault diagnosis based on the ANN is that it can learn directly from the training samples and update its knowledge when necessary. The high nonlinear mapping capability of neurons always provides a comparable and superior performance over a fuzzy system solution. However, although the computational complexity of the ANN is not too high, especially in the fault diagnostic process, it involves certain problems including slow convergence, oscillation and so on, all of which must be resolved before it is practically applied. 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 reflect the characteristics of the problem. Another major task of the ANN design is to choose network topology. Five key gases H₂, CH₄, C₂H₆, C₂H₄, and C₂H₂ are chosen as input features [22]. 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.

**TRAINING OF ANN**

The back propagation (BP)-based artificial neural nets (ANN) can identify complicated relationships among dissolved gas contents in transformer oil and corresponding fault types, using the highly nonlinear mapping nature of the neural nets in figure 1. An efficient BP-ALM (BP with Adaptive Learning Rate and Momentum coefficient) algorithm is proposed to reduce the training time and avoid being trapped into local minima, where the learning rate and the momentum coefficient are altered at iterations.

![Two-layered Feedforward Neural Network](image1)

Fig. 1. Two-layered Feedforward Neural Network [22]

**2.3 Expert Systems: [27]-[29]**

Expert system is one of the areas of Artificial Intelligence (AI) in computer application, which has moved out from research laboratory to the real word and has shown its potential in industrial and commercial applications. An expert system is a computer system which can act as human expert within one particular field of knowledge. The expert system embodies knowledge about one specific problem domain and possesses the ability to apply this knowledge to solve problem domain. Ideally the expert system can also learn from its mistakes and gain experience from its successes and failures. For the development of any expert system, there should be proper selection of a development tool. The different packages i.e., VP-Expert, Shell, Rule master, etc. can also be used for development, but these packages have their own limitations, since they use their own rules and instructions. But a computer language is more flexible and the user can develop his methodology for the program formulation. So instead of using package, we can use computer language for expert system development and a generalized flow-chart as given in figure 2. The language chosen should be simple and declarative. ‘Turbo prolog’ has these facilities. One of the major advantages of prolog is that it has its own inference engine, which facilitates easy development of expert system. Therefore, prolog has been used for the development of proposed expert system. System diagnosis is proposed to assist the situation, which cannot be handled properly by gas ratio methods. Results from the implementation of the expert system shows [29] that the expert system is a useful tool to assist human experts and maintenance engineers. The knowledge base of this expert system is incorporated within the standard interpretative methods of DGA.
The effectiveness of the state detection expert systems depends on the precision and knowledge base, which is usually very complicated and must be constricted carefully. Such an expert system can neither acquire knowledge from new data samples through self-learning process and nor can it adjust its diagnostic rules automatically.

2.4 Cerebellar model articulation controller based (CMAC) method: [30]

Albus proposed a neural model called CMAC (Cerebellar Model Articulation controller), which like the models of human memory, perform a reflexive processing. The CMAC, in a table look-up fashion, produced a vector output in response to a state vector input. Figure 3 shows a basic configuration of CMAC network for Key gas ratio. Refer to the IEC std. 60599 the key gas ratio of $C_2H_2/C_2H_4$, $CH_4/H_2$, and $C_2H_6/C_2H_4$ are used as the input states. The diagnosis system contains 9 parallel memory layers and every memory layer has one output node. Every memory layer remembers one fault type feature.

Key-gas ratio methods have been adopted by many utilities and different manufacturers. Various faults occupy different lower and upper limits are given IEC’s key-gas ratios. Crisp ranges are used to identify the thermal and electrical faults. However, it is difficult to describe the uncertain boundaries of the gas ratios such as the threshold values 0.1, 0.3, 1.0, 3.0, and 3.3 in the numerical data; determination of the specific codes of gas ratio becomes contentious. GCA is introduced to develop the diagnostic procedure. Like the Fuzzy approach, different ranges of gas ratios are described as small, medium, and large values of gas ratios by the weighted functions, the so-called “whiten-weight function”. The common ones are triangular and trapezoidal functions. The trapezoidal functions are used, which can be represented, where $a$ and $a_1$ are the lower boundaries ($a \leq a_1$); $b$ and $b_1$ are the upper boundaries ($b_1 \leq b$). According to Key –gas ratios (IEC/IEEE and CIGRE standard criteria), whiten weight functions for pre-selected key-gas ratios $x_1$ to $x_4$ are defined by boundary parameters as shown in Figure 4 below.

2.5 Grey Clustering Analysis (GCA) method: [31], [32]

To develop an assistance tool, the diagnostic algorithm must be easy to implement in the portable device and hardware device with a compact configuration. The grey theory provides the applications of clustering analysis, relational analysis, predication, and decision for the grey system. The so-called “grey” means that system information is incomplete, unclear, and uncertain. It is a useful method to deal with the problems of limited, deficient, and or no rules available for data processing. Its analysis makes use of minor data and does not demand strict statistical procedures and inference rules. GCA has a function of mathematical operation for processing numerical data without adjusting any parameter.

![Fig. 3. The Configuration of CMAC-based Fault Diagnosis of power transformer [30]](image-url)
2.7 Other Emerging Individual AI techniques:
In last two decades, some more individual AI techniques have been used by researchers. Few of them are; Particle Swarm Optimizer (PSO) [35], Wavelet Networks [36] & Evolving Wavelet Networks [37], Self-organizing Polynomial Networks [38], Organizing-map Algorithm [39], Support Vector Machine (SVM) [40], Pattern-based [41], Data Mining approach [42], Extension Theory [43], Bayesian Network [44], Kernel-based Possibilistic [45], etc. But unfortunately, these individual AI techniques could not be tested and utilized much by the researchers/users for fault diagnosis in power transformers. Their advantages have proved better results in hybrid AI techniques.

3. HYBRID ARTIFICIAL INTELLIGENCE TECHNIQUES

3.1 Neuro- Fuzzy techniques: [46]-[54]
In recent years, many researchers have studied the application of artificial intelligence, such as neural networks and fuzzy set theory to increase diagnosis accuracy and proved to be better diagnostic tools. The fuzzy systems, though good at handling uncertainties, could not learn from previous diagnosis results and, hence, are not able to adjust the diagnostic rules automatically [1]-[15]. To account for uncertainties, the artificial neural networks (ANNs) have been proposed to diagnose the transformer’s faults because of their superior learning capabilities [16]-[26]. In general, fuzzy systems and neural networks deal efficiently with two different areas of information processing. Fuzzy systems are good at various aspects of uncertain knowledge representation, while neural networks are efficient structures that are capable of learning from examples. Both techniques complement each other. A combination of neural network and fuzzy system [47]-[54] are proposed for enhancing the performance of the diagnostic system.

3.2 Reasoning Integration of Rough Set and Fuzzy Set and Bayesian Optimal Classifier:
In accordance with intelligent complementary strategies, a new transformer fault diagnosis method is proposed based on rough set (RS) and fuzzy set (FS) and Bayesian optimal classifier in this paper [55]. Through RS reduction, the diagnostic decision table is greatly simplified and fault symptoms information is compressed, dramatically, and the minimal decision rules can be obtained. In the light of the minimal decision rules, the complexity of Bayesian reasoning and difficulties of fault symptom acquisition are dramatically decreased. Moreover, probability reasoning may be realized applying Bayesian optimal classifier, it can be used to describe the characteristics of fault information and investigate the fault reasons of transformer.

3.3 Rough Set and Fuzzy Wavelet Neural Network Integrated with Least Square Weighted Algorithm:
[56] Rough set and fuzzy wavelet neural network integrated with LS weighted fusion algorithm based fault diagnosis for power transformers is proposed. Comparing the diagnosis result with the IEC three-ratio method, the accuracy of the proposed algorithm proved 88.33%, the accuracy of the three-ratio method is 71.7%, and one of the single FWNN is 81.67%, during the testing within 60 samples. The system constructing process is as follows:
(1) Preprocessing the input data, which is the application precondition of rough set and wavelet neural network.
(2) Simplifying the input of wavelet neural network, including input dimension and training sample number by rough set algorithm, i.e., the process of attribute reduction based on rough sets. The process is also regarded as the process of simplifying the input dimension and training pattern number. Deleting the same row in the decision-making table (i.e., the same rule) can simplify the training samples. Deleting the superfluous column (condition attribute) can simplify the network input dimension number.
(3) Mining the valuable rules with RS algorithm, the rules whose “confidence” and “support” are higher than the requirement are used as a diagnosis rules base for fault diagnosis and offer fault diagnosis service directly.
(4) Wavelet neural network, integrated with LS weighted fusion algorithm, is used to diagnose the case that cannot be diagnosed by rough set based mined rules. In this process, using seven trained fuzzy wavelet neural networks which have different learning rates, middle hidden layer numbers, and correlative parameters (here, each fuzzy wavelet neural network is also called one-child fuzzy neural network), to diagnose the fault, individually. The diagnosing results of the seven-child fuzzy wavelet neural network are fused with LS weighted
fusion algorithm, to identify the fault type with the fused result, finally.

3.4 Support Vector Machine (SVM) and Genetic Algorithm (GA): [57], [58]

Support vector machine (SVM) is a new machine learning method based on the statistical learning theory, which is a powerful tool for solving the problem with small sampling, nonlinearity and high dimension. The selection of SVM parameters has an important influence on the classification accuracy of SVM. However, it is very difficult to select appropriate SVM parameters.

Optimizing the SVM parameters with genetic algorithm: The free parameters C and r greatly affect the classification accuracy of SVM. Here, C and r are user-determined parameters, the election of the parameters plays an important role in the performance of SVM. However, it is not known beforehand what values of the parameters are appropriate. Therefore, GA is used to search for better combinations of the parameters in SVM. Based on the Darwinian principle of ‘survival of the fittest’, GA can obtain the optimal solution after a series of iterative computations. Figure 5 presents the process of optimizing the SVM parameters with genetic algorithm.

The whole process of applying SVM with GA needs three steps as follows;
- a. Encoding SVM parameters and initialization
- b. Calculating the fitness function
- c. GA operators

3.5 SVM and ANN with Particle Swarm Optimization (PSO): [59]

The SVM learns a separating hyperplane to maximize the boundary margin and to produce good generalization ability, which is claimed to have better generalization properties than ANN based classifiers in system fault diagnosis. PSO algorithm optimizes an object function by conducting population-based searches. The population consists of potential solutions, called particles, which are a metaphor of birds in bird flocking. These particles are randomly initialized and then freely fly across the multi-dimensional search space. During the flying, every particle updates its velocity and position based on its own experience and that of the entire population. The updating policy will drive the particle swarm to move toward the region with higher object value, and finally all particles will gather around the point with the highest object value.

[59], For the purpose of incipient power transformer fault symptom diagnosis, a successful adaptation of the particle swarm optimization (PSO) algorithm to improve the performances of Artificial Neural Network (ANN) and Support Vector Machine (SVM) is presented in figure 6. A PSO-based encoding technique is applied to improve the accuracy of classification, which removed redundant input features that may be confusing the classifier. Experiments using actual data demonstrated the effectiveness and high efficiency of the proposed approach, which makes operation faster and also increases the accuracy of the classification.

3.6 Hybrid Genetic Algorithm Evolving Wavelet Neural Network Technique: [60], [37]

The main drawbacks of a back propagation algorithm of wavelet neural network (WNN) commonly used in fault diagnosis of power transformers are that the optimal procedure is easily stacked into the local minima and cases that strictly demand initial value. A fault diagnostic method based on a real-encoded hybrid genetic algorithm evolving a WNN, which can be used to optimize the
structure and the parameters of WNN instead of humans in the same training process. Through the process, compromise is satisfactorily made among network complexity, convergence and generalization ability. Genetic algorithm (GA) is a directed random search technique that is widely applied in optimization problems. The GA can help to find out the optimal solution globally over a domain. This is especially useful for complex optimization problems in which the number of parameters is large, and analytical solutions are difficult to obtain in the fault diagnosis of power transformers. Recently, the evolution of the WNN based on a hybrid genetic algorithm (WNN_GA) to accurately capture the complicated, numerically rich input–output relationships of dissolved gases to corresponding fault types and to improve the disadvantages of existing diagnostic methods. Figure 7 is the flow chart for construction and steps of the proposed algorithm to adjust the WNN parameters and structure based on the hybrid GA optimization.

4. CONCLUSION
A thorough review has been conducted on the computer applied AI techniques used for fault diagnosis by DGA in power transformers by researchers for more than last 20 years through their researches published in different international/national research journals. Some of the well known simple individual AI techniques with their advantages are stated above. For the better performances in terms of optimum monitoring and diagnostic, most of the hybrid AI techniques with their relevant results are also mentioned. Few of them are also being tested and applied by different countries and utilities.

5. REFERENCES

Fig. 7. Flow-chart of WNN_GA [60].


