



Vaasan yliopisto
UNIVERSITY OF VAASA

OSUVA Open
Science

This is a self-archived – parallel published version of this article in the publication archive of the University of Vaasa. It might differ from the original.

Classification of PD Faults Using Features Extraction and K-Means Clustering Techniques

Author(s): Kumar, Haresh; Shafiq, Muhammad; Hussain, Ghulam Amjad; Kumpulainen, Lauri; Kauhaniemi, Kimmo

Title: Classification of PD Faults Using Features Extraction and K-Means Clustering Techniques

Year: 2020

Version: Accepted manuscript

Copyright ©2020 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Please cite the original version:

Kumar, H., Shafiq, M., Hussain, G. A., Kumpulainen, L. & Kauhaniemi, K. (2020). Classification of PD Faults Using Features Extraction and K-Means Clustering Techniques. *2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, 919-923.
<https://doi.org/10.1109/ISGT-Europe47291.2020.9248984>

Classification of PD Faults Using Features Extraction and K-Means Clustering Techniques

Haresh Kumar
School of Technology and Innovations
University of Vaasa
Vaasa, Finland
haresh.kumar@uniwaasa.fi

Muhammad Shafiq
School of Technology and Innovations
University of Vaasa
Vaasa, Finland
muhammad.shafiq@uniwaasa.fi

Ghulam Amjad Hussain
Dept. of Elect. & Comp. Eng.
American University of Kuwait
Salmiya, Kuwait
ghussain@auk.edu.kw

Lauri Kumpulainen
School of Technology and Innovations
University of Vaasa
Vaasa, Finland
lauri.kumpulainen@uniwaasa.fi

Kimmo Kauhaniemi
School of Technology and Innovations
University of Vaasa
Vaasa, Finland
kimmo.kauhaniemi@uniwaasa.fi

Abstract— Partial discharge (PD) diagnostic is a crucial tool for condition monitoring of power system equipment (e.g. switchgear, cable) in the medium voltage (MV) network, which is degraded by the gradual deterioration of insulation elements, ageing, and various operational and environmental stresses. In the MV network, different types of PD faults are generated from different sources and to know the impact of an individual PD fault on the health of MV equipment, classification plays an important role. This paper aims to provide suitable techniques for classifying PD faults. The data is collected from an experimental investigation of three different types of PD faults from MV switchgear and classified using features extraction, dimensionality reduction and clustering techniques. To identify the best classification technique, dimensionality reduction techniques (principal component analysis and t-distributed stochastic neighbour embedding) are used, and their results are compared using the confusion matrix after applying k-means clustering technique.

Keywords— partial discharge, medium voltage, classification, features extraction, dimensionality reduction techniques, k-means clustering

I. INTRODUCTION

Partial discharge (PD) is a localized electrical discharge that is caused by insulation degradation and indicates the weakness of insulation of power components and their accessories in the medium voltage (MV) network. The insulation degradation is caused by thermal, electrical, ambient/environmental, and mechanical (TEAM) stresses on power components, which leads to different kinds of faults.

The PD measurement is a practical approach for detecting insulation defects and assessing the insulation condition of MV equipment. To identify, sense and locate PDs within MV equipment, the condition-monitoring process is used. The condition monitoring process is the combination of four subsystems i-e; sensor, data acquisition, fault detection and diagnosis technique. The sensor is used to sense the physical phenomena of type of stress, while fault detection is the process of determining fault (occurrence of PDs in the signal) through the monitoring process. The data acquisition system collects PD

data with high sampling rate from the sensor and creates the link between measured data and analyses of measured data (diagnosis), and diagnosis part identifies the trends and patterns of measured data [1], [2].

In MV switchgear, PDs from multiple sources may occur simultaneously, which results in the mixture of PD signals and interpretation problems. To avoid the misinterpretation of signals, separation of defects is needed. The process of separation consists of features extraction and classification [3].

There is a variety of features extraction and classification techniques which are used in different papers for PD faults classification, such as; signal processing tools, statistical tools, image processing, principal component analysis (PCA), t-distributed stochastic neighbour embedding (t-SNE), k-means, support vector machine, decision tree and artificial neural network [4]. In this paper, different features representation and extraction techniques are discussed in the next section and based on the study of different techniques, discrete wavelet transform (DWT) combined with statistical parameters is used for features extraction from PD signals. For classification, an unsupervised machine technique (k-means) is utilized after reducing the dimensionality of extracted features.

An experimental setup was developed to generate three types of PD faults-corona, internal and surface discharge in MV switchgear. For classification of PD faults, an assumption is made that PD pulses produced by different PD faults present distinct waveform characteristics. Therefore, it can be deduced that different PD faults will show a unique pattern after classification. For identifying the unique pattern of individual PD fault, a novel algorithm is designed in this paper, which consists of signal processing based features combined with statistical parameters and machine learning techniques.

This paper is organized as follows. Section II illustrates the literature review in which different techniques of PD features representation are discussed with respect to their advantages and disadvantages. Section III presents a brief introduction about PD classification (DWT, statistical parameters, PCA, t-SNE and k-means clustering) process.

Section IV describes the laboratory experiment (data measurement) setup. The implementation of the algorithm and results are carried out in Section V, and the conclusion is presented in Section VI.

II. LITERATURE REVIEW

During the measurement process of PDs, it is hard to differentiate different types of PD faults with the naked eye on phase-resolved partial discharge analyzer. Because of that it is important to use more advanced techniques for the classification of PD faults. Moreover, for determining the harm of individual PD fault on the health of equipment or an asset, to diagnose the individual PD fault, and to predict the deterioration level of an equipment or asset, classification is necessary[5].

Features selection plays an important role in classification, and the choice of features (variables) is challenging for classifying different PD faults. For classification, a single feature (PD variable), such as; maximum applied voltage (V_{max}) or apparent current (I_a) can be taken into account to know the effect of PD on MV equipment. However, the author has observed less accurate results in [5] by applying a single feature for classifying PD faults. Considering this point, it needs more than one feature to get the accurate results of classification.

From the time domain signal of PDs, general features and pulse specific features can be obtained. The general features consist of pulse charge magnitude (for both positive and negative peaks), number of pulses, pulse repetition rate and average discharge current. On the other hand, pulse specific features consist of rising time, decay time and pulse width. From literature studies [6], it has been concluded that both types of features extraction methods are related to the pulse peak value that is affected by noise and by applying thresholding techniques, it leads to less accurate results of classification.

The classification accuracy for PD signals can be improved by frequency domain analysis. A time-domain signal can be converted into frequency domain either by fourier analysis or by fast fourier transform (FFT). The disadvantage of fourier analysis is that all the transient/temporal information is lost during the transformation process of PD signals from the time domain to the frequency domain, and this information is essential for analysing PD signals. The fourier analysis gives information about the occurrence of a particular event at a specific frequency, but it does not provide the information that when the event took place. To overcome this problem, STFT is preferred. As compared to the fourier series, STFT uses a window technique where a small section of the signal is taken into account for analysing purpose. The disadvantage of this method is that it has limited computational precision, which is dependent on the size of the window [7]. Moreover, it is hard to determine the occurrence of time localization of certain part of frequency when a time-domain signal is transformed into the frequency domain by STFT. In simple words, STFT is suitable for stationary signals and provides erroneous results for

time-varying signals. To overcome the limitations of FFT, short-time fourier transform (STFT) can be used. The advantage of STFT over FFT is that STFT uses a time-frequency window to localize transient in a signal, but it has a limitation of the fixed time-frequency window [8].

The wavelet transform is preferred to overcome these limitations of time and frequency domain signals. The advantage of the wavelet transform over frequency domain signals (fourier series, FFT, STFT) and the time-domain signal is that the wavelet transform has the capability of processing longer duration window/intervals which contain low-frequency information as well as it provides signal's transient behaviour by high-frequency information [7].

The choice of the wavelet transform, type of wavelet and decomposition of the signal into different levels is also a broad question. There are many wavelet types that are used for features generation of PD signals; Daubechies (dbN), Symlet (symN) and Coiflets (coifN) wavelet types are most preferred among them. These wavelet types are preferred because of their performance (compactness, orthogonality and asymmetry) for PD data analysis [9]. The choice of the wavelet transform, wavelet type and the number of decomposition levels is an iterative process, and it is dependent on the data that is carried out for analysis. The DWT is chosen in this paper for features generation of PD signals (that are taken into account for analysis) because of discrete nature of signals, having the capability of denoising and giving information of PD pulse into time and frequency domains.

The features extraction process is used to extract useful information from PD data with minimum loss of information. The features extraction is done for the simplicity of signal analysis instead of analysing complex data, which makes classification simpler. The statistical parameters are used in this paper to extract the distribution of PD energy from different features (wavelet coefficients) of PD data. According to research studies [6], statistical parameters can be used for different applications of features extraction irrespective of the domain (time domain, frequency domain, wavelet domain) of the signal. There are many statistical parameters, such as; mean, standard deviation, skewness, kurtosis, which are used for features extraction [5].

The dimensionality reduction techniques are widely used to make the classification simpler by reducing the dimensions of features. There are many dimensionality reduction techniques, and most common among them are PCA and t-SNE. These techniques are chosen in this paper to reduce the dimensions of features and to find the best separation technique between them for the classification of PD faults. PCA is mainly used for reducing the dimensions of features, which are linear in nature, while t-SNE is used for dimensionality reduction of non-linear or complex features. The disadvantage of PCA is that it performs poorly in case of non-linear data, and t-SNE's disadvantage is that it calculates the conditional probability for each data point which increases the computational burden and makes the process slower [10][11].

III. PD CLASSIFICATION

In this section, the features extraction, dimensionality reduction and clustering techniques are described, which are applied to the measured PD data. The details of the measurement setup for recording the PD data is provided in the IV section.

A. Discrete Wavelet Transform (DWT)

DWT is applied on discrete signals, and it can provide the information about PD signals in time and frequency domains within specific frequency ranges. The first step of DWT is to choose the wavelet type and the number of decomposition levels. In DWT, a PD signal is decomposed into different levels by a pair of complementary high pass and low pass filters into the series of detail (D) and approximate (A) coefficients [12]. This transformation is an iterative process in which approximation coefficients are used as the new input for the next decomposition level. At each level of the decomposition process, the bandwidth and length of the signal are kept half from its previous level, and the result of this decomposition is the scaled and shifted frequency components [13].

B. PCA

PCA determines how features are varying in a particular dimension (variance) and how multiple features are varying in a particular dimension (covariance). PCA works on the principle of orthogonal linear transformation and plots or projects the data with maximum variance in a particular direction [4], [11].

C. t-SNE

As compared to PCA, t-SNE is a non-linear technique which transforms the data from high dimensions to low dimensions by converting the euclidean distances between data points into conditional probabilities which shows similarities [11].

D. k-Means

The clustering technique is used to determine the similarity of how two or more data patterns are close to each other. For the grouping of PD data, it is assumed that different PD faults represent different pattern at different distances. Therefore, PD data can be separated based on the pattern of distance in the data sets, and clusters can be created by considering the separation distance between individual clusters. The k-means clustering technique separates the number of data points from the data sets of many dimensions into k-clusters (predefined clusters) based on euclidean distance as a similarity measure [14].

IV. Experimental Setup

The experimental data for PD classification was taken from the measurement setup, which was arranged in the High Voltage Laboratory of Aalto University. In this measurement setup, a commercial MV switchgear (20kV) was energized by keeping the three-phase circuit breaker in a closed position and defective insulations were attached to the outgoing terminals of circuit breaker. In this experiment, three types of PD faults (corona, PD in a void as the source of internal PDs, surface discharge due to base conductors)

were generated in the MV switchgear outgoing connection compartment. The test objects of these faults are shown in Fig. 1.

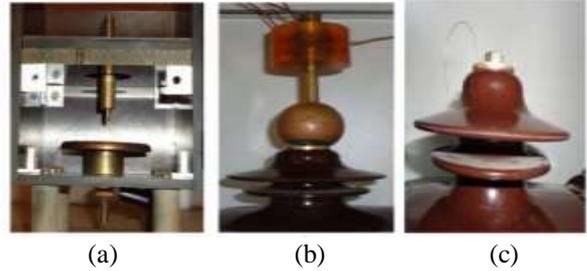


Fig. 1. PD sources: (a) corona, (b) internal, (c) surface discharge [15].

Two types of sensors; high-frequency current transformer (HFCT) and D-dot were used to record the PD signals. HFCT sensor measures the PD current pulses by electromagnetic induction, while the D-dot sensor uses an electric field to capture the PD signals. The HFCT sensor was connected around at the ground connection of switchgear, while D-dot sensor was fixed inside the upper portion of the switchgear compartment. For the recording of PD signals, a high-frequency oscilloscope with a sampling rate of 2.5 GS/s was used, and each PD pulse was captured at the length of 5 μ s [15].

V. IMPLEMENTATION OF THE ALGORITHM AND RESULTS

The proposed techniques are applied to PD data sets (114 PD signals of three types: 28 signals of the corona, 46 signals of internal and 40 of surface discharge) which were collected from the outgoing connection compartment of MV switchgear. Fig. 2 shows the flow chart of the implementation of the algorithm.

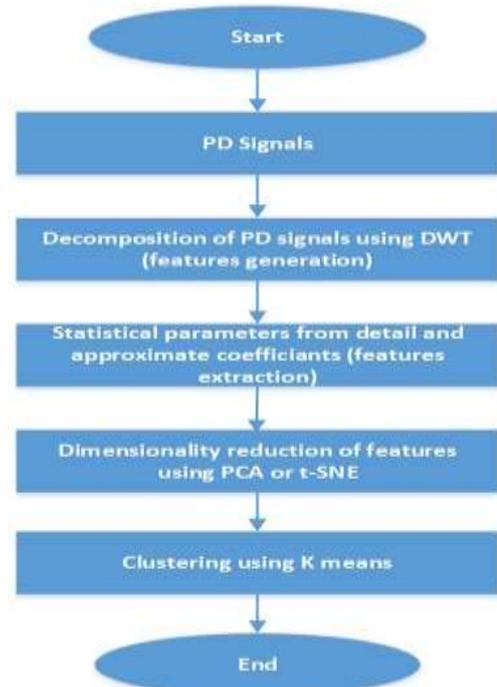


Fig. 2. Flow Chart of PD Signals Classification.

The brief description of implementation of the algorithm is as follows:

A. Features Generation and Extraction

For features generation of PD signals, mother wavelet (from DWT) type symlet of order seven is chosen for the experimental data, which has provided us interesting features (a five-dimensional vector) by decomposing the original signals until level four into series of detailed and approximate coefficients (CD1, CD2, CD3, CD4, CA4). An assumption is made that these selected coefficients represent the distinct parameters of PD signals from different discharge sources. Moreover, for determining the distribution of signal energy, symmetry and sharpness from the features of PD signals, these features or coefficients are further analysed (extracted) using statistical parameters (mean, standard deviation, skewness and kurtosis) which has provided us a vector of 20 dimensions.

B. Dimensionality reduction of features

Due to the high dimensionality of extracted features, it is not possible to visualize all features of different PD signals (PD faults) into 2D or 3D space. In addition to visualization, we need those features, which contain maximum energy or PD information for classification. For this purpose, PCA and t-SNE techniques are used in this paper to reduce the dimensions of extracted features and to visualize the pattern of PD data in 2D or 3D space. The results of PCA and t-SNE techniques are presented in Fig. 3.

Fig. 3 depicts the points in PCA and in t-SNE space that can be categorized into three different groups just by creating a separation line. From Fig. 3, it can be seen that t-SNE provides better results than PCA and three different groups of data points are clearly visible. After the reduction of dimensionality of features, the next step is to classify these features into different groups using clustering technique.

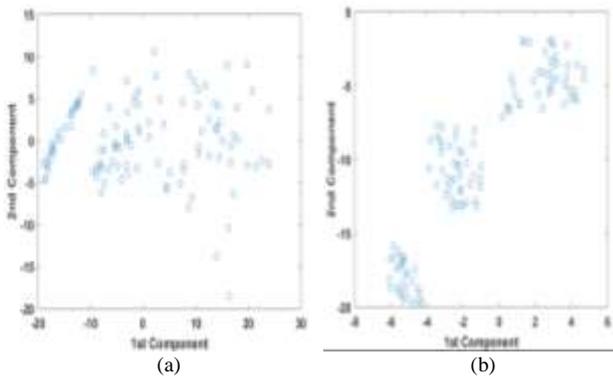


Fig. 3. Data visualization in 2-D (a) with PCA (b) with t-SNE.

C. Classification or Clustering of Features

The k-means clustering technique is chosen in this paper for PD data sets to classify them into different groups and to find the labels of each signal. The selection of assigning the number of clusters into the k-means clustering technique is challenging, and for that, different optimal solutions (silhouette method, elbow method) can be used for selecting the number of clusters. In our case, we have assigned three clusters in the k-means technique by already knowing that

PD data is taken from three different types of PD faults and by seeing the results of PCA and t-SNE techniques. The results of the k-means clustering technique are shown in Fig. 4.

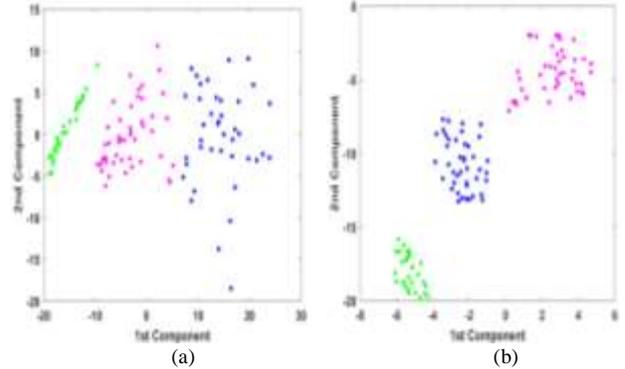


Fig. 4. Data visualization in 2D space (a) with PCA and k-means clustering (b) with t-SNE and k-means clustering.

After knowing the labels generated by the k-means clustering technique, we compared these labels with the labels of the original data sets (true labels v/s predicted labels) in order to evaluate the performance of classification techniques. This comparison was carried out for both dimensionality reduction techniques (PCA and t-SNE) after applying k-means clustering using the confusion matrix, and it was observed from the confusion matrix that t-SNE gives better classification results as compared to the PCA technique. Therefore, it can be concluded that t-SNE with k-means has provided us with more correct results in classification for PD data sets of this paper. The results of the confusion matrix are shown in Table 1 and Table 2.

TABLE I. CONFUSION MATRIX BETWEEN TRUE LABELS AND PREDICTED LABELS (PCA AND K-MEANS)

True Labels	Predicted Labels (K-Means Labels)		
	1	2	3
1	28	0	0
2	0	35	5
3	0	5	41

TABLE II. CONFUSION MATRIX BETWEEN TRUE LABELS AND PREDICTED LABELS (t-SNE AND K-MEANS)

True Labels	Predicted Labels (K-Means Labels)		
	1	2	3
1	28	0	0
2	0	39	1
3	0	5	41

It can be seen from the confusion matrix that both techniques (PCA and t-SNE) give accurate results in the classification of corona PD signals. At the same time, there is some misclassification between surface and internal PD signals. We can assume that there are some non-linear signals in surface and internal PD data sets, for that t-SNE performs better than PCA.

VI. CONCLUSION

This paper aimed to determine the suitable techniques for classifying three different types of PD faults. For this purpose, various features and classification techniques were studied, and it was found that signal processing based features combined with t-SNE and k-means techniques have provided good results in classification. Moreover, for knowing the classification results of PD faults and determining the best classification technique between t-SNE and PCA, a confusion matrix is utilized in this paper.

This paper has presented how three types of PD faults (created inside the MV switchgear compartment) can be utilized for testing the performance of various classification techniques. The classification results of PD faults may not provide information about the progression of PD faults, but changes in PD characteristics and severity of PD may be determined in future by creating a suitable prediction model.

REFERENCES

- [1] M. Shafiq, M. Lehtonen, G. A. Hussain, and L. Kütt, "Performance evaluation of PD monitoring technique integrated into medium voltage cable network for smart condition assessment," *9th Int. 2014 Electr. Power Qual. Supply Reliab. Conf. PQ 2014 - Proc.*, pp. 351–357, 2014, doi: 10.1109/PQ.2014.6866840.
- [2] M. Shafiq, G. A. Hussain, L. Kütt, and M. Lehtonen, "Electromagnetic sensing for predictive diagnostics of electrical insulation defects in MV power lines," *Meas. J. Int. Meas. Confed.*, vol. 73, pp. 480–493, 2015, doi: 10.1016/j.measurement.2015.05.040.
- [3] R. Liao, G. A. Taylor, K. Tavernier, and O. Khan, "Comparative study of feature extraction methods applied to partial discharge signals," *Proc. Univ. Power Eng. Conf.*, pp. 1–6, 2012, doi: 10.1109/UPEC.2012.6398585.
- [4] M. Wu, H. Cao, J. Cao, H. L. Nguyen, J. B. Gomes, and S. P. Krishnaswamy, "An overview of state-of-the-art partial discharge analysis techniques for condition monitoring," *IEEE Electr. Insul. Mag.*, vol. 31, no. 6, pp. 22–35, 2015, doi: 10.1109/MEI.2015.7303259.
- [5] Z. Ahmed, "Development of a continuous condition monitoring system based on probabilistic modelling of partial discharge data for polymer insulation cables," Mississippi University, 2017.
- [6] R. accurate Fernandez, "Accurate Classification of Partial Discharge Phenomena in Power Transformers in the Presence of Noise," Qatar University, 2017.
- [7] M. A. Alsaedi and M. M. Yaacob, "Partial Discharge Signal Analysis Using Wavelet Transform Technique: Review," *Int. J. Sci. Eng. Res.*, vol. 5, no. 2, pp. 553–555, 2014.
- [8] S. Khokhar, A. A. Mohd Zin, A. P. Memon, and A. S. Mokhtar, "A new optimal feature selection algorithm for classification of power quality disturbances using discrete wavelet transform and probabilistic neural network," *Meas. J. Int. Meas. Confed.*, vol. 95, pp. 246–259, 2017, doi: 10.1016/j.measurement.2016.10.013.
- [9] A. Hussain, Z. Ahmed, M. Shafiq, and M. Lehtonen, "Self-adaptive De-noising Technique based on DWT for PD Measurements and Self-healing Networks," *Int. Conf. Innov. Smart Grid Technol. ISGT Asia 2018*, pp. 1085–1090, 2018, doi: 10.1109/ISGT-Asia.2018.8467865.
- [10] H. D. Data, F. Extraction, A. Meyer-baese, and V. Schmid, "Dimensionality Reduction Data Preprocessing." 2014.
- [11] N. H. N. Ali, W. Goldsmith, J. A. Hunter, P. L. Lewin, and P. Rapisarda, "Comparison of clustering techniques of multiple partial discharge sources in high voltage transformer windings," *Proc. IEEE Int. Conf. Prop. Appl. Dielectr. Mater.*, vol. 2015-Octob, pp. 256–259, 2015, doi: 10.1109/ICPADM.2015.7295257.
- [12] 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, doi: 10.1109/TDEI.2011.6032842.
- [13] A. A. Bajwa, S. Habib, and M. Kamran, "Wavelet energy distribution with PCA & DBSCAN for partial discharge pulse extraction," *17th IEEE Int. Multi Top. Conf. Collab. Sustain. Dev. Technol. IEEE INMIC 2014 - Proc.*, pp. 422–427, 2014, doi: 10.1109/INMIC.2014.7097377.
- [14] X. Peng, C. Zhou, D. Hepburn, M. D. Judd, and W. H. Siew, "Application of K-Means method to pattern recognition in on-line cable partial discharge monitoring," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 20, no. 3, pp. 754–761, 2013, doi: 10.1109/TDEI.2013.6518945.
- [15] G. A. Hussain, D. Hummes, M. Shafiq, and M. Safdar, "Detection of multiple partial discharge faults in switchgear and power cables," *2019 IEEE Texas Power Energy Conf. TPEC 2019*, pp. 1–4, 2019, doi: 10.1109/TPEC.2019.8662173.