

## Pattern Recognition Method of PD Signals Based on Fuzzy Clustering

Zhi-Gang REN<sup>1</sup>, Wei LI<sup>1</sup>, Wen-Jie JIN<sup>2</sup>, Yao CHE<sup>2</sup>, Song-Lin ZHOU<sup>2</sup>, Dian-Chun ZHENG<sup>3\*</sup> and Ian-Xiang HE<sup>3</sup>

<sup>1</sup> State Grid Beijing Electric Power Research Institute, Beijing, China, 100075

<sup>2</sup> Beijing Electric Power Research Institute, Beijing, China, 100075

<sup>3</sup> Harbin University of Science and Technology, Key Laboratory of Engineering Dielectrics and Its Application, Ministry of Education, Harbin, China, 150080

\*Corresponding author

zhengdianchun@126.com

**Keywords:** Pattern Recognition, Fuzzy Clustering, Hierarchical Clustering Method.

**Abstract.** Fuzzy clustering is a technique which adopts the fuzzy mathematics method to build up faintness relations on the basis of signal features. In this paper, two methods adopting fuzzy hierarchical clustering and fuzzy equivalent matrix method for distinguishing the partial discharge(PD) signals are demonstrated and rather efficient for pattern recognition of the PD signals appearing in the high voltage apparatus, and having great application prospect in the future.

### Introduction

The traditional clustering analysis is a kind of hard partition, which divides every pending discriminating object strictly to a certain class, and each object is only in one class, so this division is well-defined<sup>[1]</sup>. However, in fact most of the objects do not have a strict description, which are intermediary in the sides of quality and adscription. Not only do they belong to this class but also to another class, which is suitable for flexible partition<sup>[2]</sup>. Fuzzy cluster is the combination of cluster analysis and fuzzy theory. Compared with the traditional cluster analysis, fuzzy cluster algorithm improves the priority probability of the algorithm.

### Preconditioning of the Data-Signals

The fuzzy cluster system is built using the fuzzy-tool box of MATLAB to precondition the collected data-signal of partial discharge. The original data-signals are respectively captured from the models of the needle-plate, needle-needle and the sphere-plate electrode systems in the laboratory as shown in Fig.1, and each data-signal are presented in Fig.2. Then importing 30 data signals of the each model to working space of MATLAB, and combining a  $1000 \times 30$  elementary array, in which the first 10 columns indicate the PD signal of the needle-plane model, the middle 10 columns show the PD signal of the needle-needle model, and the last 10 columns express the PD signal of the sphere-plane model. And furthermore each row elements of  $1000 \times 10$  array are averaged, and acquiring a  $1000 \times 1$  vector. The same treatment is applied to the

needle-plane, the needle-needle and the sphere-plane model, so that three vectors are combined to an elementary array as standard modes of three kinds of PD signal <sup>[3]</sup>.

According to fuzzy cluster theory, if the tested data is matched with one standard mode, then they are identified and belonged to this one <sup>[4] [5]</sup>.

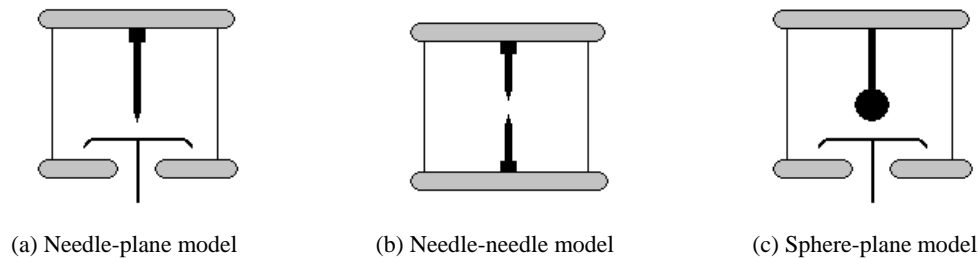


Fig.1 Models of the PD signal in the laboratory

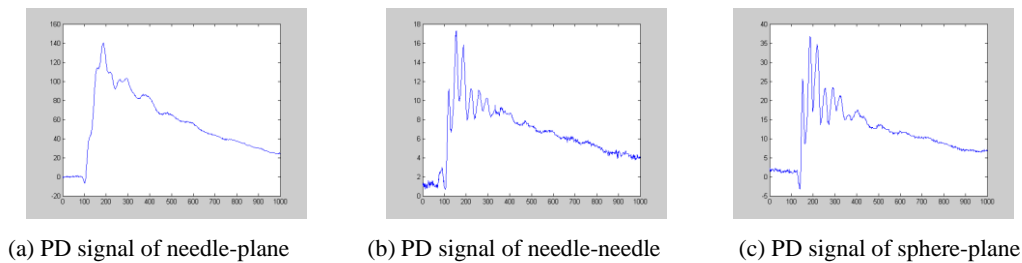


Fig.2 Signals of the PD models

## Implementation of the Fuzzy Cluster Pattern Recognition

### Fuzzy Hierarchical Clustering Method

The cluster standard has been chosen, and a new array is made up of the prepared data from data signals according with the partial discharge(PD) of the experimental system in the laboratory, therefore proceeding the fuzzy hierarchical clustering is first to be established <sup>[6]</sup>.

Treating the 30 known arrays as the data signals of the partial discharges in different kinds of PD model geometry, which are combined with the standard data to a 1000×33 array, using the extremal standardization formula to compress the standard data into [0,1]. The formula of extremal standardization as follow:

$$x''_{ij} = \frac{x'_{ij} - x'_{j\min}}{x'_{j\max} - x'_{j\min}} \quad (1)$$

Where,  $x'_{j\max}$  and  $x'_{j\min}$  are the max and min data in  $x'_{1j}, x'_{2j} \dots x'_{nj}$ . The matching distance of observed quantity is also calculated, and the output is defined in the form of a square matrix and the hierarchical clustering tree is simultaneously established, then related coefficients have gained and the dendrogram. Cutoff is the blow-threshold of discontinuous parameter, and the difference is then used to calculate discontinuous coefficients.

When cutoff = 3, the related coefficient is C=0.9302, then the dendrogram of 30 groups data and standard data is given as shown in Fig.3.

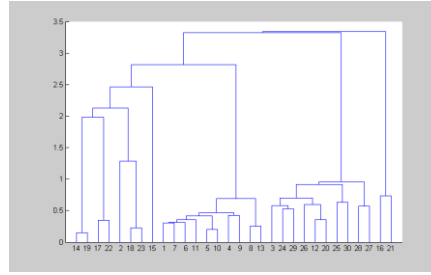


Fig. 3 Dendrogram of 30 groups data and standard data

Choosing one group or several groups data at random from the known data signals as tested data to cluster. Such as choosing the fourth group of data from electrode system of the needle-plate geometry as tested data. When cutoff=3, and C=0.9792, then the result shows that clustering is implemented as shown in the Fig.5.

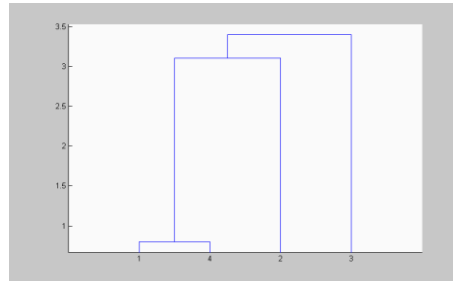


Fig. 5 Dendrogram of the fourth groups data of needle- plate and standard data

### Fuzzy Equivalence Matrix Method

The standard of clustering has been chosen, combine the tested data and standard data to a new array, then use the fuzzy equivalence matrix method to cluster<sup>[7]</sup>.

The extreme standardization formula was chosen to compress the standard data into [0, 1] space and cosine of separation angle method was used to produce fuzzy similar matrices R.

Cosine of separation angle:

$$\cos \alpha_{ij} = \frac{\sum_{k=1}^m x_{ki} \cdot x_{kj}}{\sqrt{(\sum_{k=1}^n x_{ki}^2)(\sum_{k=1}^n x_{kj}^2)}} \quad (2)$$

And fuzzy similar matrices R as follow:

$$R = \begin{bmatrix} 1.0000 & 0.9863 & 0.9753 & 0.9994 \\ 0.9863 & 1.0000 & 0.9625 & 0.9833 \\ 0.9753 & 0.9625 & 1.0000 & 0.9780 \\ 0.9994 & 0.9833 & 0.9780 & 1.0000 \end{bmatrix} \quad (3)$$

According to the method of synthesis fuzzy matrices  $R^2, R^4, \dots, R^{2k}$ , when  $k=2$ , then  $R^4=R^2$ , so acquired fuzzy equivalence matrix  $R^*$ :

$$R^* = \begin{bmatrix} 1.0000 & 0.9863 & 0.9780 & 0.9994 \\ 0.9863 & 1.0000 & 0.9780 & 0.9863 \\ 0.9780 & 0.9780 & 1.0000 & 0.9780 \\ 0.9994 & 0.9863 & 0.9780 & 1.0000 \end{bmatrix} \quad (4)$$

According to matrix  $R^*$ , acquiring different  $R_\lambda$  has been obtained for different value  $\lambda$ , therefore realizing different clustering, the dynamic clustering chart is acquired as shown Fig.6.

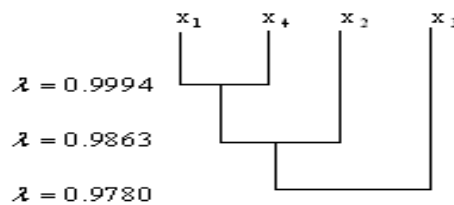


Fig. 6 Dynamic clustering chart

When  $\lambda = 0.9994$ , then

$$R_{\lambda=0.9994} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \quad (5)$$

It is seen that the fourth group tested data and the first group of standard data are clustered together, it is true that the tested data is the type of needle-plane.

## Conclusions

In this article, the technique of fuzzy clustering is applied to classify PD signal pattern appeared in the high-voltage electrical apparatus, and the fuzzy-tool box is used to standardize the original PD signals and establish the clustering standard modes. Then the fuzzy hierarchical clustering method and the fuzzy equivalence matrix method is employed to cluster and analyze PD signals, respectively. The facts are proven effectively and has great promising, but some questions are given rise to attention and required to continue research in practice.

## References

- [1] Sergios Theodoridis, Konstantinos Koutroumbas. Pattern Recognition (second edition) [M]. Greece: Publishing House of Electronics Industry, 2004:1-5
- [2] Syed Mofizul Islam, Tony Wu, Gerard Ledwich. A Novel Fuzzy Logic Approach to Transformer Fault Diagnosis. *IEEE Transactions on Dielectrics and Electrical Insulation*, 2000, 7(2):177~186
- [3] Y Nakamori. Fuzzy Clustering and Regression analysis. Journal of Japan Society for Fuzzy Theory and Systems, 1996, 8(3):431-439
- [4] Zheng Zhong, Tan Ke-xiong and Wang Meng, Wu Hao. "Partial discharge recognition base on features of pulse waveform," *Advanced technology of electrical engineering and energy*, 2001, vol.20 no.2, pp.20-24
- [5] Ahmed Rubaai, Daniel Ricketts, M. David Kankam. Development and Implementation of an Adaptive Fuzzy-Neural-Network Controller for Brushless Drives [J]. *IEEE TRANSACTIONS ON INDUSTRY APPLICATIONS*, 2002, 38(2): 4451-446.
- [6] Gao Xin-bo. "Analysis and Application of Fuzzy Clustering". *Xi Dian University publishing company* 2004, 1.
- [7] T. Tanaka, T. Okamoto. Partial discharge pattern recognition for three kinds of model electrodes with a neural network[J]. *IEE Proc.-Sri. Meas. Technol*, 1995, 142(1) : 75-83.