Automated Statistical Recognition of Partial Discharges in Insulation Systems.

Massih-Reza AMINI, Patrick GALLINARI, Florence d'ALCHE-BUC
LIP6, *Université Paris 6*, 4 Place Jussieu, F-75252 Paris cedex 5.
{Massih-Reza.Amini, Patrick.Gallinari, Florence.dAlche}@lip6.fr

François BONNARD, Edouard FERNANDEZ
Schneider Electric, Research Center, A2 plant, 4 rue Volta, F-38050 Grenoble 09.
{francois_bonnard, edouard_fernandez}@mail.schneider.fr

Abstract

We present the development of the successive stages of a statistical system for a classical diagnosis problem in the domain of power systems. It is shown how the different steps may be designed in a sound statistical way in order to develop a complete and efficient diagnosis system which provides the end user with a maximum of information about the system behaviour.

1 Introduction

We describe here the use of Neural Networks (NNs) and statistical techniques for a real life application in the domain of power systems. This application - the detection of electrical Partial Discharges (PD) in power apparatus - has been a challenging problem for more than 20 years. Efforts to automate PD detection began with expert systems [1]. Later on simple statistical models [2] and more recently Hidden Markov Models [3] and neural networks [4] have also been tested. However, most of this work has been exploratory and has addressed specific aspects of the problem. Some of the major statistical issues involved in the development of the successive processing steps were never considered. For example, in most approaches, feature selection relies on expertise, results validation is missing, data bases are too small to allow a sound comparison of different classifiers and feature sets. In the study presented here, a complete and valid approach ranging from feature extraction to statistical validation is developed. This allows us to compare both the performances and reliability of different systems and to show the interest of using flexible methods such as neural networks in real-world problems. In section 2, the PD phenomenon in insulating systems is described. Section 3 is devoted to feature extraction and selection and section 4 to the validation of the classifier. In section 5 we present experimental results.

2 Apparent charge detection

In high voltage power systems, small defects during manufacturing lead to the occurrence of Partial Discharges. Their accumulation is recognized as one of the

major sources of deterioration for power apparatus and can trigger breakdown. Commercially available PD detectors allow to measure PD pulses, to process them and display the PD patterns on a time basis. Some of these detectors also offer basic diagnosis tools. The usual representation of these discharges is via 3-d *fingerprints* defined by the phase angle φ_i, apparent charge q_j and pulse count $H_n(\varphi_i, q_j)$ coordinates (Figure. 1-left). Other useful statistics may be derived from this representation like (Figure 1-right): $H_{q\max}(\varphi)$ the maximum pulse height distribution, $H_{qn}(\varphi)$ the mean pulse height distribution, $H_n(\varphi)$ the pulse count distribution, $H(q)$ the number of discharges vs. discharge magnitude and $H(p)$ the number of discharges vs. discharge energy.

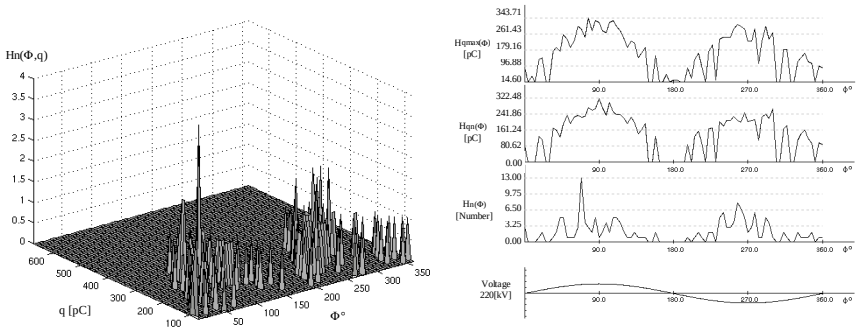

Figure 1 : left : 3-d representation of a PD, an entry n_{ij} represents the number of PD pulses having magnitude q_j and phase position φ_i. Right: some of the distributions which may be derived from the 3-d pattern, from top to bottom $H_{q\max}(\varphi)$, $H_{qn}(\varphi)$ and $H_n(\varphi)$.

PDs characterize random phenomena and are influenced by several factors like aging, amplitude of voltage, frequency of voltage applications. PD measures are corrupted by interferences and intrinsic background noise. These 3-d patterns are thus complex and their characterization is a challenging task. The problem we tackle here is the detection and classification of PDs for apparatus which use SF6 gas for insulation. More precisely we are interested into classifying PD signals into 4 classes: corona discharges (sharp points in an electric field), surface discharges (surface irregularities), floating parts (small particles in an electric field) or background measurement noises.

3 Feature extraction and selection

In the dielectric literature, features derived from the 3-d representation are used to characterize PDs. We have considered here some of the most promising feature sets in order to evaluate them, we have also performed variable selection on each set and on a combination of all these feature sets. The features we have considered are: fractal characteristics of the 3-d patterns, simple statistics over $H_{q\max}(\varphi)$, $H_{qn}(\varphi)$, $H_n(\varphi)$, $H(q)$ and $H(p)$ distributions, Fourier transform of $H_{qn}(\varphi)$. We first briefly describe these methods below and present after that variable selection.

3.1 Pre-processing

Fractal features

Fractal measures allow to compute global characteristics of an image. We have used two fractal measures on the 3-d fingerprints, the fractal dimension and the lacunarity [5] which quantify respectively the irregularity and the denseness of an image surface.

Statistical operators

Useful information from the distributions $H_{q\max}(\varphi)$, $H_{qn}(\varphi)$, $H_n(\varphi)$, $H(q)$ and $H(p)$ may be inferred from their moments. The following features have been found useful for PD detection [2], μ_i denotes the ith order moment of variable x for a given distribution: *Skewness*, $Sk = \mu_3/\mu_2^{3/2}$, which measures the degree of tilting of a distribution, *Kurtosis*, $Ku = \mu_4/\mu_2^2-3$, which measures the peakedness of a distribution, the number of modes of the distribution, and other features which characterize the asymmetry of the charges.

Fourier Analysis

The normalized charge distribution can be evaluated by the average value of its spectral components in the frequency domain. The normalized average discharge is given by $\tilde{H}_{qn}(\varphi) = \frac{a_0}{2} + \sum_{k=1}^{\infty}\left[\bar{a}_k \cos(k\varphi) + \bar{b}_k \sin(k\varphi)\right]$ where a and b are the Fourier coefficients in the frequency domain, φ is the phase angle. The first 8 spectral components a and b have been used here to approximate the charge distribution $H_{qn}(\varphi)$.

3.2 Variable selection

All measurements proposed by domain experts are not equally informative, it is thus useful to test automatic selection methods on these variable sets. Several methods have been proposed for variable selection with NNs. We have chosen here a method proposed in [6]. The relevance of a set of input variables is defined as the mutual information between these variables and the corresponding desired outputs. This dependence measure is well suited for measuring non linear dependencies as they are captured in NNs. For two variables x and d, it is defined as: $MI(x,d) = \sum_{x,d} P(x,d)\log\frac{P(x,d)}{P(x)P(d)}$. Starting from an empty set, variables are added one at a time, the variable x_i selected at step i being the one which maximizes MI(Sv_{i-1},x_i,d) where Sv_{i-1} is the set of *i-1* already selected variables and d the desired output. Selection was stopped when MI increase falls below a fixed threshold. Densities were estimated by Epanechnikov kernels.

For the classification, we have used multilayer perceptrons trained according to a MSE criterion plus a regularization term.

4 Validation

The validation of classifier decisions is important for any real-world application. It includes both performance and confidence evaluation. We focus here on the latter. There are two aspects of interest regarding the confidence, the first, is the computation of global confidence intervals which allows to calibrate a classifier, the second is the computation of local confidence measures for each classifier decision.
For computing global confidence intervals, we have used both a classical parametric estimation which relies on the hypothesis of binomial classifier outputs and a bootstrap estimate [7]. For the latter, the quantity to be estimated is the percentage of correct classification for each class. We compute : for each bootstrap replicate i and output k the percentage of correct classification $\hat{\theta}_{ik}$, the mean of these estimates $\hat{\theta}_k$ and their standard error $\hat{\sigma}_k$, $z_{ik} = \frac{\hat{\theta}_{ik} - \hat{\theta}_k}{\hat{\sigma}_k}$, $t^{(1-\alpha)}$ and $t^{(\alpha)}$ the left and right α^{th} percentiles of the $\hat{\theta}_{ik}$ distribution. The *bootstrap-t confidence interval* is then $\left[\hat{\theta}_k - \hat{t}_k^{(1-\alpha)}\hat{\sigma}_k, \hat{\theta}_k + \hat{t}_k^{(\alpha)}\hat{\sigma}_k\right]$.
Note that the confidence could be computed in the same way for the global classification performances instead of class performances.
For the local confidence, several heuristic measures have been proposed in the literature. Most of them assume that the estimates of posterior class probabilities are accurate enough. We have used here the following two estimates :

- $\max_k y_k - \max 2_k y_k$, where $\max 2_k y_k$ is the second maximum output value.
- $1 - \frac{-\sum_{k=1}^{p} \hat{P}(k/x) \log \hat{P}(k/x)}{\log p}$ where $\hat{P}(k/x)$ is the kth output normalized so that outputs sum to one.

5 Results

5.1 Experiments

Data were generated using an experiment design plan in order to insure a good representativity of the different classes and a minimum set of measures. The design factors for PD source discharges and background noise were : measure attenuation [0, 25] dB, gas pressure [1, 3] bar and test voltage [3.5 , 64.1] kV.
Simulations have been carried out to compare the different pre-processing methods and to show that the combination of these methods leads to a satisfactory discrimination. Variable selection allows to reduce the initial input dimension with a mean classification increase of about 1%. The number of variables is reduced from 28 to 12, 16 to 7 and 46 to 9 respectively for the statistical operators, Fourier components, and the combination of these two feature sets altogether with the 2 fractal measures. Performances are shown in table 1, the feature combination plus

variable selection offers a significant performance increase compared to the other feature sets.

Table 1. Performances of different NN models retrained on the selected variables and for different space features.

Space	NN Architecture	Performances (%)	
		Training	Test
Fractal Dimension	2-6-4	88.34	81.25
Statistical operators	12-16-4	97.25	93.12
Fourier components	7-16-4	98.75	94.38
Combination	9-16-4	100	**96.25**

5.2 Validation

Table 2, gives the confidence interval with 95% precision for the different classes using the 9 variables selected among the combination of the three feature sets. Intervals have been computed using the binomial assumption and the t-bootstrap interval method. With the t-bootstrap method, the confidence interval could be a singleton when a class has been recognized with a 0% or a 100% correct classification rate.

Table 2 : 95% confidence interval (on the test set) for the 9 input variables selected from the combination set. This interval has been computed with two methods (binomial assumption and t-bootstrap estimates).

	binomial distribution	t-bootstrap interval
Corona discharge	[80.12%, 95.41%]	[84.31%, 97.18%]
Surface discharge	[96.62%, 100%]	{100%}
Floating parts	[98.17%, 100%]	{100%}
Noise	[96.62%, 100%]	{100%}
Global performance	[94.54%, 97.43%]	[94.16%, 97.18%]

Additional information about the reliability of a classifier is provided by the confidence intervals computed directly on the outputs of the classifier instead of being computed on the classifier decision as in table 2. These intervals have been computed here via t-bootstrap and Figure 2 shows the corresponding box-plots for the combination of feature sets. In this case, the precision of the output values is high, it is significantly higher than for other feature sets.

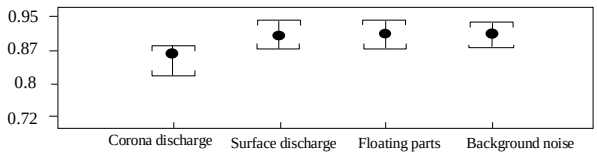

Figure 2 : Box-plots for the combination feature sets, The bold dot in each box indicates the median value and the lower and upper edges the 2.5 and 97.5% percentiles.

Local confidence intervals have been computed with the two methods described in section 4. Figure 3 shows the correct classification performances against the

percentage of reject. These experiments have been performed on the selected combination set. For both estimates of the local confidence, performances increase rapidly with the reject threshold to reach perfect recognition at 65% reject. These estimates allow the user to set up very easily the confidence level which fits the application.

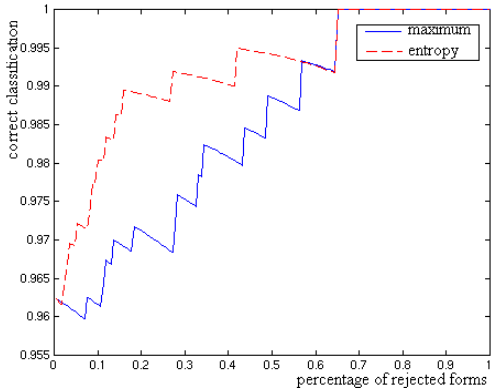

Figure 3: percentage of correct classification performances versus percentage of reject computed for two confidence measures.

6 Conclusion

We have described the development of the successive stages of a diagnosis system in the domain of power systems. Our goal was to illustrate how the different processing steps could be performed in order to develop, validate and calibrate as best as possible an operational system for a real world challenging problem. For this, we have used at the different steps (feature preprocessing and selection, classification and validation) methods which are well suited for the problem and complementary one to the other. Such software sensors could be easily implemented for several industrial diagnosis problems.

References

[1] Wootton RE. Computer Assistance for the Performance and Interpretation of H.V. ac Discharge Tests. 5th Int. Conf. On H.V. Eng., Braunschweig, 1987 ; Journal 41.12

[2] Gulski E. Computer-Aided Recognition of Partial Discharges using Statistical Tools. PhD thesis, Delft University, 1991

[3] Satish L, Gururaj BI. Use of Hidden Markov Models for Partial Discharge Pattern Classification. IEEE trans. On Electrical Insulation, 1993 ; 28 :172-182

[4] Gulski E, Krivda A. Neural Networks as a Tool for Recognition of Partial Discharges. IEEE trans. On Electrical Insulation, 1993 ; 28 :984-1001

[5] Satish L, Zaengl WS. Can Fractal Features be Used for Recognizing 3-d Partial Discharge Pattern ? IEEE trans. On Electrical Insulation, 1995 ; 2 :352-359

[6] Bonnlander BV, Weigend AS. Selecting Input Variables Using Mutual Information and Nonparametric Density Evaluation. In Procedings of ISANN'94, Taiwan, 1994 ; 42-50

[7] Efron B, Tibshirani R. An Introduction to Bootstrap, Chapman Hall, London, 1993