Journal of ELECTRICAL ENGINEERING, VOL. 57, NO. 3, 2006, 138–145

PNN AND ITS ADAPTIVE VERSION — AN INGENIOUS APPROACH TO PD PATTERN CLASSIFICATION COMPARED WITH BPA NETWORK

Balasundaram Karthikeyan [*] — **Srinivasan Gopal** [**] — **Srinivasan Venkatesh** [***] — **Subramanian Saravanan** [****]

The reliability of insulation systems is a major requirement of any power apparatus. The incidence of minor flaws and irregularities such as voids, surface imperfections *etc*, in insulation systems is however inevitable and leads to partial discharges (PD). Classification of PD patterns plays an important role during manufacturing and on-site assessment of power apparatus. The innovative trend of using artificial neural network towards classification of PD patterns is perceptible. A novel method for the classification of PD patterns using the original probabilistic neural network (PNN) and its variation has been proposed and implemented in this work. The classification of single-type insulation defects such as voids, surface discharges and corona has been considered primarily. The efficacy and merits of PNN and its adaptive version over that of the back propagation algorithm based feed forward neural network has been established through exhaustive comparisons on the performance of the neural networks in PD pattern classification task.

K e y w o r d s: partial discharge, probabilistic neural network, smoothing parameter, Bayesian strategy, back propagation algorithm, adaptive probabilistic neural network

1 INTRODUCTION

Used abbreviations: PD — partial discharge, ANN — artificial neural network, PNN — probabilistic neural network, APNN — adaptive probabilistic neural network, pdf — probability density function, BPA NN — back propagation algorithm neural network

Towards assessing the quality of the insulation system of power apparatus both during manufacture and service, tests to determine the soundness of the insulating/dielectric material are conducted which include overstressing the insulation with *ac* and/or *dc* or surge voltages. The disadvantage of these techniques is that during the process of testing the equipment may get damaged if the insulation is faulty. Partial discharge (PD) phenomena obviate this necessity of overstressing the insulation since the inherent property of the theory can be exploited to serve as a non destructive testing (NDT) technique.

The incidence of minor flaws such as voids, surface imperfections etc are inevitable in electrical insulation system of any power apparatus, leading to partial discharges. Every partial discharge event causes deterioration of the insulation material by the energy impact of high-energy electrons or accelerated ions. Since each defect has a particular deterioration mechanism, it is imperative to discern the correlation between the discharge patterns and the nature of defect in order to ascertain the quality of the insulation. ANN, which is a non-parametric method, augurs well for the PD pattern classification task since PD phenomenon is inherently a stochastic and a non-Markovian process [1, 2] in which there can be significant statistical variability. This paper suggests the approach of the novel probabilistic neural network for the classification of PD pattern of various types of defects. In order to evaluate and ascertain the efficacy of the classification of PD patterns using this novel approach, the BPA based feed forward neural network has been utilized. An exhaustive study of both the unsupervised (PNN versions) and supervised (BPANN) version of the NNs has been carried out and the merits of the proposed NN are summarized.

2 PARTIAL DISCHARGE PHENOMENA

Partial discharge is an electrical breakdown confined to the localized regions of the insulating medium of a power apparatus. PD phenomenon are inherently self quenching and stochastic. They are characterized by pulsating currents, which have a very short time period varying between a few nano seconds [1] up to a few microseconds. PD may be classified based on the site of occurrence. Voids, inclusions, occlusions occurring in the insulation material are classified as internal discharges while those accruing at the interface of the insulation (external to the surface of the material) are classified as surface discharges

[*] School of EEE, SASTRA, Deemed University, Thanjavur 613 402, India, E-Mail: b_carthik@rediffmail.com
[**] W. S. Test Systems Ltd., 27$^{\text{th}}$ kM Bellary Road, Bangalore 562 157, India, E-Mail: mwbil@bgl.vsnl.net.in
[***] Department of EEE, SRMIST, Deemed University, Kanchipuram Dist., India, E-Mail: venkatsri_73in@yahoo.co.in
[****] School of EEE, SASTRA, Deemed University, Thanjavur 613 402, India, E-Mail: saravanan.me@gmail.com

ISSN 1335-3632

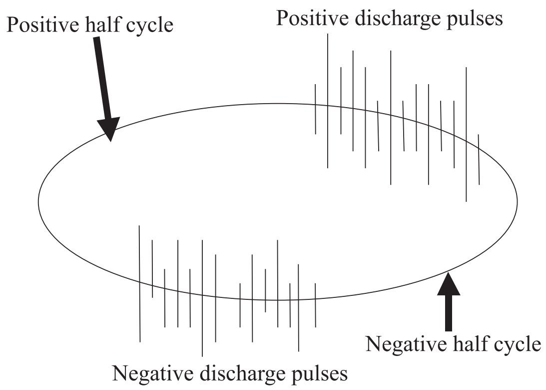

Fig. 1. PD Pulse representation on elliptical time base

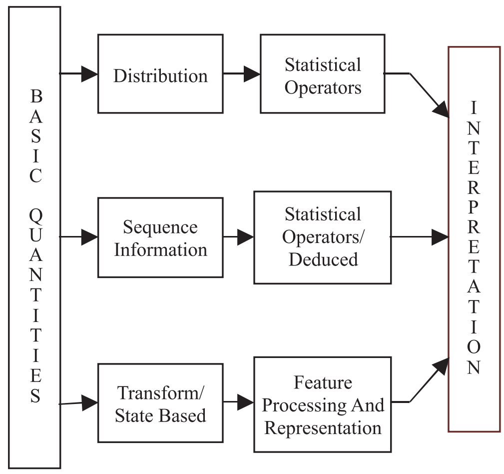

Fig. 2. Techniques used in PD analysis

and those occurring at locations of sharp edges are classified as corona discharges. The display of PD pulses in elliptical time base as indicated in Fig. 1 is usually preferred to the sinusoidal time base system for representation of PD pulses for better visualization and recognition. Figure 2 indicates the various techniques [3] attempted by several researchers to diagnose PD in order to detect and ascertain the flaws/ faults in the insulation system.

The distribution analysis [4] approach follows probabilistic models and is popularly used for ageing characterization. The pulse sequence analysis (PSA) and height distribution analysis [5] characterizes the temporal behavior of the data. This approach is useful for understanding the physical process involved in the system. Further, the characterization based on time information has distinct advantages (Pulse sequence analysis can lead to instrument independent characterization). The transform or state based techniques such as, artificial neural network [6, 7], fuzzy logic controllers (FLC), hidden Markov models [8, 9], fractal features [10], fast Fourier transform *etc*are mainly used for feature processing and classification of PD patterns.

3 PATTERN RECOGNITION AND CLASSIFICATION USING NEURAL NETWORKS

The complexity of analyzing PD patterns obtained from digital computer acquisition system is evident since PD phenomenon is a highly stochastic process as described in the previous section and hence a complex nonlinear problem. The process being stochastic, the associated effects of memory propagation with the influence of residues [11] from previous PD pulses *etc* (non-Markovian process) have made the classification of such PD patterns in terms of φ-q-n even more complex.

Pattern recognition task basically involves identification of similar data within a collection, which resembles the new input. Since artificial neural network (ANN) has the ability to learn from examples [6], generalize well from training, handle noisy data conveniently, create their own relationship amongst information, it has become an innovative technique suitable for PD pattern recognition and classification.

The major variable parameters [12, 13] which characterize the PD pulses and which define the basis of the physical phenomena of PD are the time of occurrence (φ), magnitude of discharge (q) and the number of discharges (n) which are represented as a three dimensional pattern.

Researchers have used several types of ANNs [6, 14–20] to date for the classification of PD patterns. To name a few, these include the traditional BPA neural network, adaptive resonance theory (ART) neural network, self organizing mapping (SOM) neural network and earning vector quantization (LVQ) neural network. A novel approach of using the probabilistic neural network (PNN) and its adaptive version has been proposed and implemented for classifying PD patterns due to the inherent strengths of PNN such as its simplicity, robustness and good generalization ability as compared to other neural networks.

4 PROBABILISTIC NEURAL NETWORK AND ITS ADAPTIVE VERSION

PNN [21, 22] is a network formulation of 'probability density estimation'. It is a model based on competitive learning with a 'winner takes all attitude' and the core concept based on multivariate probability estimation. The original and the adaptive versions of PNN do not have feedback paths. The development of PNN relies on the Parzen window concept of multivariate probabilities. The PNN is a classifier version, which combines the Baye's strategy for decision-making with a nonparametric estimator for obtaining the probability density function (pdf).

The PNN network as described in Fig. 3 consists of an input layer, two hidden layers (one each for exemplar/pattern and a class/summation layers) and an output layer.

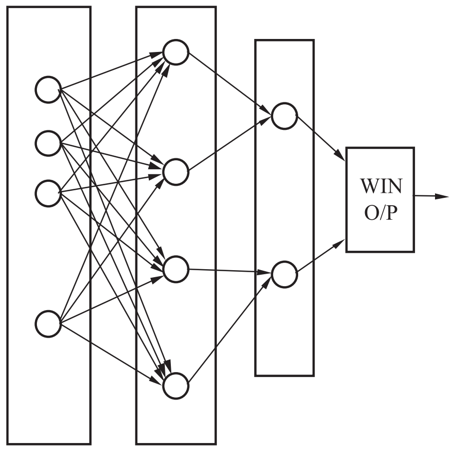

Fig. 3. Architecture of PNN

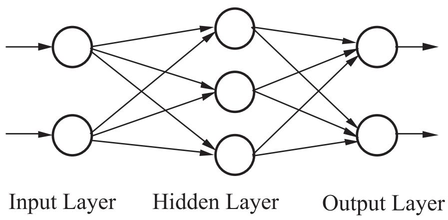

Fig. 4. The multi layer feed forward NN architecture

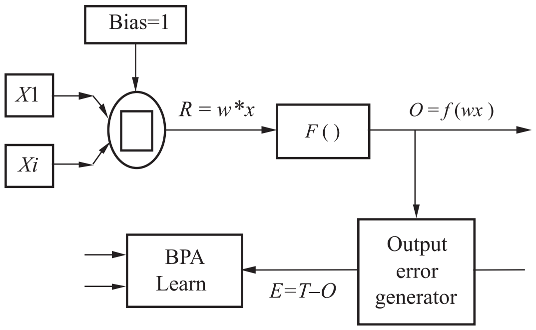

Fig. 5. Representation of back propagation algorithm NN

The PNN classifier has sometimes been accepted as belonging to the class of radial basis function (RBF). Another school of thought [23] prefers to associate RBF classifiers topologically with feed forward network having only one hidden layer as illustrated in Fig. 4. The process-based classification that differentiates PNN from RBF is that PNN works on estimation of probability density function (pdf) while RBF works on iterative function approximation.

The distinguishing feature of this NN as compared to the BPA NN [23, 24, 25] is that the computational load in the training phase is transferred to the evaluation phase. However, the speed of computation is noticeably more than the BPA feed forward NN. Another distinction between BPA neural network and PNN is that the BPA NN training phase is in terms of global basis function which is defined by the non-linear (usually sigmoidal) functions of the distance pattern vectors from the hyper-plane as illustrated in Fig. 5 while the training in PNN is in terms of local basis functions usually the exponential functions. The strength of using local basis function stems from the fact that it is possible to train a network of local basis functions in one pass of the data, by straightforwardly applying the principles of statistics.

In order to classify a feature pattern vector $\boldsymbol{X} \in R^M$, that is to assign the pattern to one among K predefined classes, the conditional probability $p\left(x|C_k\right)$ of each class C_k is estimated since it represents the uncertainty associated to class attribution. Then these estimates are combined by the rule of Bayes to yield a-posteriori class probabilities $p\left(C_k|x\right)$ that allow in making optimal decisions.

There is one pattern or exemplar node for each training example. Each pattern node forms a product of the weight vector and the given example for classification, where the weights entering a node are from a particular example. After that the product passes through the activation function $\exp\left[(x^\top w_{ki} - 1)/\sigma^2\right]$. The second hidden layer contains one summation unit for each class. Each summation unit (node) receives the output from the pattern nodes associated with a given class

$$\sum_{i=1}^{N_k} \exp\left[x^\top w_{ki} - 1)/\sigma^2\right].$$

The output layer has as many neurons as the number of data classes considered. The output nodes are binary neurons that produce the classification decision based on

$$\sum_{i=1}^{N_k} \exp\left[(x^\top w_{ki} - 1)/\sigma^2\right] > \sigma_{i=1}^{N_j} \exp\left[(x^\top w_{kj} - 1)/\sigma^2\right].$$

The network as shown in Fig. 6 indicates the pattern unit, which requires normalization of the input and exemplar vectors to unit length. This is achieved by calculating the Euclidean norm.

A variation of the PNN called the Adaptive PNN involves a change in the free parameter σ (the variance parameter) or the smoothing parameter. The original PNN involves a single value for all the classes while the adaptive PNN involves different values of σ for each class based on the calculation of the Euclidean distance and then the average distance as indicated in Fig. 7.

The other feature being used in this approach is that a simplified formula of pdf is used which obviates the necessity for normalization and hence a considerable amount of computation is reduced

5 PD PATTERN RECOGNITION AND CLASSIFICATION

Before undertaking the task of classification of PD, both of the networks (PNN and APNN) are trained to

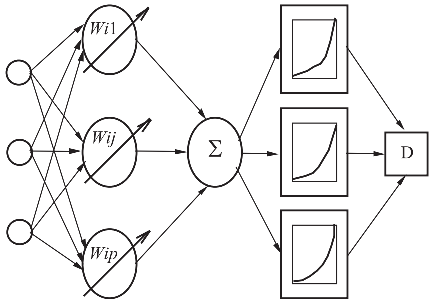

Fig. 6. Normalization of pattern unit

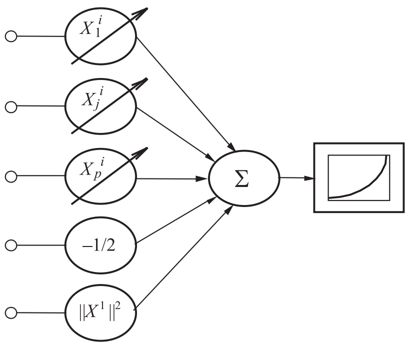

Fig. 7. Pattern unit without normalization

recognize and classify the alphabet characters (in this case the uppercase alphabets have been provided as inputs in a 7×5 matrix of pixels). Also, the classical Fisher's database is used for classification for verification.

As indicated in previous sections, the PD patterns composed of three parameters φ-q-n have been considered for the classification task. φ-q pattern indicates the mean PD magnitude of the pulses in each phase window and φ-n pattern indicates the pulse count in each phase window. The distribution of the φ-q-n parameters for a typical PD defect pattern, familiarly referred to as 'fingerprint', is shown in Fig. 8.

The φ-q-n characteristics as a distribution, which consists of phase angle, the apparent charge q indicating the pulse height and the number of pulses n is provided to the ANN black box as the input. For extensive testing and effective verification of training of the NN, the inputs to the ANN are provided using an ingenious preprocessing technique, which ensures the compactness of the characteristic input feature vector. The inputs characterizing the data are provided based on the phase window concept. The inputs based on the measures of statistical operators are: a) measures based on maximum/ minimum values of a specific parameter, b) measures of dispersion, c) measures of central tendency and d) a set of mathematical descriptors. The preprocessing ensures the compactness thus reducing the number of inputs to the ANN. The network is trained to classify the patterns of

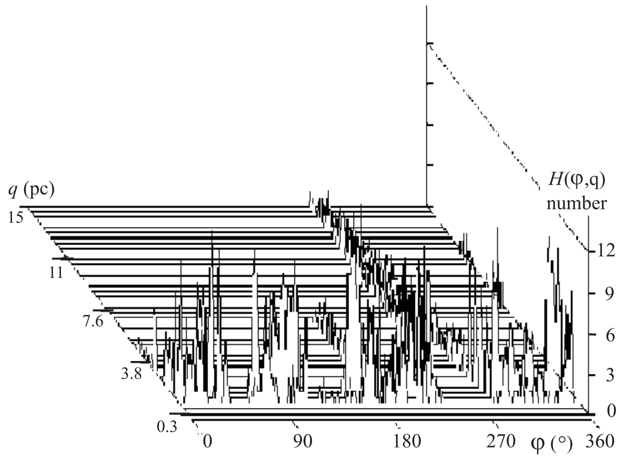

Fig. 8. Phase resolved distribution describing φ-q-n relationship (typical)

defects from various origins such as void, corona, surface discharges and oil corona.

6 PREPROCESSED INPUTS TO PNN, APNN AND BPA FFNN

A total of 20 training patterns for each type of input methodology with 5 each for every type of defect namely, void, surface discharge, oil corona and corona has been taken up for implementation and analysis of the proposed PNNs and the BPA FNN as indicated in Table 1 and 2. A total of 180 test patterns have been used in the classification task. The basic objective of using different methodologies of preprocessing [26] inputs is to ascertain the performance of the ANN (classification/misclassification) so that tangible decisions may be taken with regard to the role played by the dimensionality of the input vectors presented to the ANN.

The strength of PNN lies in providing good generalization capability notwithstanding the fact that the only tunable parameter is the smoothing parameter or the variance parameter. Further, obtaining the optimal training value of the smoothing parameter can be made relatively easily through trial and error as compared to the time consuming iterative procedures as adopted by a few other supervised ANN paradigms.

7 OBSERVATIONS, COMPARISONS AND RESULTS

The following exhaustive observations have been made and summarized:

1. It is observed in the case of the original PNN that the variation of the smoothing parameter from a value of 1 to 0.218* (*indicates that the optimum value of σ for which the number of misclassification is low) gave good results with only one misclassification. However, this was a case wherein the value of the smoothing parameter is changed on a case-to-case basis (this was checked for only one type of input sequence). However, a common value of the smoothing parameter for all the input

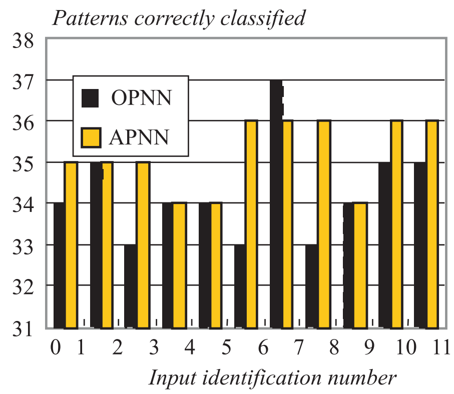

Fig. 9. Comparison of OPNN and APNN

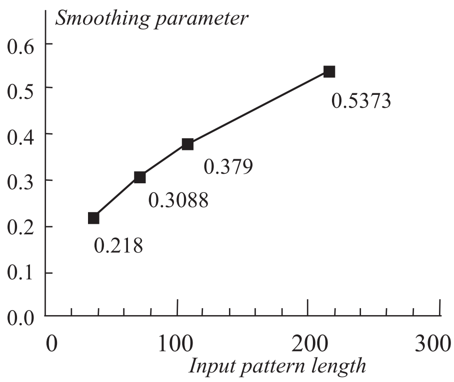

Fig. 10. Influence of value of smoothing parameter on the dimensionality of input

Table 1. Tabulation of observations of PD pattern classification using original PNN (v, s, oc, c denotes void, surface discharge, oil corona and corona defects respectively).

Input Type	Degree Width	σ	Misclassifications
$q_{\max}$-φ-n	30°	1, 0.218*	4 No & 1No* (c-10; s-7) & s-7
$q_{\max}$-φ-n	10°	1	3 No (s-6, 7)
$q_{\min}$-φ-n	30°	1	5 No (s-6,7; c-10)
$q_{\min}$-φ-n	10°	1	4 No (s-7; c-10)
$n_{\max}$-φ-q	10°	1	4 No (s-7; c-10)
$q_{\max}$-φ-n $q_{\min}$-φ-n	30°	1	5 No (s-7; c-10)
$\varphi_{\max}$-q-n/ $\varphi_{\min}$-q-n	30°	1	5 No (s-7; c-10)
$q_{\max}$-φ-n $q_{\min}$-φ-n	10°	1	1 No (s-7)
$q_{\max}$-$n_{\max}$-φ	30°	1	4 No (s-7; v-6)
Central Tendency	30°	0.3097	3 No (s- 8; c-8,9)
Dispersion	30°	0.3097	3 No (c-5,8,9)

type combinations is not possible due to the inherent structure of the original PNN. These observations are depicted in Fig. 10. and Tab. 1.

2. It is also observed and demonstrated that for the adaptive PNN with a constant value of $g = 0.05$ for all the input types considered, a relatively lesser misclassifications of about 3 with no misclassifications of the training patterns is obtained. This illustrates the capability of the use of a variable smoothing parameter in the case of the adaptive PNN as shown in Tab. 2.

Table 2. Tabulation of observations of PD pattern classification using adaptive PNN (v, s, oc, c denotes void, surface discharge, oil corona and corona defects respectively).

Type of Input	Degree Width	g	Misclassifications
$q_{\max}$-φ-n	30°	0.05	3 No (s-8,9; v-6)
$q_{\max}$-φ-n	10°	0.05	3 No (s-5,9,10)
$q_{\min}$-φ-n	30°	0.05	3 No (s-2,9,10)
$q_{\min}$-φ-n	10°	0.05	4 No (oc-3,8; s-9,10)
$n_{\max}$-φ-q	10°	0.05	4 No (oc-3,6,8; s-5)
$q_{\max}$-φ-n, $q_{\min}$-φ-n	30°	0.05	2 No (s-9, 10)
$\varphi_{\max}$-q-n/ $\varphi_{\min}$-q-n	30°	0.05	2 No (s-9, 10)
$q_{\max}$-φ-n, $q_{\min}$-φ-n	10°	0.05	2 No (s-9,10)
$q_{\max}$-$n_{\max}$-φ	30°	0.05	4 No (s-9,10; oc-3,10)
Central Tendency	30°	0.05	2 No (s-9,10; oc-3,10)
Dispersion	30°	0.05	2 No (s-9,10; oc-3,10)

3. PD pattern classification inherently involves classification of patterns that are not always same and identical. It is imperative that the ANN is able to provide a reasonable generalization of such patterns characterizing the flaw. Hence, an adequate number of training exemplars are required to describe the function. This is also made evident from Tab. 1 and 2 as this aspect is related to the dimensionality of the input vector.
4. A comparison of PNN and APNN paradigms indicating the number of classifications gives a clear indication on the advantage of using the APNN as illustrated in Tab. 2 and Fig. 9.
5. If the number of training exemplars used is not sufficient to pin down the free parameters in the network to capture the regularity in the data, the best the network can do is to assign some random component to some parameters. Thus, it is meaningless to evaluate the capability of generalization using exemplars that reflect variations not captured in the training exemplars.
6. The role of the smoothing parameter [21, 22] and the basis of the choice of the value of this parameter as a function of the dimension of the problem and the number of training patterns is markedly noticeable as shown in Fig. 10.

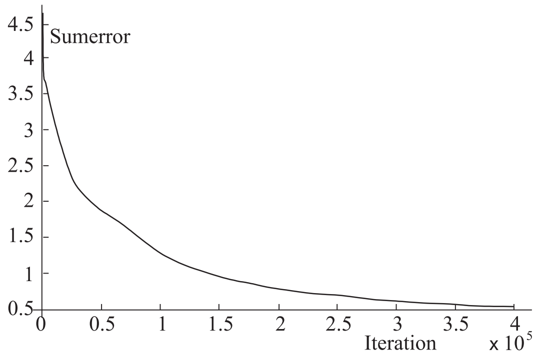

Fig. 11. Typical plot of error versus iteration — φ-$q_{\min}$-n

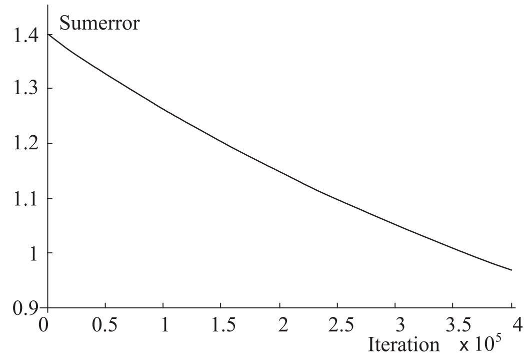

Fig. 12. Typical plot of error versus iteration — $n_{\max}$-φ-q

Table 3. Tabulation of observations of PD pattern classification using back propagation algorithm ANN (v, s, oc, c denotes void, surface discharge, oil corona and corona defects respectively).

Input Type	Width	Total Patterns Used	Misclassifications
$q_{\max}$-φ-n	30°	20	4 No (s-9,10; oc-2,3)
$q_{\max}$-φ-n	10°	20	5 No (v-1,2,3; s-2.4; c-4)
$q_{\min}$-φ-n	30°	20	8 No (s-2,4,6,9,10; oc-1,3.,6)
$q_{\min}$-φ-n	10°	20	7 No (s-2,6,9,10; oc-1,3,6)
$n_{\max}$-φ-q	10°	20	5 No (s-9; oc-3,8,10; c-8)
$q_{\max}$-φ-n, $q_{\min}$-φ-n	30°	20	5 No (v-1; s-3; oc-3,8; c-2)
$\varphi_{\max}$-q-n/ $\varphi_{\max}$-q-n	30°	20	5 No (v-1; s-3; oc-3,8; c-2)
$q_{\max}$-φ-n, $q_{\min}$-φ-n	10°	20	5 No (v-1; s-3; oc-3,8; -2)
$q_{\max}$-$n_{\max}$-φ	30°	20	4 No (s-9,10; oc-2,3)
Central Tendency Measure	30°	20	3 No (s-9; oc-3,8)
Dispersion Measure	30°	20	3 No (s-9; oc-3,8)

7. In the actual case of training and testing it has been observed that it is not difficult to find a good value of the smoothing parameter and that the misclassification rate does not change dramatically with small changes in the value of the smoothing parameter.
8. A quite obvious yet an important observation indicated the decrease in the value of the smoothing parameter in the formation of the required decision surface, while at higher values of the smoothing parameter over the actual responsive range showed insignificant changes in the classification of the input of the network.

The above listed observations have been carried out in order to verify and correlate the general observations drawn by Dr. Specht [21, 22], the founder of PNN. These observations indicated the veracity of the statements of Dr. Specht.

The same set of inputs have been used for training the BPA feed forward neural network and observation and comparison of the performance of BPA network is summarized as follows:

1. The learning rate plays an important role in the convergence of the error. Greater the learning rate lesser is the time for training though the error converges beyond the minimum error criteria. This aspect is indicated in Fig. 11.
2. The nature of the input plays an important role with regard to the training time and the number of iterations required for convergence. It is observed in the case of BPA algorithm that for lesser number of iterations during training, the number of misclassifications is high and can be improved by increasing the number of iterations [27]. This feature is indicated in Fig. 12.
3. The limitation observed in the case of the BPA NN is that the time involved for training is extremely long (it is in fact prohibitively long to the extent that the training took about 36 hours initially!). This feature indicates the limitation of the BPA NN and its inability to be used in on-line PD diagnosis.
4. The gradient descent method of supervised learning and the mechanism of weight adjustment by a magnitude proportional to the first derivative of the error is iteratively calculated to ascertain the appropriate class pertaining to the hyper plane. The objective in this process is to decrease the error function by avoiding the local minima and reaching the actual or global minimum. It is to be noted that other variations of BPA NN with different learning rules may possibly yield better performance. However, it is obvious that the non-iterative PNN versions perform the classification task swiftly yet generalize better than the BPA NN.
5. The comparison of the performance of BPA NN with the PNN versions clearly indicates the superiority of the original version of PNN and in particular its adaptive version (APNN) in the PD pattern classification task. Table 3 illustrates the aforementioned deduction.

8 INPUT VALIDATION AND INFERENCES ON THE SIGNIFICANCE OF THE SMOOTHING PARAMETER

It is mandatory that appropriate validation is performed to ascertain the veracity of the inputs presented to the ANNs prior to carrying out training of the neural networks. In this work this concept has been well taken care of using the two most established validation techniques namely the partial set training method and the hold-one-out method as suggested by Donald F. Specht [21, 22].

The following conclusions have been arrived at based on exhaustive training and testing of the PNN and APNN paradigms. They are:

1. The shape of the decision surface can be made as complex as necessary or as simple as desired by choosing an appropriate value of the smoothing parameter. However, this value affects the number of misclassifications.
2. The decision surface can approach optimal minimum risk decision surfaces.
3. Erroneous and sparse samples are tolerated.
4. Smoothing parameter can be made smaller as the number of patterns gets larger without retraining, so long as the inputs are properly validated.

9 CONCLUSIONS AND DISCUSSION

1. It is observed that both PNN and its adaptive version are effective in classifying PD patterns.
2. The impact of the value of the smoothing parameter in the case of the adaptive version and hence its superiority is also perceptibly noticeable. For highly sparse yet dense clusters of pulses of PD, the PNN version of the Neural Network gives good generalization for classifying PD patterns.
3. Use of algorithms for proper clustering of the inputs (in case of unknown input training patterns) such as the EM (expected maximization) and the much superior jack knife method may be used for solving the problems of misclassifications.
4. PNN that has been used employs the Gaussian kernel function. Alternative kernel functions such as the rotated kernel function (RKPNN) may be used in the classification of PD patterns.
5. Another robust methodology includes the use of a heteroscedastic PNN [29, 30] as a possible future trend. Validation using the training set methodology and the holdout approach has been carried out. It is however noted that further research and validation is required.

Acknowledgement

Authors are grateful to Dr P. S. Srinivasan, Dean/SEEE, SASTRA, Deemed University for periodic discussions on the topic. The authors are also thankful to all staff members for having taken part in discussion regarding the subject.

References

[1] Van BRUNT, R. J.: Stochastic Properties of Partial Discharge Phenomenon, IEEE Transaction on Electrical Insulation **26** No. 5 (October 1991), 902–948.

[2] Van BRUNT, R. J.—CERNYAR, E. W.—Von GLAHN, P.: Importance of Unraveling Memory Propagation Effects in Interpreting Data on Partial Discharge Statistics, IEEE Transaction on Electrical Insulation **28** No. 6 (December 1993), 905–916.

[3] SENTHIL KUMAR, S.—NARAYANACHAR, M. N.—NEMA, R. S.: Pulse Sequence Studies on PD Data, 11th International Symposium on High Voltage Engineering, U.K., 5.25.S1–5.25.S7, 1999.

[4] SENTHIL KUMAR, S.—NARAYANACHAR, M. N.—NEMA, R. S.: Response of Narrow Band Detector and Analyzer in Ageing Experiments, 11th International Symposium on High Voltage Engineering, U.K., 5.25.S8–5.25.S8, 1999.

[5] SENTHIL KUMAR, S.—NARAYANACHAR, M. N.—NEMA, R. S. paper Partial Discharge Pulse Sequence Analysis — A New Representation of Partial Discharge Data: 11th International Symposium on High Voltage Engineering, U.K., 5.25.S8–5.25.S8, 1999.

[6] GULSKI, E.—KRIVDA, A.: Neural Network as a Tool for Recognition of Partial Discharges,, IEEE Transaction on Electrical Insulation **28** No. 6 (December 1993), 984–1001.

[7] SATISH, L.—ZAENGL, W. S.: ANN for Recognition of 3-D PD Patterns, IEEE Transaction on Dielectrics and Electrical Insulation **1** (April 1994), 265–275.

[8] SATISH, L.—GURURAJ, B. I.: Use of Hidden Markov Models for Partial discharges Pattern classification, IEEE Transaction on Electrical Insulation **28** (1993), 178–182.

[9] ABDEL-GALIL, T. K. *et al*: Partial Discharge Pulse Pattern Recognition using Hidden Markov Models, IEEE Transaction on Dielectrics and Electrical Insulation **11** No. 4 (August 2004), 715-723.

[10] SATISH, L.—ZAENGL, W. S.: Can Fractal Features be Used for Recognizing 3-D Partial Discharge Patterns, IEEE Transaction on Dielectrics and Electrical Insulation **2** No. 3 (April 1995), 352–358.

[11] Van BRUNT, R. J.: Stochastic Properties of Partial Discharge Phenomenon, IEEE Transaction on Electrical Insulation **26** No. 5 (October 1991), 902–948.

[12] GULSKI, E.—KREUGER, F. H.: Computer Aided Recognition of Discharge Source, IEEE Transaction on Electrical Insulation **27** No. 1 (February 1992), 82–92.

[13] KREUGER, F. H.—GULSKI, E.—KRIVDA, A.: Classification of Partial Discharge, IEEE transaction on electrical insulation **28** (December 1993), 917–931.

[14] AMIRA MAZROUA, A.—SALAMA, M. M. A.: PD Pattern Recognition with Neural Network Using the Multi Layer Perceptron Technique, IEEE Transaction on Electrical Insulation **28** No. 6 (December 1993), 1082–1087.

[15] AMIRA MAZROUA, A.—SALAMA, M. M. A.: Neural Network System Using the Multi Layer Perceptron Technique for the Recognition of PD Pulse Shapes due to Cavities and Electrical Trees, IEEE Transaction on Power Delivery **10** (1995), 92–96.

[16] KARTHIKEYAN, B.—GOPAL, S.: Automated Statistical Recognition of Partial Discharges in Insulations, 14th International Symposium on High Voltage Engineering, Beijing, China, G-021, 2005.

[17] KARTHIKEYAN, B.—GOPAL, S.: Artificial Neural Network – A Non parametric Method for Partial Discharge Pattern Classification, 14th International Symposium on High Voltage Engineering, Beijing, China, G-101, 2005.

[18] KAI GAO *et al* : PD Pattern Recognition for Stator bar Models with Six Kinds of Characteristic Vectors Using BP Network, IEEE Transaction on Dielectrics and Electrical Insulation **9** No. 3 (June 2002), 381–389.

[19] HOOF, M.*et al* : PD Source Identification with Novel Discharge Parameters Using Counter Propagation Neural Network, IEEE Transaction on Dielectrics and Electrical Insulation **4** No. 1 (1997), 17–32.

[20] HAO, L.—LEWIN, P. L.—TIAN, Y. : Partial Discharge Discrimination Using a Support Vector Machine, 14^{th} International Symposium on High Voltage Engineering, Beijing, China, G-115, 2005.

[21] SPECHT, D. F. : Probabilistic Neural Networks for Classification, Mapping or Associative Memory, IEEE Conference on Neural Networks, Vol.1, No.1, 525-532, 1998.

[22] SPECHT, D. F. : Probabilistic Neural Networks and the Polynomial Adaline as Complementary Techniques for Classification, IEEE Conference on Neural Networks, Vol. 1, No. 1, March 1990.

[23] CHENG, C. : Handbook on Fuzzy Logic and Neural Networks, Chapter 3, McGraw Hill Publishing Company, New York.

[24] SPECHT, D. F.—SHAPIRO, P. D. : Generalization Accuracy of Probabilistic Neural Networks Compared with Back Propagation Networks, International IEEE Conference on Neural Networks, Vol. 1, 887–892, Seattle, WA, July 1991.

[25] SPECHT, D. F.—ROMSDAHL, H. D. : Experience with Adaptive Probabilistic Neural Network and Adaptive General Regression Neural Network,, IEEE International Conference on Neural Networks, Vol. 2, 1203–1208, Orlando, FL, June 1994.

[26] CACHIN, C. *et al* : PD Recognition with Knowledge based Preprocessing Techniques in Neural Networks, IEEE Transaction on Dielectrics and Electrical Insulation **2** No. 4 (August 1995), 578–589.

[27] SKAPURA, M. S. : Building Neural Networks, Addison Welsey Publishing Company, 1996.

[28] ANDERSON, J. A. : An Introduction to Neural Network, 3^{rd} Edition, Prentice Hall, India, 2001.

[29] ZENG RONG YANG—SHENG CHEN : Robust Maximum Likelihood Training of Heteroscedastic Probabilistic Neural Networks, International Journal on Neural Networks, neural Networks **11** (1998), 739–747.

[30] ZENG RONG YANG—ZWOLINSKI, M.—CHALK, C. D.—WILLIAMS, A. C. : A Robust Heteroscedastic Probabilistic Neural Network to Analog Fault Detection and Classification, IEEE Transaction on Computer Aided Design of Integrated Circuits and Systems **19** No. 1 (January 2000).

Received 28 September 2005

Balasundaram Karthikeyan is a senior lecturer in the School of Electrical and Electronics Engineering, SASTRA, Deemed University, India. He obtained his bachelor degree in electrical and electronics engineering from Bharathidasan University, India in 1996 and master degree in high voltage engineering from Anna University, India in 2000. He is working for a PhD in the field of partial discharge. His specific area of interest includes impulse testing of power transformers.

Srinivasan Gopal is the executive director of WS Test Systems Limited, Bangalore, India. He obtained his MIE degree from Mysore University, India and his master degree from Mysore University, India in 1978. He received a PhD from Indian Institute of Technology (IIT), Chennai, India in 1984. He underwent advanced training in high voltage engineering in Germany for a period of one year. His specific area of research includes impulse testing of power transformers and partial discharge phenomena in power apparatus.

Srinivasan Venkatesh is a senior lecturer in the Department of Electrical and Electronics, SRMIST, Deemed University, India. He obtained a bachelor degree in electrical and electronics engineering from Annamalai University, India in 1994 and master degree in technology-high voltage engineering from SASTRA, Deemed University in 2004. He was with Alstom Limited, Systems Group-India, T & D Projects Division, as assistant manager of engineering & QA. His specific area of interest includes partial discharge, substation engineering and EHV transmission systems.

Subramanian Saravanan is a lecturer in the School of Electrical and Electronics Engineering, SASTRA, Deemed University, India. He obtained his bachelor degree in electrical and electronics engineering from Bharathiyar University, India in 2001 and master degree in power electronics and drives from PSG College of Technology, India in 2003. His specific area of interest includes impulse testing of power transformers, double tuned circuits and HVDC.