IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE)
e-ISSN: 2278-1676,p-ISSN: 2320-3331, Volume 11, Issue 5 Ver. III (Sep - Oct 2016), PP 74-79
www.iosrjournals.org

An Improved SOM ForPD Pattern Recognition Using Density Evaluated Rival Penalized Competitive Learning Algorithm

Rubén Jaramillo-Vacio[1], Carlos Alberto Ochoa-Zezzatti[2], Vicente García-Jiménez[2], Fernando Figueroa-Godoy[3]
[1]*CFE-LAPEM*
[2]*University of Juarez*
[3]*ITESI*

Abstract :*In this paper a novel applicationof improving rival penalized competitive learning to solve the problem of pattern recognition of partial discharges (PD), the algorithm is robust to environmental noise and highly efficient separation and classification of partial discharge sources. The main contribution of this paper is a novel internal structure for Self-Organizing Map (SOM) obtaining an effective classification. To demonstrate the effectiveness of SOM proposed were performed cases study.*
Keywords: *Partial Discharge, Self-Organizing Maps, DERCPL*

I. Introduction

The term "Partial discharge" (PD) is defined by IEC 60270 as, 'a localized electrical discharge that only partially bridges the insulation between conductors and may or may not occur adjacent to a conductor' [1]. Today, the insulation systems can consist of combinations of solid, liquid or gaseous material. PD may happen in a cavity, in a solid insulation material, on a surface or around an edge subject a high voltage. It is widely recognized that the analysis of PD signals can be used to monitor the health of dielectric materials. The ability to accurately separate and recognize PD sources is a fundamental stage in the development of a knowledge -based system.

With the development of increasingly efficient for analysing large amounts of data [2], nowadays computer algorithms, the research are focused to select different attributes to classify and identify the sources of PD to evaluate quality on insulation systems [3]. In [4] and [5] were used multilayer perception for recognition between different cylindrical cavities, the constraint was the recognition of different sources in the same sample. Kim et al [6] made the comparison between Back Propagation Neural Network (BPNN) and Fuzzy-Neural Networks (FNN), but is necessary to improve in the multiple discharges location. Support Vector Machine (SVP) was applying for location of PD in GIS, on site application should improve performance [7]. In [8] the SOM was use for PD pattern recognition and classification without quality measurement and optimization of the structure. Jun[9] worked using Wavelet Packet Transform for detection of PD in GIS. Abdel-Galil[10] showed that Fuzzy Decision tree applied for PD classification, however, the rules are sensitive to small perturbations in the dataset.

Another line of research in fault diagnosis is using Knowledge -based System**(**KBS), the analysis presented by Calvo-Rolle et al [11] shows generalities and representations; the most important PD recognition results are presented by Rudd et al in [12], apply rules of association between data patterns and defects. In [13] the combination of algorithms in order to increase classification efficiency and improve the display of the evaluation is present. In [14] a study of the most important metrics for classification is performing using an improved SOM. The contribution of this work is KBS algorithms to potentiate dimensionality reduction and pattern recognition of improvedSOM with the efficiency of the evaluation of probabilistic neural network (PNN) to resolve the problem of partial discharge analysis.

II. Implementation

This paper proposes a novel SOM for clustering implemented on (Fig. 1). The following stages of the methodology applied in the system, which later two case studies are presented arise.DP's measurement is a nondestructive testing and can be used to condition assess of the insulation from manufacturing to commissioning. This metric is undoubtedly the most efficient way to determine the state of insulation in a laboratory for quality control of electrical products, where the levels of electromagnetic interference are controlled through external filters, etc, the backgroun noise values ranging between 0 and 2 picoCoulomb (pC). However, to PD measurement on site (transmission and distribution substations, generation facilities, equipment production factories, etc.), levels of electromagnetic interference are unquantifiable (Fig 2), there are further parameters that affect the interpretation, as the cavity size, moisture, incepton and extintion voltage, etc, this produces change and inestability of dataset adquired.

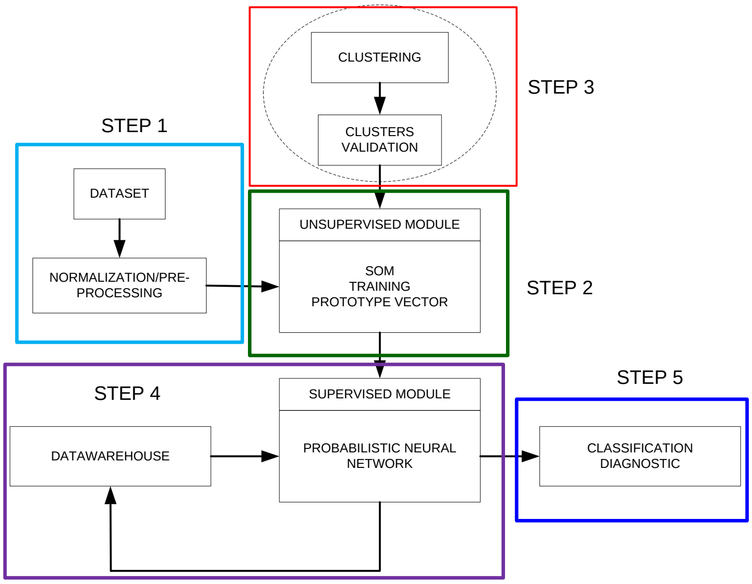
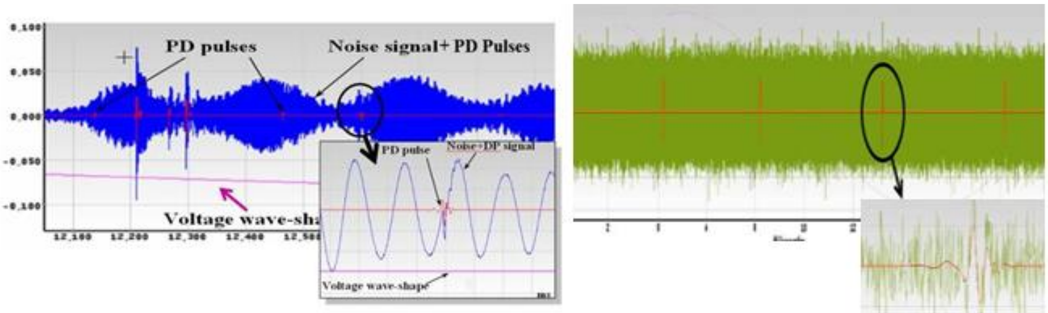


Figure 1: Conceptual diagram of the KBS



A B

Figure 2: a) PD pulses mixed with sinusoidal noise, b) PD pulses mixed with stochastic noise.

In figure 2 [15] is showed a typical dataset of PD, many signals are hidden behind background noise originate by power electronic, cellphones, corona discharge of facility, in others [15]. Small PD signals caused by an internal defect whose amplitude is several times smaller than background noise signal cannot be discriminate using conventional methods, in other words, the PD dataset aren´t linearly separable. PD patterns made up of statistical operators describing several PD distributions, known as fingerprint, were thus derive and use as the input data for PD patterns recognition purpose. For classifying the PD patterns associated with different defects, the main challenge is to decouple all source of both PD and background noise.

III. Density Evaluated Rival Penalized Competitive Learning

Clustering algorithms like k-Means is a classical competitive learning algorithm, which trains *k* seed points in a way that they converge to the data cluster centers by minimizing the mean-square-error (MSE) function. Although it has been widely used, it has two major challenges: (a) The number of data clusters must be previously known. (b) It suffers from the dead-unit problem. An improved clustering algorithm named rival penalized competitive learning (RPCL) is proposed. It pushes extra units away from the data set, while performing clustering with unknown exact cluster number. Also, it solves the dead unit problem. In this paper is applied an algorithm proposed in [16] called density evaluated RCPL (DERCPL) and is based on the principles of correct convergence.

The algorithm DERCPL is described as follows.

Randomly take an input x_μ from $X = \{x_\mu\}_{\mu=1}^{N}$ and let

$$u_i = \begin{cases} 1, if i = w, w = \arg\min_j \gamma_j \left\|x_\mu - m_j\right\|^2 \\ -1, if i = r, r = \arg\min_{j \neq w} \gamma_j \left\|x_\mu - m_j\right\|^2 \\ 0, otherwise \end{cases} \quad (1)$$

Where $\gamma_j = n_j / \sum_{i=1}^{k} n_i$ is designed to solve the dead unit problem and n_i is cumulative number of the occurrence of $u_i = 1$ in the past.

Update the winner unit m_w and its rival unit m_r by

$$m_i = m_i + u_i \alpha_i (x_\mu - m_i), i = m, r \ (2)$$

Where $0 \leq \alpha_r \leq \alpha_w \leq 1$.

The desired winner set is defined by

$$W_d = \left\{m_i^{(s)} \middle| m_i^{(s)} \in \Re, \forall i\right\} \text{n=s} (3)$$

The loser set as

$$L^{(s')} = \left\{m_i^{(s')} \middle| m_i^{(s')} \in \Re, \forall i\right\} s' \ll s \quad (4)$$

Randomly choose $C^{(n)}$ inputs from the dataset X an set

$$C^{(n)} = D \times k^{(n)} (5)$$

Compute de data density $D_i^{(n)}$ by

$$D_i^{(n)} = \sum_{\mu=1}^{C^{(n)}} n_{i\mu}, \ \ i = 1, 2, \ldots, k^n (6)$$

With

$$n_{i\mu} = \begin{cases} 1, if i = \arg\min_j \left\|x_\mu - m_j^{(n)}\right\|^2 \\ 0, othewise \end{cases} \quad (7)$$
$$m = 1,2, \ldots, C^{(n)}$$

Compute the value N_r, that will influence the decision of the numbers of the redundant units, by

$$N_R = \sum_{i=1}^{k^{(n)}} r_i (8)$$

With

$$r_i = \begin{cases} 1, if \delta_U < \frac{D - D_i^{(n)}}{D} \\ \frac{1}{k^{(n)}}, if \delta_L < \frac{D - D_i^{(n)}}{D} \leq \delta_U \\ o, otherwise \end{cases} (9)$$
$$i = 1,2, \ldots, k^{(n)}$$

Where δ_U and δ_L are pre-assigned parameters between 0.6 and 0.2.

With this DERPCL has four scenarios to update $M^{(n)}$ and $k^{(n)}$:

1. If $N_R \geq 1$ then $M^{(n+1)} = M^{(n)} - L^{(n)}, k^{(n+1)} = k^{(n)} - [N^R]$and goto equations (1) and (2) with $L^{(n)} = \left\{m_i^{(n)} \middle| m_i^{(n)} with r_i = 1, \forall i\right\}$.
2. If $0 < N_R < 1$and $t < \bar{t}$, then $t = t + 1$ and go to equation (5), where $\bar{t}$ is a positive integer.
3. If $0 < N_R < 1$and $t = \bar{t}$, then$M^{(n+1)} = M^{(n)} - L^{(n)}, k^{(n+1)} = k^{(n)} - 1$ and go to equations (1) and (2) with $L^{(n)} = \left\{m_i^{(n)} \middle| 1 = \arg\max_{\forall j} \left(D - D_j^{(n)}\right)\right\}$.
4. If $N_R = 0$, then the algorithm terminates.

IV. Experimental results

PD measurements for training are generated and recorded through laboratory tests. Corona was produced with a point to hemisphere configuration: needle at high voltage and hemispherical cup at ground. Surface discharge XLPE cable with no stress relief termination applied to the two ends. High voltage was applied to the cable inner conductor and the cable sheath was grounded, this produces discharges along the outer insulation surface at the cable ends. Internal discharge was used a power cable with a fault due to cavity in a joint. The dataset was used for training of the KBS and start the data warehouse. Figure 7 shows the analysis of internal discharge dataset using the KBS.

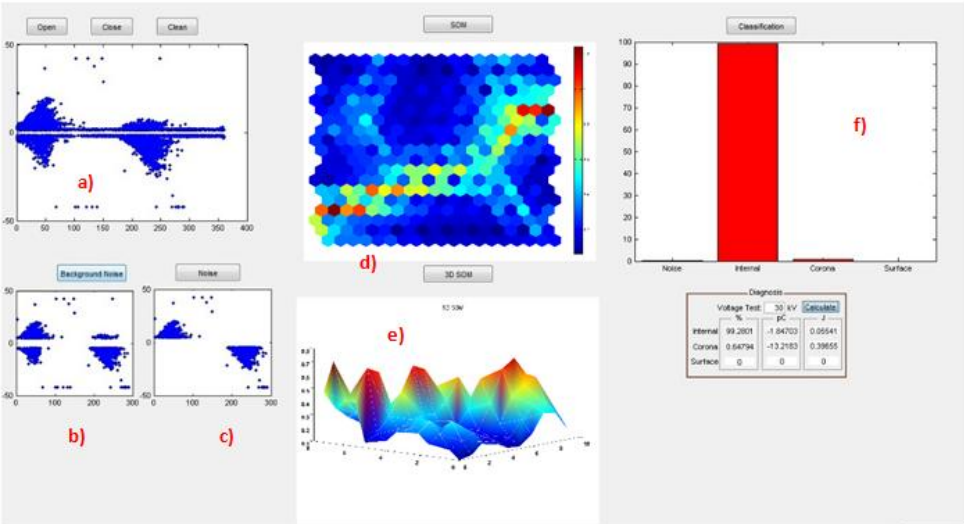


Figure 7: Output of KBS for the analysis of internal discharge dataset a) Original dataset, b) Background noise filtered, c) Eliminate inconsistent data, d) Representation of clustering using U-matrix, e) Visualization of clustering using 3D-SOM, f) Separation and classification of data according a partial discharge kind.

A. Case Study: Accessories for underground networks

The experiments were developed in accessories for underground networks like T connectors, elbows, caps, and the other, that require very high sensitivity, because the criterion evaluation according to IEEE Std. 404-2006 is that the system must not exceed 3 pC, higher levels of internal or surface PD implies a gradual failure. If the background noise levels are greater, or there corona measurement, also imply a wrong diagnosis is given. In figure 8 and 9 the effectiveness of the classification is shown when the dataset includes several PD sources.

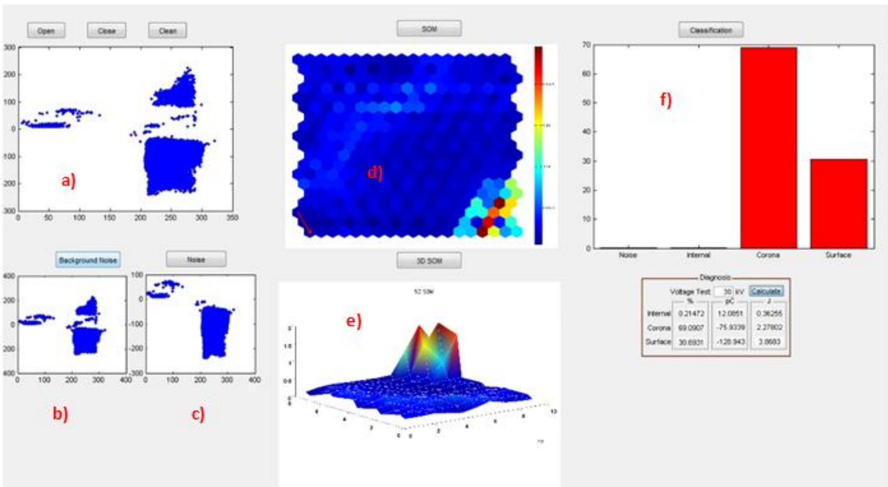


Figure 8: Output of KBS for the analysis of power cable measurement dataset a) Original dataset, b) Background noise filtered, c) Eliminate inconsistent data, d) Representation of clustering using U-matrix, e) Visualization of clustering using 3D-SOM, f) Separation and classification of data according a partial discharge kind.

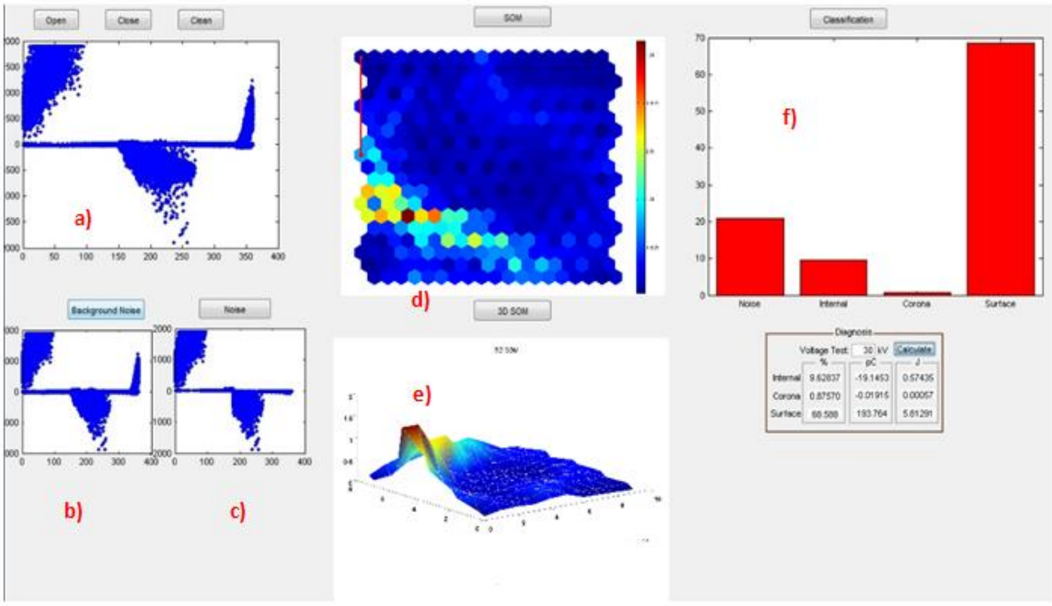


Figure 9: Output of KBS for the analysis of power cable measurement dataset a) Original dataset, b) Background noise filtered, c) Eliminate inconsistent data, d) Representation of clustering using U-matrix, e) Visualization of clustering using 3D-SOM, f) Separation and classification of data according a partial discharge kind.

B. Case Study: Test of error-containing dataset

In this experimental validation of KBS, is obtained a set of training that contains 200 training instances, and the testing data set is equal to the training data set, containing 200 training instances. The input data of a PD recognition algorithm would unavoidably contain some noise and uncertainties. The sources of error include high background noise, transducers and sensors, connection mistakes, etc., which could lead to data uncertainties. To take into account the noise and uncertainties, 200 datasets for testing data were created by adding 5% to 30% of random, uniformly distributed, error to the training data to evaluate the fault-tolerant abilities of the KBS proposed. The evaluation was compared with independent systems of SOM using winner takes all, rival penalized and DERPCL algorithms.

Table 6: Comparative evaluation of different competitive learning algorithms

Error Percentage (%)	**Accuracy of different algorithms**		
	SOM (WTA)	**SOM (RPCL)**	**SOM (DERPCL)**
± 0%	100%	100%	100%
± 5%	99%	100%	100%
± 10%	95%	100%	100%
± 15%	90%	98%	100%
± 20%	88%	95%	100%
± 25%	85%	90%	98%
± 30%	79%	90%	97%

This table shows that these methods have remarkable tolerance to the errors contained in the dataset. The KBS proposed has significantly higher recognition accuracy, but the accuracy of scheme III is lower than the other schemes. Contrarily, the accuracy of the SOM and BPNN algorithms are only 79% and 90% respectively under the same conditions.

V. Conclusion

This paper proposes a novel SOM used in knowledge based expert system for PD recognition. According experimental results, the KBS is adequate due to the higher accuracy and error tolerances. The KBS proposed consist of five stage evaluated for excellent results. The combination of SOM with PNN potentiates the performance of both algorithms. From the tested examples, the KBS proposed has a significantly high degree of recognition accuracy and shows good tolerance to errors added. This new knowledge based expert system merits more attention; we hope this paper will lead to a further investigation for industrial applications.

References

[1] Partial Discharge Measurements, IEC Publication 60270, 2000.

[2] Worapradya, K. Thanakijkasem, P. Optimising steel production schedules via a hierarchical genetic algorithm. *South African Journal of Industrial Engineering*, 25(2), 2014. pp 209-221.

[3] **Tozzi, M**. “Partial discharge in power distribution electrical system: Pulse propagation models and detection optimization” Director: Cavallini A. University of Bologna. Department of Electrical Engineering. Faculty of Engineering. 2010.

[4] **Mazroua, A**. PD pattern recognition with neural netwoks using the multilayer perception technique. *IEEE Transactions on Electrical Insulation, 28*. 1993. 1082-1089.

[5] **Krivda, A**. Automated Recognition of Partial Discharge. *IEEE Transactions on Dielectrics and Electrical Insulation*, 28. 1995. 796-821.

[6] **Kim, J. Choi, W. Oh, S. Park, K. Grzybowski, S**. Partial Discharge Pattern Recognition Using Fuzzy-Neural Networks (FNNs) Algorithm. I*EEE International Power Modulators and High Voltage Conference*. 2008. 272-275.

[7] **Khan, Y. Khan, A.A. Budiman, F.N. Beroual, A. Malik, N. H.** Al Arainy, A.A. Partial discharge pattern analysis using support vector machine to estimate size and position of metallic particle adhering to spacer in GIS. Electric Power Systems Research. Elsevier. 116. 2014. 391-398.

[8] **Chang, W. Yang, H**. Application of Self Organizing Map Approach to Partial Discharge Pattern Recognition of Cast-Resin Current Transformers. *WSEAS Transaction on Computer Research. Issue 3. Volume 3*. 2008.142-151.

[9] **Chang. C.S. Jun, J. Kumar, S. Su, Qi. Hoshino, T**. Denoising of Partial Discharge Signals in Wavelet Packets Domain. IEE Science, Measurement and Technology. Vol. 152. 2005. 129-140.

[10] **Abdel-Galil, T. K. Sharkawy, R. M. Salama, M. Bartnikas, R**. Partial discharge pattern classification using the fuzzy decision tree approach. IEEE Transaction on Instrumentation and Measurement. 54(6). 2005. 2258-2263.

[11] **Calvo-Rolle J, Ferreiro-Garcia R, Alonso-Alvarez A**. "Sistema basado en conocimiento para análisis de vibraciones en mantenimiento predictivo" DYNA. Vol. 83-8. 2008. 484-492.

[12] **Rudd S, McArthur S, Judd D**. A generic Knowledge-based Approach to the Analysis of Partial Discharge Data. IEEE Transactions on Dielectrics and Electrical Insulation. Vol. 17 Issue: 1. 2010. p. 149-156. ISSN: 1070-9878.

[13] **Jaramillo-Vacio R, Ochoa-Zezzatti A, Jöns S, Ledezma-Orozco S, Chira C**. Diagnosis of Partial Discharge Using Self Organizing Maps and Hierarchical Clustering – An Approach. 6th International Conference HAIS 2011. 2011. p. 91-98. ISSN: 0302-9743

[14] **Jaramillo-Vacio R, Ochoa-Zezzatti A, Rios-Lira A**. Optimal tuning parameters of Self Organizing Maps for classification of Partial Discharges using Response Surface Methodology. International Journal of Engineering Research and Applications. Vol. 3, Issue: 4. 2013. p. 974-984. ISSN: 2248-9622.

[15] **Garnacho, F. Sánchez-Urán, M.A. Ortego, J, Álvarez, F. Perpiñan, O**. New procedure to determine insulation condition of high voltage equipment by means of PD measurements in sevice. CIGRE 2012. D1-309. 01-10.

[16] **Chang, Y., Sheng-Sung Y. Chia-Lu H**. Improving rival penalized competitive learning using density-evaluated mechanism. *Journal of the Chinese Institute of Engineers*. Vol. 33, Nr. 7, 2010. pages 1083-1088.