Journal of Applied Sciences

ISSN 1812-5654

Journal of Applied Sciences 14 (14): 1496-1505, 2014
ISSN 1812-5654 / DOI: 10.3923/jas.2014.1496.1505

Mixture Density Estimation Clustering Based Probabilistic Neural Network Variants for Multiple Source Partial Discharge Signature Analysis

[1]S. Venkatesh, [2]S. Jayalalitha and [3]S. Gopal
[1]Department of Electrical and Electronics Engineering,
[2]Department of Electronics and Instrumentation Engineering,
School of Electrical and Electronics Engineering, SASTRA University, Thanjavur, 613401, India
[3]W.S. Test Systems Pvt. Limited, 27th km Bellary Road, Doddajala Post, 562 157, Bengaluru, India

Abstract: A gamut of insulation diagnostic methods is being practiced. Amongst them Partial Discharge (PD) detection, measurement and analysis is an inherently non pervasive-test test procedure. Hence, it is being considered as a crucial methodology. Over the last three decades attempts were made to discriminate single and partially overlapped PD sources have yielded moderate success. In the above process techniques like Fractal Features, Mixed Weibull Function, Neural Networks (NN) and Wavelet Transformation have been implemented. However, intricacies involved in discriminating abstruse overlapped signatures, aspects concerning training of neural networks for large and ill-conditioned data, complications related to varying applied voltages during measurement etc., continue to confront the research community. Since schemes for large dataset training based on arbitrarily chosen centers are found to be rather impractical and not tenable during discrimination, mixture density clustering technique that utilizes an Expectation Maximization with Maximum Likelihood strategy is implemented for training Homoscedastic and Heteroscedastic Probabilistic Neural Network (PNN) variants. Detailed analysis of the ability of the PNN variants is performed to determine the proposition of utilizing various preprocessing techniques in discriminating the PD signatures. In addition, studies are carried out on the PNN variants to determine the ability of the deterministically and autonomously created Probability Density Functions (PDF) in recognition and classification of substantially big dataset multi-source PD fingerprints due to varying levels of applied voltages.

Key words: Partial, probabilistic neural network, homoscedastic probabilistic neural network, heteroscedastic probabilistic neural network, expectation maximization, maximum likelihood, algorithm

INTRODUCTION

With the dawn of the recent deregulation strategies in power sector, competition among utilities has made it inevitable for power utilities to slash-down the cost related to operation and maintenance of equipment associated with energy utilization in power systems. Since the reliability of electrical equipment is related to the quality of its insulation, on-line diagnostic techniques have become essential and hence have grabbed the focus of researchers worldwide developing predictive diagnostic tools used for monitoring and assessment of insulation system. Recognition and classification of the source of Partial Discharge (PD) and subsequent analysis of discharge signature patterns is of immense importance, since it is a fundamental yet a key pre-requisite for credible diagnosis (Vahedi *et al*., 2012) of the insulation system. Partial Discharge (PD) is due the consequence of enhancement of electric field stresses in the restricted portion of a dielectric (BSI, 2000) bounded by electrodes. Since PD in insulation may be due to defects such as void, cavities, fissures, blow-holes, exterior imperfections etc., an array of divergent signature patterns that describe the complex nature of physics related to insulation degradation characteristics is exhibited. PD signature patterns invariably reveal a complex stochastic and non-Markovian process (Van Brunt, 1991) with substantial variability which may be ascribed to characteristics such as memory propagation effect (Van Brunt *et al*., 1993), temperature, role of the initiatory electron in the gap etc.

As an outcome of innovations in digital signal processing techniques and associated hardware, developments in high speed processors and evolution of associated data acquisition systems there has been a surge in eagerness among scientists and power

Corresponding Author: S. Venkatesh, Department of Electrical and Electronics Engineering,
School of Electrical and Electronics Engineering, SASTRA University, Thanjavur, 613401, India

equipment operators in analysing PD signatures and source discrimination. Further, since real-time PD measurements invariably involve discrimination of complex overlapped multi-source signatures, attention among peer researchers has now changed to identification of multi-source PD (Lalitha and Satish, 2000; Lee *et al*., 2000). During the past few decades a range of scientific and soft-computing techniques such as Neural Networks (NNs) (Lotfi and Benyettou, 2011; Tortoe *et al*., 2011; Gulski, 1995; Gulski and Krivda, 1993; Mazroua *et al*., 1994; Satish and Zaengl, 1994), Fuzzy Logic (Abdel-Galil *et al*., 2005; Contin *et al*., 2002), Fractal Features (Satish and Zaengl, 1995; Lalitha and Satish, 1998) and Contour Mapping, Hidden Markov Models (HMM) (Satish and Gururaj, 1993; Abdel-Galil *et al*., 2004), Wavelet Transformation (Ma *et al*., 2002) etc., were implemented for the discrimination of single source PD and partially superimposed multi-defect PD patterns with moderate results.

A few noteworthy contributions by researchers involved in PD discrimination studies including those of the authors of this research in PD pattern recognition utilizing NNs comprise the Back Propagation Network (BPN) (Mazroua *et al*., 1993; Karthikeyan *et al*., 2006a), Radial Basis Function (RBF), Self Organizing Map (SOM) (Satish and Zaengl, 1994), Adaptive Resonance Theory (ART) (Karthikeyan *et al*., 2006b), Counter Propagation Network (CPN) (Hoof *et al*., 1997) and rudimentary versions of Probabilistic Neural Network (Karthikeyan *et al*., 2005; Venkatesh *et al*., 2007; Karthikeyan *et al*., 2008), Nonetheless, problems concerning abstruseness of fully overlapped PD signature patterns, complicatedness in training large ill-conditioned PD data signatures obtained from real time and on site monitoring system, detrimental consequence of outliers during large dataset training, intricacies related to discrimination of PD finger prints as a consequence of changes in applied voltages during testing etc., persist.

This research work analyzes aspects related to classification of multi-source PD pattern from the standpoint of the function of initial seed (center selection) during clustering and subsequent training (Alamelumangai and Devishree, 2012) and complexities involved in discerning PD pulses. Further it is also the objective of this research to establish the classification capacity of the mixture density estimation based clustering algorithms that utilize the Expectation Maximization (EM) in conjunction with Maximum Likelihood (ML) strategy owing to different applied voltages since several research studies in associated fields of engineering have clearly demonstrated that training techniques that rely on arbitrarily selected prototype centers from a substantial volume of data is observed to be rigid and in many studies unfavorable during classification. Consequently, an EM-ML algorithm is implemented during the training phase for acquiring suitably initialized seed vectors which would serve as prototype centers for training the two modified versions of the Probabilistic Neural Network (PNN) namely the Homoscedastic PNN (HOPNN) and the Heteroscedastic PNN (HRPNN). The performance of PNN variants is analyzed to determine the effectiveness of the preprocessing techniques, to establish the impact of the smoothing parameter in the PNN versions during discrimination and to evaluate the ability of the deterministically and autonomously initialized PNN versions in classifying large dataset on-line multiple source PD signatures.

Further, the objective of the authors of this research is to extend their earlier study pertaining to multi-source PD pattern recognition (Venkatesh and Gopal, 2011). A wide range of similar convergent NN tools were utilized in order to ascertain the difficulties during the classification task. The same benchmark models that were made use of in the previous study is also adapted in this research.

EXPERIMENTATION AND BENCHMARK LABORATORY MODELS METHODOLOGY

Experimental test setup: Extensive laboratory analysis has been performed in this research in the lines of the stipulations and requirements laid down in IEC 60270 for acquiring and assimilating benchmark PD signature. Such stipulations and requirements would serve as a basic yet vital pointer in validating the proposed algorithms. Hence, it provides reliable and repeatable classification capability. The direct detection system and test setup procedure for measurement of PD signature patterns are obtained as laid down in IEC 60270 (BSI, 2000). In order to ensure enhanced detection and measurement of pulse signals transfer characteristics, a 1 nF (nano-Farad) coupling capacitor is appended to the test circuit. A digital PD Measurement System (Model No. DTM-D®) which comprises a Digital Storage Oscilloscope (TDS 2002B®) for measurement and built with a tunable variable filter-insert module (Model: DFT-1®) for acquiring pulses in the range 2-5000 pC whose center frequency is variable in the range 600-2400 kHz at a bandwidth of 9 kHz is utilized. The measured PD magnitude is quantified and exhibited in pico-coulomb (pC) or in milli-volt (mV). Figure 1 shows a typical layout of PD test arrangement utilized in this research.

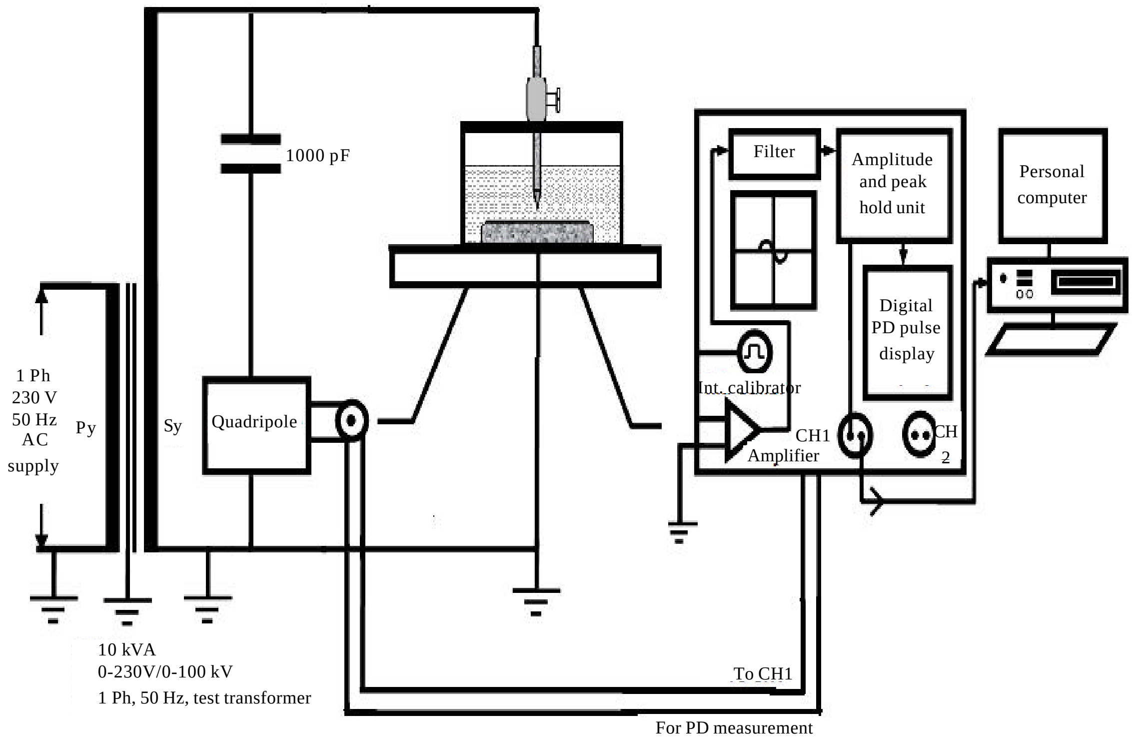

Fig. 1: Typical layout of PD testing and measurement setup

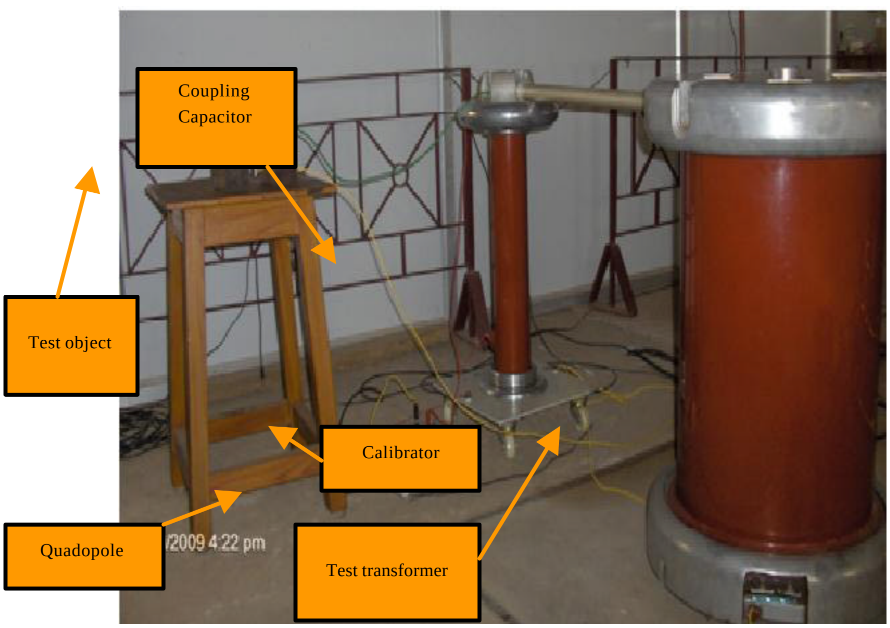

Fig. 2: Photograph of laboratory test setup with electrode bounded cavity benchmark model

Figure 2 and 3 indicate a photograph of the laboratory test layout and the PD detection, measurement and acquisition setup.

In addition, the PD test system is equipped with noise window-gating facility which mitigates the continuous noise during PD (Alesaadi *et al*., 2012)

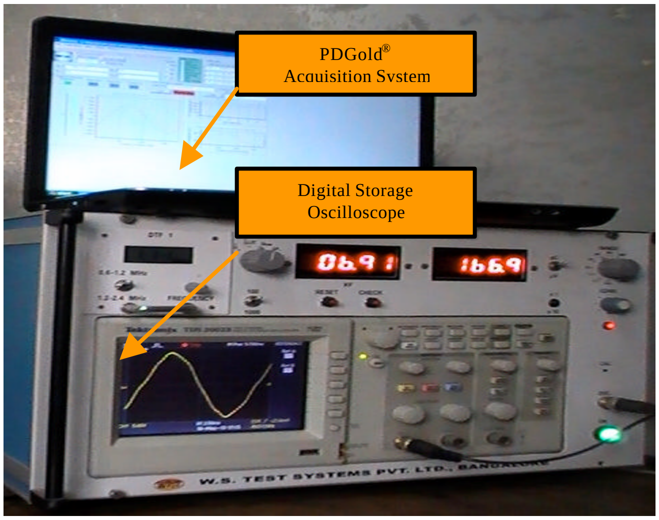

Fig. 3: Digital PD detection, measurement and acquisition module

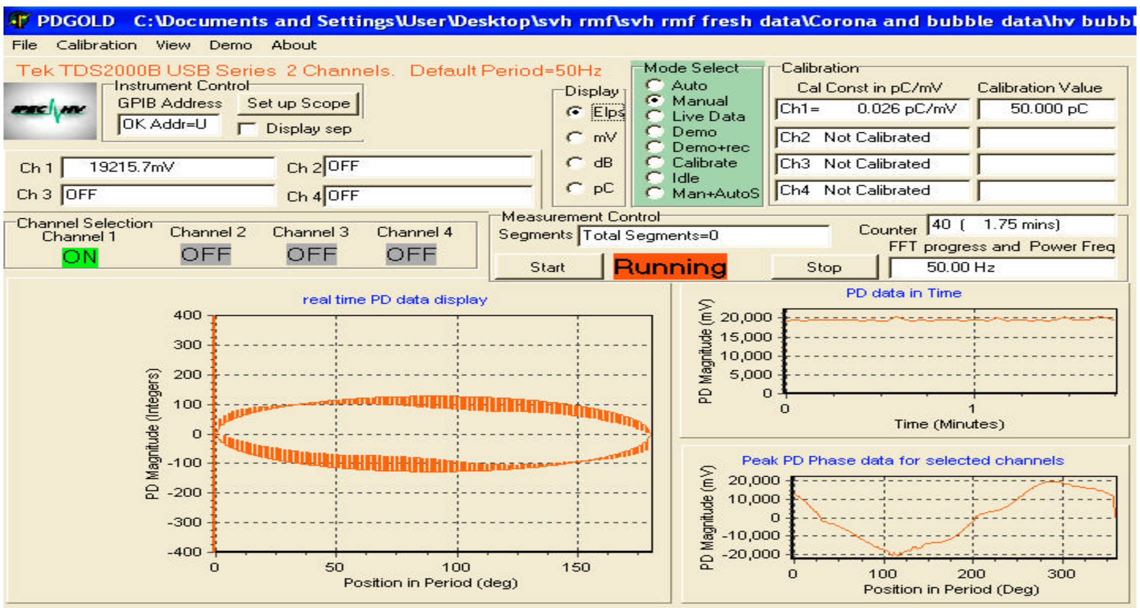

Fig. 4: Typical signature patterns representing multiple source PD (electrode bounded cavity with air corona sample) captured from PD gold©

measurement and acquisition. PD Gold© software (product of HV Solution Inc, UK) is interfaced to capture the PD pulses from the PD detection, measurement and acquisition system to in turn enable obtaining PD fingerprints. Test setup is calibrated according to the standard recommendations of IEC 60270 utilizing a digital reference calibrator (Model: PDG®). PD pulses obtained at power frequency (50 Hz) by PD Gold© acquisition software displays the pulse signature in sinusoidal and elliptical time-base which can be chosen in auto or manual category. The operator can acquire and record pulse data for a period of usually 10 min which is obtained from waveforms which are recordable for a range of 240 to 750 cycles. In this research study PD fingerprints are acquired and captured during testing for a duration of 5 min to ensure considerably preconditioned datasets that truly reflect the source of PD. Figure 4 replicates PD signature pattern representing multi-source PD fingerprints.

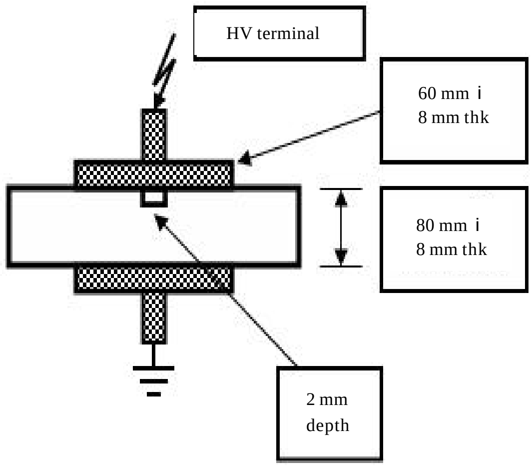

Fig. 5: Laboratory benchmark model simulating electrode bounded cavity

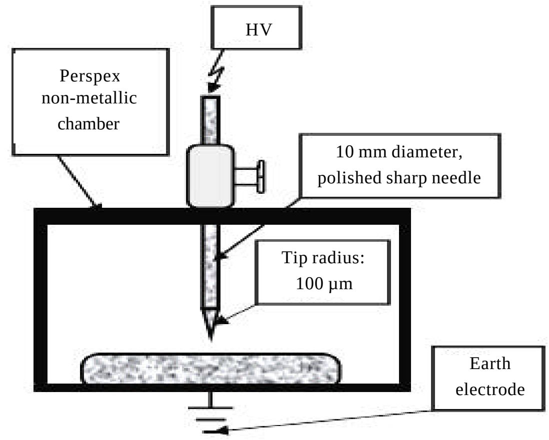

Fig. 7: Benchmark model duplicating air-corona discharges

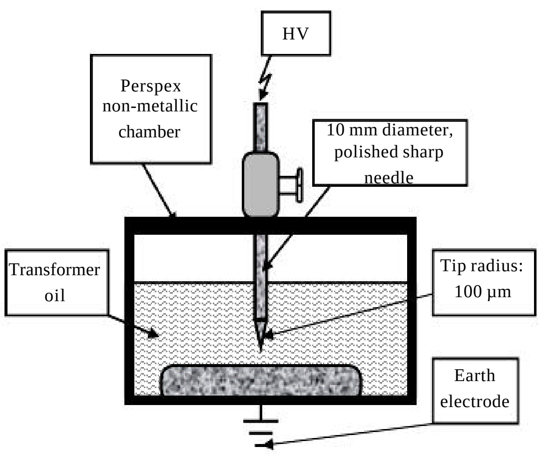

Fig. 6: Laboratory benchmark model indicating oil-corona discharges

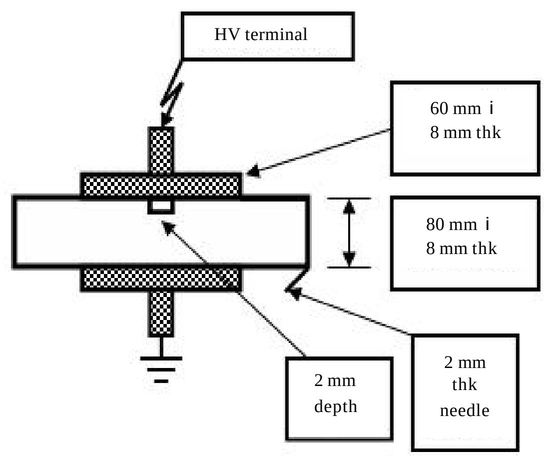

Fig. 8: Multiple source PD (electrode bounded cavity with air-corona)

Benchmark laboratory models for PD pattern classification: Five distinct sources of PD are replicated by fabricating laboratory benchmark models which comprises single and multi-source PD namely electrode bounded cavity, corona discharges in air, corona discharges in transformer oil, electrode bounded cavity overlapped with corona discharges in air and multiple source electrode bounded cavity of varying dimensions. Internal discharges are replicated by creating an electrode bounded cavity (labeled 'EC') of depth 1.5 mm with a diameter of 1 mm in sample insulation system fabricated from Poly Methyl Metha Acrylite (PPMA) with a diameter 80 andl2 mm thick as shown in Fig. 5. A second source of internal discharges called the 'corona discharges in oil' (tagged 'OC') is reproduced with a point-plane electrode configuration filled with transformer oil as shown in Fig. 6.

A important type of external PD labelled 'air-corona' (designated 'AC') is simulated using a rod electrode (with an apex angle 85°) connected to the High Voltage (HV) terminal as shown in Fig. 7. One among frequently encountered form of multi-source PD termed as 'electrode bounded cavity with corona discharges in air (labelled 'ECAC') is simulated by introducing a thin rod electrode of 2 mm diameter from the HV electrode. The other source of defect includes a 2 mm depth electrode-bounded cavity on the solid dielectric system which is in turn connected to the high voltage electrode. This arrangement is indicated in Fig. 8.

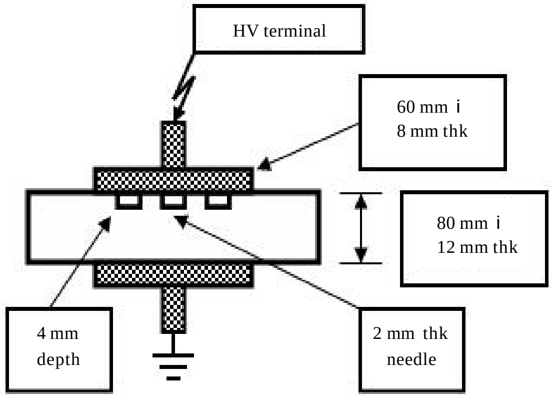

Fig. 9: Benchmark model replicating multiple electrode bounded cavity discharges

Another type of multiple source PD is comprises multi-source electrode bounded cavity discharges (labeled 'MEBC" are replicated by 3 sets of cavities wherein the outer cavities are of 4 mm depth and inner cavities are 2 mm depth placed at the high voltage terminal. The setup is shown in Fig. 9.

PREPROCESSING PD SIGNATURE PATTERNS AND FEATURE EXTRACTION

Though several established preprocessing and feature extraction strategies have been attempted by researchers (Sahoo *et al*., 2005), varied yet straight forward preprocessing techniques based on statistical measures are exploited for pattern discrimination task so as to determine the performance capability associated with the various vital trainable parts and characteristics of the PNNs such as the spread parameter, 'curse of dimensionality' (Bishop, 1995), quantification of the mean vector which is estimated based on probability density estimates etc. Incidentally, the phase-window methodology of preprocessing utilizing various statistical operators which has been adopted earlier by the authors of this research has also been taken up in this work. These measures include. (1) Maximum values of q (10° and 30°), (2) Minimum values of q (10° and 30°) and (3) Central Tendency (10° and 30°). Further, since several researchers (Gulski, 1995; James and Phung, 1995; Krivda, 1995) had earlier utilized the traditional statistical operators (operators that include mean, skewness, kurtosis, cross-correlation etc.) for the classification of PD signatures with good accuracy, research work carried out in this study focuses on establishing the capability of these operators in recognizing and discriminating substantially large dataset multi-source PD signatures.

Table 1: Multi-source large dataset pd pattern signatures for classification

Type/source of PD	Label	Applied voltage (kV)	Total No. of training patterns
Electrode bounded cavity	EC	7.3	90
		9.1	
		9.6	
Corona discharges in air	AC	13.7	90
		19.0	
		23.0	
Corona discharges in oil	OC	21.0	90
		29.0	
		32.0	
Void with corona discharges in air	ECAC	9.1	90
		9.3	
		14.0	
Multiple electrode bounded cavity	MEBC	7.0	90
		10.0	
		13.0	

Table 1 indicates the total number of fingerprints taken up for analysis. The database comprises large sets of PD signature patterns comprising of each PD source for varying applied voltages. It is pertinent to note from Table 1 that for classifying 660 sets of divergent patterns, two sets (36 sets relating to each defect category and 27 sets of fingerprint signatures related to every discharge category) are utilized for obtaining the appropriate centers chosen from the proposed clustering algorithm which are further utilized for training and testing of the PNN versions It is observed during this research that the centers of PD patterns serve as a reasonably good representative codebook.

HOMOSCEDASTIC AND HETEROSCEDASTIC PNN FOR MULTI-SOURCE PD PATTERN RECOGNITION

General aspects and architecture of PNN versions: PNN (Specht, 1988) is a representation based on the competitive learning strategy with a 'winner-takes-all attitude'. PNN is a classifier adaptation, wherein decision making is based on the Bayesian scheme which is integrated with a non-parametric estimation technique (Parzen window) for computing the Probability Density Function (PDF). The Bayesian model provides class conditional probability decision for pattern classification which is utilized for obtaining an appropriate optimal estimate of PDF. The basic (or) original version of the PNN (OPNN), makes use of a Parzen window estimator with a mixture of Gaussian kernels and has no feedback path. OPNN takes all the sample vectors during training as centres 'c' of Gaussian kernel function with only the tunable part of the network to be adjusted during the training stage being a common variance (σ). Though it is apparent from the discussion that more the exemplars the better is the classification rate of the OPNN, it is crucial that a thorough study is carried out to analyze the

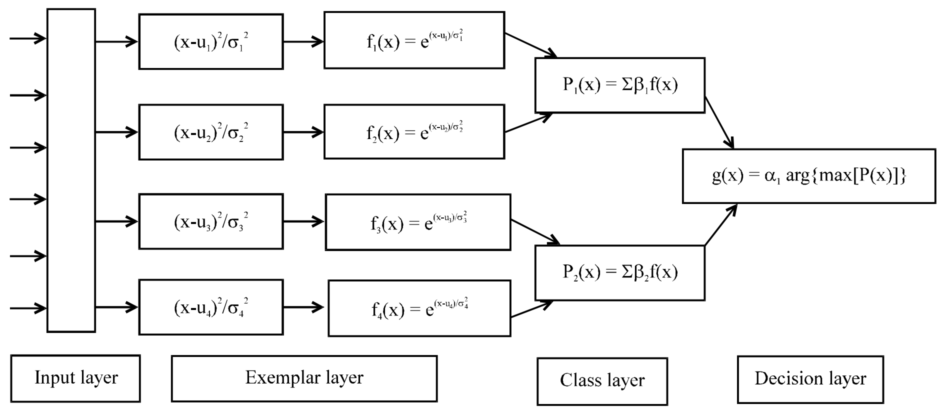

Fig. 10: Architecture of hetroscedastic PNN (HRPNN)

deleterious influence of over-training and catastrophic forgetting are analyzed during discrimination in the context of large dataset training. It is hence evident that the training scheme would be only be reasonable if a smaller set of kernel functions that reflect a meaningful depiction of the training data is obtained without conceding on the accuracy of discrimination capability. The Expectation Maximization (EM) with Maximum Likelihood (ML) training algorithm implementation for mixture of Gaussian PNN as modelled (Streit and Luginbuhl, 1994) with a common covariance ('identical spread') is termed as Homoscedastic PNN (HOPNN). Uncorrelated Gaussian kernel functions provided with the capability to obtain different variances ("varying spread") called the yield comparatively improved sets of adequate and parsimonious centers that characterize the class conditional PDFs could be utilized. Thus employing uncorrelated Gaussian kernel functions with different variance ("different scatter") is called Heteroscedastic PNN (HRPNN). The proposed PNN variants provide suitable multivariate density estimation clustering strategies that are inherently robust in handling large datasets comprising outliers. Further since the EM-ML algorithms have rapid training speed and good convergence (Xu and Jordan, 1996), the proposed versions of PNN provide viable opportunity for training and testing large dataset on-line PD signature patterns.

The architecture of HRPNN (Streit and Luginbuhl, 1994) is essentially similar to that of the OPNN with only a few modifications in the pattern layer and the approach by which decision is obtained in the output layer. The architecture of HRPNN is shown in Fig. 10.

The first layer receives the input patterns. In the second layer node (pattern unit), the ith kernel in the jth group, is defined as a Gaussian basis function:

$$p_{i,j}(x) = \frac{1}{\left(2\pi\sigma_{i,j}^2\right)^{\frac{d}{2}}} e^{-\frac{\left\|x-c_{i,j}\right\|^2}{2\sigma_{i,j}^2}}$$

where $c_{i,j}$ is the centre or the mean vector and $\sigma_{i,j}^2$ is the variance or smoothening parameter. The third layer has k nodes:

$$f_j(x) = \sum_{i-1}^{R_{i,j}} \beta_{i,j} p_{i,j}(x), 1 \le j \le R$$

and each node estimates a class conditional PDF (f_j) using a mixture of Gaussian kernels. The fourth layer (decision layer) makes decision according to the equation:

$$g(x) = \arg\left\{\max\left[\alpha_j f_j(x)\right]\right\}$$

where α_j refers to the class a-priori probability.

Learning algorithm-expectation maximization with maximum likelihood (EM-ML) estimation: EM is an iterative system that utilizes two major steps namely the expectation process (E-step) and a maximization process (M-step) to obtain the Maximum Likelihood (ML) estimation of a set of paramter. EM algorithm is assured to converge to an ML estimate (Streit and Luginbuhl, 1994; Xu and Jordan, 1996) at a rapid rate. Every sequence calculates an expectation (mean) value of a group of unobserved data using the present estimated value of the data and the observed value. In both the cases an iterative procedure is employed to minimize the variations in the log-posterior likelihood function. Each M-step takes the data from the E-step presuming it to be the truly measured data to acquire the likelihood function and thus determine the estimate of the desired parameter. From the

Table 2: Classification capability of HOPNN and HRPNN versions for discrimination of multi-source PD signatures

Feature extracted Input based on Phase window	No. of tuples	Total No. of testing datasets	Randomly chosen initial centers-144 No. (36 sets for each class): No. of PDF labeled sequentially as EC, AC, OC and ECC, MEC (Set 1 only)		Randomly chosen initial centers-144 No. (36 sets for each class): No. of Iterations		Classification capability EM-ML algorithm (%): HOPNN		Classification capability EM-ML algorithm (%): HRPNN	
			HOPNN	HRPNN	Set 1	Set 2	Set 1	Set 2	Set 1	Set 2
φ-q_{max}-n (30°)	36	624	24,22,20,21,23	20,18,18,20,17	32	18	83.4	79.1	86.2	88.4
φ-q_{min}-n (30°)	36	624	26,21,25,18,26	24,17,21,23,22	38	15	80.1	78.3	82.8	80.8
φ-q_{max}-n (10°)	108	660	16,18,16, 17,21	13,15,16,12,16	42	15	85.6	80.2	87.5	81.6
φ-q_{min}-n (10°)	108	660	19,17,18,17, 23	15,14,19,18,24	48	27	82.7	80.6	82.2	81.8
Traditional statistical operators (30°)	48	624	16,16, 17,13,15	14,16, 15,12,16	17	14	91.6	89.1	92.2	90.1
Traditional statistical operators (10°)	144	624	13,16, 18,14,16	12,11, 15,13,14	19	16	92.1	90.2	92.8	90.6

point of view of training PNN variants, it is appropriate to note that in each E and M-step, the mean and variance parameter is adjusted till the log-posterior likelihood estimate is brought to an optimal minimum. Since the deliberation in this implementation is related to obtaining the estimates for mean and variance for the PNN versions, the algorithm computes the weights which is in addition updated during the every step. E-step adapts the estimated PDF in the exemplar layer along with the mixing coefficient (β) to obtain the estimated value of the weights. Subsequently the M-step computes estimate and the weights in the E-step to create a likelihood function which in turn serves in obtaining the overall maximized likelihood estimate of the parameters. Thus, the final values for the centers ($c_{E,i}$), variances ($\sigma^2_{E,i}$) and the mixing coefficients ($\beta_{E,i}$) are computed.

ANALYSIS OF RESULTS PERFORMANCE OF PNN VERSIONS IN CLASSIFYING MULTIPLE PD PATTERNS

Comparison of classification capability of PNN versions based on randomly selected initial seed vectors: The proposed PNN versions have been implemented using MATLAB version 7.1, Release 14 and during the training phase, randomly selected preprocessed vectors for the initialization of representative centers are taken. For the purpose of comparison, 'Set 1' centers is formulated such that the vectors are alternatively odd and even numbered samples of PD signatures pertaining to each applied voltage. However, in order to verify perceptible variations in discrimination of the proposed PNN versions in discerning 'Set 2', centers pertaining to each applied voltage are sampled such that the last two signature vectors are taken up. Table 2 summarizes the performance of the proposed PNN versions. The noteworthy aspects observed in Table 2 are summarized:

- Misclassification rate invariably is doubled for each type of randomly selected centers. It is also pertinent to note that a more frugal set of centers is achieved while employing the HRPNN with superior discrimination. However, a few facets related to the effect of outliers in discriminating inter-class clusters are observed to have created problems during classification of fully overlapped patterns
- It is also clear from Table 2 that more frugal sets of centers are obtained utilizing the HRPNN algorithm with superior classification capability
- It is obvious from Table 2 that the number of PDFs produced in both PNN variants has a wide variation which may be attributed to the choice of the initialization in the input seed vectors
- It is worth noting that a perceptible reduction in the number of representative PDFs was observed when the autonomously acquired centers are utilized A sample analysis and study indicated that in the case of the measures based on traditional statistical operators a more prudent set of centers (sequence of the PDF is 12-11-15-13-14) is obtained in the case of the HRPNN version

CONCLUSION

Based on the comprehensive analysis the significant merits of utilizing the Gaussian mixture density based clustering algorithms for obtaining initialized seed vectors for training and subsequent classification of multiple source PD patterns by HOPNN and HRPNN versions are summarized:

- In spite of the fact that training PNN versions utilizing the EM-ML algorithm provide as a realistically good scheme for training the networks, it is lucid from the exhaustive studies that the discrimination rate of the PNN variants may be modified to perform with superior performance by utilizing the labelled clustering algorithms to obtain enhanced prototype initial cluster seed vectors

- Since rapid computational speed and convergence is observed with both the proposed PNN versions (for all the types of preprocessed feature vectors) this methodology provides an exciting prospect for implementing a modular PNN framework for implementation in real-time and on-line condition monitoring system for classifying insulation faults and diagnosis of insulation of power components. In this context it is pertinent to note that since the major intention of this research is to assess the performance of the homosedastic and heteroscedastic PNN versions in classifying multi-source PD signatures due to randomly initialized centers obtained from the proposed maximum likelihood based algorithm procedure and its subsequent PDF estimates, the analysis has been taken up with fundamental yet effective statistical measures only. The 'K-fold' and the 'One-Hold-One-Out' methodology have been performed to check the authenticity of the input in this research. Yet it is appropriate to note that more research and validation is indispensable in the case of large dataset PD signature analysis
- Further it is also important to note that during the course of the entire study, a fixed value of the smoothing parameter (σ = 0.0001) as a common spread parameter value has been used in this research analysis in the case of HOPNN training since it is observed that for almost all the preprocessing measures taken up, the class decision hyper-boundaries are reasonably well distinguished and uniquely separated

ACKNOWLEDGMENTS

This study was funded by the Research and Modernization Fund (RMF) Grant, Project No. 6, constituted by SASTRA University. The first author is greatly indebted to Prof. R. Sethuraman, Vice-Chancellor, SASTRA University and Dr. S. Swaminathan, Dean-Sponsored Research and Director-CeNTAB, SASTRA University for sanctioning the grant and for the unwavering support and motivation extended to him during the course of the research work. The authors are also grateful for the motivation and support extended by Dr. B. Viswanathan, Dean/SEEE, SASTRA University.

REFERENCES

Abdel-Galil, T.K., R.M. Sharkawy, M.M.A. Salama and R. Bartnikas, 2005. Partial discharge pattern classification using the fuzzy decision tree approach. IEEE Trans. Instrum. Meas., 54: 2258-2263.

Abdel-Galil, T.K., Y.G. Hegazy, M.M.A. Salama and R. Bartinikas, 2004. Partial discharge pulse pattern recognition using hidden markov models. IEEE Trans. Dielectr. Electr. Insul., 11: 715-723.

Alamelumangai, N. and J. Devishree, 2012. Automated clustering of cancer cells using fuzzy C means with repulsions in ultrasound images. J. Artif. Intel., 55: 14-25.

Alesaadi, A., A. Ghaedi, A. Ghaffari and V. Parvin, 2012. De-noising of online PD signals in power transformers using the Bhattacharyya distance. Trends Applied Sci. Res., 7: 813-828.

BSI, 2000. High-voltage test techniques-partial discharge measurements. BS EN 60270:2001, IEC 60270:2000. http://www.samgor.com/pdf/IEC60270.pdf.

Bishop, C.M., 1995. Neural Networks for Pattern Recognition. Oxford University Press, Oxford, UK., ISBN: 9780198538646.

Contin, A., A. Cavallini, G.C. Montanari, G. Pasini and F. Puletti, 2002. Digital detection and fuzzy classification of partial discharge signals. IEEE Trans. Dielectr. Electr. Insul., 9: 335-348.

Gulski, E. and A. Krivda, 1993. Neural networks as a tool for recognition of partial discharges. IEEE Trans. Electr. Insul., 28: 984-1001.

Gulski, E., 1995. Digital analysis of partial discharges. IEEE Trans. Dielectr. Electr. Insul., 2: 822-837.

Hoof, M., B. Freisleben and R. Patsch, 1997. PD source identification with novel discharge parameters using counterpropagation neural networks. IEEE Trans. Dielectr. Electr. Insul., 4: 17-32.

James, R.E. and B.T. Phung, 1995. Development of computer-based measurements and their application to PD pattern analysis. IEEE Trans. Dielectr. Electr. Insul., 2: 838-856.

Karthikeyan, B., S. Gopal and S. Venkatesh, 2005. Probabilistic neural network and its adaptive version-A stochastic approach to PD pattern classification task. Int. J. Inform. Acquisit., 2: 333-344.

Karthikeyan, B., S. Gopal and S. Venkatesh, 2006a. ART-2: An unsupervised neural network for PD pattern recognition and classification. Int. J. Expert Syst. Appl., 31: 345-350.

Karthikeyan, B., S. Gopal, S. Venkatesh and S. Saravanan, 2006b. PNN and its adaptive version-an ingenious approach to PD pattern classification compared with BPA network. J. Electr. Eng., 57: 138-145.

Karthikeyan, B., S. Gopal and S. Venkatesh, 2008. Partial discharge pattern classification using composite versions of probabilistic neural network inference engine. Int. J. Expert Syst. Appl., 34: 1938-1947.

Krivda, A., 1995. Automated recognition of partial discharges. IEEE Trans. Dielectr. Electr. Insul., 2: 796-821.

Lalitha, E.M. and L. Satish, 1998. Fractal image compression for classification of PD sources. IEEE Trans. Dielectr. Electr. Insul., 5: 550-557.

Lalitha, E.M. and L. Satish, 2000. Wavelet analysis for classification of multi-source PD patterns. IEEE Trans. Dielectr. Electr. Insul., 7: 40-47.

Lee, J.H., T. Okamoto and C.W. Yi, 2000. Classification of PD patterns from multiple defects. Proceeding of the 6th International Conference on Properties and Applications of Dielectric Materials, Volume 1, June 21-26, 2000, Xi'an, China, pp: 463-465.

Lotfi, A. and A. Benyettou, 2011. Using probabilistic neural networks for handwritten digit recognition. J. Artif. Intell., 4: 288-294.

Ma, X., C. Zhou and I.J. Kemp, 2002. Interpretation of wavelet analysis and its application in partial discharge detection. IEEE Trans. Dielectr. Electr. Insul., 9: 446-457.

Mazroua, A., R. Bartnikas and M.M.A. Salama, 1994. Discrimination between PD pulse shapes using different neural network paradigms. IEEE Trans. Dielectr. Electr. Insul., 1: 1119-1130.

Mazroua, A.A., M.M.A. Salama and R. Bartnikas, 1993. PD pattern recognition with neural networks using the multilayer perceptron technique. IEEE Trans. Electr. Insul., 28: 1082-1089.

Sahoo, N.C., M.M.A. Salama and R. Bartnikas, 2005. Trends in partial discharge pattern classification: A survey. IEEE Trans. Dielectr. Electr. Insul., 12: 248-264.

Satish, L. and B.I. Gururaj, 1993. Use of hidden Markov models for partial discharge pattern classification. IEEE Trans. Electr. Insul., 28: 178-182.

Satish, L. and W.S. Zaengl, 1994. Artificial neural networks for recognition of 3-D partial discharge patterns. IEEE Trans. Dielectr. Electr. Insulation, 1: 265-275.

Satish, L. and W.S. Zaengl, 1995. Can fractal features be used for recognizing 3-D partial discharge patterns. IEEE Trans. Dielectr. Electr. Insulation, 2: 352-359.

Specht, D.F., 1988. Probabilistic neural networks for classification, mapping, or associative memory. Proceedings of the IEEE Conference on Neural Networks, Volume 1, July 24-27, 1988, San Diego, CA., USA., pp: 525-532.

Streit, R.L. and T.E. Luginbuhl, 1994. Maximum likelihood training of probabilistic neural networks. IEEE Trans. Neural Networks, 5: 764-783.

Tortoe, C., J. Orchard, A. Beezer and J. Teetheh, 2011. Application of radial basis function network with a gaussian function of artificial neural networks in osmo-dehydration of plant materials. J. Artif. Intell., 4: 233-244.

Vahedi, H., M. Yazdani-Asrami and M. Abedi, 2012. Detection of partial discharge location in power generator windings by means of frequency response analysis. Trends Applied Sci. Res., 7: 87-95.

Van Brunt, R.J., 1991. Stochastic properties of partial-discharge phenomena. IEEE Trans. Electr. Insul., 26: 902-948.

Van Brunt, R.J., E.W. Cernyar and P. von Glahn, 1993. Importance of unraveling memory propagation effects in interpreting data on partial discharge statistics. IEEE Trans. Electr. Insul., 28: 905-916.

Venkatesh, S. and S. Gopal, 2011. Robust heteroscedastic probabilistic neural network for multiple source partial discharge pattern recognition-significance of outliers on classification capability. Int. J. Expert Syst. Appl., 38: 11501-11514.

Venkatesh, S., S. Gopal, P.S.S. Srinivasan and B. Karthikeyan, 2007. Identification of multiple source partial discharge patterns using bayesian classifier based neural networks: A comparison of supervised and unsupervised learning techniques. Proceedings of the 15th International Symposium of High Voltage Engineering, August 27-31, 2007, Ljubljana, Slovenia.

Xu, L. and M.I. Jordan, 1996. On convergence properties of the EM algorithm for Gaussian mixtures. Neural Comput., 8: 129-151.