International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-9 Issue-2, December, 2019

Performance of Clustering Techniques of Multiple Partial Discharge Sources in High Voltage Transformer Windings

N H Nik Ali, A Mohd Ariffin, P L Lewin

Abstract: *There are numerous of clustering techniques that have been exploited by researchers in many applications such in medical application, image processing application as well as in high voltage application. Clustering technique is an unsupervised learning algorithm used to identify group structure in a set of data that contain different characteristics. Nowadays, within the latest HV insulation system, there are more than one dielectric media, which contribute to multiple source of partial discharge (PD). Therefore, data identification for PD is significantly vital to discover the kinds of faults that inducing discharges in a HV insulation system. Nevertheless, it is critical that the methodology used for further investigation such as phase-resolved partial discharge (PRPD) analysis is capable of producing a sufficient separation between the clustered data. An experiment was performed to generate a pair of PD sources simultaneously within a winding of the HV transformer. The PD pulses were collected from two measuring points measured by two wideband radio frequency current transformers (RFCTs) at the bushing tap-point to earth (BT) and the neutral to earth-point (NE).The performance oft-Distributed Stochastic Neighbour Embedding (t-SNE), Principle Component Analysis (PCA) and time-frequency mapping based on sparsity roughness at distinguishing multiple PD sources is determined and presented.*

Keywords: *Partial Discharge; Clustering Techniques; High Voltage; Transformer.*

I.INTRODUCTION

Condition Monitoring (CM) of HV equipment, i.e. transformer is the procedure of gaining information from various parameters of the equipment with the aim to detect any failure at the early stage. Early precautionary measures can provide protection from defects when faults are in early stages. Identifying these incipient faults requires a comprehensive analysis and interpretation of their characteristics [1]. If the faults can be detected at the early stage, the transformer can be repaired accordingly before they lead to damages which could cause power failure and then can affect homes or on industry facilities. Hence, developing methodologies that can be used to detect the early signs of any transformer faults is very important in order to achieve the main objectives of CM of HV equipment such as reduce maintenance costs, reduce adverse impacts of maintenance activities, improve HV asset reliability as well as the their availability.

Revised Manuscript Received on December 30, 2019.
* Correspondence Author
N H Nik Ali*, Institute of Power Engineering, Universiti Tenaga Nasional, Jln IKRAM-UNITEN, 43000, Kajang, Selangor, Malaysia
A Mohd Ariffin, Institute of Power Engineering, Universiti Tenaga Nasional, Jln IKRAM-UNITEN, 43000, Kajang, Selangor, Malaysia
P L Lewin, Tony Davies High Voltage Laboratory, University of Southampton, Southampton, UK

Nowadays, there are several techniques for transformer CM for example based on dissolved gas analysis (DGA) [2,3], thermal analysis[4], PD analysis[5] and frequency response analysis (FRA)[6]. In addition, one of the well-known techniques of CM for evaluating the condition of power transformer is PD analysis. Within a power transformer, multiple PD sources can be activated simultaneously[7–9], therefore the capability to precisely classify different signals from different types of PD sources is required. By using any clustering technique, a data point that share similar attributes can be grouped into different clusters based on their similarities. Therefore, an experiment was carried out to produce various kinds of PD sources commonly activated within power transformers – floating and surface discharge. The PD pulses were extracted from two measurement points recorded by two wideband RFCTs positioned at the BT and the NE. The proposed clustering strategies rely on an initial assumption that different signal characteristics are produced by the signals produced by various PD sources. After extracting the PD pulses from the raw data, each single pulse was decomposed. This signal decomposition is very important as it allows PD signals to be examined one sample at a time. The performance of t-Distributed Stochastic Neighbour Embedding (t-SNE), Principle Component Analysis (PCA) and time-frequency mapping based on sparsity roughness at distinguishing between the PD sources is assessed and presented.

II.EXPERIMENTAL ARRANGEMENT

Two types of PD sources were artificially initiated under a 20 kV AC voltage – floating and surface discharge which were simultaneously injected into various locations along a HV transformer winding. According to Bewley [10], PD generates an electrical signal that will propagate to both ends of the transformer. Hence, in this investigation, the PD pulses are obtained from two measuring locations measured by two RFCTs positioned at the BT and the NE measurement points. The experiment setup is as shown as in Fig. and the details of the experiment can be found in [11].

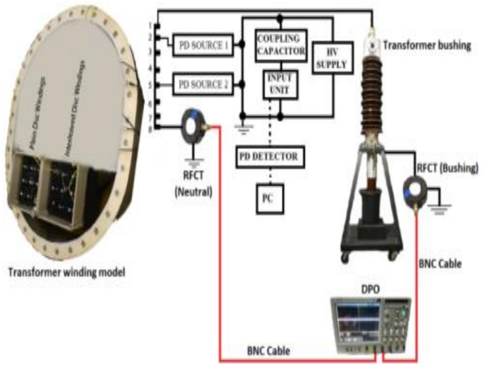


Fig. 1 Experiment Setup

III.SIGNAL PROCESSING

The signal decomposition technique is also necessary with the aim of determining the intrinsic components and to generate energy vectors related to the measured PD signals. Using the decomposition technique of Mathematical Morphology (MM), the measured PD signals either from both measurement points were decomposed. The MM decomposition technique is comparable to applying a sequence of low pass zero phase filters. As published in [12], the signal energy distribution for each SE employed in MM were implemented in this study to calculate energy difference at both measuring locations. The energy distribution can also be exploited to determine the energy different produced from different terminals of the transformer winding generated from different PD sources. After clustering technique was applied, the Density-based Spatial Clustering of Applications with Noise (DBSCAN) algorithm[7,13] which is a well-known data clustering algorithm that is commonly used in data mining and machine learning was implemented in this study to determine groups of signals with similar calculated energy distributions.

IV.CLUSTERING TECHNIQUES

Clustering is a method to identify data points that have similar features and/or properties and then groups them into specific groups called clusters. It is a technique of unsupervised learning algorithm and one of the most popular techniques for statistical data analysis which widely exploited in many applications. The use of clustering techniques in this paper is to identify, extract and visualize the important features of a large PD data set. In this investigation, three clustering techniques are used to determine specific features of the individual PD pulse that have similar characteristics which can be attributed to specific PD sources and their performance at distinguishing multiple PD sources are evaluated.

Principle Component Analysis

A number of linear transformation techniques, for example, linear discriminant analysis or factor analysis that have been employed in pattern identification to minimize the dimensions and feature extraction [14]. However, in this investigation, PCA was applied to reduce the number of dimensions as well as for visualization while minimizing the information lost in the data reduction [7]. PCA extracts the key features from a data set and characterize it as a set of new orthogonal variables known as principle components and visualize the variables as points in the principle component space [15].The PD data sets (M) considered are expressed as N vectors, each comprising dimension column of energy level, where N represents the number of measured PD signals. PCA solves the eigen problems that consist of:

- In order to provide zero mean and unity variance for each dimension, standardizing the data is done via subtraction of the mean value.
- Obtaining the covariance matrix C_m using:

$$C_m = (M.M^T)/(N-1) \qquad \textbf{(1)}$$

- Determination of the eigenvalues (λ) and eigenvectors (υ), which then sorted in descendant order.
- The matrix is transposed into a projection P_i via the eigenvectors defined as:

$$P_i = \upsilon^T \times M \qquad \textbf{(2)}$$

- The new representation of the data can be plotted in three dimensions using three lower order of the principle components.

Therefore, the measured PD signals can be visually presented and the performance in separating multiple PD sources can be verified.

t-Distributed Stochastic Neighbour Embedding

As for comparison purposes for minimizing the number of dimensions of the energy distribution data, t-SNE[16] is used as another dimension-reduction tool in this study. The main objective of t-SNE and also PCA is to keep the important information from high dimension data into lower dimensional space. In order to represent similarities, SNE converts the Euclidean distances between points in the high dimensional space into conditional probabilities $p_{j|i}$ which can be calculated by:

$$P_{(j|i)} = \frac{\exp\left(-\frac{\|x_i-x_j\|^2}{2\sigma_i^2}\right)}{\sum_{k\neq i}\exp\left(-\frac{\|x_i-x_j\|^2}{2\sigma_i^2}\right)} \qquad \textbf{(3)}$$

Where ∂_I represents the variance of the Gaussian distribution that is centered on data point x_i. For the low dimensional counter parts y_i and y_j of the high dimensional data point x_i and x_j, its conditional probability $q_{j|i}$ is calculated using:

$$q_{(j|i)} = \frac{\exp\left(-\|y_i-y_j\|^2\right)}{\sum_{k\neq i}\exp(-\|y_i-y_j\|^2)} \qquad \textbf{(4)}$$

Then t-SNE was introduced to enhance the algorithm of SNE by employing a different cost function [16].

Time-frequency Sparsity Roughness Mapping

In terms of multiple PD sources, TF mapping is commonly employed in the application of identification as well as classification. This mapping algorithm is depend on the hypothesis that different types of PD sources generate different shapes of PD signals[9, 17, 18]. A time reference for a PD pulse denoted by p with N samples is defined as:

$$t_o = \frac{\sum_{i=0}^{N} t_i p_i (t_i)^2}{\sum_{i=0}^{N} p_i (t_i)^2} \qquad \textbf{(5)}$$

Where t_i denotes the time instant of obtaining the i[th] sample. The equivalent time length of a PD pulse can be calculated using:

$$T^2 = \frac{\sum_{i=0}^{N} (t_i - t_o)^2 \, p_i (t_i)^2}{\sum_{i=0}^{N} p_i (t_i)^2} \qquad \textbf{(6)}$$

Then, the equivalent bandwidth of a PD pulse is described as:

$$F^2 = \frac{\sum_{i=0}^{\frac{N}{2}} {f_i}^2 \, |P_i(f_i)|^2}{\sum_{i=0}^{\frac{N}{2}} |P_i(f_i)|^2} \qquad \textbf{(7)}$$

where P represents the frequency value from Fourier transform while f_i denotes the frequency at the i[th] sample [19].

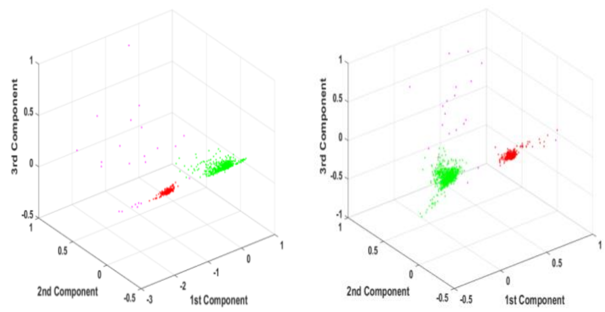
For each PD pulse, T^2 and F^2 were projected onto a two dimensional space to establish the TF map. The assumption was that, theoretically PD waveforms that have similar characteristics will form a specific group or cluster within the space. However, in some cases, PD pulses that have similar characteristic can also be dispersed into a few groups. Therefore, the TF sparsity roughness map is employed was introduced to solve the problems by considering the roughness and sparsity values within the signals that have been decomposed. The TF sparsity roughness map performance was reported to be better than a TF map in terms of PD pulses representation [19].

V.RESULT AND DISCUSSION

In order to minimize the dimensions of the processed data, PCA and t-SNE are individually used in the data sets produced using the decomposition of MM energy data. Whereas, the algorithms of time frequency were implemented to the data set derived from the decomposition of MM sparsity. The approaches mentioned are implemented for assessing their performance at discriminating multiple PD sources as well as for visualization. The assumption of this process is that PD waveforms generated from different PD sources produced instinctive features and form specific groups within their space. The performance of the proposed approaches in terms of sensitivity and suitability is assessed by applying each technique to the same data set. The generated surface and floating discharge which were both generated into terminal 5 along the transformer winding are used as the illustration of the performances of the proposed techniques. This is because of, for example, when the PD sources were injected into terminal 1 and terminal 5 of the winding, the differences within the energy of pulses between two expected clusters (two different PD sources) at two different measurement points are very significant. The energy at the BT measurement point had significantly greater than the energy of the source that was injected into terminal 5 due to propagation path taken which have different levels of dispersion and attenuation and vice versa at the NE measurement point.

Injection at terminals 1, 2, 3, 4, 6, 7 and 8 produce significant separations of the generated clusters except for terminal 5, as when both PD sources were injected into terminal 5, the propagation path taken by the pulses to reach both measurement points are almost identical. Therefore, only data generated when both PD sources were injected into terminal 5 are considered to assess the performance of clustering techniques at discriminate two different PD sources. The DBSCAN was then used to group the processed data automatically based on the density concept where the minimum number of data points used was set to two within the algorithm. This is to guarantee that it is feasible to group the least number of points that belong to the same PD source appropriately. The analysis of the performance of the techniques used have been done by visualization of the generated 3D plots by using PCA as well as the histogram pattern for each cluster within the space.

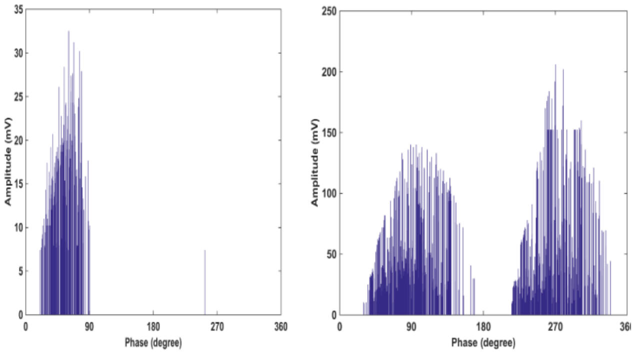
The lengths of the SE used in MM are altered from one to twenty to give twenty different levels of frequency bands of PD pulses[12]. For each SE length, the energy of the recorded signals is determined, producing twenty different energy element levels for each signal.PCA and t-SNE were then employed to the output of the MM algorithms to minimize the number of MM energy distribution levels from twenty to three element vectors for visualization. Fig shows the 3D plots generated by PCA and DBSCAN of the MM energy data set of terminal 5 for both measurement points respectively.



(a) BT (b) NE

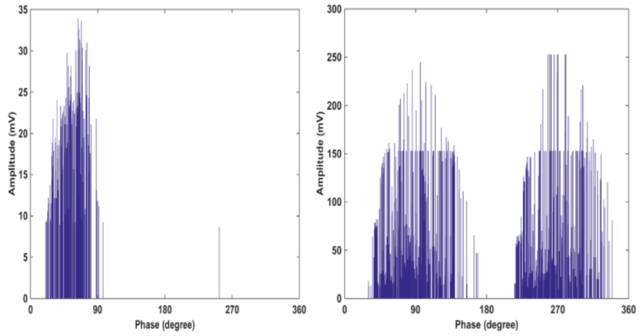
Fig. 2 MM energy data representation using PCA and DBSCAN at terminal 5

Fig illustrates that although there are outliers produced, DBSCAN is still capable to cluster the data into two expected specific groups. Fig and Fig show the histogram patterns correlated with the two obvious clusters in Fig for both measurement points respectively.



(a) Red cluster(b) Green cluster

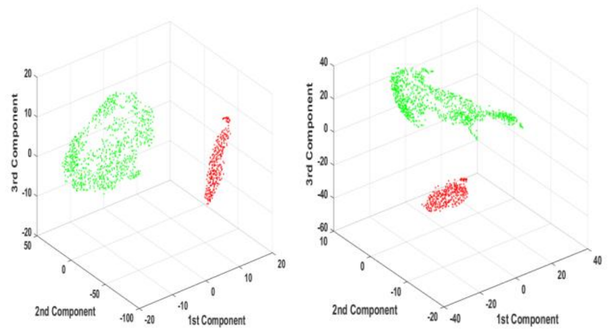
Fig. 3 Histogram pattern correlated with the two obvious clusters in Fig(a)



(a) Red cluster (b) Green cluster

Fig. 4 Histogram pattern correlated with the two obvious clusters in Fig(b)

Fig and Figillustrate that the data comprise of two different kinds of pulses, which stand for two different types of PD sources. While the 3D representation of MM energy data generated by t-SNE and DBSCAN algorithms is shown in Figure 5 on the same data set of PD pulses from both measuring points.

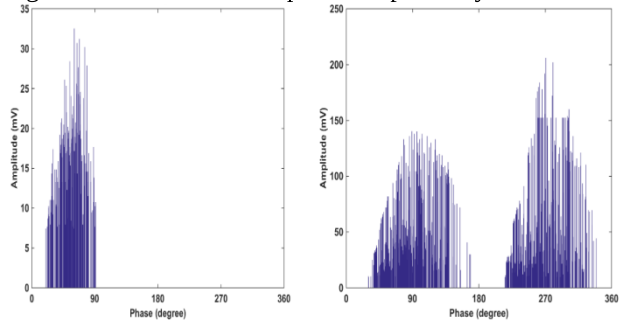


(a) BT (b) NE

Fig. 5 MM energy data representation using t-SNE and DBSCAN at terminal 5

The clusters created through the use of t-SNE with DBSCAN within the MM energy data set were divided into two main clusters by visual confirmation. The separation between clusters produced by t-SNE is very significant compared by using PCA, which is easier for DBSCAN to cluster them. Figand Fig show the histogram pattern correlated with the two obvious clusters in (a) **BT**

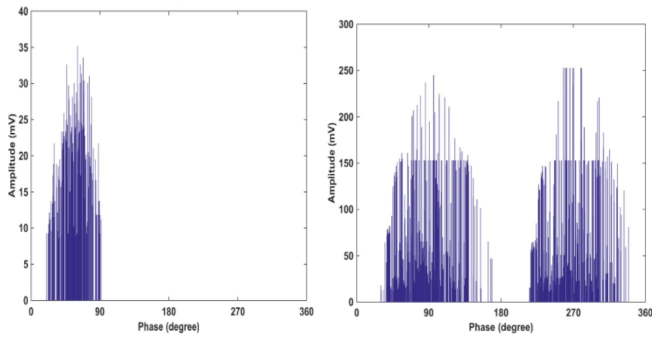
(b) NE Fig for both measurement points respectively.



(a) Red cluster (b) Green cluster

Fig. 6 Histogram pattern correlated with the two obvious clusters in (a) BT (b) NE

Fig(a)



(a) Red cluster (b) Green cluster

Fig. 7 Histogram pattern correlated with the two obvious clusters in (a) BT (b) NE

Fig(b)

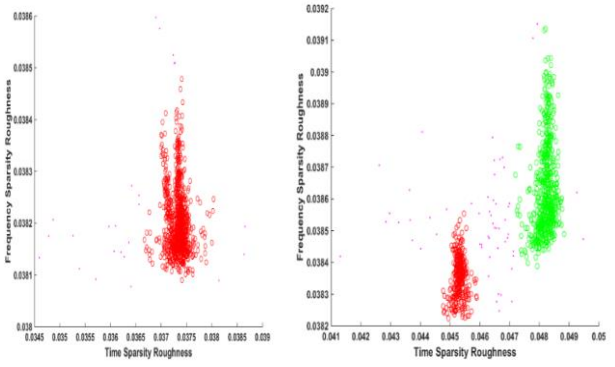
Two groups of cluster were identified by both PCA and t-SNE techniques with the application of DBSCAN. However, upon visual inspections of the clusters within the plots, t-SNE is seen to generate well defined boundaries or borders between the clusters compared to clusters that were produced by PCA. PCA generates scattered points in the space which lead to the production of outliers within the principle component space. This denotes that within a similar group of data there are some data correlated with a different characteristic that may affect the process of clustering. The t-SNE demonstrates that there are no misclassified or outliers pulses produced within its space. As mentioned earlier, the algorithms of time frequency were used as a dimension-reduction method for the data set generated by the decomposition of MM sparsity. To illustrate the variation of sparsity trends in time and frequency domains, roughness values (average absolute sparsity values) are calculated to represent the PD pulse from various types of PD sources [19]. Roughness, R that can represent by:

$$R = \frac{1}{L}\sum_{i=1}^{N} |S(i)| \qquad \textbf{(8)}$$

Retrieval Number: B4930129219/2019©BEIESP
DOI: 10.35940/ijeat.B4930.129219
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Where S represents the sparsity value for each pulse and L represents the number of length of SE. Fig shows the maps with application of DBSCAN for both measurement points.



(a) BT **(b) NE**

Fig. 8 Time-frequency sparsity roughness map using DBSCAN measured at terminal 7

In Fig(a), the expected two clusters for BT measurement point are seen to overlap each other within the TF space and DBSCAN considered the two clusters as a single cluster. This is due to the calculated values of the sparsity trends for each type of PD source are similar which are not enough to characterize the difference between them. Therefore, multiple PD sources are difficult to be recognized when they are overlapped to each other within the map.

VI.CONCLUSIONS

An experiment was developed to generate various kinds of PD sources commonly activated within power transformers and the performance of PCA, t-SNE and time-frequency mapping based on sparsity roughness at distinguishing between the PD sources are assessed and presented in this study. Analysis from of the results with the application of DBSCAN on the MM energy data sets via PCA show it is hard to gain confidence with the technique. Produced scattered data points and outliers are the disadvantage of the PCA, which lead PCA to ineffective and less robust in comparison to t-SNE. The clusters produced from t-SNE are linearly separated with a significant separation between them, which is easier for DBSCAN to group them into specific clusters. Even though t-SNE has computational drawback compared to PCA, it is seen to be acceptable given its superior performance.

REFERENCES

1. A. A. Alkahtani, F. H. Nordin, Z. A. Sharrif, and M. F. Rahmat, "Electrical Equipment Incipient Faults Simulation using Electromagnetic Field Emission," Int. J. Pure Appl. Math., vol. 119, no. 10, pp. 1213–1237, 2018.
2. "IEEE Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers," IEEE Std C57.104-2008 (Revision IEEE Std C57.104-1991), pp. 1–36, 2009.
3. J. I. Aizpurua et al., "Power Transformer Dissolved Gas Analysis through Bayesian Networks and Hypothesis Testing," IEEE Trans. Dielectr. Electr. Insul., vol. 25, no. 2, pp. 494–506, 2018.
4. D. Martin, J. Wijaya, N. Lelekakis, D. Susa, and N. Heyward, "Thermal analysis of two transformers filled with different oils," IEEE Electr. Insul. Mag., vol. 30, no. 1, pp. 39–45, 2014.
5. S. Ganeshan, J. Murugesan, A. Cavallini, F. Negri, B. Valecillos, and U. Piovan, "Identification of partial discharges in power transformers: An approach driven by practical experience," IEEE Electr. Insul. Mag., vol. 33, no. 5, pp. 23–31, 2017.
6. Bagheri, M. Lu, M. S. Naderi, and B. T. Phung, "Transformer frequency response: a new technique to analyze and distinguish the low-frequency band in the frequency response analysis spectrum," IEEE Electr. Insul. Mag., vol. 34, no. 5, pp. 39–49, 2018.
7. L. Hao et al., "Discrimination of Multiple PD Sources Using Wavelet Decomposition and Principal Component Analysis," IEEE Trans. Dielectr. Electr. Insul., vol. 18, no. 5, pp. 1702–1711, 2011.
8. A. Contin and S. Pastore, "Classification and separation of partial discharge signals by means of their auto-correlation function evaluation," IEEE Trans. Dielectr. Electr. Insul., vol. 16, no. 6, pp. 1609–1622, Dec. 2009.
9. A. Cavallini, G. C. Montanari, F. Puletti, and A. Contin, "A new methodology for the identification of PD in electrical apparatus: properties and applications," IEEE Trans. Dielectr. Electr. Insul., vol. 12, no. 2, pp. 203–215, Apr. 2005.
10. L. V. Bewley, "Traveling Waves on Transmission Systems," Trans. Am. Inst. Electr. Eng., vol. 50, no. 2, pp. 532–550, 1931.
11. N. H. Nik Ali, J. A. Hunter, P. Rapisarda, and P. L. Lewin, "Identification of Multiple Partial Discharge Sources in High Voltage Transformer Windings," IEEE Conf. Electr. Insul. Dielectr. Phenom., pp. 188–191, 2014.
12. N. H. Nik Ali, P. Rapisarda, and P. L. Lewin, "Multiple Partial Discharge Source Discrimination in a High Voltage Transformer Winding," in 13th International Electrical Insulation Conference (INSUCON), 2017.
13. M. Ester, H. P. Kriegel, J. Sander, and X. Xu, "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise," Int'ernational 2nd Conf. Knowl. Discov. Data Min., pp. 226–231, 1996.
14. T. Babnik, R. K. Aggarwal, and P. J. Moore, "Principal Component and Hierarchical Cluster Analyses as Applied to Transformer Partial Discharge Data With Particular Reference to Transformer Condition Monitoring," IEEE Trans. Power Deliv., vol. 23, no. 4, pp. 2008–2016, Oct. 2008.
15. H. Abdi and L. J. Williams, "Principal Component Analysis," WIREs Comp Stat, vol. 2, pp. 433–459, 2010.
16. L. Van Der Maaten and G. Hinton, "Visualizing Data using t-SNE," vol. 9, pp. 2579–2605, 2008.
17. A. Contin, A. Cavallini, G. C. Montanari, G. Pasini, and F. Puletti, "Digital detection and fuzzy classification of partial discharge signals," IEEE Trans. Dielectr. Electr. Insul., vol. 9, no. 3, pp. 335–348, Jun. 2002.
18. P. H. Morshuis, "Assessment of Dielectric Degradation by Ultrawide-band PD Detection," IEEE Trans. Dielectr. Electr. Insul., vol. 2, no. 5, 1995.
19. J. C. Chan, H. Ma, and T. K. Saha, "Time-frequency sparsity map on automatic partial discharge sources separation for power transformer condition assessment," IEEE Trans. Dielectr. Electr. Insul., vol. 22, no. 4, pp. 2271–2283, 2015.

AUTHORS PROFILE

N H Nik Ali, Institute of Power Engineering, Universiti Tenaga Nasional, Jln IKRAM-UNITEN, 43000, Kajang, Selangor, Malaysia

A Mohd Ariffin, Institute of Power Engineering, Universiti Tenaga Nasional, Jln IKRAM-UNITEN, 43000, Kajang, Selangor, Malaysia

P L Lewin, Tony Davies High Voltage Laboratory, University of Southampton, Southampton, UK