Expert Systems With Applications 55 (2016) 417–428

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Multiple partial discharge source discrimination with multiclass support vector machines

Guillermo Robles[a,*], Emilio Parrado-Hernández[b], Jorge Ardila-Rey[c], Juan Manuel Martínez-Tarifa[a]

[a] Department of Electrical Engineering, Universidad Carlos III de Madrid, Av. Universidad 30, 28911 Leganés, Madrid, Spain
[b] Department of Signal Processing and Communications, Universidad Carlos III de Madrid, Av. Universidad 30, 28911 Leganés, Madrid, Spain
[c] Department of Electrical Engineering, Universidad Técnica Federico Santa María, 8940000 Santiago de Chile, Chile

ARTICLE INFO

Keywords:
Support vector machine
Partial discharges
Electric maintenance
Machine learning
Condition monitoring
Risk assessment

ABSTRACT

The costs of decommissioning high-voltage equipment due to insulation breakdown are associated to the substitution of the asset and to the interruption of service. They can reach millions of dollars in new equipment purchases, fines and civil lawsuits, aggravated by the negative perception of the grid utility. Thus, condition based maintenance techniques are widely applied to have information about the status of the machine or power cable readily available. Partial discharge (PD) measurements are an important tool in the diagnosis of power systems equipment. The presence of PD can accelerate the local degradation of insulation systems and generate premature failures. Conventionally, PD classification is carried out using the phase resolved partial discharge (PRPD) pattern of pulses. The PRPD is a two dimensional representation of pulses that enables visual inspection but lacks discriminative power in common scenarios found in industrial environments, such as many simultaneous PD sources and low magnitude events that can be hidden below noise. The literature shows several works that complement PRPD with machine learning detectors (neural networks and support vector machines) and with more sophisticated signal representations, like statistics captured in several modalities, wavelets and other transforms, etc. These methods improve the classification accuracy but obscure the interpretation of the results. In this paper, the use of a support vector machine (SVM) operating on the power spectrum density of signals is proposed to identify different pulses what could be used in an online tool in the maintenance decision-making of the utility. Particularly, the approach is based on an SVM endowed with a special kernel that operates in the frequency domain. The SVM is previously trained with pulses of different PD types (internal, surface and corona) and noise that are obtained with several test objects in the laboratory. The experimental results demonstrate that this technique is highly effective in identifying PD for cases where several sources are active or when the noise level is high. Thus, the early identification of critical events with this approach during normal operation of the equipment will help in the decision of decommissioning the asset with reduced costs and low impact to the grid reliability.

1. Introduction

Electrical insulation is reported to be the weakest point of electrical assets in high voltage engineering (Kuffel, Zaengl, & Kuffel, 2000). The daily operation of power cables and electrical machines leads to high electrical, thermal and mechanical stresses which tend to degrade the insulation systems until a complete short-circuit between metallic electrodes takes place. In the previous stages before total breakdown, it is usual to observe low energy ionisation processes that occur within small volumes where highly divergent electrical fields are present, (Okabe, Ueta, Wada, & Okubo, 2010; Stone & Warren, 2004; Wang, Cavallini, Montanari, & Testa, 2012). These processes, called partial discharges (PD) are a consequence of different degradation mechanisms and may be useful for the diagnosis of electrical equipment at high rated voltages. However, not all PD may be equally harmful for dielectric materials, and their particular source needs to be identified if a proper analysis is intended to be made, (Stone & Warren, 2004). Under this assumption, intense work has been made to identify PD sources from the discharge magnitude represented superimposed to the applied voltage, (CIGRE, 1969; IEC, 2012). These so-called

* Corresponding author. Tel.: +34916245922.
E-mail addresses: grobles@ing.uc3m.es (G. Robles), eparrado@ing.uc3m.es (E. Parrado-Hernández), jorge.ardila@usm.cl (J. Ardila-Rey), jmmtarif@ing.uc3m.es (J.M. Martínez-Tarifa).

http://dx.doi.org/10.1016/j.eswa.2016.02.014
0957-4174/

phase resolved partial discharge (PRPD) patterns are very useful to characterise the status of electrical machines and power cables.

The interpretation of the results is sometimes complex if several PD sources are simultaneously active, (Venkatesh & Gopal, 2011b) or if the measurements are made in sites where low signal to noise ratio (SNR) is present, which is especially noticeable in on-line PD detection, (IEC, 2012; Montanari & Cavallini, 2013). In order to face these problems, high bandwidth detectors such as high frequency resistors, high frequency current transformers (HFCT) and inductive loops have been used following the guidelines of the standards measuring PRPD patterns (IEC, 2000, 2012) though, additionally, they can detect the discharge waveform for further signal processing, (Martínez-Tarifa, Robles, Rojas-Moreno, & Sanz-Feito, 2010; Montanari & Cavallini, 2013).

Thus, PD pulse waveform analysis can be used in the recognition of the sources of discharges and in the filtering of noise. Several techniques, such as time-frequency (T-F) maps, (Allahbakhshi & Akbari, 2011; Cavallini, Montanari, Contin, & Pulletti, 2003), power ratio (PR) maps (Ardila-Rey, Martínez-Tarifa, Robles, & Rojas-Moreno, 2013) and three-dimensions maps (Hao et al., 2011) permit to group pulses in clusters based on their source. A wider approach consists in processing these pulses through neural networks (NN) and machine learning techniques, (Chen, Gu, & Wang, 2012; Kuo, 2010; Venkatesh & Gopal, 2011a, 2011b). In particular, neural networks have been used in Majidi and Oskuoee (2015) to classify PD based on their apparent charge and phase referred to the voltage grid using 18 test objects to gather the events and test the classification.

Perhaps the closest work in the literature to our approach is (Hao, Lewin, & Swingler, 2008), in which the support vector machines (SVMs) are applied to a particular industrial environment where partial discharge signals are generated with a commercial calibrator. This work concluded that the SVM applied to the wavelet coefficients achieve almost a perfect classification, while the spectral response, implemented as the Fast Fourier Transform (FFT) of the pulses, yielded a very poor generalisation capability. Another interesting technique based on the study of the wavelet transform using SVMs was presented in de Oliveira Mota, da Rocha, de Moura Salles, and Vasconcelos (2011) though it has only been tested to separate noise from partial discharges and not to separate different types of discharges. SVMs have also been applied to PRPD patterns which can be used as inputs in the training process; (Hao & Lewin, 2010) applied a wavelet transform to reduce the information of the pulses to two parameters, the phase referred to the grid voltage and the average charge amplitude. Alternatively, the information of the type of discharge is determined using an SVM with an RBF kernel trained with individual PRPD patterns. The tests were also done individually and combined manually to have several simultaneous PD sources. They obtained good results in the identification of PD sources though, in a practical application of the method, the identification through these patterns needs that the testing data set characterises just one PD source which, in turn, requires a reliable previous pulse source separation.

In Wang et al. (2015) the authors use particle swarm optimisation (PSO) to fit the parameter σ and the penalty factor of an SVM with RBF kernel. Then, a group of statistics containing information about the shape of PRPD patterns are used to train the SVM and classify the events. In Hunter, Lewin, Hao, Walton, and Michel (2013) a partial discharge is described by 20 features including characteristics in the time domain such as peak amplitude, phase angle, rise and fall times and the definite integral of the discharge; statistical parameters of the pulses such as mean, standard deviation, skewness, kurtosis; peak frequency in the FFT and energy ratios from a 9 levels Wavelet decomposition. Then, the SVM is applied to classify signals from four different defective power cable samples.

In any approach, independently of the applied technique or the parameters used, the training of SVMs or NNs require a wide and reliable data base of already identified PD sources chiefly through their PRPD patterns which could be very difficult to obtain in high-voltage equipment. This is a common drawback that is always present in the identification of partial discharges with machine learning.

In summary, the current trend in PD classification involves the use of combinations of sophisticated features extracted from different modalities (wavelets, high order statistics, measurements from signals in the time domain, PRPD patterns, etc) as input data to neural networks or SVM classifiers with RBF kernels (Raymond, Illias, Bakar, & Mokhlis, 2015). In general, these approaches achieve very good classification rates in discriminating PD from noise as well as in detecting PD sources, but in exchange of obscuring the interpretation of the results. The neural networks and the RBF kernel melt all the input features in a more or less complex classification function, which makes practically impossible to interpret the results of the classification in the sense that one cannot analyse the contribution nor the relevance of each individual input feature to the classification.

The approach presented henceforth is more straightforward than the methods used in the reviewed bibliography and the following features can be considered as original contributions and inherent strengths of the method:

- The algorithm is based on the use of the Power Spectrum Density (PSD) instead of a collection of statistics to capture the characteristics of the partial discharge type. During the process, the pulse is not modified by filtering so there is no loss of information and we can easily accommodate theory and results in the discussion or explanation of the classification.
- The great advantage of SVMs over NNs is the automatic determination of the architecture of the classifier. NNs need an a priori determination of the number of layers and number of neurons per layer using domain knowledge. SVMs can be regarded as a single hidden layer RBF NN in which each Support Vector becomes a neuron. The SVM global optimisation automatically determines the number of support vectors, whilst NNs need a much more intense training to find the architecture of the classifier and learn its parameters; these optimisations are greedy and prone to local minima (Bishop, 2006).
- The kernel selected for the SVMs in this paper (termed KL-kernel, see Section 2.3) operates on the frequency domain and posses a physical meaning. The KL-kernel measures the overlapping of the PSDs shapes and the shape of the PSD is influenced by the source of the PD. In this scenario, the most significant qualitative advantage of the KL-kernel over the ubiquitous RBF kernel is that the former distinguishes frequencies with high energy from frequencies with low energy in the construction of the similarity. In more detail, first notice that the input to both kernels are the normalised PSDs of the pulses. The KL-kernel computes the similarity between two pulses by multiplying their normalised energies in each frequency; this results in two pulses being similar when their frequencies with higher energy level coincide. However, the RBF kernel bases the similarity measure on the difference of the normalised energies in each frequency; this biases the total similarity towards frequencies with low energy, where the differences between the two PSDs will be smaller. Since the physical meaning of the PSD is that the energy distribution is related to the nature of the PD (or the noise signal), the KL-kernel is more suited for our purposes since captures similarities in the frequencies of high energy. This way there is a theoretical link between the classifications and the nature of the pulses as the former are based on computing similarities among the spectral densities of the