Contents lists available at ScienceDirect

Electric Power Systems Research

A cognitive system for fault prognosis in power transformers

Fernando Cortez Sica [a,b], Frederico Gadelha Guimarães [c,*], Ricardo de Oliveira Duarte [d], Agnaldo J.R. Reis [e]

[a] Graduate Program in Electrical Engineering – Federal University of Minas Gerais – Av. Antônio Carlos 6627, 31270-901 Belo Horizonte, MG, Brazil
[b] Department of Computer Science, Federal University of Ouro Preto (UFOP), Ouro Preto, Brazil
[c] Department of Electrical Engineering, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Brazil
[d] Department of Electronics, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Brazil
[e] Department of Control Engineering and Automation, Federal University of Ouro Preto (UFOP), Ouro Preto, Brazil

A R T I C L E   I N F O

A B S T R A C T

The power transformer is one of the most critical and expensive equipments in an electric power system. If it is out of service in an unexpected way, the damage for both society and electric utilities is very significant. Over the last decades, many computational tools have been developed to monitor the 'health' of such an important equipment. The classification of incipient faults in power transformers via Dissolved Gas Analysis (DGA) is, for instance, a very well known technique for this purpose. In this paper we present an intelligent system based on cognitive systems for fault prognosis in power transformers. The proposed system combines both evolutionary and connectionist mechanisms into a hybrid model that can be an essential tool in the development of a predictive maintenance technology, to anticipate when any equipment fault might occur and to prevent or reduce unplanned reactive maintenance. The proposed procedure has been applied to real databases derived from chromatographic tests of power transformers found in the literature. The obtained results are fully described showing the feasibility and validity of the new methodology. The proposed system can help Transformer Predictive Maintenance programmes offering a low cost and highly flexible solution for fault prognosis.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Power transformers are considered key-elements for the Electric Utilities (EU). When those equipments fail, households, industries, and hospitals, to name a few, are prone to suffer somehow. Besides, an unplanned interruption in the power supply can be translated into heavy fines for the EU. Hence, tools for diagnosis, fault detection and fault prognosis are required. In the context of power transformers, several studies are noteworthy regarding the aspects of protection, monitoring and diagnostics, see for instance [1–5].

For many years, preventive maintenance programmes in power transformers consisted of inspections, tests and actions in periodic time intervals usually suggested by the manufacturers or determined through practical experience. It was also common the application of routine tests and procedures such as: measurement of dielectric losses, insulation resistance and winding resistance; physical–chemical analysis and chromatographic oil analysis; manual or automatic monitoring of temperature [6]. Such analyses allowed the operators to verify if a given transformer was operating normally or if there were evidences of thermal and/or electrical failures, for instance. These kind of failures stem from natural wear, environmental actions and overloads, among other causes. Reference work in this area can be found in [6–8].

Among several fault detection methods, many faults that occur in power transformers can be detected if one measures the gases concentrations in their insulating oil. This procedure is known as fault detection via Dissolved Gas Analysis (DGA) [6]. Usually, DGA can be carried out in two modes: off-line and on-line modes. In the off-line mode, the power transformer has to be disconnected from the power system and an oil sample is collected and taken to a laboratory where it will be analysed via a gas chromatography technique. Yet in the on-line mode, the power transform is kept connected to the power system and the DGA is performed in loco with a determined time interval (e.g. every 2 h) using, e.g., a compact closed-loop gas chromatograph unit, which is mounted on or near the monitored transformer. Techniques such as optical and chromatography [9–11], electrical–chemical systems [12,13] and
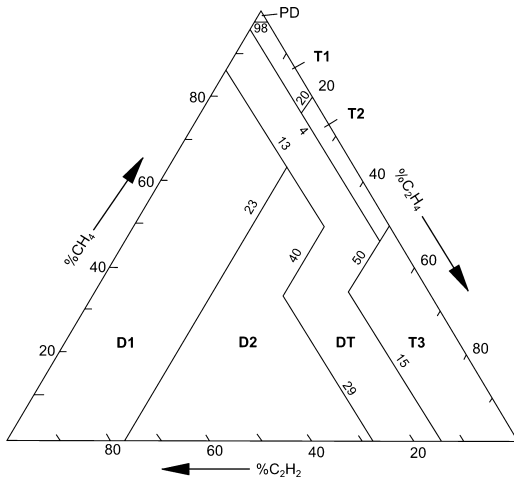
Fig. 1. Duval's Triangle Method.

near-IR absorption spectroscopy [14] can be employed in the DGA as well.

A method for fault classification in power transformers named *Duval's Triangle Method (DTM)* is introduced in [6] and described in IEEE C57.104 [7] and in IEC 60599 [8] (see Fig. 1). DTM is based on the levels of the three dissolved gases found in the power transformer oil: ethylene (C_2H_4), methane (CH_4) and acetylene (C_2H_2). The DGA by DTM involves the percentages of the aforementioned gases ratios, see Eqs. (1)–(3), presented in a graphical form. The three sides of the triangle in cartesian coordinates *x*, *y* and *z* represent the relative proportion of CH_4, C_2H_4 and C_2H_2, from 0 to 100% for each gas. The intersections of all three gases ratios indicate the kind of fault presented by the transformer. All types of faults in the Duval triangle can be found in Table 1.

In Eqs. (1)–(3), *x* is the concentration of CH_4 in ppm, *y* is the concentration of C_2H_4 in ppm and *z* is the concentration of C_2H_2 in ppm.

$$CH_4\,(\%) = 100\frac{x}{(x+y+z)} \tag{1}$$

$$C_2H_4\,(\%) = 100\frac{y}{(x+y+z)} \tag{2}$$

$$C_2H_2\,(\%) = 100\frac{z}{(x+y+z)} \tag{3}$$

Taking into consideration that there are not efficient mathematical models to describe the relationship between the rate of evolution of these concentrations and the failures, and the process of gathering historical data is a common practice nowadays, the development of pattern classifiers based on Support Vector Machines [15–17], Fuzzy Logic approach [18], Neuro-Fuzzy models [19,20], Wavenets [21,22], Neural Networks [23,24], stochastic Petri net based methodology [25], probabilistic classifier based on particle swarm optimizer [26], Decision Trees [27–29], and bootstrap and Genetic Programming [30] has received a great deal of

Table 1
Faults described by the Duval's Triangle Method (DTM).

Nomenclature	Type of fault
PD	Partial discharge
D1	Low energy discharge
D2	High energy discharge
DT	Mix of thermal and electrical faults
T1	Thermal fault with $T < 300\,^\circ C$
T2	Thermal fault with $300 < T < 700\,^\circ C$
T3	Thermal fault with $T > 700\,^\circ C$

attention. More recently, Duval himself and his colleagues published a new fault detection method based on the application of a gas-ratio combination [31].

The aforementioned works have at least two common points:

- the vast majority of papers in this field deal with off-line DGA.
- the authors focused on fault diagnosis in power transformers. Fault prognosis was not a major concern. Incidentally, this latest strategy is relatively new. Refs. [32–36] are few exceptions in this particular subject.

It is in this context that we present a new computational system capable of carrying out fault prognosis in power transformers. The new approach is based on a hybrid system formed by connectionist and evolutionary mechanisms, which was applied to real databases derived from chromatograph tests of power transformers found in the literature [37]. Instead of informing the current state of the transformer only, e.g. *Normal Condition* (NC), *Electrical Fault* (EF), *Thermal Fault* (TF), the fault prognosis tool will estimate – from a given current transformer state – if the transformer under analysis is likely to present, for instance, a thermal failure in six months and how the failure will evolve through time. Such tool is of interest of the EU because it can become an essential tool for the Transformer Predictive Maintenance (TPM). TPM can be seen as a maintenance strategy that provides transformer failure early detection and is able to recognize conditions that lead to defects. Ideally, TPM should allow the maintenance frequency to be as low as possible preventing unplanned reactive maintenance and reducing unnecessary preventive maintenance. The proposed prognosis system is an intelligent system that was implemented with an architecture for modelling cognitive systems.

The main contributions of this work are: (i) proposition of an on-line intelligent fault prognosis system for power transformers that is capable of predicting when and which type of fault is expected to occur. The proposed system can help TPM programmes offering a low cost and highly flexible solution for fault prognosis; additionally the prognosis system manipulates the gases concentrations at the same time as multiple inputs to the system. (ii) Proposition of a general architecture for cognitive and knowledge based systems, on which the fault prognosis system is implemented.

2. Fault prognosis system in power transformers

Prognosis (also Prognostics) is an engineering discipline involving the prediction of when a system or component will no longer perform as intended and, if possible, classify which kind of fault will occur [38–41]. The predicted time is termed Remaining Useful Life (RUL), which is an important concept in contingency and in System Health Management (SHM). Prognosis systems can be categorized into data-driven methods and model-based methods [43]. In any case, the method requires some initial knowledge about a model of the system or component; or knowledge in the form of information about previous fault cases and conditions.

Data-driven prognosis systems usually employ pattern recognition and machine learning techniques to detect changes in the states of the system of interest or changes in the monitored variables. They are recommended when there is no appropriate understanding of the physics behind the operation of the system, or the principles relating its operation to fault conditions.

Industrial applications of Prognosis are diverse, ranging from manufacturing, automotive and aerospace industries and also power generation and distribution. In the power generation industry, commercial applications of SHM include rotating machinery [43] and wind turbines [44]. Renewable energy applications, such as wind turbines, can also benefit from SHM technology based on

Fig. 2. Cognitive process.

Fig. 3. Basic architecture of the system.
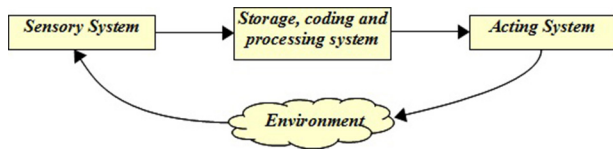
prognosis systems. Approaches based on statistical methods and machine learning can be found in [45,46].

In [47], the authors propose an ensemble of data-driven prognostic algorithms for robust prediction of RUL by the weighted sum of the outputs of five different methods: a Similarity-Based Interpolation (SBI) approach with Relevance Vector Machine (RVM) as the regression technique, SBI with Support Vector Machine (SVM), SBI with the least-square exponential fitting, a Bayesian linear regression with the least-square quadratic fitting, and a Recurrent Neural Network (RNN) approach. Specifically on fault prognosis in power transformers, it is possible to find some few recent studies in the literature, as the ones described in [35,36]. In [35], the author utilizes Bayesian Particle Filtering to estimate the RUL of the transformer, considering the degree of polymerization of the insulating material and temperature measurements in specific points of the transformer (hotspots). In [36], an Adaptive Network-based Fuzzy Inference System (ANFIS) is used to infer changes in the dissolved gas concentrations. One ANFIS is defined for each gas such that specific fuzzy rules are applied for the dynamics of each gas considering also the operation time of the power transformer.

3. Proposed cognitive system for fault prognosis

3.1. Cognitive and intelligent systems

When considering knowledge based computational models and systems, as the fault detection and prediction system proposed in this paper, the first step is the description and modelling of the knowledge and consequently, how this knowledge is actually represented and stored. It is interesting to notice that the cognitive process is a dynamic one. This dynamics is related to its interaction with the environment: collecting information, coding, storing and processing this information such that some action in the environment can be performed (Fig. 2).

Artificial Cognitive Systems are systems designed with machine learning techniques. They have the capacity of representing and manipulating knowledge to assist or improve human expertise and cognition, to perform complex decision-making and action in the environment. An architecture for cognitive systems is actually a blueprint for intelligent agents. Any proposed architecture should provide functionalities that facilitate the proposition of a model for a computational system to act as an intelligent system in some sense.

Therefore, in general, the modelling of a cognitive and intelligent system should involve the following basic components:

• Sensory system;
• Storage, coding and processing system;
• Actuation system;
• Communication system.

The goal of this section is to describe the proposed system modelling and implementation aspects. The modelling is described as generic as possible, making it useful to other applications involving knowledge based systems, although the focus here is on the power transformer fault prognosis system.
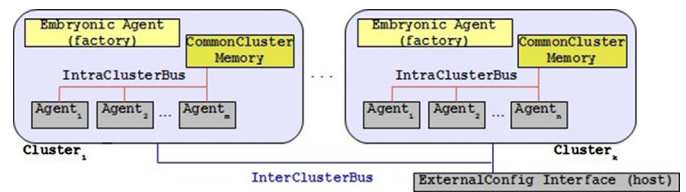
3.2. Proposed architecture

The fault prognosis system (FPS) is developed on a general proposed architecture for modelling cognitive systems. The designed system combines connectionist and evolutionary models into a knowledge based system, specifically, an automatic data-driven prognosis system.

In general terms, the proposed architecture for the development of cognitive systems uses the idea of autonomous intelligent agents for clustering of information and processing of classes. This clustering allows a better representation and organization of the knowledge incorporated and manipulated by the agents and provides better scalability. Another reason for the clustering of information is to allow the separation of agents in terms of functional classes, considering the kind of knowledge to be represented and the type of actuation required. In this way, higher system granularity can be achieved for the implementation of more complex systems applications, such as for instance *health monitoring systems* as described in [47]. The proposed architecture is a multi-agent system that allows the use of different techniques together in an ensemble approach, for example, employing Bayesian networks, fuzzy models and statistical models. The system is flexible enough in terms of scalability to include new methods and hybrid approaches for complex tasks.

Fig. 3 shows a macroscopic view of the system with multiple clusters instantiated. Each cluster has three basic and fixed elements: (i) Embryonic Agent, (ii) CommonCluster Memory and (iii) IntraClusterBus, with their respective agent instances.

The Embryonic Agent manages the creation or removal of agent instances. This mechanism allows for the self-adaptation of the multi-agent system given the environment dynamics. The CommonCluster Memory is a common area for long term information storage (long term memory), that is, information such as configuration, characterization of incorporated knowledge and codification related to the actuation process. The IntraClusterBus and the InterClusterBus are communication elements represented by storage structures of *mailbox* type in the case of software systems or based on the Wishbone standard [48] in the case of hardware systems.

Still regarding Fig. 3, it is possible to notice the presence of instances of agents in the clusters. Each agent, operating autonomously and cooperatively, is responsible for: (i) acquisition of sensory information (related to the sensory system); (ii) encoding, storing and processing information; and (iii) actuation on the environment.

In order to implement such functions, each agent has a functional structure as illustrated in Fig. 4. In this figure, one can see the basic elements that comprise the cognitive agent: the sensory system; the storage, coding and processing system (SCP System); the Agent Core and the Agent Memory. These elements will be described next.

3.3. Sensory system

The sensory system must be able to pick up stimuli (information) from the environment to feed the SCP System. The main function of this unit is to process collected information such that it can be
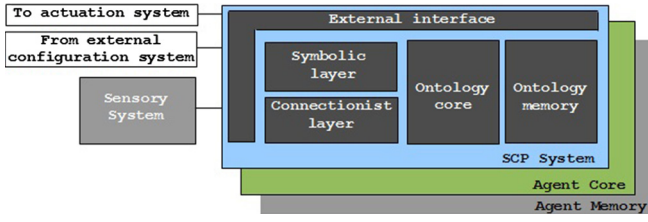
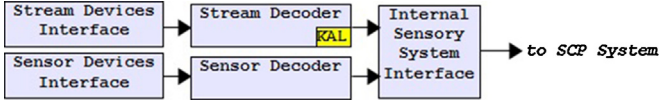**Fig. 4.** Functional structure of each agent.

Fig. 5. Functional structure of the sensory system.

manipulated by other functional units of the system and to facilitate the actuation on the external environment. These two processing types represent the deliberative and reactive behaviour, respectively [8].

The sensory deliberative system is responsible for interfacing the system with the environment for capturing, for example, image, sound or other waveforms (stream devices). Regardless of the device type used to capture data, the goal of the sensory deliberative system consists in encoding external stimuli in connectionist mappings. The connectionist maps will be the cognitive system basis.

The structure of the sensory system is presented in Fig. 5. As can be seen, the sensory system integrates mechanisms for capturing both signals of the type stream and sampled data from sensors/transducers.

In the fault prognosis system, it is assumed that the Sensor Device Interface receives information from an on-line gas chromatograph unit that provides the gas concentrations for H_2, C_2H_4, C_2H_2, CH_4, C_2H_6.

3.4. Storage, coding and processing system

In order to allow greater modularity and reuse of the structures that represent knowledge, the SCP System is based on a hybrid topology, adopting, in this case, both the connectionist and symbolic models.

The SCP internal elements form the basis of the implemented knowledge based system, which can be replicated in clusters for better information representation (classes of knowledge) and to provide greater dynamism to processing. Each cluster can evoke functions for the creation or removal of agents or even other clusters. These functions are exported by the Embryonic Agent Layer (factory), as shown in Fig. 3.

For a better description of the system modelling, the internal components of the SCP System will be summarized next.

3.4.1. Connectionist layer

The connectionist layer is related to the reactive functionality, but also providing the basis for the deliberative functionality. The connectionist mappings form the basis of the cognitive system.

For the connectionist layer, we adopt an implementation based on *ensemble learning*, in which each operation status of the power transformer (*Normal Condition* (NC), *Electrical Fault* (EF), *Thermal Fault* (TF)) is abstracted by a specific group of neural networks. In our implementation, we have adopted the General Regression Neural Network (GRNN), originally proposed in [49].

The GRNN belongs to the class of probabilistic neural networks. Its main advantage is requiring less training data samples than

a back-propagation neural network would. The prediction of the GRNN for a given input $\mathbf{x}$ is:

$$g(\mathbf{x}) = \frac{\sum_{i=1}^{n} y_i \exp(-(((\mathbf{x} - \mathbf{x}_i)^T(\mathbf{x} - \mathbf{x}_i))/(2\sigma^2)))}{\sum_{i=1}^{n} \exp(-(((\mathbf{x} - \mathbf{x}_i)^T(\mathbf{x} - \mathbf{x}_i))/(2\sigma^2)))} \quad (4)$$

In the GRNN model, $\mathbf{x}$ represents an input, the training data samples are represented by $\{\mathbf{x}_i ; y_i\}$. When the input is close to the sample $\mathbf{x}_j$, the weight of the prediction y_j is close to one and the contribution of the other training samples becomes smaller. With this model it is possible to predict the behaviour of systems with few samples and interpolate smoothly between training data samples. The smoothness parameter σ can be determined with an iterative method [49].

Ensemble learning methods involve employing a finite set of different classifiers together in order to achieve better predictive performance. Bootstrap aggregation was adopted for the ensemble classification [50]. In this approach, each model in the ensemble vote with equal weight, thus the output of the ensemble is the simple average of the output of each classifier. There is one ensemble of N neural networks trained for recognizing electrical fault, another ensemble of N neural networks trained for recognizing thermal fault and another ensemble for normal operation. These networks form the connectionist layer.

$$E^{NC}(\mathbf{x}) = \frac{1}{N} \sum_{s=1}^{N} g_s^{NC}(\mathbf{x}) \quad (5)$$

$$E^{TF}(\mathbf{x}) = \frac{1}{N} \sum_{s=1}^{N} g_s^{TF}(\mathbf{x}) \quad (6)$$

$$E^{EF}(\mathbf{x}) = \frac{1}{N} \sum_{s=1}^{N} g_s^{EF}(\mathbf{x}) \quad (7)$$

E^{NC} is the output of the ensemble associated with normal operation, E^{TF} is the output of the ensemble associated with thermal fault and E^{EF} is the output of the ensemble associated with electrical fault.

In the prognosis system for power transformers, considering the Dissolved Gas Analysis (DGA), the input $\mathbf{x}$ is composed of the gas concentrations for H_2, C_2H_4, C_2H_2, CH_4, C_2H_6 and the labelled output is equal to one when belonging to the class, and equal to zero when not belonging to the class.

3.4.2. Symbolic Layer

The Symbolic Layer implements an evolutionary mechanism for evolving candidate conditions that maximize the probability of the connectionist layer to detecting faults. The evolutionary algorithm is based on Evolution Strategies [51]. Prognosis is a dynamic process where predictions change with time and as new observation data is available. The integration of the connectionist and the evolutionary model is the basis of the prognosis system.

The initial population of the evolutionary mechanism is generated based on the current sample of gas concentrations from the sensory system and a statistical model of the variation of these gases with time. The statistical model of the dynamics of these concentrations is developed from studies performed by [52]. The statistical model tries to predict the next configuration of gases based on the current one. Therefore, given the current sample of gas concentrations, a population of μ individuals is generated with the statistical model, providing possible future scenarios for the current concentration. The algorithm is detailed below.

1. Given the current sample of gas concentrations from DGA, generate an initial population $\{\mathbf{p}_{t,k}\}$, $t = 1, \ldots, T_{\max}$, $k = 1, \ldots, \mu$, with

μ individuals by applying the statistical model of the variation of these gases with time.

2. Classify the μ individuals $\{\mathbf{p}_{t,k}\}$ with the neural networks in the connectionist layer in terms of normal operation, thermal fault or electrical fault. The fitness values of the individuals correspond to the maximum value among the outputs of the ensemble networks, which are values between zero and one. The classification is determined by the biggest output:

$$\phi(\mathbf{p}_{t,k}) = \max[E^{NC}(\mathbf{p}_{t,k}), E^{TF}(\mathbf{p}_{t,k}), E^{EF}(\mathbf{p}_{t,k})]$$

3. Generate λ offspring $\{\mathbf{o}_{t,j}\}, j = 1, \ldots, \lambda$ by selecting one individual with binary tournament and applying recombination and Gaussian mutation to its values. The best parents have higher probability of being selected for reproduction, hence producing more offspring.
4. Classify the λ individuals, $\{\mathbf{o}_{t,j}\}$, with the neural networks in the connectionist layer in terms of normal operation or fault. The fitness values of the individuals correspond to the maximum value among the outputs of the ensemble networks, which are values between zero and one.
5. Apply ES($\mu + \lambda$) selection, i.e. select the μ best individuals based on the fitness values. These individuals will form the parent population of the next iteration.
6. Apply the statistical model over the μ parent individuals, in order to predict their possible next state.
7. Return to step 2 until convergence.

Notice that the evolutionary model is actually evolving the population in order to assess future scenarios for the current concentration of gases. Those future scenarios that are likely to lead to a recognition condition are selected and reproduced. Therefore, the prognosis is done along the evolutionary process. The method converges when the entire population is classified as fault or when the maximum number of iterations is reached.

Each iteration of the evolutionary algorithm corresponds to one year, making the bridge between the model and the prognosis of the power transformer. When the whole population is classified as fault, the method stops and the number of iterations executed corresponds to the number of years that a fault (electrical or thermal) is predicted to occur with 100% (in other words, the RUL). Otherwise, if the method stops by reaching the maximum number of iterations ($T_{\max}$) without having the entire population classified as fault, then the prognosis corresponds to a given condition based on the average fitness of the population for each ensemble, which gives a perspective of the proportion of individuals in the population that can be classified as fault or normal operation:

$$r^{NC}(t) = \frac{1}{\mu} \sum_{k=1}^{\mu} E^{NC}(\mathbf{p}_{t,k}) \tag{8}$$

$$r^{TF}(t) = \frac{1}{\mu} \sum_{k=1}^{\mu} E^{TF}(\mathbf{p}_{t,k}) \tag{9}$$

$$r^{EF}(t) = \frac{1}{\mu} \sum_{k=1}^{\mu} E^{EF}(\mathbf{p}_{t,k}) \tag{10}$$

The prognosis is Normal Condition (NC) in the number of years t' corresponding to the condition:

$$(r^{NC}(t') > r^{TF}(t')) \wedge (r^{NC}(t') > r^{EF}(t')) \wedge (r^{NC}(t') > 0.5)$$

The prognosis is Thermal Fault (TF) in the number of years t' corresponding to the condition:

$$(r^{TF}(t') > r^{NC}(t')) \wedge (r^{TF}(t') > r^{EF}(t')) \wedge (r^{TF}(t') > 0.5)$$
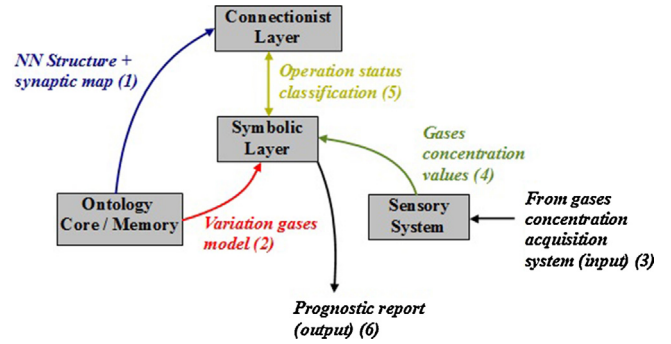

Fig. 6. Prognosis steps.

The prognosis is Electrical Fault (EF) in the number of years t' corresponding to the condition:

$$(r^{EF}(t') > r^{NC}(t')) \wedge (r^{EF}(t') > r^{TF}(t')) \wedge (r^{EF}(t') > 0.5)$$

3.5. Ontology Core

The basic function of the Ontology Core is to execute the XML parser, which performs search and manipulation operators over the XML structures (with XML denoting the ontologies themselves). The ontologies, in our system, provide the representation of the knowledge under manipulation, e.g., grammatical rules, a given knowledge specification or an initial Bayesian network [53]. The Ontology Memory module has the function of storing ontologies manipulated by each agent cluster.

3.6. Agent Core

The Agent Core has the function of performing functions inherent to the agents. Unlike the Ontology Memory, the Agent Memory aims to store agents code and attributes (e.g., their states).

3.7. Actuation System

For initial and prototyping purposes, the Actuation System is modelled as a simple text output stream.

3.8. Communication System

In addition to the need to model mechanisms related to the knowledge representation and the cognition environment, it is also necessary to have infrastructure for communication between agents and between agent clusters. Joining the cognitive functionalities, where each cluster stores part of the knowledge, intra and inter communication mechanisms are also required. In Fig. 3, these two communication levels are represented by IntraClusterBus and InterClusterBus respectively. To facilitate the initial prototype, we employ a model based on shared buffers of *mailbox* type.

3.9. Fault diagnosis system on the proposed architecture

To better illustrate the power transformer fault prognosis system using the proposed architecture, Fig. 6 shows the subsystems interactions and their functionalities. The numbers in the figure indicate the sequence of actions during the system utilization, as follows:

1. The *Ontology Core & Memory* provides to the *Connectionist Layer*, as described in Section 3.4.1, the neural network settings and the

Table 2
Test cases.

Case	H_2	CH_4	C_2H_4	C_2H_6	C_2H_2	op	Date
	29	110	53	290	10	N	02/03/1983
1	26	51	52	530	5	N	17/08/1984
	64	180	65	770	5	N	10/10/1986
	620	2200	2000	1100	13	T	05/08/1987
	29	02	16	1	0.4	N	11/01/1983
2	27	5	12	2	0.4	N	13/02/1984
	18	3	80	0.4	0.4	E	08/06/1989

op, operation; N, normal; T, thermal fault; E, electrical fault.

Table 3
Test environment.

Item	Description
Processor	Intel Core i3-380M
Memory	3 GB DDR3
Clock frequency	2.53 GHz
Operating system	CygWin emulated on Windows 7
C Compiler	GNU C++ vs 4.5.3
Mathematical analysis	Scilab 5.4.1

synaptic weights after the training stage. It is important to high-light that the training of the models in the Connectionist Layer occurs off-line using data samples taken by the power company or databases available from the academic community. This data is used to train the neural network in the task of recognizing normal and fault conditions from the gas concentration values.

2. The *Ontology Core & Memory* provides to the *Symbolic Layer* the statistical model of the gases variation with time. The statistical model can be built from analyses carried out in the specialized literature or from observations built by the power company. Thus, the use of this system does not become infeasible in the case of utilities which do not have historical data. Additionally, other models for forecasting the gases variation could be easily adopted in the system.

3. Sampled gas concentrations are presented to the *Sensory System* to initiate the fault prognosis.

4. The information of gases concentrations, after decodification by the *Sensory System*, is passed on to the *Symbolic Layer*, starting the evolutionary algorithm described in Section 3.4.2.

5. For each iteration of the evolutionary algorithm in Section 3.4.2, the *Conexionist Layer* is triggered to perform the analysis and classification of the offspring generated by the evolutionary model.

6. After the evolutionary algorithm stops, the prognosis report is generated.

4. Results

4.1. Experimental data

The neural networks in the connectionist layer were trained with the data set IEC TC 10 from the work of Duval as cited in [54]. There are 73 instances of electrical fault, 34 instances of thermal fault and 48 instances of normal operation.

Obtaining data to test and validate the proposed Prognosis system is not an easy task because utility companies do not usually label the measurement data of power transformers with the date it was collected. Nonetheless, we utilize information available from [37], with data provided by an utility company which performed the monitoring of its transformers, registering the gas concentrations and the dates of collecting this data. Two test cases are summarized in Table 2. In the first test case, by analysing the data, there is a thermal fault after 4 years and 5 months (4.42 years). In the second case, there is an electrical fault after 6 years and 5 months (6.42 years).

4.2. Configuration of experiments

The fault prognosis system was implemented on a PC-based computer using GNU C++ compiler. Table 3 details the characteristics of the environment used for implementation and testing.

4.3. Results

The Prognosis system was applied to these test cases using the concentration data collected on the first date in Table 2. Based on this initial sample, the Prognosis system is started. The symbolic layer generates a population whose individuals are classified as normal operation, thermal fault or electrical fault by the connectionist layer. After the evolutionary process converges, we obtain the type of fault and the prediction of when it is expected to occur. Since the method is stochastic, we run our method 120 times for each case, generating different values for the Remaining Useful Life (RUL) in years. The parameters used in the experiment are: $\mu = 10$, $\lambda = 100$, $N = 5$, $T_{max} = 15$. The results are summarized in Fig. 7, which shows the box plot of the RUL for all the independent runs for each test case.

As we can see, the system correctly converges to an estimate that is on average very close to the actual time of occurrence of the fault. The predicted RUL for the case 1 is 4.75 ± 1.18, while the predicted RUL for the case 2 is 6.125 ± 2.88. Figs. 8 and 9 show the histograms of the results for each case.

Finally, we perform an additional experiment, in which the initial state is randomly selected from one of the entries in Table 2. Then the method is applied based on this initial state, thus reaching a prognosis. The obtained RUL is compared with the actual time for the fault, and the error is calculated. After 120 runs, we obtained the following results: 20% of runs corresponded to the exact fault time, 33.33% of runs corresponded to an error equal to one year, 31.67% of runs corresponded to an error equal to two years and 15% of runs corresponded to an error equal or greater than three years. These outcomes suggest that the proposed method here is reliable for this application.

The computational cost is low: one independent run of the proposed fault prognosis system takes on average 6 ms in the test environment, therefore, it is available as soon as a sample of gas concentration is provided. Of course the run time depends on the
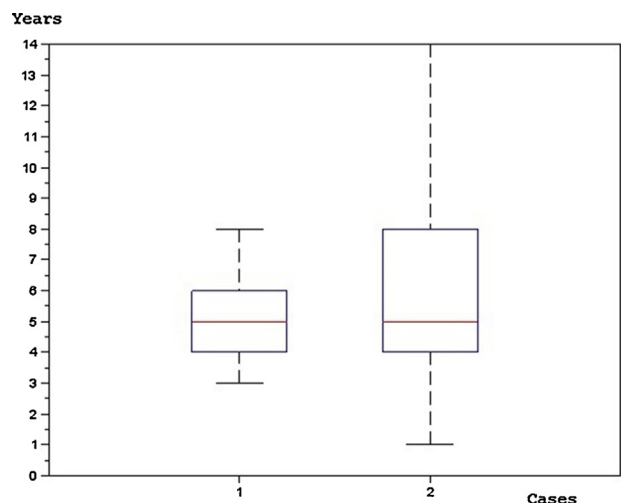

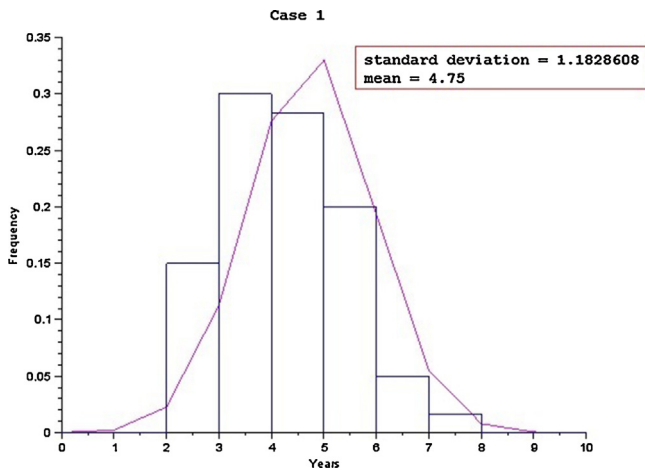**Fig. 7.** Box plot of the RUL results over 120 independent runs.

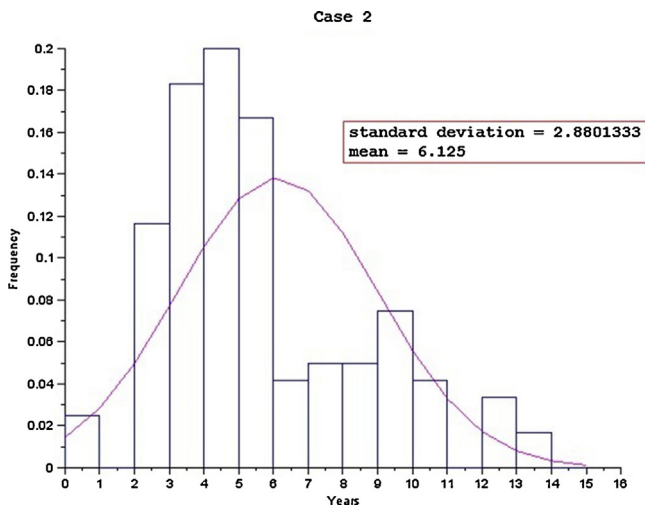**Fig. 8.** Histogram of the RUL for the case 1 over 120 independent runs.

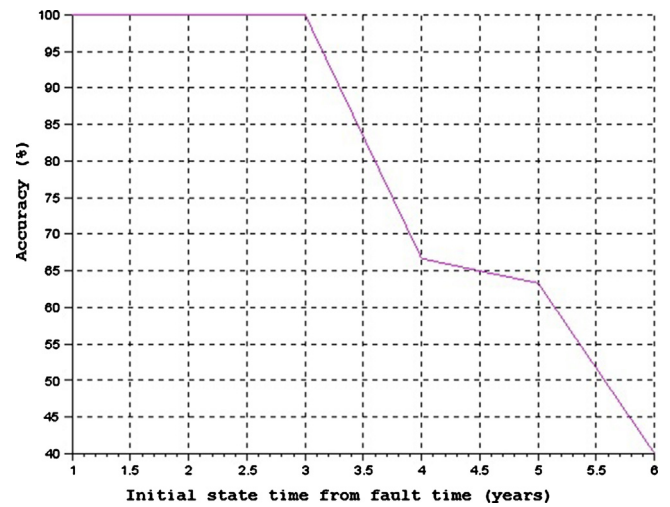**Fig. 9.** Histogram of the RUL for the case 2 over 120 independent runs.

Fig. 10. Accuracy of the prognosis with respect to the time between the time of the initial state and the time of failure.

FP imply unnecessary preventive maintenance action while FN represent corrective or reactive maintenance, because a fault was not predicted correctly. Although undesirable, FP are less critical than FN. TN would represent a situation where the system does not predict a fault and no fault actually happens.

From the confusion matrix described in Table 4, one can establish the accuracy obtained through the relation

$$ACC = \frac{TP + TN}{P + N}.$$

Fig. 10 shows the accuracy in relation to the proximity of fault time. We have observed in our experiments that when the initial state is closer to when the fault actually occurred, the proposed system is more accurate.

5. Conclusions

In this paper we introduced a data-driven intelligent fault Prognosis system based on cognitive systems for health monitoring in power transformers. The proposed system is implemented on a generic architecture for modelling cognitive systems. A fault prognosis system provides many benefits for the Electrical Utilities because it can assist in power transformer preventive maintenance, consequently, benefiting consumers such as households, industries and hospitals.

The generic architecture allows the implementation of other knowledge based applications and is flexible enough to allow for changes or improvements in the techniques employed in the symbolic and connectionist layers, if required for the application. The proposed fault Prognosis system combines evolutionary and connectionist mechanisms into a hybrid system. The connectionist layer employs ensemble learning for its potential in achieving higher accuracy and robustness. Ensemble learning can combine together different classifiers based on machine learning that are available in the literature and statistical methods. The symbolic layer employs an evolutionary method and a statistical model for the dynamics of the variation of the gas concentrations.

The results reported show that the system was able to correctly predict not only the type of fault but also when that fault is expected to occur. This kind of prediction is very useful for System Health Management. This is one important contribution of our work, since it goes one step further than other research in the literature regarding fault classification in power transformers.

parameters of the evolutionary algorithm μ, λ, and T_{max} and on the type of machine learning technique and number of neural networks in the ensemble in the connectionist layer.

In order to give a better idea about the accuracy of the method, Table 4 presents the confusion matrix of case 1 and case 2. It is considered that an acceptable prognosis error is within one year ($\delta = 1$).

In Table 4, considering $\delta = 1$ and $\Delta = 5$, let RT represent the real fault time and PT represent the fault time predicted by the prognosis system. We have:

- True Positive (TP): when $RT - \delta \leq PT \leq RT + \delta$, the predicted time is within the margin of error.
- False Positive (FP): when $PT < RT - \delta$, the real fault occurred after the time estimated by the prognosis system.
- False Negative (FN): when $RT + \delta \leq PT \leq RT + \Delta$, the real fault occurred before the time estimated by the prognosis system.
- True Negative (TN): when $PT > RT + \Delta$, the real fault occurred much earlier than the time estimated by the prognosis system.

Table 4
Confusion matrix.

Case	TP	FP	FN	TN
Case 1	50.00%	2.50%	30.83%	16.67%
Case 2	53.33%	45.00%	1.67%	0.00%

F.C. Sica et al. / Electric Power Systems Research 127 (2015) 109–117

One of the advantages of our method is that knowledge about the age of the power transformer as well as its history is not required. The methodology relies on DGA and neural networks to recognize the condition of the transformer.

Regarding costs for deployment of the proposed system, they could be estimated as a minimum since the proposed methodology can be implemented as a basic embedded system. The most expensive part is related to the DGA equipment itself.

Since the proposed architecture is flexible, one can use it for various applications. Given these characteristics, the applicability of the architecture for other cognitive systems applications is wide. Moreover the developed fault prognosis application is of general use in electric power systems.

As future work, we intend to improve the accuracy of the connectionist layer by testing other machine learning methods in the ensemble. Another important follow-up of this work is the hardware implementation, leading to an electronic system that could be actually embedded in power transformers, thus integrating system health monitoring and management to the power system equipment.

Acknowledgements

References

[1] P.S. Georgilakis, Spotlight on Modern Transformer Design, Springer, London, UK, 2009.
[2] S. Singh, M. Bandyopadhyay, Dissolved gas analysis technique for incipient fault diagnosis in power transformers: a bibliographic survey, Int. IEEE Electr. Insul. Mag. 26 (6) (2010) 41–46.
[3] P.S. Georgilakis, Condition monitoring and assessment of power transformers using computational intelligence, Int. J. Electr. Power Energy Syst. 33 (10) (2011) 1784–1785.
[4] P.S. Georgilakis, A.G. Kagiannas, A novel validated solution for lightning and surge protection of distribution transformers, Int. J. Electr. Power Energy Syst. 63 (2014) 373–381.
[5] W. Zhan, A.E. Goulart, M. Falahi, P. Rondla, Development of a low-cost self-diagnostic module for oil-immerse forced-air cooling transformers, IEEE Trans. Power Deliv. 30 (1) (2015) 129–137.
[6] M. Duval, A review of faults detectable by gas-in-oil analysis in transformers, IEEE Electr. Insul. Mag. 18 (3) (2002) 8–17.
[7] IEEE, IEEE guide for the interpretation of gases generated in oil-immersed transformers, IEEE standard c57.104, 2008.
[8] IEC, IEC mineral oil-impregnated electrical equipment in service – guide to the interpretation of dissolved and free gases analysis, IEC 60599, 2008.
[9] F. Wan, Q. Zhou, J. Zou, Z. Gu, Using a sensitive optical system to analyze gases dissolved in samples extracted from transformer oil, IEEE Electr. Insul. Mag. 30 (5) (2014) 15–22.
[10] V.G. Arakelian, The long way to the automatic chromatographic analysis of gases, IEEE Electr. Insul. Mag. 20 (6) (2004) 8–25.
[11] Y. Goldstein, N.B. Grover, C. Chang, A. Jelli, J. Andre, P. Mark, A. Goodwin, A gas chromatographic air analyzer fabricated on a silicon wafer, IEEE Trans. Electron Dev. 26 (12) (1979) 1880–1886.
[12] A. Depari, A. Flammini, E. Sisinni, A. De Marcellis, G. Ferri, P. Mantenuto, Fast, versatile, and low-cost interface circuit for electrochemical and resistive gas sensor, Electr. Eng. 14 (2) (2004) 315–323.
[13] P. Zylka, Electrochemical gas sensors can supplement chromatography based DGA, Electr. Eng. 87 (3) (2005) 137–142.
[14] L. Dong, J. Wright, B. Peters, B.A. Ferguson, F.K. Tittel, S. McWhorter, Compact QEPAS sensor for trace methane and ammonia detection in impure hydrogen, Appl. Phys. B 107 (2) (2012) 459–467.
[15] R. Liao, J. Bian, L. Yang, S. Grzybowski, Y. Wang, J. Li, Forecasting dissolved gases content in power transformer oil based on weakening buffer operator and least square support vector machine-Markov, IET Gener. Transm. Distrib. 6 (2) (2010) 142–151.
[16] S.W. Fei, Y. Sun, Forecasting dissolved gases content in power transformer oil based on support vector machine with genetic algorithm, Electr. Power Syst. Res. 78 (2008) 507–514.
[17] M.Y. Cho, T. Lee, S. Kau, C. Shieh, C. Chou, Fault diagnosis of power transformers using SVM/ANN with clonal selection algorithm for features and kernel parameters selection, in: Proc. International Conference on Innovative Computing, Information and Control, 2006, pp. 26–30.
[18] Y.C. Huang, H.C. Sun, Dissolved gas analysis of mineral oil for power transformer fault diagnosis using fuzzy logic, IEEE Trans. Dielectr. Electr. Insul. 20 (2013) 974–981.
[19] M. Rajabimendi, E.P. Dadios, A hybrid algorithm based on neural-fuzzy system for interpretation of dissolved gas analysis in power transformers, in: Proc. TENCON 2012–2012 IEEE Region 10 Conference, 2012, pp. 1–6.
[20] R. Naresh, V. Sharma, M. Vashisth, An integrated neural fuzzy approach for fault diagnosis of transformers, IEEE Trans. Power Deliv. 23 (4) (2008) 2017–2024.
[21] C. Pan, W. Chen, Y. Yun, Fault diagnostic method of power transformers based on hybrid genetic algorithm evolving wavelet neural network, IET Electr. Power Appl. (2008) 71–76.
[22] W. Chen, C. Pan, Y. Yun, Y. Liu, Wavelet networks in power transformers diagnosis using dissolved gas analysis, IEEE Trans. Power Deliv. 24 (1) (2009) 187–194.
[23] S. Seifeddine, B. Khmais, C. Abdelkader, Power transformer fault diagnosis based on dissolved gas analysis by artificial neural network, in: Proc. First International Conference on Renewable Energies and Vehicular Technology (REVET), 2012, pp. 230–236.
[24] L.G. Castanheira, J.A. Vasconcelos, A.J.R. Reis, P.H.V.M. Raes, S.A.L. Silva, Application of neural networks in the classification of incipient faults in power transformers: a study of case, in: Proc. International Joint Conference on Neural Networks 2011 (IJCNN 2011), 2011, pp. 3099–3104.
[25] P.S. Georgilakis, J.A. Katsigiannis, K.P. Valavanis, A.T. Souflaris, A systematic stochastic petri net based methodology for transformer fault diagnosis and repair actions, J. Intell. Robot. Syst. 45 (2) (2006) 181–201.
[26] W.H. Tang, J.Y. Goulermas, Q.H. Wu, Z.J. Richardson, J. Fitch, A probabilistic classifier for transformer dissolved gas analysis with a particle swarm optimizer, IEEE Trans. Power Deliv. 23 (2) (2008) 751–759.
[27] F. Zhao, H. Su, A decision tree approach for power transformer insulation fault diagnosis, in: Proc. 7th World Congress on Intelligent Control and Automation (WCICA), 2008, pp. 6882–6886.
[28] K.X.P. Lai, T.R.B.T. Blackburn, Descriptive data mining of partial discharge using decision tree with genetic algorithm, in: Proc. Australasian Universities Power Engineering Conference (AUPEC'08), 2008, pp. 1–6.
[29] L.G. Castanheira, J.A. Vasconcelos, A.J.R. Reis, P.H.V.M. aes, S.A.L. Silva, A comparative study of the application of neural networks and decision trees in the classification of incipient faults in power transformers, in: Proc. X Brazilian Conference on Computational Intelligence (X CBIC), 2011, pp. 1–6.
[30] A. Shintemirov, W. Tang, Q.H. Wu, Power transformer fault classification based on dissolved gas analysis by implementing bootstrap and genetic programming, IEEE Trans. Syst. Man Cybern. – Part C: Appl. Rev. 39 (1) (2009) 69–79.
[31] S. Kim, S. Kim, H. Seo, J. Jung, H. Yang, M. Duval, New methods of DGA diagnosis using IEC TC 10 and related databases. Part 1: application of gas-ratio combinations, IEEE Trans. Dielectr. Electr. Insul. 20 (2) (2013) 685–690.
[32] D. Ozgur-Unluakin, T. Bilgiç, Predictive maintenance using dynamic probabilistic networks, in: Probabilistic Graphical Models'06, 2006, pp. 239–246.
[33] T. Brescia, S. Bruno, M. La Scala, S. Lamonaca, G. Rotondo, U. Stecchi, A fuzzy-logic approach to preventive maintenance of critical power transformers, in: 20th International Conference on Electricity Distribution, CIRED, 2009, pp. 1–5.
[34] D. Zhou, A. Wang, C. Li, Data requisites for transformer statistical lifetime modelling – Part I. Aging-related failures, IEEE Trans. Power Deliv. 28 (3) (2013) 1750–1757.
[35] V.M. Catterson, Prognostic modeling of transformer aging using Bayesian particle filtering, in: Proc. IEEE Conference on Electrical Insulation and Dielectric Phenomena (CEIDP), 2014, pp. 413–416.
[36] J. Zarei, M. Shasadeghi, A. Ramezani, Fault prognosis in power transformers using adaptive-network-based fuzzy inference system, J. Intell. Fuzzy Syst. 26 (2014) 2577–2590.
[37] D.R. Morais, J.G. Rolim, A hybrid tool for detection of incipient faults in transformers based on the dissolved gas analysis of insulating oil, IEEE Trans. Power Deliv. 21 (2) (2006) 673–680.
[38] G. Vachtsevanos, F.L. Lewis, M. Roemer, A. Hess, B. Wu, Intelligent Fault Diagnosis and Prognosis for Engineering Systems, Wiley, 2006.
[39] L. Liao, Design of a reconfigurable prognostics platform for machine tools, Expert Syst. Appl. 37 (2009) 240–252.
[40] A.N. Srivastava, J. Han, Machine Learning and Knowledge Discovery for Engineering Systems Health Management, in: Chapman & Hall/CRC Data Mining and Knowledge Discovery Series, CRC Press, 2011.
[41] A. Mosallam, K. Medjaher, N. Zerhouni, Data-driven prognostic method based on Bayesian approaches for direct remaining useful life prediction, J. Intell. Manuf. (2014) 1–12.
[43] M. Schwabacher, K. Goebel, A survey of artificial intelligence for prognostics, in: Proc. AAAI Fall Symposium, 2007, pp. 107–114.
[43] A. Heng, S. Zhang, A. Tan, J. Mathew, Rotating machinery prognostics: state of the art, challenges and opportunities, Mech. Syst. Signal Process. 23 (3) (2009) 724–739.
[44] B.C.P. Lau, E.W.M. Ma, M. Pecht, Review of offshore wind turbine failures and fault prognostic methods, in: 2012 IEEE Conference on Prognostics and System Health Management (PHM), IEEE, 2012, pp. 1–5.
[45] A. Kusiak, W. Li, The prediction and diagnosis of wind turbine faults, Renew. Energy 36 (1) (2011) 16–23.
[46] A. Kusiak, A. Verma, A data-driven approach for monitoring blade pitch faults in wind turbines, IEEE Trans. Sustain. Energy 2 (1) (2011) 87–96.
[47] N. Julka, A. Thirunavukkarasu, P. Lendermann, B.P. Gan, A. Schirrmann, H. Fromm, E. Wong, Making use of prognostics health management information for aerospace spare components logistics network optimisation, Comput. Ind. 62 (2011) 613–622.

[48] M. Sharma, D. Kumar, Wishbone bus architecture – a survey and comparison, CoRR abs/1205.1860, 2012.

[49] D.F. Specht, A general regression neural network, IEEE Trans. Neural Netw. 2 (6) (1991) 568–576.

[50] L. Breiman, Bagging predictors, Mach. Learn. 24 (2) (1996) 123–140.

[51] T. Back, Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms, Oxford University Press, 1996.

[52] T.K. Sloat, J.L. Johnson, G.M.L. Sommerman, Gas evolution from transformer oils under high-voltage stress, IEEE Trans. Power Appar. Syst. 86 (3) (1967) 374–384.

[53] J. Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1988.

[54] M. Duval, A. de Pablo, Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases, IEEE Electr. Insul. Mag. 17 (2) (2001) 31–41, DEIS Feature Article.