A NOVEL HYBRID CONTINUOUS DENSITY HIDDEN MARKOV MODEL- PROBABILISTIC NEURAL NETWORK FOR MULTIPLE SOURCE PARTIAL DISCHARGE PATTERN RECOGNITION

S. Venkatesh[1*], S. Gopal[2] and K. Kannan[1]
[1]SASTRA University, Thanjavur, Tamil Nadu, India
[2]W.S. Test Systems, 27th km Bellary Road, Bangalore, India
*Email: <venkatsri73in@gmail.com>

Abstract: Partial Discharge (PD) is inherently a non-Markovian process exhibiting significant statistical variability in the correlated patterns that describe the nature of discharges. Hence, the utility of Hidden Markov Model (HMM) as a tool to recognize spatio-temporal sequences of dynamically varying patterns becomes perceptible. Though a few researchers have attempted PD pattern recognition utilizing the standard version of HMM, such studies are confined to stationary Discrete Density HMM (DDHMM) versions which reported a relatively lower success rate. Since PD patterns exhibit dynamic behaviour, a non-stationary Continuous Density HMM (CDHMM) which best describes the hidden state transition probabilities as time dependant estimates coupled with multivariate Gaussian observation densities is taken up for study. This research focuses on implementing a novel hybrid HMM-Probabilistic Neural Network (HMMPNN) PD pattern recognition system which utilizes the complementary advantages of both the models i.e., ability of HMM in recognizing spatio-temporal sequences and capability of NN in classifying patterns. Exhaustive studies and analysis is carried out on laboratory models and industrial objects to determine the effectiveness of classification of the proposed hybrid system to cater to large training dataset. Comparison of performance of the algorithms with various pre-processing schemes that utilize statistical operators has also been carried out.

1 INTRODUCTION

Though there have been rapid advances in design, manufacture, quality control etc of power apparatus, instances of failure of equipment due to insulation are still being reported since minor flaws such as blow-holes, voids, surface imperfections etc are practically inevitable leading to partial discharges (PD). PD phenomenon which incidentally serves also as a non-destructive technique is an electrical breakdown which is essentially confined to the localized region of dielectric system of a power apparatus. Since identification and classification of the source of PD is fundamental to diagnosis, recently the focus of researchers has shifted to recognition of defects due to multiple sources of PD. Recent research indicates that recognition of multiple source PD patterns is yet a challenging facet though techniques such as Mixed Weibull functions, Neural Networks (NN) [1], Wavelet Transformation, Hidden Markov Models (HMM) [2-3], etc have reported significant success for single source PD only. Moreover, since modern digital PD acquisition systems record data for a stipulated duration, the database is substantially large leading to difficulties in clustering and categorizing PD sources.

Incidentally, a few researchers [3-4] have attempted HMM for PD pattern recognition which was confined to investigations based on stationary discrete density HMM (DDHMM) versions and reported a moderate degree of recognition capability. Since PD is inherently a non-Markovian and a complex stochastic process wherein the pulse patterns that describe the major attributes of the discharge mechanism has significant statistical variability, a non-stationary continuous density HMM (CDHMM) [4] which best describes the hidden state transition probabilities as time dependant estimates coupled with multivariate Gaussian observation densities is proposed as a model that describes the dynamics in PD patterns. In addition, since several researchers in various allied fields have also reported on the limitations of HMM in discriminating patterns [5], a novel hybrid HMM-Probabilistic Neural Network (HMMPNN) PD pattern recognition system which utilizes the complementary advantages of both the models is implemented. Detailed studies are carried out to ascertain the efficacy of classification of the proposed hybrid system in catering to large multiple source PD training dataset. Comparison of the performance of the algorithm with various preprocessing schemes that utilize statistical operators has also been studied.

2 DEVELOPMENT OF HYBRID CDHMM-PNN FOR PD PATTERN RECOGNITION

2.1 Non-stationary CDMM for Dynamic PD Pattern Recognition

A HMM is defined as a doubly stochastic process, comprising an underlying stochastic process that is not directly observable, but can only be visualized through another set of stochastic processes that produce a sequence of observations. Since the state of the model at an instant is not directly observable, the model takes the name 'Hidden Markov Model'. Specific details are provided in the following sections. HMMs can be classified based on their topology into two major categories as either an ergodic HMM (fully connected HMM) or as a non-ergodic HMM (left-to-right HMM). Since the left-right model is characterized by an increase in the state index for increasing time, it inherently attempts to map the spatio-temporal behaviour of the time varying (for varying applied voltages) PD patterns and hence has been considered in this research work.

HMMs are also classified based on the procedure of obtaining the density estimates of the observations as discrete and continuous density HMMs. A HMM comprises 'N' states wherein some form of physical significance may be attributed to them. In a Markov process, a new state is given by $s_t \in \{1,2,...,N\}$ in steps of t= 1, 2 … T. The initial state distribution matrix is labelled 'π' wherein $\prod = \{\pi_i\}, \pi_i = P_r(s_1 = i)$. Computation of the state transition probability matrix 'A' involves calculation of $A = \{aij\}, aij = P_r(s_{t+1} = j \mid s_t = i)$ for i=1,2,…,N. The resulting state sequence is denoted by $S = \{s_t, t = 1,2,...T\}$ and the probability obtained is given by $P_r(S \mid A, \prod) = \pi_{s1} \prod_{t=2}^{T} a_{s_{t-1}s_t}$.The emission or output vector is described by a sequence $O = \{o_1, o_2, ..., o_T\}$ where $o_t \in \Re^M$ with M dimension for the 'O' observation sequences. The output vector is produced according to the probability distribution called the observation probability distribution 'B' given by $B = \{b_j(o_t)\}$ where $b_j(o_t) = P(o_i = k \mid s_t = j)$. A model of this sort is called the DDHMM. If the emission probability distributions are described as Gaussian mixture model with a conditional mean vector μ_j and a covariance matrix V_j obtained for density related to 'j^{th}' state the density is described as $N(\mu_j, V_j)$. This model is called the continuous density HMM (CDHMM). The parameters that specify the CDHMM are A, π, B.

CDHMM, models the transition probabilities as one in which from t to t+1, the transition is dependent on the state't' only. Though this may provide classification/ recognition capability, the dynamic behaviour of the PD patterns may not always provide satisfactory results. Hence, to enhance the dynamic behaviour of the model a non-stationary component is added into the model which aims at capturing the time-dependence in the state transition probabilities. The state transition matrix is now generated to have T-1 number of matrices and is modified to A_t where t =1,2,…,T-1. Another major aspect of implementation of the non-stationary (dynamic) CDHMM is in ascertaining the number of hidden states. In the case of PD pattern recognition the number of states are obtained from the relative durations of the zones where PD pulses are discerned and where the 'background' (no PD) zones are observed [2]. It is obvious from such correlations that the number of states for appropriate PD pattern classification is 5 in most studies [3]. The training phase of the dynamic HMM involves obtaining the state optimized likelihood function (using the maximum likelihood algorithm) pertaining to the parameters A, Π, B related to a class of PD source. Maximization of the state optimized likelihood for each training sequence of a set of observations is obtained by utilizing the Viterbi algorithm. At the completion of the training stage, the basic dynamic patterns pertaining to the source of PD is deemed to have been learnt.

2.2 PNN as a Post- Processor for Hybrid CDHMM in Classifying PD Patterns

The a-posteriori probabilities obtained as optimized state estimates of the CDHMM algorithm forms the weight vector for further training by the neural network. Since Probabilistic Neural Network (PNN) has its inherent strength of utilizing a probabilistic framework in addition to implementing a Bayesian strategy during decision making (obtaining the class conditional probability) process, the PNN augurs well for the classification task.

PNN [6] is a formulation based on the concept of a non-parametric estimator for obtaining multivariate probability density estimates. It is a model based on competitive learning with a 'winner takes all attitude'. The fundamental version does not comprise any feedback path. PNN is essentially a classifier, which combines the Bayesian strategy for decision-making as a part of the decision layer of the NN along with utilizing a non-parametric estimator (Parzen window) for obtaining estimates of conditional probabilities. This basic version of PNN uses all the training samples as centers 'c' or mean vectors of the Gaussian function with only the trainable part of the NN namely the mixing coefficients (β) and a common variance (σ) or smoothing parameter to be estimated. PNN is a four layer neural network. It comprises an input layer, two hidden layers (exemplar/ pattern layer and class/ summation layer) and one decision layer.

Further, since studies by several researchers in allied fields and by the authors of this research [7] have clearly demonstrated the classification capability of PNN and more so in condition monitoring applications, this network paradigm has been taken up for study in this research. As hindsight, though the standard version of PNN may result in a huge number of hidden nodes during training, since the optimal state estimates that form a part of the CDHMM trained output now becomes the training dataset, frugal sets of feature vectors are obtained.

3 PD TEST SETUP AND LABORATORY MODELS FOR PATTERN CLASSIFICATION

3.1 Generic Arrangement of PD Measurement and Acquisition System

Investigations have been carried out using a 10kVA, 100kV, 50Hz test transformer with associated accessories and a W.S Test Systems make Digital PD Measurement System Model No.: DTM-D®) with a measurement range of 2-5000pC. A built-in oscilloscope (TDS 2002B) alongside a tunable adjustable filter-insert (Model: DFT-1) with a center frequency variable from 600 kHz- 2400 kHz at a bandwidth of 9 kHz is utilized for acquiring PD pulses. The direct detection and measurement test setup as recommended in IEC has been utilized in carrying out all the tests in this research since the tests have been carried out in a controlled laboratory environment thus obviating the need for alternative strategies for noise suppression. Notwithstanding, the PD measurement comprises window gating facility to mask and suppress unwanted background noise during measurement. The arrangement of the test circuit is carried out so as to comply with the stipulations laid down in IEC 60 270 with regard to the various requirements of the test procedure. In order to improve the transfer characteristics of the test circuit a 1nF coupling capacitor is also included in the test setup. Calibration of the test setup is carried out using a reference calibrator (Model: PDG®) and in line with the requirements of IEC 60 270. Fig. 1 and Fig. 2 show a typical generic arrangement of the laboratory test setup for PD pattern recognition studies and the PD measurement and acquisition system used in this research.

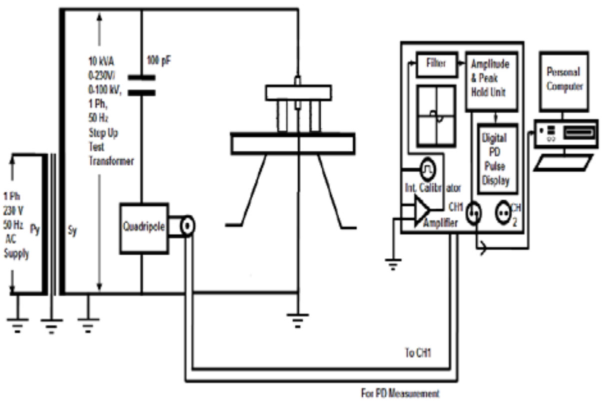

Figure 1: Generic Arrangement of PD Measurement and Acquisition System

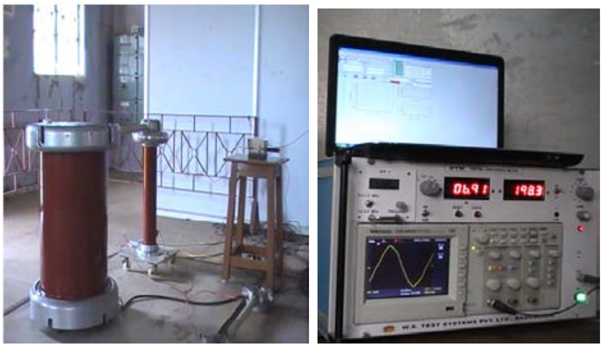

Figure 2: Typical Layout of Laboratory Setup and Digital PD Measurement and Acquisition System

PD Gold® acquisition software is interfaced with the PD detection system for acquiring the PD patterns. The unit also detects PD for 50Hz power cycle which allows the user to observe the shape of the PD pulses detected in addition to acquiring the phase resolved PD (PRPD) patterns in real-time. The pulses are displayed in both sinusoidal and elliptical forms selectable in auto or manual mode.

3.2 PD Experimental Studies and Laboratory Models

In order to ascertain the recognition (identification) and classification (discrimination) capability of the proposed model two exhaustive case studies have been taken up for studies. The first study comprises complex multiple source PD patterns wherein three models pertain to single source discharge patterns i.e., electrode bounded cavity (EC), air corona (C) and oil corona (OC) while the fourth model involves another major form of complex multiple source PD (overlapped patterns) namely the electrode bounded cavity with air corona (ECC). The second comprises a study pertaining to dynamically varying pattern due to surface tracking initiated dry band formation in ceramic disc insulators which serves as a precursor to pollution initiated flashover.

3.2.1 Case Study 1: Laboratory Models simulating Multiple Source PD Patterns Laboratory models have been fabricated to replicate PD patterns that are representative of the source of discharge displayed in the oscilloscope in line with the recommendations of [8]. A 12mm thick, 20 mm diameter perspex with electrode-bounded cavity of 2 mm depth simulates internal PD. Electrode bounded cavity with air-corona is produced by inserting a needle configuration of 2 mm diameter from the HV electrode in addition to a 2 mm electrode-bounded cavity on perspex in the high voltage electrode. These are shown in Figure. 3. Corona discharges in air and oil are replicated with a point electrode initiated from a stainless steel rod with an approximate angle of 85° and a plane electrode of 10mm thick, 60mm diameter as shown in Fig. 4.

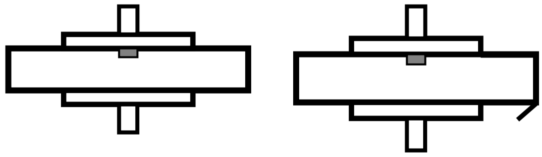

Figure 3: Laboratory models simulating Electrode bounded cavity and multiple source discharges

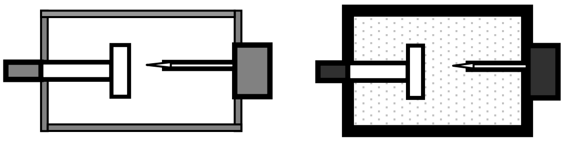

Figure 4: Laboratory models simulating Air and Oil Corona discharges

PD database comprises hundred patterns of each type of the defect for various applied voltages. It is to be noted that these patterns exhibit the statistical variations in the pulse patterns for each cycle of the sinusoidal voltage indicating the inherent non–Markovian nature of PD thus making obvious the difficulty encountered during classification. The task becomes even more demanding due to varying applied voltages. Table 1 shows the patterns acquired for large dataset from various sources of PD.

Table 1: Database for multiple source PD patterns for varying applied voltages

Class Label of PD for Identification	Type of PD	Applied Voltage (kV)	Total Patterns
EC	Electrode Bounded Cavity	7.28 9.1 9.6	120
C	Air-Corona	13.65 20.93 22.75	120
OC	Oil-Corona	20.93 29.12 31.85	120
ECC	Electrode Bounded Cavity with Air-Corona	9.1 9.6 13.6	120

3.2.2 Case Study 2: PD Pattern Approach for Pollution Severity Initiated Flashover in Ceramic Insulators Since contamination flashover has become a vital aspect in the design of high-voltage outdoor insulation, this research envisages studies pertaining to severity associated flashover prediction in ceramic insulators using partial discharge patterns as a tool for diagnosis. Since this research focuses on conducting predictive tests to determine the performance of polluted insulator due to dynamic changes in the PD pattern due to varying salinity, an artificial pollution test is conducted with equivalent salt as the pollutant to analyze the performance of the insulators. Since the clean fog test reflects the contamination mechanism occurring in industrial locations, the test is carried out by is dipping the insulator into slurry consisting of 40g of kaolin with varying levels of salinity (39 gm/l, 57 gm/l, 67 gm/l and 91 gm/l). It is observed during studies that except for the clean and dry case of leakage current waveforms the remaining cases are similar in spite of the varying polluting conditions. On the contrary, in the case of PD patterns, it is revealed that there exist significant differences between phenomena that do or do not affect insulator flashover performance and arcing due to polluted surfaces. It is observed that PD on polluted surfaces clearly exhibits a large distribution and scattered number of PD pulses as compared to dry and wetted insulator before the initiation of scintillation led flashover. Due to the afore-mentioned reasons, PD pulses are analyzed to assess the pollution initiated flash over in insulators. A series of experiments are performed on four different porcelain insulator samples of varying dimensions to assess the performance of pollution severity initiated flashover due to power frequency voltages. Measurement of correlated PD and leakage current were studied for the following cases: 1. pin of the ceramic insulator, 2. cap of the ceramic insulator and 3. both pin and cap of ceramic insulator. Table 2 indicates study carried out for the samples.

Table 2: Database of PD patterns during to pollution performance studies in ceramic insulators

Category of PD	Type of PD	**Applied Voltage (kV)**	**No. of Patterns**
1	PD (wetted without Dry band arcing)	6.1	40
2	PD during Dry band (at pin end)	2.8 3.6	40
3	PD during Dry band (at cap end)	6.4 10.7	80
4	PD during Dry band (at pin and cap)	12.4 16.3	40

3.3 Feature Extraction and pre-processing of PD Patterns

The raw data obtained from the PD measurement and acquisition system is in the form of φ-q-n characteristics which describes the PD source pattern. Though several researchers have utilized either the phase resolved or the time resolved representation of PD patterns for diagnosis of insulation, this research resorts to the former approach since it augurs well for pattern recognition. This aspect has also been reiterated by several researchers carrying out studies pertaining to pattern recognition since it has been inferred that each discharge pulse reflects the physical process at the discharge site and a strong relationship has been substantiated between the type of patterns and the type of defect.

The role of pre-processing is to ensure compactness of the input without compromising the loss of unique features. In this research the data describing the source of PD is characterized based on phase window technique into three categories: 1. measures based on maximum/ minimum values of apparent charge and number of discharge pulses, 2. measures based on types of mean measures and 3. measures based on mean-slope-angle. The objective of using various methods of pre-processing is to ascertain the performance of the proposed hybrid system in recognizing and classifying the patterns so that tangible decisions may be taken on the role played by the various key parameters of the HMM-PNN such as smoothing parameter, influence of curse of dimensionality, change of state sequence that represents the dynamic behaviour of the patterns etc.

4 OBSERVATIONS AND ANALYSIS

4.1 Classification Capability of Hybrid CDHMM –PNN

It is evident from the tests as stated in 3.2.1 and as mentioned in Table 1 that for all the types of pre-processing and feature extraction techniques taken up for study, the hybrid classifier resulted in excellent classification rate. Table 3 and Table 4 show the performance of the hybrid HMM-PNN in discriminating the multiple source PD patterns and dynamically varying patterns pertaining to pollution led flashovers in insulator samples.

Table 3: Performance of Hybrid CDHMM-PNN for Multiple Source PD Laboratory Models

Feature Extraction Scheme	Phase Window	No. of Tuples	Training Patterns	Classification (%)
Measures based on Maximum Value	Φ-q_{max}-n (10º)	36	108	96
Measures based on Minimum Value	Φ-q_{min}-n (10º)	36	108	94.2
Measures based on Types of Mean	AM-GM-HM-RM (10º)	36	108	96.2
Measures based on Mean-Slope-Angle	Mean-Slope-Angle (10º)	36	108	97.44

It is also observed that the measures based on mean-slope-angle had better classification rate as compared to the other statistical measures. It is also pertinent to note that only the simple statistical operators have been utilized as a pre-processing method in this research while more advanced measures based on cross-correlation, auto-correlation, statistical moments etc could provide enhanced classification rate.

Table 4: Performance of hybrid CDHMM-PNN for pollution performance studies in ceramic insulators

Feature Extraction Scheme	**Phase Window**	**No. of Tuples**	**Training Patterns**	**Classification (%)**
Measures based on Maximum Value	Φ-q_{max}-n (10º)	36	360	91.67
Measures based on Mean-Slope-Angle	Mean-slope-angle (10º)	36	360	94.4

4.2 Dynamic Nature of Patterns during the training phase and its relevance

Though it has been clarified by researchers utilizing HMM that the state transition representation of the pattern do not directly provide inferences pertaining to the change in the physical phenomena of the system under study, yet it is worth mentioning that the changes in the state labels (state transition) reflect the change in the dynamics of the system in turn relating to the plausible change in the physics governing the nature of such dynamics. This aspect is evident in both the studies. Table 5 shows this aspect.

Table 5: Optimized State Transition Labels

PD Source	Applied Voltage (kV)	Optimized State Sequence										
PD (wetted without Dry band arcing)	6.1	S.L.	2	3	4	5	4	3	2	1		
		N.R.	3	2	4	12	2	3	3	7		
PD during Dry band (at pin end)	2.8	S.L.	1-2	2-3	3-4	4-5	5	4-5	5	3-4	2	1
		N.R.	3	1	3	4	7	3	2	2	2	7
	3.6	S.L.	1	1-2	3	4	5	4	3	2-3	1-2	
		N.R.	1	2	2	4	12	2	3	3	8	
PD during Dry band (at cap end)	20.93	S.L.	2-1	3-2	2-1	3-2	4-5	5	4-5	3-2	2	1-2
		N.R.	2	3	3	2	2	8	1	8	1	5
	29.12	S.L.	2-1	3	4	5	4	4-3	2-1			
		N.R.	3	1	5	12	3	4	8			
PD during Dry band (at pin and cap)	9.1	S.L.	2	3	4	5	4	3	2	1		
		N.R.	4	1	6	10	3	3	2	7		
	9.6	S.L.	2	3	4	5	4	3	2	1		
		N.R.	3	3	4	12	2	4	2	6		

It is perceptible from Table 5 that the nature of PD patterns though vary considerably during the dry-band formation, are rapidly transformed to closely resemble patterns that is prior to the advent of such scintillation led discharges. Such subtle changes in states are perceived in such discharges in the non-stationary CDHMM scheme. However, since the changes in the multiple source PD pattern studies are conducted for observing the changes in the patterns pertaining to one type of defect rather than a source of PD that transitions from one type to another, it is evident that the state transitions are distinctly different.

4.3 Dimensionality reduction due optimal state and impact on classification capability

Since the optimal states have been obtained for a phase window of 10º a total of 36 tuple vectors are obtained as the optimal state density estimates that describes a pattern pertaining to a PD source as against 108 tuple vector that is obtained from pre-processing methods. This reduction in the dimensionality of the classifier provides a mechanism to address the major issue pertaining to the ‘curse of dimensionality’ usually encountered in NN architectures.

It is observed that this reduction in the dimensionality of the input feature vectors, in fact, does not lead to misclassification and on the contrary represents the appropriate changes that reflect the features of the system dynamics.

4.4 Number of Iterations and States in Training a CDHMM

It is also worth mentioning from Table 5 that the number of iterations for obtaining the optimal state estimates is reasonably low varying from 10- 20. This aspect of rapid training augurs well with regard to utilizing the non-stationary CDHMM classifier since research taken up earlier is based on only stationary DDHMMs.

It is also observed that the proposed HMM classifier requires 5 states to model and describe the transition probabilities. This aspect of having about 4 or 5 states have been observed and utilized by researchers [2-3] who have utilized HMM for the PD pattern recognition task.

5 CONCLUSIONS AND DISCUSSIONS

It is evident from the aforesaid studies that for divergent types of pre-processing and feature extraction techniques taken up for study, the hybrid non-stationary CDHMM-PNN classifier provides an excellent recognition and classification mechanism. Further, the non-stationary representation provides an excellent framework for plausible means to understand the dynamics of the physics of discharge mechanism in real time applications. This aspect is made evident in this study as shown in Table 5. However the following issues are also worth mentioning: 1. Enhanced PNN versions such as Adaptive Time Varying PNN, Recursive PNN, Stofoscedastic PNN etc may be utilized in the classification phase for superior classification capability. Research is being carried out by the authors in this aspect which is presently ongoing. 2. Training real time PD patterns involve large datasets which may hence necessitate appropriate center-selection procedures such as Orthogonal Least Square (OLS) algorithm, Recursive Orthogonal Least Square (ROLS) algorithm etc.

5 ACKNOWLEDGEMENTS

This research was supported by the Research and Modernization Fund (RMF) Grant, constituted by SASTRA University. The first author is extremely grateful to Prof. R. Sethuraman, Vice-Chancellor, SASTRA University, Dr. S. Swaminathan, Dean-Sponsored Research and Dr. S. Vaidhyasubramaniam, Dean- P & D., SASTRA University for awarding the grant and for the unstinted support and motivation extended to him during the course of the project.

6 REFERENCES

[1] N. C. Sahoo, et al: "Trends in partial discharge pattern classification: A survey", IEEE Trans. on Dielectrics & Electrical Insulation, Vol.12, No.2, pp. 248–264, 2005

[2] L. Satish, B. I. Gururaj: "Use of Hidden Markov models for partial discharges pattern classification", IEEE Trans. on Electrical Insulation, Vol. 28, No. 2, pp. 178–182, 1993

[3] T.K. Abdel-Galil, et al: "PD pulse pattern recognition using Hidden Markov models", IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 11, No. 4, 715–723, 2004

[4] Yaman Barlas and Korhan Kanar: "A Dynamic Pattern-oriented Test for Model Validation", Proc. 4th Systems Science European Congress, Valencia, Spain, pp. 269-286, 1999

[5] Piotr Wilinski, et al: "Toward the Border between Neural and Markovian Paradigms", IEEE Trans. on Systems, Man, and Cybernetics, Vol. 28, No. 2, 1998

[6] D.F. Specht; "Probabilistic neural networks for classification, mapping or associative memory", IEEE International Conference on Neural Networks, No.1, Vol.1, pp.525–532, 1988

[7] S.Venkatesh, S.Gopal: "Robust Heteroscedastic Probabilistic Neural Network for multiple source partial discharge pattern recognition- Significance of outliers on classification capability", International Journal on Expert Systems with Applications, 38, 11501- 11514, 2011

[8] CIGRE Working Group Report 21.03:"Recognition of discharges", Electra No. 11, pp. 61–98, 1969