

LOCATING PD IN TRANSFORMERS THROUGH DETAILED MODEL AND NEURAL NETWORKS

Hamed Nafisi — Mehrdad Abedi — Gevorg B. Gharehpetian *

In a power transformer as one of the major component in electric power networks, partial discharge (PD) is a major source of insulation failure. Therefore the accurate and high speed techniques for locating of PD sources are required regarding to repair and maintenance. In this paper an attempt has been made to introduce the novel methods based on two different artificial neural networks (ANN) for identifying PD location in the power transformers. In present report Fuzzy ARTmap and Bayesian neural networks are employed for PD locating while using detailed model (DM) for a power transformer for simulation purposes. In present paper PD phenomenon is implemented in different points of transformer winding using three-capacitor model. Then impulse test is applied to transformer terminals in order to use produced current in neutral point for training and test of employed ANNs. In practice obtained current signals include noise components. Thus the performance of Fuzzy ARTmap and Bayesian networks for correct identification of PD location in a noisy condition for detected currents is also investigated. In this paper RBF learning procedure is used for Bayesian network, while Markov chain Monte Carlo (MCMC) method is employed for training of Fuzzy ARTmap network for locating PD in a power transformer winding and results are compared.

Key words: Bayesian network, detailed model (DM), fuzzy ARTmap (FAM) neural network, partial discharge (PD), transformer

1 INTRODUCTION

Partial discharges (PD) are well known as a source of insulation degradation and the major sources for insulation failure in power transformers, which play important role in electric power system [1, 2]. The capital cost of a power transformer is relatively high and economic penalty due to transformer failure and consequent outage is remarkable. Thus deterioration of insulated material caused by PD activity can be detected in early stage, then incipient insulation failure can be identified and preventive maintenance measures can be done [3]. PD detection technique is classified into acoustic and electrical methods. Electrical method is based on detecting of created impulses in the cavity of transformer insulation. Assessment of PD in electrical method is possible by using current transducers, which are connected to measuring terminals. In this method different procedures such as tip-up, dielectric loss analysing, inductive probes, pulse detecting and analysing, or other methods can be employed [4].

The advantage of acoustic relative to electrical technique for PD detection is its simplicity. However acoustic method has low sensitivity. On the other hand complicated structure of power transformer causes difficulty due to propagation velocity of acoustic waves associated with PD [5]. Therefore in recent years most reports are available concentrated on electrical methods [3–7]. Most of these reports deal with discharge between transformer winding and ground, and discharge between coil to coil has received little and incomplete attentions.

In this paper, partial discharge in the insulation between coil to coil is considered With EMTP simulation tools and DM of transformer. Then the current of neutral point of winding was measured when PD model was located at different positions in the winding and is used to finding location of PD using Fuzzy ARTmap neural network and Bayesian network. Simulated results must contain measurement noise for approximate to the truth. For mentioned reason in the last section simulated currents is changed to new one with considering measurement noise. Then the corrected currents are used for determination of the location of PD in transformer with aforesaid neural networks.

2 PARTIAL DISCHARGE MODEL

PD is localized ionization within insulator caused by high strength electric field. PD occurs in the part of insulation and is limited to some extends. Therefore PD does not cause full insulation breakdown immediately [4]. In this paper three-capacitor model shown in Fig. 1 is employed for PD modeling, which its accuracy is verified by EMTP software [5]. In this model we have

- C_g is the capacitance of the region in which discharge occurs,
- C_b is the capacitance of region located in series with C_g ,
- C_a is the capacitance of the other region in dielectric.

If discharge happens in C_g a current (I_d) flows from external terminals through C_a and C_b .

* Electrical Engineering Department, Amirkabir University of Technology, Tehran, No. 424, Hafez Ave, 15914, Tehran, Iran, nafisi@aut.ac.ir, abedi@aut.ac.ir, grptian@aut.ac.ir

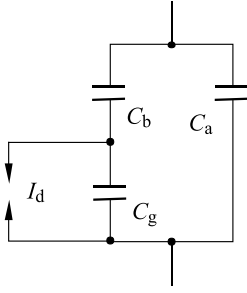


Fig. 1. Three-capacitor model for PD

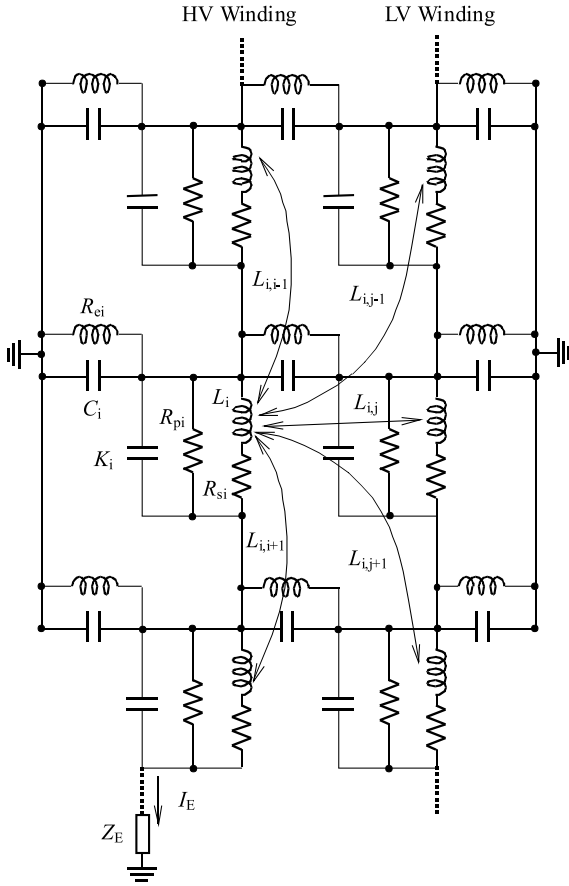


Fig. 2. DM of a two winding transformer

3 DETAILED MODEL

The equivalent circuit diagram of the test objects beyond 10 kHz is shown in Fig. 2. A winding unit can contain one disk, two disks or several numbers of turns. The number of units is a modeling parameter and the chosen value is a compromise between the accuracy and the complexity. For the sake of simplicity only three winding units of the double disk high voltage winding are shown in Fig. 2. Only one layer with three winding units have been shown for the low voltage winding in Fig. 2, too. This model is called detailed model [8].

The elements of the circuit diagram are defined in [8]. Using this model, it is possible to calculate node voltages

and branch currents in the time as well as in the frequency domain. Due to the frequency dependence behaviour of the resistive elements (R_{pi} , R_{ei} and R_{si}) the calculation in the frequency domain is preferable.

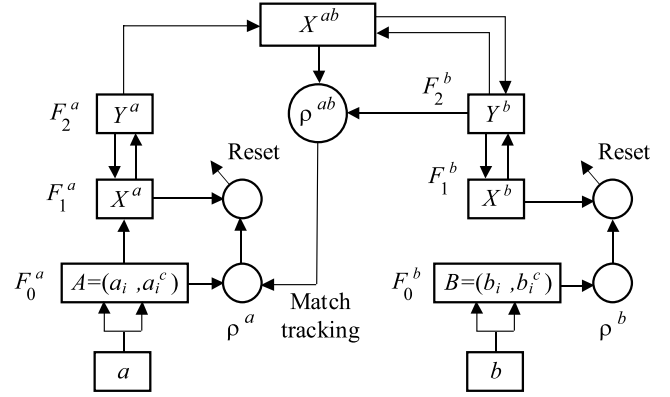


Fig. 3. Typical fuzzy ARTmap

Model parameters are calculated analytically after some simplifications of the geometrical structure of the winding. Determination of self inductance, mutual inductance, capacitances, and resistances is described in [8].

4 FUZZY ARTMAP NEURAL NETWORK

For the purpose of training and testing, in this paper Adaptive Resonance Theory (ART) neural networks have been used. In general, this family of neural networks include ART1, ART2 [9], ART3 [10], ARTmap [11], Fuzzy ART [12] and Fuzzy ARTmap [13]. ART1 and ARTmap categorize the binary input patterns while, Fuzzy ARTmap are also capable to categorize analogue patterns.

Fuzzy ARTmap is an incremental supervised learning algorithm which combines fuzzy logic and Adaptive Resonance Theory (ART) neural network for recognition of pattern categories and multidimensional maps in response to input vectors presented in an arbitrary order. It realizes a new minmax learning rule which conjointly minimizes predictive error and maximizes code compression, and therefore gives generalization. This is achieved by a match tracking process that increase the ART vigilance parameter (fuzzy degree of membership of the input with respect to the category templates) by the minimum amount needed to correct a predictive error (PE). The Fuzzy ARTmap neural network is composed of two Fuzzy ART modules [13], *ie* fuzzy ART_a and fuzzy ART_b , which are depicted in Fig. 3 and are essentially the same as those described by Carpenter *et al.*

The interactions mediated by the map field F^{ab} operationally characterized as follows.

4.1 ART_a and ART_b

Inputs to ART_a and ART_b are in the complement code form: for ART_a $I = A = (a, a^c)$ and for ART_b $I = B = (b, b^c)$ (See Fig. 3). Variables in ART_a or ART_b are designated by subscript “a” and “b” respectively. For ART_a , let $x^a = \{x_1^a, \dots, x_{2M_a}^a\}$ denote the F_1^a output vector, let $y^a = \{y_1^a, \dots, y_{N_a}^a\}$ denote F_2^a , and let $w_j^a = \{w_{j1}^a, \dots, w_{j2M_a}^a\}$ denote the j^{th} ART_a weight vector. For ART_b , let $x^b = \{x_1^b, \dots, x_{2M_b}^b\}$ denote the F_1^b output vector and let $y^b = \{y_1^b, \dots, y_{N_b}^b\}$ denote F_2^b . And let $w_k^b = \{w_{k1}^b, \dots, w_{k2M_b}^b\}$ denote the k^{th} ART_b weight vector. For the map field, let $x^{ab} = \{x_1^{ab}, \dots, x_{N_a}^{ab}\}$ denote the F^{ab} output vector, and let $w_j^{ab} = \{w_{j1}^{ab}, \dots, w_{jN_b}^{ab}\}$ denote the weight vector from the j^{th} F_2^a node to F^{ab} . Vectors x^a , y^a , x^b , y^b , and x^{ab} are set to 0 between input presentations.

4.2 Map Field Activation

The map field F^{ab} is activated whenever one of the ART_a or ART_b categories is active. If node J of F_2^a is chosen, then its weights w_j^{ab} activate F^{ab} . If K in F_2^b is active, then node K in F^{ab} is activated by one-to-one pathways between F_2^b and F^{ab} . If both ART_a and ART_b are active, then F^{ab} becomes active only if ART_a predicts the same category as ART_b via the weights w_j^{ab} . The F^{ab} output vector x^{ab} obeys the following

$$\left\{ \begin{array}{ll} y^b \wedge w_j^{ab} & \text{If the } j^{\text{th}} F_2^a \text{ is active} \\ & \text{and } F_2^b \text{ is active.} \\ w_j^{ab} & \text{If the } j^{\text{th}} F_2^a \text{ is active} \\ & \text{and } F_2^b \text{ is inactive.} \\ y^b & \text{If the } j^{\text{th}} F_2^a \text{ is inactive} \\ & \text{and } F_2^b \text{ is active.} \\ 0 & \text{If the } j^{\text{th}} F_2^a \text{ is inactive} \\ & \text{and } F_2^b \text{ is inactive.} \end{array} \right. \quad (1)$$

From (1), $x^{ab} = 0$ if the prediction w_j^{ab} is disconfirmed by y^b . Even such a mismatch triggers and ART_a search for a better category, as follows.

4.3 Match Tracking

At the start of each input presentation, the ART_a vigilance parameter ρ_a equals to baseline vigilance ρ_a . The map field vigilance parameter is ρ_{ab}

$$\text{If } x^{ab} < \rho_{ab} \cdot |y^b|. \quad (2)$$

The ρ_a is increased until it is slightly larger than $|A \wedge w_j^a| |A|^{-1}$, where A is the input to F_1^a , in complement coding form, and

$$|x^a| = |A \wedge w_j^a| < \rho_a |A| \quad (3)$$

where J is the index of active F_2^a node.

When this occurs, ART_a search leads either to activation of another F_2^a node J with

$$|x^a| = |A \wedge w_j^a| \geq \rho_a |A| \quad (4)$$

and

$$|x^a| = |y^b \wedge w_j^{ab}| \geq \rho_a |y^b|. \quad (5)$$

Or, if no such nodes exist, *ie* the input pattern to F_2^a layer does not match any pattern: the input pattern is classified as a new pattern.

4.4 Testing of Network

In this testing method, neural network have no surveillance on foreseen output. This means for each patterns of inputs one pattern is offered in output and this proffer downright after passing from condition (minimum error with existing output patterns) is accomplished which is caused faster response but reduction in precision because errors may so classified that clusters are sorely close to each other. In these cases resemblance of estimated pattern to each output pattern may cause accuracy descend in estimation, thereupon existence probability of error is considered.

Block diagram of accustomed testing method is shown in Fig. 4.

Also test algorithm for FAM network is depicted in Fig. 5.

5 FUNDAMENTAL OF BAYESIAN NETWORK

Bayesian networks are usually used to model the situations (*eg*, medical diagnosis) in which causality plays a role but where the understanding of what is actually going on is incomplete. That is, a Bayesian network for the domain represents a joint probability distribution over a set of variables (*ie*, chance nodes) [14].

Bayesian network is a directed acyclic graph that consists of single-evidence, multiple-evidence, and multiple-layer probabilistic relationships among the variables. For detailed description about DAGs see [15]. Thus, Bayesian network expresses the global joint distribution with a set of local distributions and relates only the neighboring nodes. Figure 6 illustrates the basic structure of a Bayesian network.

The Bayesian network has been successfully applied in many fields such as medical diagnosis [16], equipment diagnosis [17], and mineral exploration [18]. Extensive review of Bayesian networks can be found, for example, in [19].

The Bayesian network is a directed acyclic graph in which the following holds.

- A set of random variables makes up the nodes of the network.

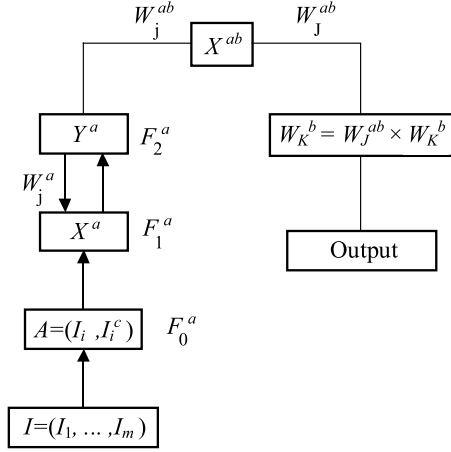


Fig. 4. Typical fuzzy ARTmap test method

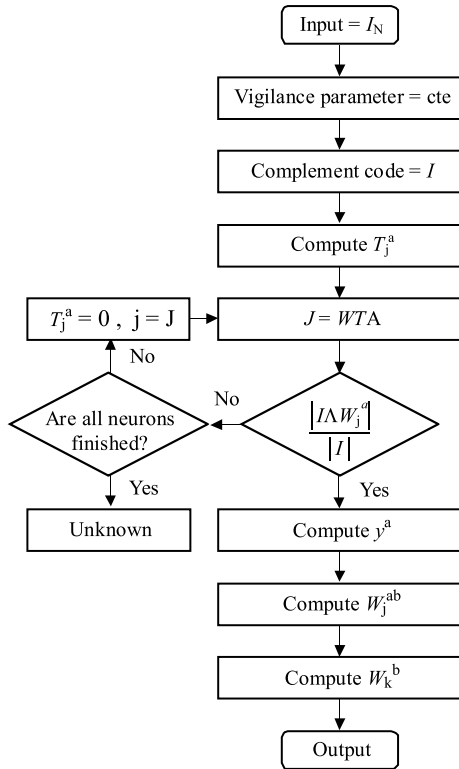


Fig. 5. Algorithm of FAM network test method

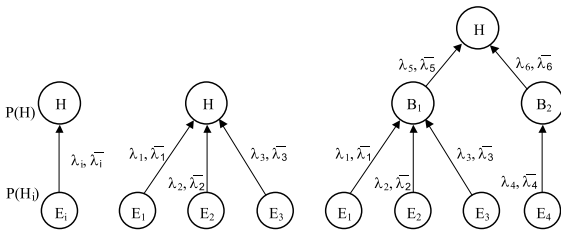


Fig. 6. Basic structures of Bayesian network (DAG) [15]

- A set of directed links or arrows connects pairs of nodes.

- Each node has a conditional probability table that quantifies the effects that the parents have on the node. The parents of a node are all those nodes that have arrows pointing to it.
- The graph has no directed cycles (hence is a directed, acyclic graph or DAG).

A Bayesian network provides a complete description of the domain. Every entry in the joint probability distribution can be calculated from the information in the network. A generic entry in the joint is the probability of a conjunction of particular assignments to each variable. The value of this entry is given by [20]

$$P(x_1, \dots, x_n) = \prod_{i=1}^n P(x_i \mid \text{parents}(X_i)). \quad (6)$$

We use the notation $P(x_1, \dots, x_n)$ as an abbreviation for this. Thus, each entry in the joint is represented by the product of the appropriate elements of the conditional probability tables (CPTs) in the belief network. The CPTs therefore provide a decomposed representation of the joint.

$$P(x_1, \dots, x_n) =$$

$$P(x_n \mid x_{n-1}, \dots, x_1) P(x_{n-1} \mid x_{n-2}, \dots, x_1) \dots P(x_1) = \prod_{i=1}^n P(x_i \mid x_{i-1}, \dots, x_1) \quad (7)$$

Then we repeat this process, reducing each conjunctive probability to a conditional probability and a smaller conjunction. We end up with one big product. Comparing this with Equation (6) and (7), we see that the specification of the joint is equivalent to the general assertion that.

5.1 Markov Chain Monte Carlo Blanket

A node is conditionally independent of its non-descendants, given its parents. A node is conditionally independent of all other nodes in the network, given its parents, children, and children's parents that is, given its Markov blanket.

From these conditional independence assertions and the CPTs, the full joint distribution can be reconstructed; thus, the “numerical” semantics and the “topological” semantics are equivalent.

According to the theory of Markov blanket, the nodes for inference, such as the fault node or protection node, are chosen first.

5.2 The MCMC Algorithm

The MCMC generates each event by making a random change to the preceding event. It is therefore helpful to think of the network as being in a particular current

state specifying a value for every variable. The next state is generated by randomly sampling a value for one of the non-evidence variables X_i , conditioned on the current values of the variables in the Markov blanket of X_i . MCMC therefore wanders randomly around the state space—the space of the possible complete assignments—flipping one variable at a time, but keeping the evidence variables fixed. The algorithm is that:

Let $q(x \rightarrow x')$ be the probability that the process makes a transition from states x to state x' . This transition probability defines what is called a Markov chain on the state space. Now suppose that we run the Markov chain for t steps, and let $P_t(x)$ be the probability of being in state x at time t . Similarly, let $P_{t+1}(x')$ be the probability of being in state x' at time $t+1$. Given $P_t(x)$, we can calculate $P_{t+1}(x')$ by summing, for all states the system could be in at time t , the probability of being in that state times the probability of making the transition to x'

$$P_{t+1}(x') = \sum_x P_t(x) q(x' \rightarrow x). \quad (8)$$

We will say that the chain has reached its stationary distribution if $P_t(x) = P_{t+1}(x')$. Let us call this stationary distribution P ; its defining equation is therefore

$$P(x') = \sum_x P(x) q(x' \rightarrow x) \quad \text{for all } x'. \quad (9)$$

Under certain standard assumptions about the transition probability distribution q , there is exactly one distribution P satisfying this equation for any given q .

Equation (8) can be read as saying that the expected “outflow” from each state (*ie*, its current “population”) is equal to the expected “inflow” from all the states. One obvious way to satisfy this relationship is if the expected flow between any pair of states is the same in both directions. This is the property of detailed balance

$$P(x') q(x \rightarrow x') = \sum_x P(x) q(x' \rightarrow x) \quad \text{for all } x, x'. \quad (10)$$

5.3 Radial Basis Function (RBF)

In order to assess the input data Radial Basis Function (RBF) method is used. A radial basis function (RBF) is a real-valued function whose value depends only on the distance from the origin, so that [21]

$$\varphi(x) = \varphi(\|x\|). \quad (11)$$

Or alternatively on the distance from some other point c , called a *center*, so that

$$\varphi(x, c) = \varphi(\|x - c\|). \quad (12)$$

Any function φ that satisfies the property $\varphi(x) = \varphi(\|x\|)$ is a radial function. The norm is usually *Euclidean distance*.

Radial basis functions are typically used to build up function approximations of the form

$$y(x) = \sum_{i=1}^N \omega_i \varphi(\|x - c_i\|) \quad (13)$$

where the approximating function $y(x)$ is represented as a sum of N radial basis functions, each associated with a different center c_i , and weighted by an appropriate coefficient ω_i . Approximation schemes of this kind have been particularly used in time series prediction and control of nonlinear systems exhibiting sufficiently simple chaotic behaviour.

The sum can also be interpreted as a rather simple single-layer type of artificial neural network called a radial basis function network, with the radial basis functions taking on the role of the activation functions of the network. It can be shown that any continuous function on a compact interval can in principle be interpolated with arbitrary accuracy by a sum of this form, if a sufficiently large number N of radial basis functions is used.

There are some commonly used types of radial basis functions include $r = \|x - c_i\|$.

- Gaussian

$$\varphi(r) = \exp(-\beta r^2) \quad \text{for some } \beta > 0. \quad (14)$$

- Multi-quadric

$$\varphi(r) = \sqrt{r^2 + \beta^2} \quad \text{for some } \beta > 0. \quad (15)$$

- Polyharmonic spline

$$\varphi(r) = \begin{cases} r^k, & k = 1, 3, 5, \dots \\ r^k \ln r, & k = 2, 4, 6, \dots \end{cases} \quad (16)$$

- Thin plate spline (a special polyharmonic spline)

$$\varphi(r) = r^2 \ln r. \quad (17)$$

These polyharmonic splines (which include the thin-plate spline) minimise certain energy semi-norms and are therefore the “smoothest” interpolators. Note that the associated basic functions are not compactly supported — they grow as r increases from the origin.

RBFs are popular for interpolating scattered data as the associated system of linear equations is guaranteed to be invertible under very mild conditions on the locations of the data points. For example, the thin-plate spline only requires that the points are not co-linear while the Gaussian and multi-quadric place no restrictions on the locations of the points. In particular, RBFs do not require that the data lie on any sort of regular grid.

In this paper Radial Basis Function method with Gaussian type is considered for assessment of data which is depicted in Fig. 7.

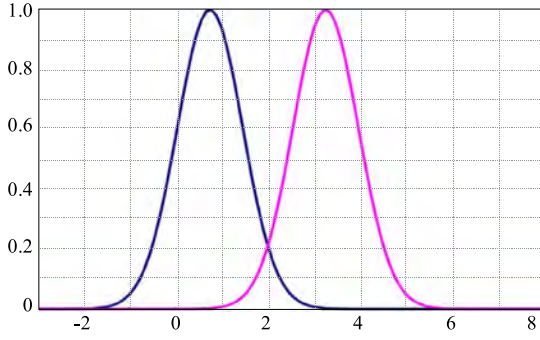


Fig. 7. Un-normalized radial basis functions with $c_1 = 0.75$ and $c_2 = 3.25$

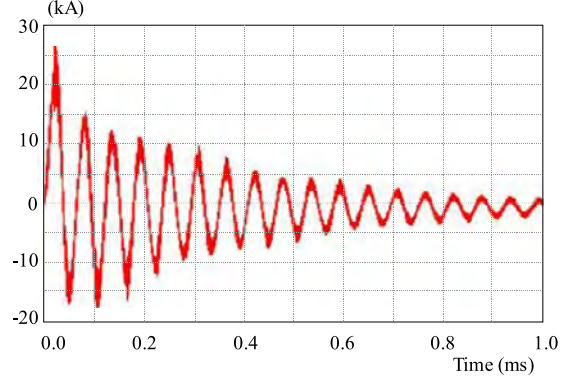


Fig. 8. Current waveform in neutral point resulting from PD model in 1st node of DM

5.4 Estimating the Weights

The approximant $y(x)$ is differentiable with respect to the weights ω_i . The weights could thus be learned using any of the standard iterative methods for neural networks. But such iterative schemes are not in fact necessary because the approximating function is *linear* in the weights ω_i , the ω_i can simply be estimated directly, using the matrix methods of linear least squares.

The input assessment has been shown in this section. In the next section required input-output data and modeling of power system will be presented.

6 CASE STUDY

DM of transformer is used for simulation PD mechanism in EMTP and this PD model is considered between coil to coil, as mentioned in previous sections.

6.1 Transformer Specification

The used transformer is 35 kV/220 kV and 50 MVA. Windings dimensions and model of the transformer tank are detailed in [4].

HV winding of the simulating transformer consist of 56 discs. The first 6 two-discs is interleaved type and the 22 other two-disc is inverted type. Dimensions of all two-discs are presented in Tab. 1.

Table 1. Technical Specification of HV Winding

Winding type	Num of discs	Disc number from-to	Num. of turns in each disc	height \times width
Interleaved	4	01-05	$13\frac{18}{20}$	3.0×15.0
Interleaved	8	05-13	$12\frac{18}{20}$	2.8×13.2
Inverted	4	13-17	$15\frac{18}{20}$	2.5×17.0
Inverted	28	17-45	$19\frac{17}{20}$	2.0×23.0
Inverted	9	45-54	$19\frac{18}{20}$	2.0×17.0
Inverted	3	54-56	$15\frac{18}{20}$	2.5×17.0
Total	56		1005	

Table 2. Results of neural network testing

Training rate	Vigilance parameter	Num. of clusters	Accuracy (%)
0.95	0.9	35	83.3
	0.95	62	66.7
	0.98	208	66.7
0.93	0.9	50	83.3
	0.95	89	100
	0.98	288	83.3
0.9	0.9	69	83.3
	0.95	103	66.7
	0.95	301	83.3

6.2 simulation Results

PD model is placed in different points of the winding and steep impulse current applied to the winding as input signal. The current at the other terminal is measured as output. Produced current in neutral point of the winding is recorded. For instance Figs. 8, 9 and 10 show the current in neutral point of the winding when PD in two-disk 1, 18 and 28 was modelled, respectively.

6.3 Test result of Fuzzy ARTmap Neural Network

In order to train the neural networks there is a need for measured training patterns. The current neutral point is used for learning of Fuzzy ARTmap neural network. This neural network is implemented in MATLAB software. Used DM consists of 28 nodes. Thus all simulated states are 28. For training of neural network 22 of them is used. This low number of data sets is because of the EMTP limitation for number of the mutual inductances.

As mentioned, necessary simulations of transformer are done in EMTP and the results used for input data for the Fuzzy ARTmap neural network in MATLAB. Result of the Fuzzy ARTmap neural network testing is shown in Tab. 2.

As it is shown in Tab. 2, the best PD location accuracy is reached to 100 for 0.93 of training rate, 0.95 for mapfield vigilance parameters. The obtained number of

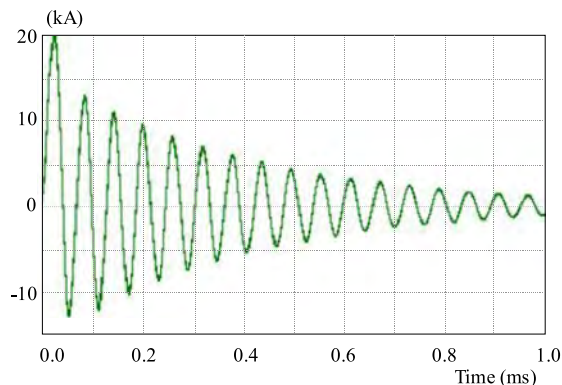


Fig. 9. Current waveform in neutral point resulting from PD model in 18th node of DM

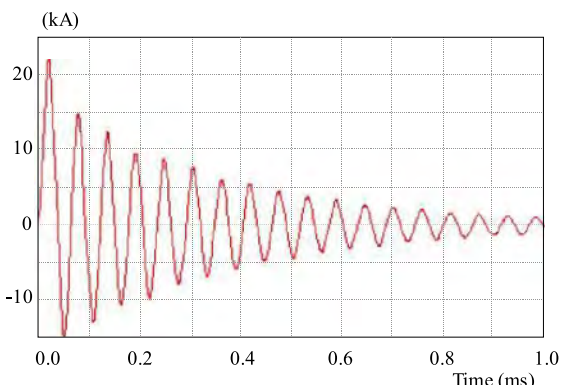


Fig. 10. Current waveform in neutral point resulting from PD model in 28th node of DM

Table 3. Results of neural network testing with noise

Training rate	Vigilance parameter	Num. of clusters	Accuracy (%)
0.93	0.9	85	75.3
	0.95	268	86.2
	0.98	971	68.8

Table 4. Results of Bayesian Network

PD located in	Output of Bayesian network	Rounded value
1	1.42	1
10	9.89	10
18	18.37	18

clusters is 89. The reason of this matter is lack of the parameters which are used for training and test of Fuzzy ARTmap neural network because of EMTP limitation. But because of sufficient difference among current waveform Fuzzy ARTmap neural network can determine the correct position of PD in power transformer.

With regard to this point that in truth and actual measured current waveforms contain measurement noise and different from simulated values, white noise is added to simulated currents. Amplitude of the white noise considered form 1 to 10 percent of peak value of simulated current amplitude in 1 percent steps. Therefore, neural network response to noisy inputs with best training rate and mapfield vigilance parameters, obtained from Tab. 2, is shown in Tab. 3.

According to Tab. 3, accuracy of Fuzzy ARTmap neural network with considering noise is about 86.2%. The best output of FAM network is a network with vigilance parameter equal to 0.95 and training rate of 0.93.

6.4 Test result of Bayesian Network

In Bayesian network such as FAM neural network, EMTP software is used for simulations of the transformer. The obtained current waveforms used for input data for the Bayesian network in MATLAB. Some result of the Bayesian network test result is shown in Tab. 4.

As it is shown in Tab. 2, a wide range of outputs are similar to the targets and if the output of the Bayesian network round, the best PD location accuracy is reached. As the result of the simulation, Bayesian network has 100 % accuracy for determination of partial discharge location as FAM neural network.

Actually measured currents have measurement noise component and is not identical as the simulated values, hence white noise is added to simulated current waveforms. Amplitude of the white noise considered form 1 to 10 percent of peak value of simulated current amplitude in 1 percent steps. Accuracy of this network is 91 percent which is better than FAM network. The excellence of Bayesian network respect to FAM neural network is the better accuracy to finding the PD location in power transformer with presence of measurement noise.

7 CONCLUSION

In this paper Fuzzy ARTmap neural network and Bayesian network are proposed for locating of PD in power transformers winding. DM of transformer is used for PD simulation in EMTP and three-capacitor PD model is considered between coil to coil of power transformer winding. A study on DM parameters selection is presented in this paper and the results show highly dependency of results and prediction of PD location to these parameters. It has been shown that Fuzzy ARTmap neural network response to the signals with considering noise is acceptable and the results have good accuracy. But Bayesian network have better accuracy in noisy environment. These methods can be used as a general solution for locating of partial discharge in power transformers, provided the parameters value of detailed model of transformers to be selected truly.

REFERENCES

[1] NADERI, MOHAMMAD S.—VAKILIAN, M.—BLACKBURN, T. R.—PHUNG, B. T.—NADERI, MEHDI S.—NASIRI, A.: A Hybrid Transformer Model for Determination of Partial Discharge Location in Transformer Winding, IEEE

- Transactions on Dielectrics and Electrical Insulation **14**, No. 2 (Apr 2007).
- [2] MOORE, P. J.—PORTUGUES, I. E.—GLOVER, I. A.: Partial Discharge Investigation of a Power Transformer using Wireless Wideband Radio-Frequency Measurements, IEEE Transaction on Power Delivery **21** No. 1 (Jan 2006).
 - [3] WANG, Z. D.—CROSSLEY, P. A.—CORNICK, K. J.—ZHU, D. H.: An Algorithm for Partial Discharge Location in distribution Power Transformers, Power Engineering Society Winter Meeting, vol. 3, 23–27 Jan 2000, pp. 2217–2222.
 - [4] NAFISI, H.—DAVARI, M.—GHAREHPETIAN, G. B.—ABEDI, M.: Using Fuzzy ARTmap Neural Network for Determination of Partial Discharge Location in Power Transformers, in Proc. of IEEE, Power Tech, 2009.
 - [5] NAFAR, M.—ABEDI, M.—GHAREHPETIAN, G. B.—TAGHIPOUR, S.—YOUSEFPOUR, B.: Locating Partial Discharge in Transformer by Wavelet, WSEAS Transactions on circuits and systems **3** No. 6 (Aug 2004), 1499–1503.
 - [6] AKBARI, A.—WERLE, P.—BORSI, H.—GOCKENBACH, E.: Transfer Function-based Partial Discharge Localization in Power Transformers: a Feasibility Study, IEEE Electrical Insulation Magazine **18** No. 5 (Sep/Oct 2002), 22–32.
 - [7] SMITH, K. N.—PEREZ, R. A.: Locating Partial Discharges in a Power Generating System using Neural Networks and Wavelets, Annual Report Conference on Electrical Insulation and Dielectric Phenomena, 2002, pp. 458–461.
 - [8] DAVARI, M. ALE-EMRAN, S. M.—MOBARHANI, A. R.—NAFISI, H.—SALABEIGI, I.—GHAREHPETIAN, G. B.: A Novel Approach to VFTO Analysis of Power Transformers Including FVL based on Detailed Model, In Proceeding of IEEE International Conference on Industrial Technology (ICIT), Australia, February, 10–13, 2009.
 - [9] CARPENTER, G. A.—GROSSBERG, S.: ART₂: Self-Organization of Stable Category Recognition Codes for Analog Input Patterns, Applied Optics **26**(23) (1987), 4919–4930.
 - [10] CARPENTER, G. A.—GROSSBERG, S.: ART₃: Hierarchical Search using Chemical Transmitters in Self-Organizing Pattern Recognition Architectures, Neural Networks (Publication) **3** (1990), 129–152.
 - [11] CARPENTER, G. A.—GROSSBERG, S.—REYNOLDS, J. H.: ARTMAP: Supervised Real-Time Learning and Classification of Nonstationary Data by a Self-Organizing Neural Network, Neural Networks (Publication) **4** (1991), 565–588.
 - [12] CARPENTER, G. A.—GROSSBERG, S.—ROSEN, D. B.: Fuzzy ART: Fast Stable Learning and Categorization of Analog Patterns by an Adaptive Resonance System, Neural Networks (Publication) **4** (1991), 759–771.
 - [13] CARPENTER, G. A.—GROSSBERG, S.—MARKUZON, N.—REYNOLDS, J. H.—ROSEN, D. B.: Fuzzy ARTMAP: A Neural Network Architecture for Incremental Supervised Learning of Analog Multidimensional Maps, IEEE Transactions on Neural Networks **3** (1992), 698–713.
 - [14] HAMED NAFISI—ABDI, B.—AGHAKHANI, A.: Comparison of Bayesian and Fuzzy ARTmap Networks in HV Transmission Lines Fault Diagnosis, In Proceeding of WSEAS International Conference on Selected Topics in Mathematical Methods and Computational Techniques in Electrical Engineering (MMACTEE '10), Politehnica University of Timisoara, Romania, Oct 21–23, 2010.
 - [15] CHEN-FU CHIEN—SHI-LIN CHEN—YIH-SHIN LIN: Using Bayesian Network for Fault Location on Distribution Feeder, IEEE Transactions on Power Delivery **17** No. 13 (July 2002).
 - [16] SZOLOVITS, P.—PAUKER, S.: Categorical and Probabilistic Reasoning in Medical Diagnosis, Artif. Intell. **1** (1978), 115–144.
 - [17] SOVARONG, L.—COSTAS, J. S.: A General Equipment Diagnostic System and its Application on Photolithographic Sequences jour IEEE Trans. Semi-conductor Manufact..
 - [18] DUDA, R.—GASCHNING, J.—HART, P.: Model Design in the Prospector Consultant System for Mineral Exploration, in Expert Systems in the Microelectronic Age (D. Michie, ed.), Edinburgh Univ. Press, Edinburgh, U.K., 1979, pp. 153–167.
 - [19] BUNTINE, W.: A Guide to the Literature on Learning Probabilistic Networks from Data, IEEE Trans. Knowl. Data Eng. **8** No. 2 (1996), 195–210.
 - [20] WEI ZHAO—XIAOMIN BAI—JIAN DING—ZHU FANG—ZAIHUA LI—ZIGUAN ZHOU: A New Uncertain Fault Diagnosis Approach of Power System based on Markov Chain Monte Carlo Method, IEEE International Conference on Power System Technology, 2006, 1-4244-0111.
 - [21] DENISON—HOLMES—MALLICK—SMITH: Bayesian Methods for Nonlinear Classification and Regression, Wiley, 2002.

Received 18 March 2012

Hamed Nafisi was born in Tehran in the Iran, on April 31, 1984. He received his BSc and MSc in electrical engineering in 2006 and 2008 from the Amirkabir University of Technology (AUT), Tehran, Iran. Now, he is the Ph.D. student in the Amirkabir University of Technology (AUT), Tehran, Iran. His research interests include Smart Grid, Power System Protection, Power Electronic, Using Artificial Intelligence in Power System and Distributed Generation (DG).

Mehrdad Abedi (Prof) was born in Tehran, Iran in 1948. He received his BS degree in electrical engineering from Tehran University in 1970. He also received his MS and PhD in electrical engineering from the University of London and Newcastle University, respectively, in 1973 and 1977. In 1978 he joined Amikabir University of Technology where he is now full professor of the Power System Center of Excellency. Professor Abedi is the author of 14 books in the field of electrical machines and power systems. He also has published 140 national and international conference and journal papers.

Gevorg B. Gharehpetian (Prof) was born in Tehran, in 1962. He received his BS and MS degrees in electrical engineering in 1987 and 1989 from Tabriz University, Tabriz, Iran and Amirkabir University of Technology (AUT), Tehran, Iran, respectively. In 1989 he joined the Electrical Engineering Department of AUT as a lecturer. He received the Ph.D. degree in electrical engineering from Tehran University, Tehran, Iran, in 1996. As a PhD student he has received scholarship from DAAD (German Academic Exchange Service) from 1993 to 1996 and he was with High Voltage Institute of RWTH Aachen, Aachen, Germany. He held the position of Assistant Professor in AUT from 1997 to 2003, and has been Associate Professor since 2004. Now he is professor in AUT. Since 2004 he is the Editor-in-Chief of the Journal of Iranian Association of Electrical and Electronics Engineers (IAEEE). The power engineering group of AUT has been selected as a Center of Excellence on Power Systems in Iran since 2001. He is a member of this center. He is the author of more than 300 journal and conference papers. His teaching and research interest include power system and transformers transients, FACTS devices and HVDC transmission.