

펄스 분석 기법 및 데이터 마이닝 기법을 이용한 부분방전 패턴인식에 대한 연구

김정태*, 이육*, 김지홍**, 구자윤**
 * 대전대학교 전기정보시스템공학과 ** 한양대학교 전기공학과

A Study on the PD Pattern Recognition using the Pulse Analysis Method and Data Mining Methods

Jeong-Tae Kim*, Wook Rhee*, Ji-Hong Kim**, Ja-Yoon Koo**
 * Daejin University ** Hanyang University

Abstract - Recently, the noise discrimination method using PD pulse waveshape analysis has been suggested to be very effective method which can improve the reliability of the on-site PD measurement. In this method, the data clusters due to PD pulses or noises can be distinguished on the PA map. And for the automatic recognition of the PD clusters, it is necessary to adopt the adaptable pattern recognition method. In this study, as for the algorithm which can recognize data clusters, the data mining method has been adopted and the result of the analysis has been reported.

1. 서 론

최근 부분방전 펄스와 노이즈 펄스의 파형이 다르다는데 착안하여 측정되는 펄스의 특성을 추출하여 구분하는 방법에 대한 연구가 많이 이루어지고 있다.[1-3] 이러한 경우 부분방전의 패턴인식 방법으로 기존의 PRPD Map에 대해 수행하는 BP 또는 SOM과 같은 신경회로망(Neural Network) 기법이나 퍼지 이론을 적용할 수도 있다.

그러나, 기본적으로 신경회로망 기법은 단정적이라는 문제를 안고 있다. 즉, BP-NN 기법의 경우 사전에 정해진 패턴 이외의 패턴은 인식할 수 없으며 어느 결합일 확률적인 개념의 적용도 미흡하다. 아울러, SOM-NN 기법은 새로운 패턴일 가능성에 대해서는 정보를 제공하지만 흑백 논리의 결과를 주기 때문에 치명적인 오판의 가능성을 갖고 있다.

또한 펄스의 특성을 추출하여 부분방전을 구분하는 방법의 결과인 TF Map 또는 PA Map의 경우에는 여러 그룹의 데이터 클러스터가 형성되므로, 이 클러스터들이 어느 결합이나 노이즈에 해당하는가를 인식하기 위해서는 그 클러스터의 중심이나 범위 등을 고려하여 분석하는 것이 보다 적절하다고 할 수 있다.

이러한 측면에서 EM 알고리즘이나 k-means 기법과 같은 데이터 마이닝 기법은 주어진 데이터 클러스터의 중심을 추적해 들어가는 방법으로 PA map에 대한 분석으로는 신경회로망 기법보다 적절하다고 판단된다. 따라서, 본 연구에서는 측정 펄스파형의 특성을 이용한 노이즈 제거 기법의 효과를 높이고 패턴인식의 정확도를 제고시키기 위한 패턴인식 기법으로 데이터 마이닝 기법을 선정하고 이에 대해 검토하였다.

2. 본 론

2.1 분석 방법

2.1.1 PA 데이터 도출

PA를 수행하기 위해서는 각종 결합에서 발생하는 부분방전 펄스 파형의 특성 및 방전의 양상이 어떠한 차이가 있는지 파악하여야 한다. 이에 따라 본 연구에서는 TDS 744A 500MHz 오실로스코프와 제작한 HFDP 진단장치 및 컴퓨터를 이용하여 100kHz~20MHz 범위에서 3dB 특성을 갖고 있는 상용 HFCT에서 입력되는 부분방전 펄스 파형을 측정하고 분석하였다. 본 연구에서 같이 를 사용하였다.

아울러, 본 연구에서 데이터 마이닝 기법의 적용가능성을 파악하기 위해 채택한 모의 결합으로는 전력기기의 절연 시스템의 대표적인 결합인 절연체 내 보이드(에폭시 절연체 내 직경 1 mmφ 길이 1 mm의 원통형 보이드) 및 이를 절연유 속에 넣은 경우와 유중 표면방전(절연유 속에서 직경 2 mmφ 길이 10 mm의 강선 조각을 두께 3 mm의 프레스보드와 상부 전극 사이에 설치)을 채택하였다. 아울러, 이들 결합에서 발생하는 부분방전과 구분되어야 할 노이즈로는 기중 코로나를 선정하였고 곡률반경 0.5mm인 침을 모의결합 상부 전극에 같이 설치하여 모의하였다.

한편 PA Map은 펄스 파형을 캡처한 뒤 최대값과 최소값의 비율과 이에 대한 인덱스의 합(적분값)을 이용한 MMR-S Map을 적용하였다.

2.1.2 데이터 마이닝 기법

데이터 마이닝이란 각종 데이터 속에 숨어있는 패턴, 규칙, 관계 등의 정보를 자동으로 탐색하고 데이터 간의 연관을 발굴(mining)해 내는 인공지능, 공학 및 통계 기법을 연구하는 학문 분야이다. 여기에서는 데이터 마이닝의 시작이라고 할 수 있는 같은 패턴끼리 모아주는 백터 양자화 혹은 클러스터링 방법으로 EM 알고리즘과 k-means 알고리즘을 적용하였다.

EM 알고리즘이란, 숨겨진 정보를 포함하고 있는 문제에서 숨겨진 정보의 최적 해를 찾아내는데 매우 유용한 알고리즘이다. 예를 들어, 어떠한 데

이터를 클러스터링하는 문제에서, 만약 각 데이터 점들이 어떠한 집합 혹은 클러스터에 속하는 것을 알 경우에는 최적의 코드 벡터들은 간단히 이들 클러스터들의 중심이 될 것이다. 그러나 불행하게도 우리는 어떤 데이터 점들이 어느 그룹에 속해 있는지 모른다. 결국 이것이 감추어진 정보가 된다. EM 알고리즘은 지금까지 최상의 해결법으로 숨겨진 정보를 일단 추정(expectation 단계)하여 얻어진 해로부터 이 추정치를 개선하기 위해서 다시 이것을 이용하는 절차(Maximization 단계)를 해가 수렴할 때 까지 번갈아 행하는 방법이다. 결국 이 과정을 통하여 국부적인 최적치로 수렴하게 되는 것이다. EM 최적화 알고리즘 중에서 가장 간단한 예가 k-means 알고리즘이다.

EM 알고리즘의 경우, 초기값 선택이 최적화에 결정적인 요소가 되는데 k-means 알고리즘은 임의의 초기 값에서부터 추정(E)-최대화(M)과정을 수렴할 때까지 반복시키면서 중심을 찾는다. 즉, E 단계에서 중심으로부터 클러스터를 선택하여 결정하고, M 단계에서 거꾸로 클러스터로부터 중심을 결정하는 과정을 반복하는 추정 알고리즘이다.

이 때, 왜곡이 안정되었는지 아닌지를 확인하는 효과적인 검증(test)은 이전 반복시의 왜곡과 현재 왜곡에 대한 상대적인 감소치를 점검하면 된다. 이를 수식으로 표현하면 다음과 같다.

$$\Delta D = \frac{D_{prev} - D_{curr}}{D_{prev}} < 10^{-4}$$

이 알고리즘은 일반적으로 100번 반복을 하기 전에는 수렴해야만 한다.

k-means 알고리즘은 가장 가까운 양자화된 중심을 찾는 과정에서 모든 k개의 코드 벡터들에 대하여 데이터의 수와 같은 크기의(uniform) 검색을 함으로 좋은 코드북을 얻을 수 있지만 속도가 느린 단점이 있다. 비균일 이진 분할법은 계층적 클러스터링 방법으로 순환적으로 데이터 집합을 클러스터와 부 클러스터로 나누는 방법이다. 이 방법은 부분 최적화 방법이라고 할 수 있으며 k-means 보다 훨씬 빠르다.

이진 분할 알고리즘은 초기에 데이터 집합을 두 개의 클러스터로 나눈다. 그리고 k 개의 클러스터가 남을 때까지 반으로 클러스터를 나누는 과정을 반복한다. 즉, logK 번의 이진 검색을 통하여 가장 가까운 중심을 찾게 되는 것이다. 물론 코드북이 크면 클수록 계산상에서 더 큰 개선이 있을 것이다. 이방법의 가장 큰 장점은 분할 속도가 빠르다는 것이고, 반면에 코드북의 질은 k-means에 비교하여 떨어지는 단점이 있다.

2.2 데이터 마이닝 기법을 이용한 부분방전 패턴인식 결과

2.2.1 EM 알고리즘 적용

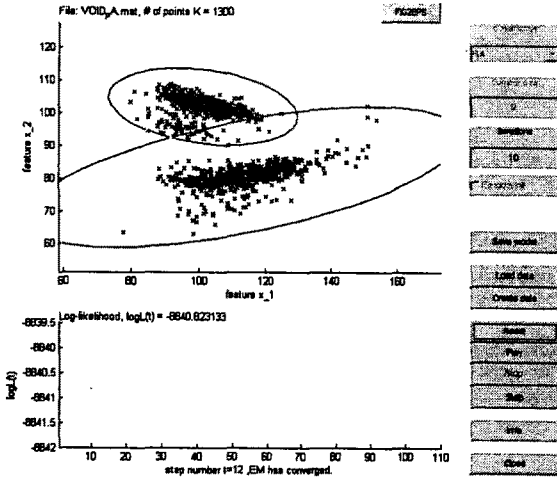
그림 1은 유중 고체절연체 내 보이드 결합에서 발생시킨 정극성 및 부극성 보이드 방전의 PA mapping 데이터에 EM 알고리즘을 적용한 것이다. 상부에 위치한 부극성 보이드 방전 클러스터 패턴과 하단에 위치한 정극성 보이드의 클러스터가 명확하게 형성되어 타원형의 범위로서 나타남을 알 수 있다.

그림 2는 유중 표면방전과 노이즈로서 고려한 기중 코로나의 다중방전에 대한 PA map 클러스터링 분류를 위한 EM 알고리즘 적용 결과이다. 기본적으로 정극성과 부극성의 데이터는 명확히 구분되어지고, 또한 맨 아래 위치한 정극성의 유중 표면방전과 상부 왼쪽에 위치한 부극성 기중 코로나의 패턴도 제대로 클러스터 형성됨을 볼 수 있다. 하지만 상부 오른쪽에 위치한 정극성 유중 표면방전인 경우 그 중심점을 제대로 인식하지 못해 러프하게 큰 범위로 클러스터가 형성되었다.

그림 3은 유중 절연체 내 보이드 방전과 부극성 기중 코로나 방전의 다중 방전에 대해 PA 알고리즘을 적용한 mapping의 EM 패턴인식 알고리즘 적용 결과이다. 이 경우, 그림 2에 비해 매우 패턴분류가 명확히 잘 되었음을 알 수 있다. 학습 과정시 상부 왼쪽에 위치한 부극성 기중 코로나의 클러스터는 데이터 밀집으로 인하여 그 중심의 해를 찾는 과정에서 모든 합산의 값을 최소화하는데 가장 학습율이 좋기 때문에 쉽게 중심점을 찾는 것을 볼 수 있다. 이러한 과정은 다른 여러 가지 결합의 부분방전 데이터에 PA mapping을 적용할 때도 동일하게 나타나고 있다. 따라서 학습에 따른 클러스터 구분의 성공률은 먼저 cluster mapping 알고리즘에 의한 명확한 분류에 좌우됨을 알 수 있다.

2.2.2 k-means 알고리즘 적용

그림 4는 유증 고체절연체 내 보이드 방전의 PA mapping 데이터에 k-means 알고리즘 적용한 결과이다. 그림 1과 같은 EM 알고리즘 경우나 k-means 알고리즘인 경우 명확히 데이터 클러스터를 구분하는 것을 알 수 있다. 이 경우에 k-means 중심점 결과 값을 나타내었다.



<그림 1> 유증 고체절연체 내 보이드 방전에 대한 EM 알고리즘 적용

CBout =	
114.0033	80.5460
101.2920	101.7394

<그림 4> 유증 고체절연체 내 보이드 방전의 k-means 알고리즘 적용

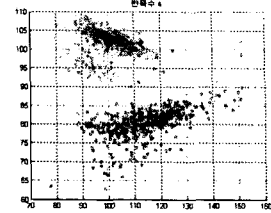
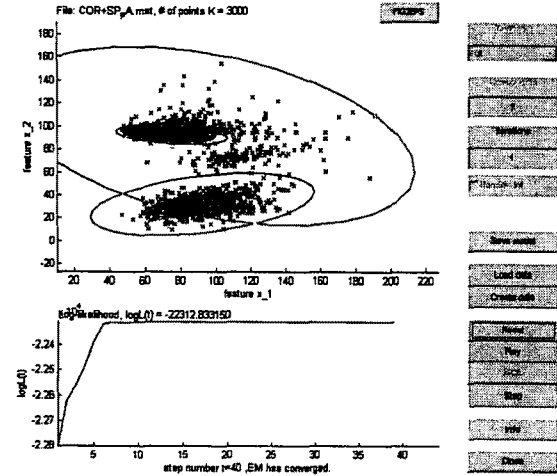
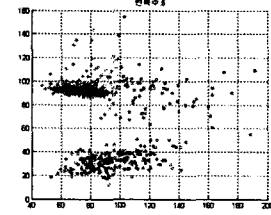


그림 5는 유증 표면방전과 부극성 기증 코로나 다중방전의 PA mapping 데이터를 k-means 알고리즘을 적용한 패턴 인식이다. 이에 대한 k-means의 중심점 결과 값을 나타내었다. 하단에 위치하는 정극성 표면방전의 데이터와 오른쪽 상부에 위치한 부극성 표면방전의 데이터를 잘 추출하고 있으며 상부 좌측의 강하게 군집된 부극성 기증 코로나의 클러스터가 명확하게 분류되는 것을 알 수 있다. 이것은 동일한 다중방전으로 분석한 그림 3.3.14의 EM 알고리즘에서는 정극성 계면방전의 데이터 추출과정에서 많은 오차를 보여주는데 반하여 명확한 결과를 볼 수 있다.

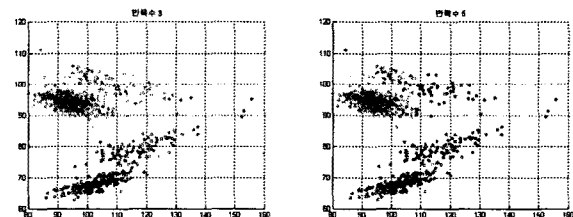
CBout =	
93.3182	32.0320
76.2381	93.0391
126.1128	91.9085

<그림 5> 유증 표면방전과 부극성 기증 코로나 방전의 패턴 인식

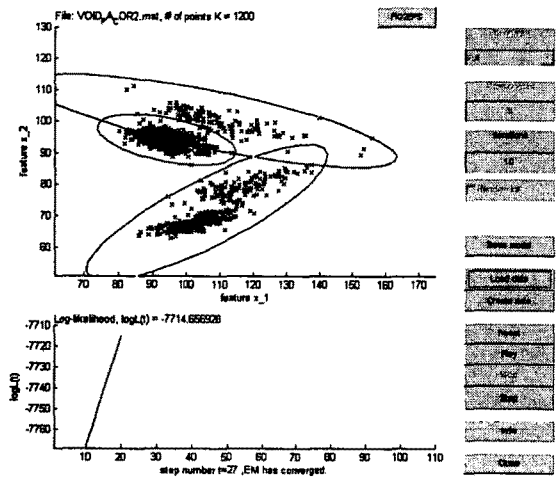


<그림 2> (유증 표면방전 + 기증 코로나) 다중방전의 PA mapping에 대한 EM 알고리즘 적용

그림 6은 PA mapping 데이터에서 k-means를 적용한 유증 고체절연체 내 보이드 방전 및 부극성 기증 코로나 방전의 다중방전에 대한 패턴 인식이다. 이에 따른 각각의 k-means 중심점의 결과 값을 나타내었다. 그림 6.(a)와 같이 2 클러스터 분류를 사용할 경우 정확히 패턴이 구분되지만 부극성 보이드 방전과 부극성 기증코로나 방전의 구분이 어려운 것을 알 수 있다. 또한, 그림 6.(b)와 같이 3 개의 클러스터를 형성할 경우 하단의 정극성 보이드 방전 우측 부분에서 오차가 발생되고 있다. 이에 반해, 상기한 그림 3의 EM 알고리즘을 적용한 경우 보이드와 기증 코로나 방전을 명확한 데이터를 구분하였다.



<그림 6> k-means를 적용한 (보이드 + 기증코로나) 다중방전의 패턴 인식



<그림 3> (보이드 방전 + 부극성 기증 코로나) 다중 방전의 PA mapping에 대한 EM 알고리즘 적용 결과

2.2.3 EM 알고리즘과 k-means 알고리즘의 복합 적용

그림 7은 유증 고체절연체 내 보이드 방전 데이터에 대해 EM 알고리즘과 k-means 알고리즘을 같이 적용한 경우의 결과이다. 그림 1 및 그림 4에서 나타난 EM 알고리즘과 k-means 알고리즘의 결과가 복합적으로 나타나 명확한 클러스터 구분을 보여주고 있다.

그림 8은 유증 표면방전과 부극성 기증 코로나의 다중 방전 데이터에 대해 EM 및 k-means 복합 알고리즘을 적용한 분석 결과이다. 다중 방전에 따른 복잡한 양상에 대해, 그림 2에서 EM 알고리즘만으로 분석할 경우의 부극성 유증 표면방전과 부극성 기증 코로나 방전의 중첩으로 인한 오차 부분을 명확히 k-means가 보완하여 패턴을 분류한 것을 볼 수 있다. 이에 대한 k-means 중심점 결과 값을 나타내었다.

그림 9는 PA mapping된 (보이드 + 기증코로나) 다중방전에 대한 EM 및 k-means 복합 알고리즘을 적용한 경우의 결과이다. 데이터의 구분이 그림 6의 k-means 알고리즘만 적용한 경우에 비하여 데이터의 구분이 보다 용이함을 알 수 있다.

3. 결 론

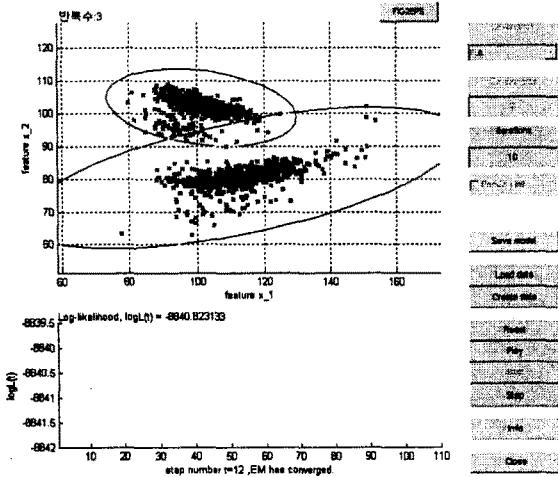
이상 언급한 바와 같이 데이터 마이닝 기법은 PA mapping된 데이터 클러스터를 구분하는데 효과적임을 알 수 있었다. 단, 어느 하나만의 패턴 인식 기법을 사용하는 것보다 EM 알고리즘과 k-means 알고리즘을 복합적

로 적용하여 상호 단점을 보완할 수 있는 알고리즘이 보다 적절한 것으로 판단된다.

아울러, 패턴인식 기법에 따른 효과보다는 기본적인 PA map의 데이터 클러스터 자체가 구분 가능하도록 알고리즘을 개선하는 것이 부분방전 진단의 노이즈 제거 효과를 극대화시키는 방안으로 고려된다.

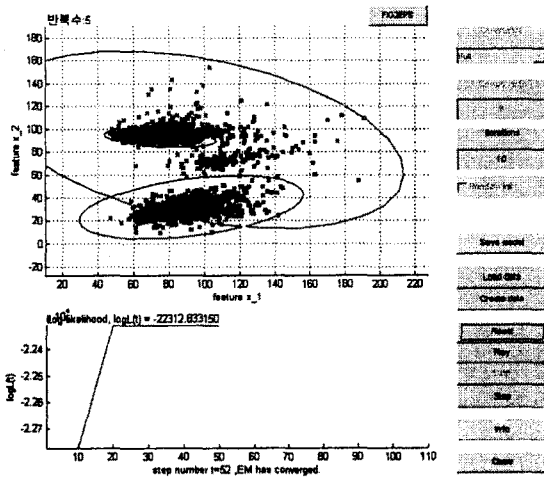
하지만, 향후 PA mapping에 대한 데이터 마이닝 분석 기법에 기존의 PRPD 패턴을 복합적으로 적용하여 패턴인식을 수행할 경우 보다 효과적인 결과를 도출할 수 있을 것으로 사료된다.

CBout =		CBout =	
106.8221	71.1661	103.2984	69.0872
96.1558	95.1156	94.8821	94.9956
		122.6521	86.8488

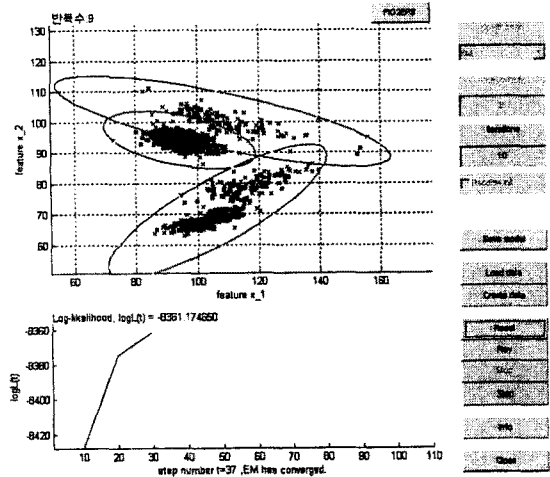


〈그림 7〉 유종 고체절연체 내 보이드 방전에 대한 EM 및 k-means 알고리즘의 복합 적용 결과

CBout =	
92.6800	31.7545
75.5637	93.0971
116.9245	82.4271



〈그림 8〉 (유종 표면방전+기중 코로나) 다중방전의 PA mapping에 대한 EM 및 k-means 복합 알고리즘 적용



〈그림 9〉 보이드 + 계면 방전의 EM 알고리즘과 k-means 알고리즘

감사의 글
본 연구는 산업자원부의 지원에 의하여 기초전력연구원 (R-2004-B-123) 주관으로 수행된 과제임.

〔참 고 문 헌〕

- [1] A.Cavallini et al, "A New Approach to the Dianosis of Solid Insulation Systems Based on PD Signal Inference", IEEE Electrical Insulation Magazine, Vol.19, No.2, pp.23~30, 2003
- [2] H.J. van Breen, E. Gulski, J.J. Smit, H.F.A. Verhaart and W. de Leeuw, "Performance analysis of different on-line PD measurement techniques for usage on Turbo Generators", XIIIth International Symposium on High Voltage Engineering (ISH 2003), Diagnostics and on-line monitoring for CBM Session, Delft, Netherlands, August 25-29, 2003
- [3] W. Krause, H.-G. Kranz, "Pulse Preclassification for online PD Noise Suppression based on realtime Pulse Shape Identification", 14th International Symposium on High Voltage Engineering, G-002, Tsinghua University, Beijing, China, August 25-29, 2005