ABNORMALITY DIAGNOSIS OF GIS USING ADAPTIVE RESONANCE THEORY

Hiromi OGI Hideo TANAKA
Yoshiakira AKIMOTO
Computer & Communication Research Center
Tokyo Electric Power Company (TEPCO)
1-4-10, Irifune, Chuo-ku, Tokyo 104 JAPAN
ogi@aisun.tepco.co.jp

Yoshio IZUI
Industrial Systems Laboratory
Mitsubishi Electric Corporation
8-1-1,Tsukaguchi Honmachi,Amagasaki
Hyogo,661 JAPAN
izui@soc.sdl.melco.co.jp

Abstract

The paper presents an artificial neural network(ANN) approach using ART2(Adaptive Resonance Theory 2) to a diagnostic system for a Gas Insulated Switchgear(GIS). To begin with, we will show the background of abnormality diagnosis of GISs from the view point of predictive maintenance of them. Then, we will discuss the necessity of ART-type ANNs, as an unsupervised learning method, in which neuron(s) are self-organized and self-created when detecting unexpected signals even if un-trained by ANNs through a sensor. Finally, we will present our brief simulation results and their evaluation.

Keywords : Gas Insulated Switchgear, Adaptive Resonance Theory, Artificial Neural Network, Abnormality Diagnosis

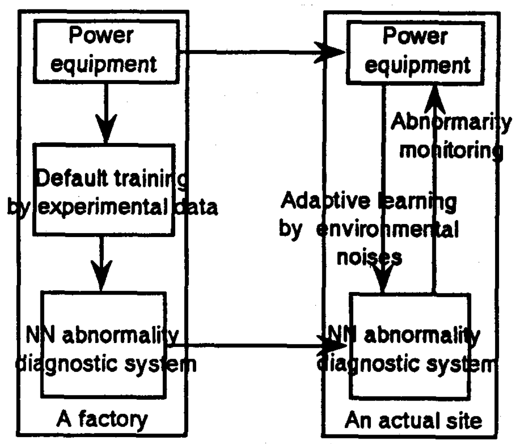

Figure 1: The Idea of the Adaptive learning

1 Introduction

Abnormality diagnosis of power equipment such as GISs, before a fatal fault occurs in a power system in particular, is playing an important role in keeping the reliability of electricity supply. It requires the advanced information processing technologies either to process analog signals attached to the equipment or to identify the abnormality more quickly. There have been a lot of sensing devices or systems developed and attached to power equipment so far. The principle of them, however, is a kind of level-detection method, by which a monitoring system only catches the magnitude of a signal and transfers it to a maintenance personnel at the control center. Such a diagnostic system has little abilities of noise tolerance, location and cause of abnormality as well as precise and quick identification.

On the other hand, ANNs that simulate human nervous system may have new abilities such as learning, adaptation, self-organization that have big potential in abnormality diagnosis. Among the variety of ANN architectures, the need for self-organized and self-enhanced ANNs like ART2 arises as a function of the diagnostic system to adapt ANN itself to the equipment variations, actual environmental variations and unexpected abnormalities.

As power equipment under operation in an actual site has very little fault possibilities, we have no way obtaining abnormality data for ANN training except through factory experiments. As for the equipment variations, we have several kinds of types of GISs that each training set may be required for every type of GIS. However, if we can assume a typical GIS and make the diagnostic system adapt to the other GISs, no training data for them may be needed any more.

As for the actual environmental variations, the environment in which GISs are equipped is different from the one in the factory. In addition, seasons and weather make the environmental variations. Unless the diagnostic system adapts itself to the environmental variations, it may misclassify an environmental noise as an abnormal status. The diagnostic system must adapt itself to the actual environment through self-training the external noises under operation of equipment.

As for the unexpected abnormality problem, this means that there may not be included in the already-trained abnormality-cause set of data. Even in this situation we require the diagnostic system that can classify the unexpected abnormality as not abnormal status but unusual status. In the paper the abnormal status means just abnormal identified by inspection of the equipment, and the no-normal status means the one including whether an abnormality or a noise.

We call the concepts above adaptability of the diagnostic system in the paper. ANNs that have static characteristics never resolve those problems, because they learn only pre-obtained training set and lack the adaptability. ART2-type ANNs, however, may have the potential of the adaptability to resolve those problems because of its self-organization and self-enhancement capabilities.

2 Predictive Maintenance for GISs

2.1 Predictive Maintenance

The fundamental idea of predictive maintenance is that detecting signals through sensors as symptoms of an abnormality before a fatal damage of equipment happen, allows us to keep the normal operation for the equipment. Even if a small signal of the symptom is found in the equipment, with the signal magnitude of the phenomena becoming larger, that may lead to a fatal damage of the equipment. That is, it is important for the predictive maintenance system that it has to detect smaller symptoms before a fatal fault. To get better performance of diagnosis, we have to consider both sensing devices and signal processing technologies.

2.2 Sensing devices for GISs

It is popular to catch the symptoms from the partial discharge phenomena in order to detect abnormalities of GISs. The symptoms include sound, flush, heat, gas contamination, electro-magnetic waves and leakage current etc. occurred inside an enclosure of the GIS. Sensing devices must catch these physical signals out of the enclosure. There are some devices such as acceleration sensor detecting mechanical vibration, supersonic sensor, differential voltage sensor, gas checker detecting dissolved insulation gas, and so on. It is said that sensing devices themselves have higher performance.

2.3 Signal Processing of Sensor Output

The principle of conventional method was level-detection of signal magnitude. That is, if the output of a sensor is over a threshold level set in advance, the judgement is abnormal. On the other hand, if it is under the threshold, the judgement is considered as normal. Only whether normal or abnormal can be identified by the method. Precise judgement is also very difficult in the method. We have a premise that abnormal conditions has always bigger signals than threshold level manufacturers usually decided. Sometime, misclassification may happen because an environmental noise makes the signal output distorted.

3 Adaptive Learning

3.1 Why Adaptive Learning?

There are problems discussed so far in developing GIS diagnostic system even if ANN technologies are employed. The problems are related to;

- Unexpected abnormality
- Environmental noise variations
- Equipment structure variations

The problems above may be expected to be resolve by the adaptive learning of self-organized and self-enhanced ANNs.

3.1.1 Unexpected Abnormality

Generally speaking, few abnormality in the equipment has been found out in its operation. Imaginary abnormal conditions have been pre-defined and experimented to obtain some abnormal data for NN training in a manufacturer's factory of GISs.[5] If these abnormal data are trained by an ANN, whether supervised-training or unsupervised-training, correct classifications would be provided according to their abnormality-cause training set. Unexpected abnormality or no-normal status, however, would not be guaranteed to identify the cause of it even if any conventional technology is used. Although ANN may not identify what the cause is, ANN is expected to classify it as the other abnormality-cause pattern which has not been trained so far.

3.1.2 Environmental Noise

The data for training have been acquired in a manufacturer's factory, where several experiments are conducted on the target equipment with some sensing devices and under better condition in general than an actual site. The training data are obtained with some noises in factory environment, which may not be the same as the ones in a site environment. To solve the problem training data should be directly measured in the site, after that, additional abnormality-cause data set should be trained to the ANN. The actual environment, however, varies from hour to hour, for example, weather variation such as fine and rainy, calm and windy. Some are the sounds of circuit breakers and disconnecting switches, Some are the vibrations from steps of inspection personnel and from collisions between the enclosure and rainfalls. We can not obtain data of all those variations for ANN training. One of the compromising approaches is the combination of default training in the factory and self-training during a certain period in the site, where a new pattern not including the default training patterns, is regarded as normal status temporally.

3.1.3 Equipment Variation

ANN approaches require sufficient and non-partial data for training, which come from a fixed-object environment. The fixed-object means that training data obtained from a specified GIS are available to the very GIS which has the same specifications on its physical structures. It is expected that typical GISs should be experimented with an ANN for the default training, the ANN trained should be built in the other GISs.

3.2 Method for Adaptive Learning

Figure 1 shows the method for the default training and adaptive learning. In the manufacturer's factory, a diagnostic with an ANN is trained by artificially-generated abnormality data from a standard equipment (Default training). In a site of a power company, the equipment is transported to the site and equipped with an environment under operation, the diagnostic system is also attached to the equipment, and monitoring the external noises for certain period as well as the internal status of the equipment.

3.3 Constraints for Adaptive Learning

One of the easiest way of adaptive learning is to continue to re-train on-line data which are obtained and restored in memories as comparing them with already-stored data. The way has at least two constraints, there are;

- As the way has to restore the past training data, a large amount of memories will be required and that lineally increased according to time in the diagnostic system.
- Strictly speaking, the way has no supervised signal of training, so that a human has to teach what the signal means and why it comes from after the system detected an unexpected data.

3.4 Behavior of Adaptive Learning NN

We considered that the current training data are only used in the default training stage without storing the past-trained data. The diagnostic system should be able to identify either the abnormality cause which was already-trained during default training or no-abnormal status that is unexpected abnormality which was no alternative in the training pairs.

If the hierarchical NN is used in the diagnostic system, it may judge the identification well for well-trained pairs, but a large amount of memory may be required in the case of hierarchical NN, because all the training data in the past must be restored. In order to cope with the constraints above, the output neurons of the adaptive learning NN should be labeled, when one of them is already-trained as default training is invoked, the label of the neuron is displayed, when there happens an input pattern which is not included in the default learning pairs, an additional neuron of the output will be generated with a label of input signal or pattern not already-trained so far. As in this stage it is difficult for the system to identify the abnormality-cause of the generated neuron, a human expert may have to teach the cause of no-normality to the system.

4 Basis of ART2

4.1 Why ART2?

To meet the constraints with NN architectures, ART2(Adaptive Resonance Theory 2) with analog data is examined in the beginning of the study, because it has features of self-organization for unexpected changes in an external environment and that autonomously and in real-time. In the basic idea of ART2, a bottom-up information is focused on an expectation based on a top-down information, so that past memory would not be erased by new learning events. If a learning event is not stored in a database, the data will be stored in the database without inconsistency.

4.2 Outline of ART2

ART2 was proposed by Carpenter&Grossberg of Boston University in 1986.[6, 7, 8] There are some versions like ART1 for binary data, ART2 for analogous data and ART3 refined by biological findings. The ART2 features sensory system for creatures which can simulate to autonomously learn and identify patterns from the continuous signals including noise input under on-line and real-time conditions. The structure of the ART2 is two-layered network which consists of input neurons and output neurons labeled by invoked patterns. The ART2 is one of the unsupervised and feedback-type NNs which means that information of output neurons affects the input neurons. We focus on the features capable of on-line, real-time and self-trained pattern classifier.

The learning algorithm of ART2 is based on the Hebbian learning rule which is the connections of neurons enhanced each other when they are excited at the same time. The ART2 is a self-created neural network model, which uses a clustering algorithm. The algorithm is that the similarity of input patterns is compared with already-learned LTM(Long term Memory)s, the most similar LTM is revised by the input pattern. If the degree of the similarity is less than the pre-defined threshold, that is the input pattern is not similar than any other past-trained patterns, then the new input pattern will be stored in the LTMs. There are two differences between the ART2 and conventional clustering. In the conventional clustering, clustering center and input pattern are normalized by unit vector, on the other hand, in the ART2, they are not always normalized because of biological findings. The ART2 takes the approach of resonance between an input pattern and the LTM, which means the ART2 revises the input pattern based on the LTM after it remembers the most similar LTM.

4.3 Dynamics of ART2

Figure 2 illustrates some ART2 dynamics features. There are two principal fields of ART2, one is an attentional subsystem which contains an input representation field F_1 and a category representation field F_2, and an orienting subsystem which interacts with the attentional subsystem. The two fields are linked by both a bottom-up adaptive filter and a top-down adaptive filter. A path from the ith F_1 node to the jth F_2 node contains a long term memory(LTM) trace, or adaptive weight $\mathbf{z}_{ij}$. The $\mathbf{I}$ is an input vector of a signal, the vector $\mathbf{w}$ is the sum of the input vector $\mathbf{I}$.

Each rectangle shows a short term memory(STM), in which internal feedback signal vector $a\mathbf{u}_i$ and activity vector $\mathbf{x}_i$ are calculated.

$$\mathbf{w}_i = \mathbf{I}_i + a\mathbf{u}_i \tag{1}$$

$$\mathbf{u}_i = \frac{\mathbf{v}_i}{\|\mathbf{v}\|} \tag{2}$$

$$\mathbf{x}_i = \frac{\mathbf{w}_i}{\|\mathbf{w}\|} \tag{3}$$

At the top F_1 layer $\mathbf{p}$ sums both the internal F_1 signal $\mathbf{u}$ and all the $F_2 \rightarrow F_1$ filtered signals. That is,

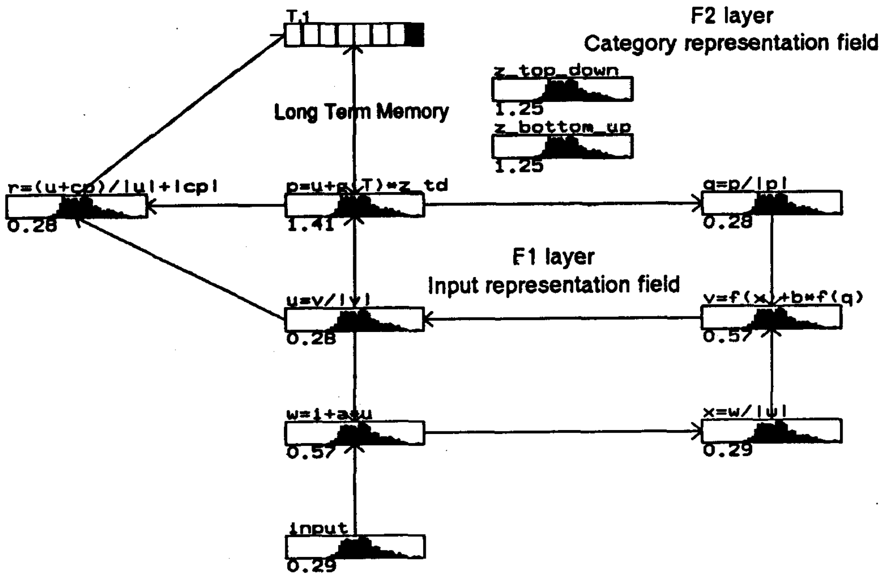

Figure 2: Dynamics of ART2

$$\mathbf{p_i} = \mathbf{u_i} + \sum_j g(\mathbf{y_j})\mathbf{z_{ji}} \tag{4}$$

where $g(\mathbf{y_j})$ is the output signal from the jth F_2 node and $\mathbf{z_{ji}}$ is the LTM trace in the path from the jth F_2 node to the ith F_1 node.

$$\mathbf{q_i} = \frac{\mathbf{p_i}}{||\mathbf{p}||} \tag{5}$$

$$\mathbf{v_i} = f(\mathbf{x_i}) + bf(\mathbf{q_i}) \tag{6}$$

$$f(x) = \begin{cases} 0 & if \quad 0 \le x \le \theta \\ \theta & if \quad x \ge \theta \end{cases} \tag{7}$$

There is a resonance in the LTM between $\mathbf{T_j}$ and $\mathbf{p}$. The vector $\mathbf{r}$ monitors the degree of match between F_1 bottom-up input $\mathbf{I}$ and the top-down input $d\mathbf{z_j}$. System reset occurs if $||\mathbf{r}|| < \rho$, where ρ is a dimensionless vigilance parameter between 0 and 1. Vector $\mathbf{r}$ obeies the following equation.

$$\mathbf{r_i} = \frac{\mathbf{u_i} + c\mathbf{p_i}}{||\mathbf{u}|| + ||c\mathbf{p}||} \tag{8}$$

$$\rho > ||\mathbf{r}|| \tag{9}$$

Suffix shows the number of a neuron(the element number of vectors), $|| * ||$ is a norm and a, b, c, d are appropriate parameters.

$$\frac{d\mathbf{zij}}{dt} = g(\mathbf{y_j})[\mathbf{p_i} - \mathbf{z_{ji}}] \tag{10}$$

$$\frac{d\mathbf{zji}}{dt} = g(\mathbf{y_j})[\mathbf{p_i} - \mathbf{z_{ji}}] \tag{11}$$

$$g(\mathbf{y_j}) = \begin{cases} d & if \quad \mathbf{T_j} = max_j \mathbf{T_j} \\ 0 & otherwise \end{cases} \tag{12}$$

$$\mathbf{T_j} = \sum_i \mathbf{p_i}\mathbf{z_{ij}} \tag{13}$$

In the short term memory, although we have to note that differential equations(8) and (9) should be calculated as well as LTM, so that we can get calculation time shorten by calculating the stable status. Equation(7) means the value less than θ is regard as noise, there is no problem even if it is ignored.

In the long term memory, $\mathbf{z_{ij}}$ or $\mathbf{z_{ji}}$ is equivalent to a kind of template or prototype. Equation(10) and (11) are learning equations. The most similar template is selected by equation(12) and (13). After the template is compared with the evaluation function $\mathbf{r}$, a new neuron is generated in the long term memory if the value is less than ρ.

In the actual ART2 diagnostic system implementation, abnormality-cause label should be assigned to the neuron $\mathbf{T_j}$ in the LTM when the cause is already known.

5 Simulation

5.1 Data Acquisition

We had conducted a field experiment for data acquisition from the 275kV GCB(Gas-insulated Circuit Breaker) to evaluate the methodology of identifying abnormalities during April 19— April 28,1991 at the TAMAHARA Powerstation in TEPCO. The result showed that there was no abnormality during the experiment. The data were acquired from one acceleration sensor fixed by metallic enclosure and protected out of external vibrations, so that in fact only data we got during the period was the ON/OFF switching signal from the circuit breaker.

On the other hand, we have also carried out a preliminary test generating abnormal conditions at the factory of the GIS Manufacturer during November — December,1992 to get continuous signal of two kinds of sensors which were

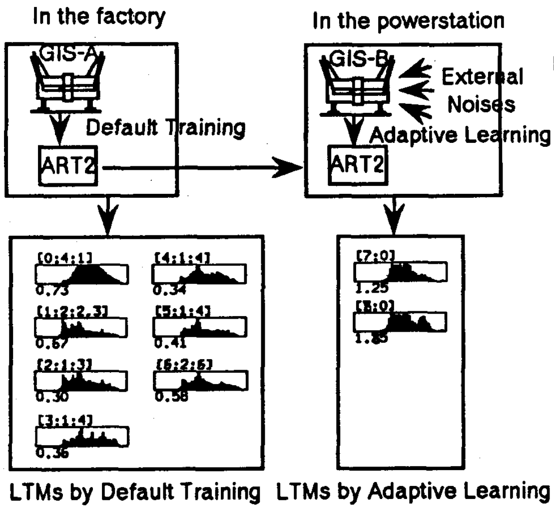

Figure 3: Default training and Adaptive training

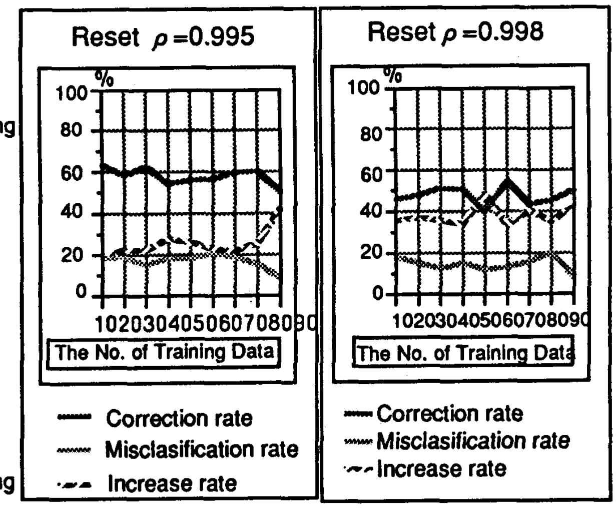

Figure 4: Reset threshold

originally mounted on the enclosure of GIS, that is, an acceleration sensor and a differential voltage sensor.

In the first step we assumed six patterns of experimental conditions of abnormalities which include particles inside an enclosure of GIS and bad contacts modeled by a small gap between conductors and a loose-coupled ring around a conductor. A particle made of a thin aluminum wire of 1mm diameter and a ring around a conductor were used in the experiment. The abnormalities caused by particles were modeled by particle fixed on the conductor, the one fixed on the enclosure and floating one within the enclosure. These particles and bad contacts make partial discharge inside the enclosure under high voltage, which vibrates the insulated gas of SF_6 and the enclosure itself. The acceleration sensor mounted on the outside of the enclosure detects the acceleration change of the vibration.

An example of the conditions were set in the test as follows:

- Insulated gas used: SF_6 $4kg/cm^2$
- Test voltage: 0 — 2× nominal voltage
- Particle model: A thin aluminum wire
- The length of a particle: X=28 mm, Y=40 mm
- Abnormal conditions tested:
  1. No particle(i.e. normal status)
  2. Fixed particle on the conductor
  3. Fixed particle on the enclosure
  4. Floating particle within the enclosure
  5. A small gap between conductors
  6. An insufficiently-fixed metal fitting

5.2 Outline of the Experiment

In the application of ART2 to abnormality diagnosis of GISs, learning stages are separated from default training and adaptive learning. GIS-A in the factory and GIS-B in the powerstation have been experimented by the ART system, they have almost same specifications.

1. Default training in the factory
   In the manufacturer's factory shown in left in Figure 3, we have obtained approximately 100 data for training including the six abnormalities above from the GIS-A. Then the ART2 was trained by the data as varying the number of them to check the dependency of relations between correction rate and the number of training data. ART2 learns the abnormality status with little noises derived from pre-assumed abnormal causes so that it predicts perfectly pre-assumed abnormalities.

2. Adaptive learning in the Powerstation
   In the stage of adaptive learning shown right in Figure 3, ART2 system was put in the GIS site of a powerstation, and had been collected the environmental noises so that it performed classifications of them. If one of them is classified to one of pre-assumed abnormal causes, ART2 modifies a LTM in the already-learned LTMs in the system. If it is not classified to any LTM, ART2 creates one more new neuron and classifies it as a new no-normal status. In fact, the ART2 system trained by default six abnormality-cause pairs were implemented at the Protective Relay Room in the TAMAHARA Powerstation, where an output cable from an acceleration sensor mounted on the enclosure of GCB(GIS-B) was connected to the system. The system had been monitoring the GCB during almost 10 days. For the convenience after analysis of the experiment, raw signals from the sensor were also recorded by DAT(Digital Audio Tape)s.

6 Evaluation

The sample data of the spectrum patterns were selected in random under nominal voltage to train the ART2 by data out of about 100 samples. Both correction rate and misclassification rate were evaluated in the GIS-A.

6.1 Default training in the factory

Figure 4 show the correction rate and misclassification rate. 10% to 90% out of the 100 data were selected, each percentage of the data was trained on the ART2, rest of the data were input to the ART2 to evaluate the correction rate and misclassification rate, that is considered trained data as assumed abnormalities, rest of the data as unexpected abnormalities.

The correction rate means that when a pattern is input to the ART2, the cause of it is already trained and is equivalent to the LTM invoked.

The misclassification rate means that the cause of it is already trained but is not equivalent to the LTM invoked. Increase rate means that the cause of it was not trained by the training data. The left hand figure is a threshold of the Reset as $\rho = 0.995$, the right hand one is that as $\rho = 0.998$ which is tighter than the left, in other word, it is difficult to classify the unexpected abnormalities correctly, so that right one shows there are more neurons self-generated than the left one.

6.2 Adaptive learning in the Powerstation

Figure 3 show the LTMs stored by default training in GIS-A and that of by adaptive learning in GIS-B. After the diagnostic system with the ART2 was trained by data from the GIS-A, the system was provided with the GIS-B in the TAMAHARA Powerstation and input by on-line signals from the acceleration sensor mounted on the GIS-B. Adaptive learning was conducted for several days in the site, that is in a new environment. In fact, in the actual field, there was no abnormality except environmental noise. The pattern of the right hand figure shows FFT patterns come from abnormal status with the environmental noise and switching status of the circuit breaker of the GIS-B.

7 Conclusion

We have applied ART2 as self-trained NN, on-line and in real-time, to abnormality diagnosis of GISs. It is considered that signal processing technologies in the power industry are now less under development than sensor technologies are.

GISs have been operated for over 20 years in the past, and its rate of faults and malfunctions are fairly lower than those in the other conventional equipment partly because of the protection of charged section against external environment. But invisibility inside GISs prevents us from inspecting them easily through daily patrol or quick-locating an abnormality before fatal damage occurs. We believe that gas-insulated substations bring their power facilities maintenance-free out of weathering and contamination problems and reduced size of equipment and devices, but locating (of faults and malfunctions) problems such as internal corona activity or any breakdown of the insulation system, have been unsatisfactory. Therefore further R&D improvement for predictive diagnostic technologies should be done including interpreting online signal data from single or multiple sensors. We expect that pattern classification capability of ANNs with learning function might bring one of the solutions in the near future. We will investigate any other adaptive type ANNs as well as the ART2 and develop the prototype of diagnostic system with adaptive type NN in a couple of years.

Acknowledgement

The authors wish to express gratitude to those who works in the ITAMI WORKS of Mitsubishi Electric Corporation and in the TAMAHARA Powerstation of the TEPCO for their great contributions to the approaches.

References

[1] L.Lundgaad et al., "Acoustic Diagnoses of Gas Insulated Substations; A Theoretical and Experimental Basis", *IEEE PES Winter Meeting*, Paper no.90 WM 133-9 PWRD, Atlanta, 1990

[2] T.Yamagiwa et al., "Development of Preventive Maintenance System for Highly Reliable Gas Insulated Switchgear", *IEEE PES Winter Meeting*, 90 WM 160-2 PWRD, Feb.1990.

[3] H.Tanaka,et al., "Design and Evaluation of Neural Network for Fault Diagnosis", *Proc. of ESAP'89*, pp.378-384, July 1989.

[4] D.E.Rumelhart et al., "Learning representations by back-propagating errors", *Nature*, Vol.323,9, pp.533-536, Oct. 1986.

[5] H.Ogi et al., "Preventive Maintenance System for Gas Insulated Switchgear Using an Artificial Neural Network", *Proc. of ESAP'91*, pp.627-633, April 1991.

[6] G.A.Carpenter and S.Grossberg, "ART2: self-organization of stable category recognition codes for analog input patterns", *Applied Optics*, vol.26, No.23, pp.4919-4930, 1987.

[7] G.A.Carpenter and S.Grossberg, "A Massively Parallel Architecture for a Self-Oeganizing Neural Pattern Recognition Machine", *Computer Vision Graphics Image Processing*, vol.37, No.54, 1987.

[8] G.A.Carpenter and S.Grossberg, "ART3: Hierarchical Search Using Chemical Transmitters in Self-Organizing Pattern Recognition Architectures", *Neural Networks*, Vol.3, No.2, pp.129-152, 1990.

[9] D.L.Lubkeman et al., "Unsupervised Learning Strategies for the Detection and Classification of Transient Phenomena on Electric Power Distribution Systems", *IEEE Applications of Neural Networks to Power Systems(ANNPS91)*, p.107-p.111, July 1991.

[10] H.Ogi et al., "Fault Diagnosis System for GIS using an Artificial Neural Network", *IEEE Applications of Neural Networks to Power Systems(ANNPS91)*, p.112-p.116, July 1991.

[11] Mori et al., "Power System Voltage Diagnosis by an Artificial Neural Network", *IEE Electric&Energy*, 138, 1992.