

<http://dx.doi.org/10.17703/JCCT.2023.9.5.569>

JCCT 2023–9–71

DBSCAN과 통계적 검증 알고리즘을 사용한 배터리 열폭주 셀 탐지

Battery thermal runaway cell detection using DBSCAN and statistical validation algorithms

김진근*, 윤유림**

Jingun Kim*, Yourim Yoon**

요약 납축전지는 가장 오래된 충전식 배터리 시스템으로 현재까지 충전식 배터리 분야에서 자리를 지키고 있다. 이 배터리는 다양한 이유로 열폭주 현상이 생기는데 이는 큰 사고로 이어질 가능성이 있다. 그렇기 때문에 열폭주 현상을 예방하는 것은 배터리 관리 시스템의 핵심부분이다. 최근에는 열폭주 위험 배터리 셀을 기계학습으로 분류하는 연구가 진행 중이다. 본 논문에서는 비지도학습인 DBSCAN 클러스터링과 통계적 방법을 사용하여 열폭주 위험 셀 탐지 및 검증 알고리즘을 제안하였다. BMS에서 측정된 lead-acid 배터리의 저항 값을 사용하여 열폭주 위험 셀 분류 실험을 진행하였고 본 논문에서 제안한 알고리즘이 열폭주 위험 셀을 정확히 검출해 냄을 보여주었다. 또한 본 논문에서 제안한 알고리즘을 사용하여 배터리 내 열폭주 위험이 있는 셀과 노이즈가 심한 셀을 분류할 수 있었으며 그리드 서치를 통한 DBSCAN 파라미터 최적화를 통해 열폭주 위험 셀을 초기에 검출해 낼 수 있었다.

주요어 : 클러스터링, DBSCAN, 기계 학습, 납축전지, 열폭주 현상

Abstract Lead-acid Battery is the oldest rechargeable battery system and has maintained its position in the rechargeable battery field. The battery causes thermal runaway for various reasons, which can lead to major accidents. Therefore, preventing thermal runaway is a key part of the battery management system. Recently, research is underway to categorize thermal runaway battery cells into machine learning. In this paper, we present a thermal runaway hazard cell detection and verification algorithm using DBSCAN and statistical method. An experiment was conducted to classify thermal runaway hazard cells using only the resistance values as measured by the Battery Management System (BMS). The results demonstrated the efficacy of the proposed algorithms in accurately classifying thermal runaway cells. Furthermore, the proposed algorithm was able to classify thermal runaway cells between thermal runaway hazard cells and cells containing noise. Additionally, the thermal runaway hazard cells were early detected through the optimization of DBSCAN parameters using a grid search approach.

Key words : Clustering, DBSCAN, Machine learning, Lead-acid Battery, Thermal runaway

*준회원, 가천대학교 IT융합학과 석사 과정 (제1저자)
**정회원, 가천대학교 컴퓨터공학과 부교수 (교신자)
접수일: 2023년 7월 12일, 수정완료일: 2023년 8월 16일
게재확정일: 2023년 9월 5일

Received: July 12, 2023 / Revised: August 16, 2023

Accepted: September 5, 2023

**Corresponding Author: yryoon@gachon.ac.kr

Dept. of Computer Engineering, Gachon Univ, Korea

I. 서 론

배터리는 처음 발명된 이후 많은 장치들에게 전원을 공급해주는 역할을 해왔다. 그러나 일부 예상치 못한 고장으로 배터리 관리의 중요성이 부각되면서 배터리 관리 시스템(BMS)이나 에너지 관리 시스템이 개발되었다 [1]. 이러한 시스템을 통해 얻은 데이터를 사용하여 시스템의 문제를 예측하고 계획되지 않은 사고를 사전에 방지하는 것은 유지관리의 핵심 부분이다 [2].

납축전지는 가장 오래된 충전식 배터리 시스템으로 다소 제한적인 저장능력을 가졌지만 100년 이상 충전식 배터리 분야에서 자리를 지키고 있다 [3,4].

납축전지는 다른 2차 전지보다 저렴한 가격과 낮은 온도에 민감성이 작아 주요 에너지 저장 및 공급 역할을 해왔지만 배터리의 결함, 과충전 그리고 과방전 등 다양한 원인에 의해서 열폭주 현상이 생길 수 있다 [5,6]. 열폭주 현상이 일어나면 전류와 온도, 저항이 모두 높은 값에 도달할 때까지 온도가 상승하여 배터리 폭발의 위험성이 있다 [7-10].

열폭주 위험 배터리 셀 진단을 위해 기계학습 중 지도 학습인 서포트 벡터 머신(SVM), 의사결정트리 그리고 인공신경망을 사용하기도 하지만 본 논문에서는 클러스터링을 사용하여 진단하는 방법을 제안한다 [11-13]. 비지도 학습의 종류인 클러스터링 방식 중 클러스터 개수를 정해줄 필요가 없고 시계열 데이터에 적합한 DBSCAN을 선택하였다 [14-16]. 또한 통계적 방법을 사용한 검증 알고리즘을 사용하여 정확하게 분류하도록 시도했다.

최근 열폭주 위험 배터리 셀 진단에 기계학습을 사용하는 연구가 진행되고 있다 [17-19]. 특히 최근 한 논문 [20]에서는 DBSCAN과 임계값을 사용하여 실제 BMS에서 이상 배터리 셀을 검출하는 연구가 진행되었다. 그러나 실험에 사용된 데이터의 개수가 작아 DBSCAN과 임계값의 성능을 명확하게 판단할 수 없었다. 또한 임계값을 설정하는 기준에 있어 타당한 근거가 없다는 한계가 있었다.

본 논문에서는 이전 연구의 문제를 보완하기 위해 더 많은 실험 데이터를 사용하였다. 그리고 3 시그마

규칙을 임계값의 기준으로 설정하여 타당성을 높였으며, 특히 그리드 서치를 통한 DBSCAN 파라미터 최적화를 통해 이상 배터리 셀을 빨리 분류할 수 있도록 설계하였다.

본 논문에서는 총 96개의 배터리 셀을 가진 배터리 데이터 셋 여섯 개를 사용한다. 다섯 개는 실제 BMS에서 측정된 데이터이고 한 개는 인위적으로 열폭주 위험 셀을 포함하도록 만든 데이터이다. DBSCAN 모델의 입력 값으로는 각각 배터리 셀의 내부 저항 값만을 사용하며 논문에서 제시하는 분류 방법과 전문가의 진단 결과를 비교하여 모델의 성능을 검증한다. DBSCAN 모델은 Python으로 구현하였다. 본 논문은 총 5장으로 구성된다. 2장에서는 연구 배경에 대해 기술한다. 3장에서는 데이터 전처리, 위험 셀 감지 알고리즘 그리고 검증 알고리즘에 대해 설명한다. 제 4장에서는 실험결과에 대해 설명한다. 제 5장에서는 결론과 향후 연구방향을 제시한다.

II. 배 경

1. 납축전지

납축전지는 에너지를 저장하기 위해 가역 화학 반응을 사용하는 전기 저장 장치이다. 이것은 납판 또는 그리드와 희석된 황산으로 구성된 전해질을 결합하여 전기 에너지를 잠재적 화학 에너지로 변환한 후 다시 되돌린다 [21]. 각 셀들은 약 37% w/w (5.99 Molar) 황산(H₂SO₄)의 전해액에 납 금속(Pb)과 산화 납(IV) 전극(PbO₂)을 포함하고 있다.

납축전지 사이에는 전극을 위한 두개의 납판이 있는데 그림 1과 같이 양극판과 음극판 사이 덴드라이트가 있어 양 전극의 물리적 접촉을 방지한다. 그리고 활성 물질의 방출을 통해 합선을 방지하기 위한 분리기는 플레이트 사이의 이온 흐름을 방지하고 셀의 내부저항을 증가시킨다. 방전된 상태에서 황산납으로 들어가는 전극과 전해질은 용해된 황산을 잃고 주로 물이 되며, 물의 어는점 저하로 인해 배터리는 방전되고 황산의 농도가 줄어들어 전해질이 얼기 쉬워진다. 이때 과도한 충전 전압으로 과충전을 하게면 물을

전기분해하여 산소와 수소 가스가 생성되어 폭발적 혼합물을 형성한다 [22]. 이러한 메커니즘을 가진 납축전지의 특성을 고려하여 본 논문에서는 셀의 내부저항을 사용하여 열폭주 위험이 있는 셀을 분류한다.

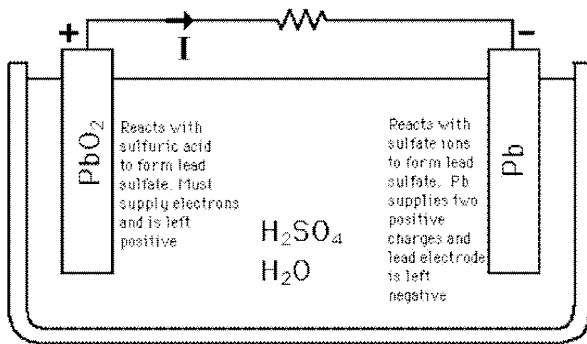


그림 1. 납축전지 [22,23]
 Figure 1. Lead-acid battery [22,23]

2. 기계학습

기계학습이란 인공지능의 한 분야로서 컴퓨터가 데이터로부터 학습하여 실행되는 알고리즘과 기술을 연구하는 분야이다 [24]. 기계학습은 레이블 유무에 따라 지도학습과 비지도 학습으로 나눌 수 있다 [25]. 지도학습은 예시를 통해 기계를 훈련한다. 주어진 특성을 사용하여 레이블을 예측하는 방식이고 대표적으로 서포트 벡터 머신, 의사결정트리, 인공신경망이 이에 해당한다 [26]. 비지도 학습은 레이블 없이 입력 값만을 사용하여 데이터들의 상관관계를 인식하여 훈련한다. 비지도 학습에는 차원 축소, 연관 규칙 학습법 그리고 클러스터링이 있다 [27,28].

3. DBSCAN

DBSCAN이란 K-평균 클러스터링 그리고 계층 클러스터링과 같이 군집간의 거리를 이용하여 클러스터링하는 방법이 아닌 밀도 기반의 클러스터링 방법이다. 최초의 밀도 기반 클러스터링 알고리즘으로 1996년 Martin Ester 등이 제안했으며 공간 및 비공간 고차원 데이터 베이스에서 노이즈가 존재하는 경우 임의의 형태의 데이터를 클러스터링 하도록 설계되었다 [29].

DBSCAN의 핵심 아이디어는 클러스터링 각 객체에 대해 주어진 반지름(Eps)안에 있는 이웃은 최소 개체 수(minPts)이상 이어야 한다는 것이다, 이는 이웃의

카디널리티가 일부 임계값을 초과해야 한다는 것을 의미한다 [30].

DBSCAN은 데이터 공간 안에서 클러스터를 만들기 위해 공간 상의 밀도를 활용한다. 예를 들어 그림 2에 표시된 데이터 집합에서는 A, B, C1, C2, C3를 동시에 클러스터링 할 수 없다. 공간 상의 밀도를 활용하여 클러스터 A, B, C 또는 클러스터 C1, C2, C3로 분류할 수 있다. 클러스터 C1, C2, C3로 클러스터링하는 경우는 A와 B를 노이즈로 분류할 수 있다.

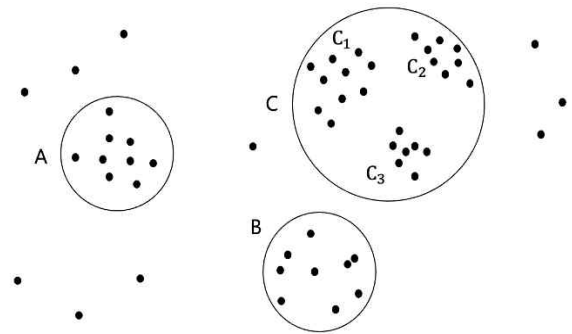


그림 2. 서로 다른 밀도에 대한 클러스터 [31,32]
 Figure 2. Clusters for different densities [31,32]

논문에서는 sklearn의 DBSCAN 라이브러리를 사용하여 클러스터링 모델을 구축하였다. DBSCAN의 파라미터는 데이터 셋에 적합한 값을 찾기 위해 실험을 통해 설정하였다.

III. 연구방법

시계열 데이터를 k-shape 클러스터링, KNN 클러스터링, k-평균 클러스터링 그리고 계층적 클러스터링로 분류한 연구들이 있다 [33-36].

본 논문에서 사용한 BMS 데이터에는 많은 양의 노이즈가 포함되어 있다. k-shape 클러스터링, KNN 클러스터링, k-평균 클러스터링 그리고 계층적 클러스터링 등의 알고리즘은 몇 개의 클러스터로 분류할 지에 대해 미리 설정해야 하는데 이러한 설정이 단순한 노이즈를 이상이 있다고 판단하여 정상 셀과 다른 클러스터로 분류하는 경우가 있다. 따라서 클러스터의 개수를 지정하여 분류하는 방식의 알고리즘은 노이즈가 많은 데이터를 분류할 때 적합하지 않다 [37]. 그렇기 때문에 클러스터의 개수를 지정하지 않는 DBSCAN을

사용하여 이 문제를 해결하였다. Mean-shift 또한 클러스터의 개수를 지정하지 않고 클러스터링하는 알고리즘이지만 DBSCAN에 비해 이상치의 영향을 더 받기 때문에 노이즈가 많은 데이터에는 적합하지 않다.

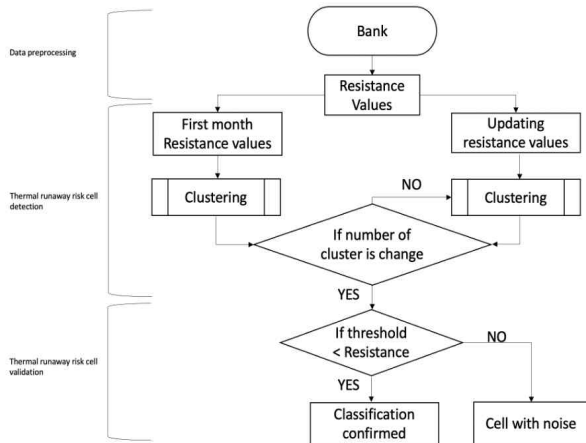


그림 3. 제안된 열폭주 위험 셀 탐지 알고리즘과 열폭주 위험 셀 검증 알고리즘 순서도

Figure 3. Flowchart of the proposed thermal runaway hazard cell detection algorithm and thermal runaway hazard cell verification algorithm

그림 3은 DBSCAN과 3 시그마 규칙을 사용하여 배터리 뱅크 내 열폭주 위험 셀을 검출하는 과정을 나타낸 순서도이다. 그림 3과 같이 BMS에서 측정된 96개의 셀 데이터가 저장된 뱅크에서 저항 값만 추출한다. 그 후 월 별로 데이터 셋을 만들어 클러스터링한다. 이전 달의 클러스터링 결과와 비교했을 때 클러스터의 개수가 달라지면 열폭주 위험이 있는 셀이 새롭게 생긴 클러스터에 있다고 가정한다. 3 시그마 규칙을 임계값으로 설정하여 열폭주 위험이 있다고 탐지된 셀의 평균 저항값이 임계값보다 높으면 열폭주 위험 셀이 맞다고 검증한다.

1. 데이터 전처리

Watson의 BMS에서 배터리의 저항, 온도 그리고 내부 저항을 측정된 데이터를 사용하였다. 본 논문에서는 각각 다른 여섯 개의 데이터 셋을 사용하였다. 데이터 셋 01, 데이터 셋 04 그리고 데이터 셋 06은 25개월간 측정된 데이터이고 데이터 셋 02와 데이터 셋 03은 33개월간 측정된 데이터이다. 그리고 데이터 셋 05는 37개월간

측정된 데이터이다. 각 데이터 셋에는 내부저항이 급격하게 상승하는 열폭주 위험 셀이 포함되어 있다.

모든 데이터 셋은 하루에 두 번 간격으로 측정되었다. 그림 3과 같이 내부저항을 DBSCAN의 입력 값으로 사용하기 때문에 내부저항 외의 값은 삭제한다. 또한 월 별로 열폭주 위험 셀을 탐지하기 때문에 한 달 단위로 데이터 셋을 분리했다.

2. 열폭주 위험 셀 탐지 알고리즘

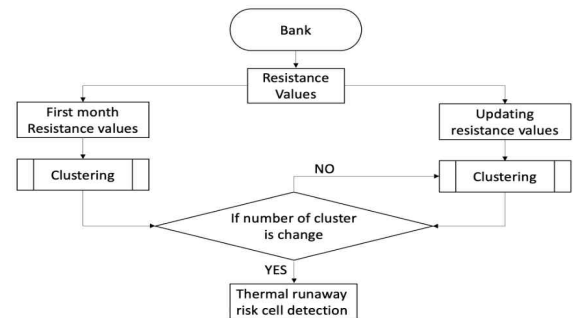


그림 4. 열폭주 위험 셀 탐지 알고리즘 순서도

Figure 4. Thermal runaway hazard cell detection algorithm flowchart

열폭주 위험 셀 탐지 알고리즘은 그림 4과 같이 한 달 단위로 자른 데이터 셋들을 순서대로 입력하여 첫 번째 달의 클러스터링 결과를 기준으로 다음 달의 클러스터링 결과를 비교한다. 예를 들어 지난달은 모든 배터리 셀이 한 개의 클러스터로 분류 되었지만 이번 달은 배터리 셀이 두 개의 클러스터로 분류 되었으면 새로 생긴 클러스터에 열폭주 위험 셀이 존재한다고 판단한다.

데이터에 적합한 DBSCAN의 파라미터를 찾기 위해 그리드 서치를 활용하였다. 표 1과 표 2는 데이터 셋 01과 데이터 셋 02를 DBSCAN으로 클러스터링하여 이상 셀을 탐지한 시점을 나타낸 것이다. 표1에 따르면 eps가 1이고 min_sample이 20인 DBSCAN은 데이터 셋 01의 이상 셀을 22개월 만에 탐지하였으며 eps가 0.5이고 min_sample이 5 또는 10인 DBSCAN은 데이터 셋 01의 고장 난 셀을 21개월 만에 탐지하였다. 또한 표2에 따르면 eps가 0.5이고 min_sample이 10인 DBSCAN은 데이터 셋 02의 고장 난 셀을 27개월 만에 탐지하였다. 두 데이터 셋을 빠른 시점에 탐지하는 모델의 파라미터 값을 찾아야하기 때문에 eps는 0.5, min_sample은 10으로 설정했다.

표 1. Dataset 01에 대한 DBSCAN 매개 변수별 탐지 시점
 Table 1. Detection point by DBSCAN parameter for Dataset 01

eps	min_sample	state	Detection point
0.5	5	O	after 21month
0.5	10	O	after 21month
0.5	15	X	-
0.5	20	X	-
1	5	O	after 23month
1	10	O	after 23month
1	15	O	after 23month
1	20	O	after 22month
1.5	5	O	after 25month
1.5	10	O	after 25month
1.5	15	O	after 25month
1.5	20	O	after 25month
2	5	X	-
2	10	X	-
2	15	X	-
2	20	X	-

표 2. Dataset 02에 대한 DBSCAN 매개 변수별 탐지 시점
 Table 2. Detection point by DBSCAN parameter for Dataset 02

eps	min_sample	state	Detection point
0.5	5	X	-
0.5	10	O	after 27month
0.5	15	O	after 27month
0.5	20	X	-
1	5	O	after 32month
1	10	X	-
1	15	X	-
1	20	X	-
1.5	5	X	-
1.5	10	X	-
1.5	15	X	-
1.5	20	X	-
2	5	X	-
2	10	X	-
2	15	X	-
2	20	X	-

3. 열폭주 위험 셀 검증 알고리즘

열폭주 위험 셀 탐지 알고리즘을 통해 분류된 셀이

열폭주 위험이 있는 셀이 아닌 노이즈가 심한 정상 셀일 가능성도 존재한다. 그렇기 때문에 그에 대한 검증과정이 필요하다. 본 논문에서 열폭주 위험 셀 검증의 기준은 3 시그마 규칙을 사용한다. 데이터 전처리를 통해 얻은 첫 번째 달부터 클러스터링을 진행한 달까지 내부저항의 평균과 표준편차를 구하여 평균으로부터 3σ 만큼의 값을 임계점으로 설정하였다. 셀이 열폭주 위험 셀로 감지된 달의 월 평균 내부 저항 값과 임계값을 비교하여 임계점보다 낮을 경우 노이즈에 의해 새 클러스터로 분류되었고 열폭주 위험 셀이 아니라고 판단한다. 반면 임계점보다 높으면 열폭주 위험 셀이 맞다고 검증한다.

IV. 실험 결과

본 논문에서 사용되는 데이터 셋은 총 6개이며 데이터의 길이는 약 2년에서 3년이다. 데이터 셋 01, 데이터 셋 02, 데이터 셋 03, 데이터 셋 04, 데이터 셋 05는 실제 BMS에서 추출한 데이터이고 데이터 셋 06은 인위적으로 열폭주 위험 셀과 정상 셀을 만든 데이터이다. 이러한 데이터 셋을 사용하여 열폭주 위험 탐지 알고리즘과 열폭주 위험 검증 알고리즘의 성능을 판단하고 다양한 데이터 셋에 대해 성능이 균일한가를 알아본다.

1. Dataset 01

약 2년간 실제 BMS를 통해 측정된 데이터 셋 01을 전처리한 후 논문에서 제안한 두 가지 알고리즘을 실행했다. 데이터 셋 01을 전처리하여 만들어진 25개월의 데이터 셋 중 첫 번째 달 데이터에 대한 클러스터링 결과는 그림 5와 같이 모든 셀들이 한 개의 클러스터로 클러스터링 되었다. 이는 처음 가동된 배터리 셀들은 모두 정상 값을 가지기 때문이다.

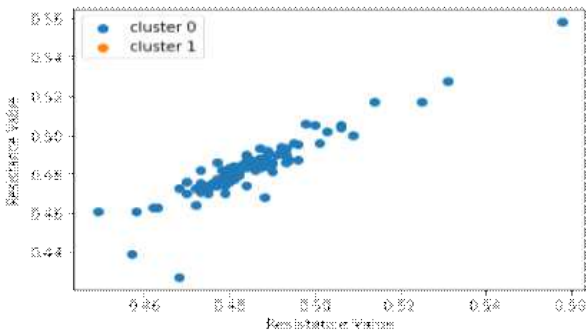


그림 5. 첫 달 데이터 셋에 대한 클러스터링 결과
Figure 5. The result of clustering for the first month dataset

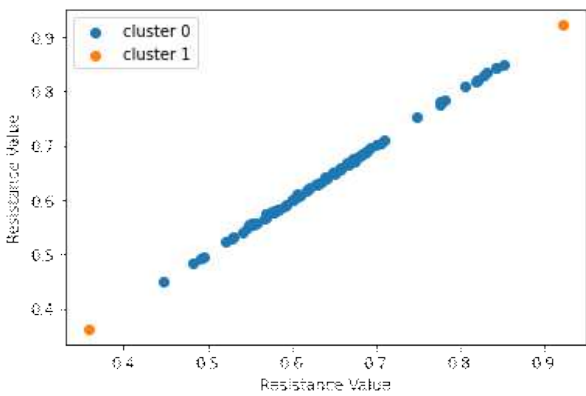


그림 6. 21번째 달 데이터 셋에 대한 클러스터링 결과
Figure 6. The result of clustering of 21st month dataset

그 후 첫 달의 클러스터링 결과를 기준으로 한 달 간격으로 데이터를 업데이트하며 매달 클러스터링 결과를 비교한다. 첫 달부터 20번째 달의 클러스터링 결과가 모두 하나의 클러스터로 분류되었다. 그러나 그림 6과 같이 21번째 달의 결과부터 2개의 클러스터로 나누어지는 것을 확인할 수 있었다. 이때 하나의 클러스터에서 다른 클러스터로 클러스터링 된 77번셀과 96번 셀은 열폭주 위험 셀로 탐지된다. 열폭주 위험 셀을 탐지한 이후로는 해당되는 셀들이 정말 열폭주 위험 셀인지 검증한다.

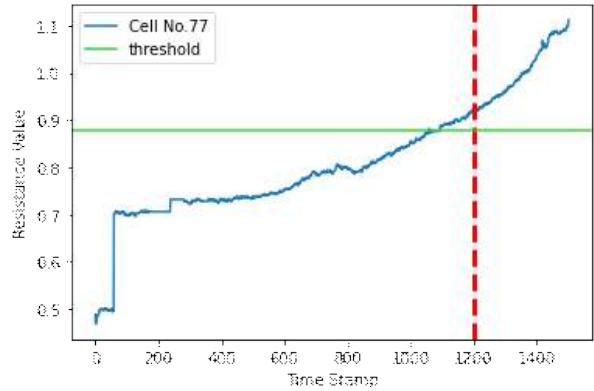


그림 7. 77번 셀의 저항 값, 임계값 그리고 탐지된 시점
Figure 7. Resistance values, threshold and time of detection of Cell 77

그림 7의 초록 실선은 3 시그마 규칙에 의해 설정된 임계값을 나타내고 빨간 점선은 열폭주 위험 셀 탐지 알고리즘이 77번 셀을 열폭주 위험 셀로 탐지한 시점을 나타낸다. 열폭주 위험 셀로 탐지되었던 시점에서 77번 셀의 평균 내부저항 값이 임계값보다 높기 때문에 열폭주 위험 셀로 분류된다.

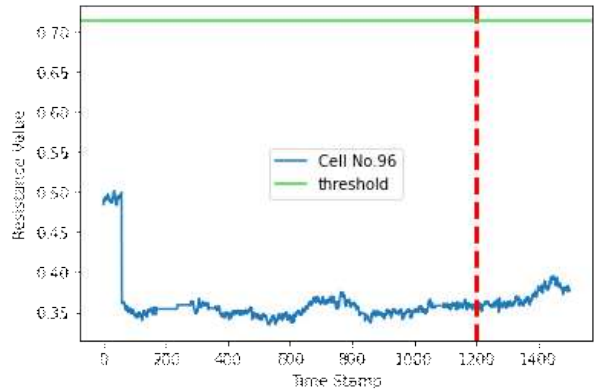


그림 8. 96번 셀의 저항 값, 임계값 그리고 탐지된 시점
Figure 8. Resistance values, threshold and time of detection of Cell 96

그림 8의 경우 열폭주 위험 셀로 탐지된 시점에 셀의 내부저항 값이 임계값보다 낮기 때문에 열폭주 위험 셀이 아니라고 분류했다. 즉 96개의 배터리 셀 중 열폭주 위험 셀이었던 77번 셀과 96번 셀 중 77번 셀은 열폭주 위험 셀 탐지 알고리즘과 열폭주 위험 셀 검증 알고리즘에 의하여 열폭주 위험 셀로 분류했다.

데이터 셋 01의 경우 77번 셀이 열폭주 위험 셀이라는 전문가의 진단이 있었다. 그렇기 때문에 본 논문에서 제안한 알고리즘이 열폭주 위험 셀을 잘 검출하였음을 확인하였다.

2. Dataset 02

약 3년간 BMS를 통해 측정된 데이터 셋 02를 전처리한 후 논문에서 제안한 두가지 알고리즘을 실행했다. 데이터 셋 02을 전처리하여 만들어진 33개의 데이터 셋 중 첫 번째 데이터 셋을 클러스터링 한 결과는 그림 9와 같이 모든 셀들이 한 개의 클러스터로 클러스터링 되었다.

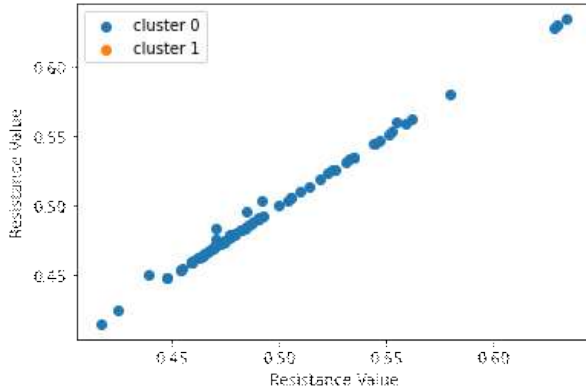


그림 9. 첫 달 데이터 셋에 대한 클러스터링 결과
 Figure 9. The result of clustering for the first month dataset

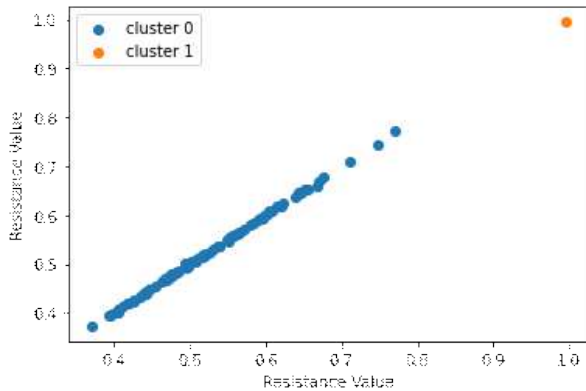


그림 10. 8번째 달 데이터 셋에 대한 클러스터링 결과
 Figure 10. The result of clustering for the 8th month dataset

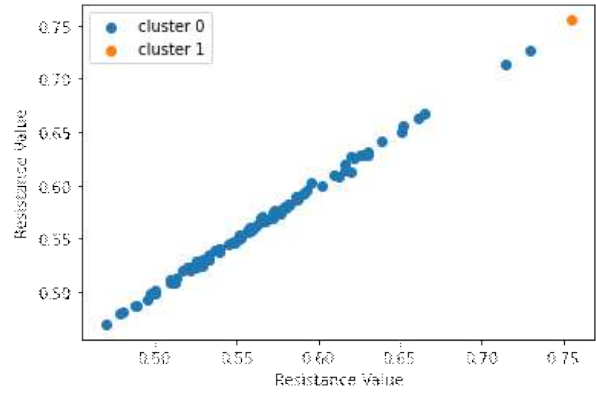


그림 11. 28번째 달 데이터 셋에 대한 클러스터링 결과
 Figure 11. The result of clustering for the 28th month dataset

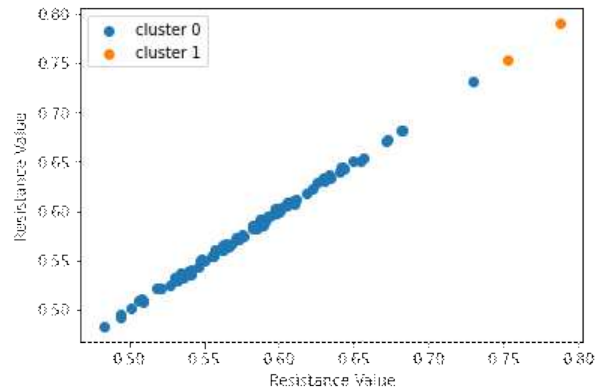


그림 12. 30번째 달 데이터 셋에 대한 클러스터링 결과
 Figure 12. The result of clustering for the 30th month dataset

그림 10은 8번째 달 데이터 셋 02를 DBSCAN으로 클러스터링한 결과다. 첫 번째 달부터 7번째 달까지 한 개의 클러스터로 클러스터링 되었지만 8번째 달부터 두 개의 클러스터로 클러스터링 되었다. 이 때 다른 클러스터로 클러스터링 된 45번째 셀은 열폭주 위험 셀 탐지 알고리즘에 의해 열폭주 위험 셀로 탐지되었다.

또한 그림 11은 28번째 달 데이터 셋을 클러스터링 한 결과이다. 9번째 달부터 27번째 달까지 하나의 클러스터로 클러스터링 되었지만 28번째 달부터 85번 셀에 의해 두 개의 클러스터로 클러스터링 되었으니 85번 셀은 열폭주 위험 셀 탐지 알고리즘에 의해 열폭주 위험 셀로 탐지되었다.

28번째 달부터 29번째 달까지 85번 셀이 다른 클러스터로 클러스터링 되었으며 그림 12와 같이 30번째 달에는 29번 셀과 85번 셀이 함께 클러스터링 되었다. 즉 열폭주 위험 셀 탐지 알고리즘에 의해 45번

셀, 85번 셀 그리고 29번 셀이 열폭주 위험 셀로 탐지되었다.

데이터 셋 02에는 노이즈가 심한 셀들이 많기 때문에 세 개의 셀이 열폭주 위험 셀 탐지 알고리즘에 의해 열폭주 위험 셀로 탐지되었고 실제 열폭주 셀이 맞는지 열폭주 위험 셀 검증 알고리즘을 사용하여 확인해야 한다.

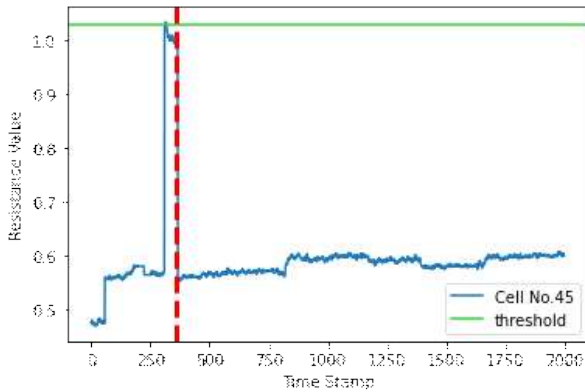


그림 13. 45번 셀의 저항 값, 임계값 그리고 탐지된 시점
Figure 13. Resistance values, threshold and time of detection of Cell 45

그림 13의 초록 실선은 임계값을 의미하며 빨간 점선은 45번 셀이 열폭주 위험 셀로 탐지된 시점을 나타낸다. 45번 셀은 열폭주 위험 셀로 탐지된 시점에 내부저항이 임계값보다 높지 않고 평균 내부저항 역시 높지 않았기 때문에 열폭주 위험 셀 검증 알고리즘에 의해 열폭주 위험 셀이 아니라고 판단했다.

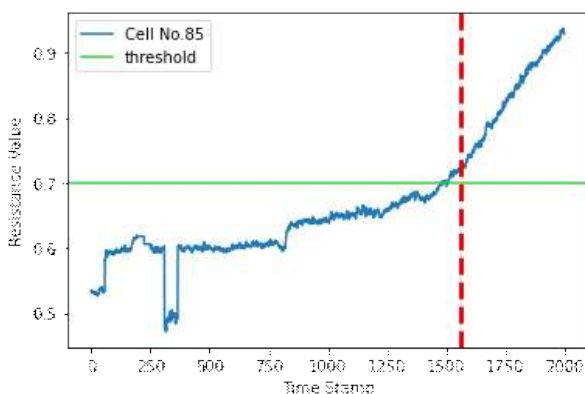


그림 14. 85번 셀의 저항 값, 임계값 그리고 탐지된 시점
Figure 14. Resistance values, threshold and time of detection of Cell 85

28번째 달에 85번 셀이 열폭주 위험 셀로 탐지되었던 시점, 임계값 그리고 85번 셀의 내부저항을 그림 14에

나타냈다. 85번 셀 같은 경우 열폭주 위험 셀로 탐지된 시점에 내부 저항 값이 임계값보다 높고 평균 내부 저항 또한 임계값보다 높았기 때문에 열폭주 위험 셀 검증 알고리즘에 의해 열폭주 위험 셀이 맞다고 분류했다.

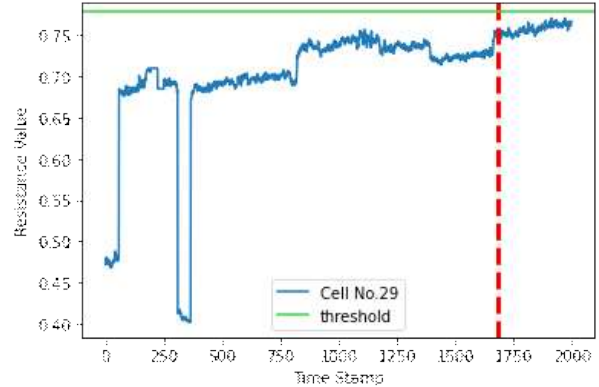


그림 15. 29번 셀의 저항 값, 임계값 그리고 탐지된 시점
Figure 15. Resistance values, threshold and time of detection of Cell 29

30번째 달에 열폭주 위험 셀로 탐지된 29번 셀의 저항 값, 임계값 그리고 탐지된 시점을 그림 15으로 나타냈다. 29번 셀의 경우 내부 저항 값이 임계값보다 높지 않고 그렇기 때문에 평균 내부 저항도 임계값보다 낮았다. 따라서 열폭주 위험 셀 검증 알고리즘에 의해 29번 셀은 열폭주 위험 셀로 탐지되었음에도 불구하고 열폭주 위험 셀이 아니라고 판단했다.

데이터 셋 02의 경우 85번 셀이 열폭주 위험 셀이라는 전문가의 진단이 있었다. 그렇기 때문에 본 논문에서 제안한 알고리즘이 노이즈가 있는 정상 셀과 열폭주 위험 셀을 잘 분류하여 검출했다고 볼 수 있다.

3. Dataset 03

데이터 셋 03은 약 3년간 실제 BMS를 통해 측정된 데이터이며 본 논문에서 제안한 두 가지 알고리즘을 통해 열폭주 위험 셀을 탐지하는 실험을 진행하였다.

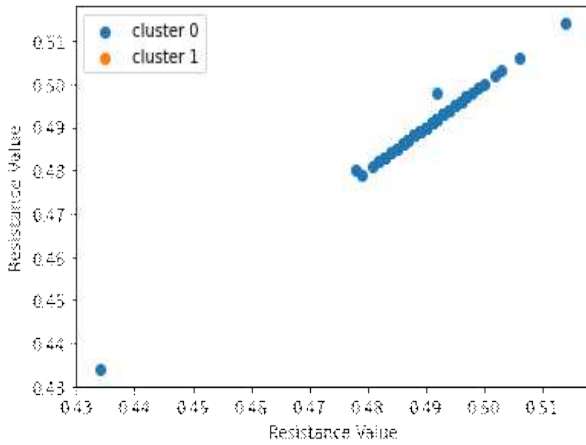


그림 16. 첫 번째 달 데이터 셋에 대한 클러스터링 결과
 Figure 16. The result of clustering for the 1st month dataset

그림16은 첫 번째 달 데이터 셋을 DBSCAN으로 클러스터링 했을 때 결과를 나타낸 것이다. 첫 번째 달 데이터 셋에는 모든 셀의 내부저항 값이 정상 범위 안에 있기 때문에 하나의 클러스터로 클러스터링 되었다.

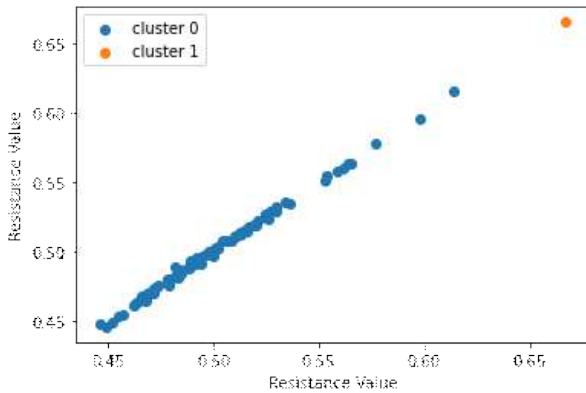


그림 17. 31 번째 달 데이터 셋에 대한 클러스터링 결과
 Figure 17. The result of clustering for the 31th month dataset

그림17은 31번째 달 데이터 셋을 DBSCAN으로 클러스터링한 결과다. 달별로 클러스터링 했을 때 30 번째 달까지 클러스터 수의 변화가 없었지만 31번째 달부터 변화가 생기기 시작했다. 따라서 다른 클러스터로 클러스터링된 22번 셀을 열폭주 위험이 있는 셀이라고 판단한다. 다음으로는 22번 셀에 대해 열폭주 위험 셀 검증 알고리즘을 실행하여 실제 열폭주 위험 셀인지에 대하여 검증한다.

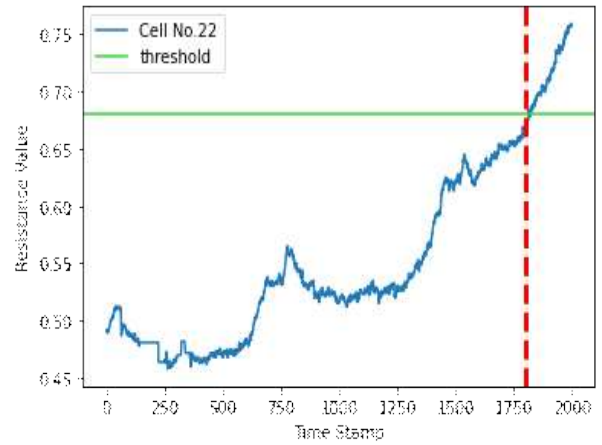


그림 18. 22번 셀의 저항 값, 임계값 그리고 탐지된 시점
 Figure 18. Resistance values, threshold and time of detection of Cell 22

그림18은 30 번째 달에 열폭주 위험 셀로 탐지된 22 번 셀의 저항 값, 임계값 그리고 탐지된 시점을 나타낸 그래프이다. 열폭주 위험 셀로 탐지된 시점에서 월 평균 내부저항 값이 임계값보다 높았기 때문에 22번 셀은 열폭주 위험 셀이 맞다고 판단했다.

데이터 셋 03에 포함된 열폭주 위험 셀은 22번 셀이라는 전문가의 진단이 있었고 본 논문에서 제안한 두가지 알고리즘으로 잘 분류하였다.

4. Dataset 04

데이터 셋 04은 약 2년간 BMS를 통해 측정된 데이터이며 본 논문에서 제안한 두 가지 알고리즘을 통해 열폭주 위험 셀을 탐지하는 실험을 진행하였다.

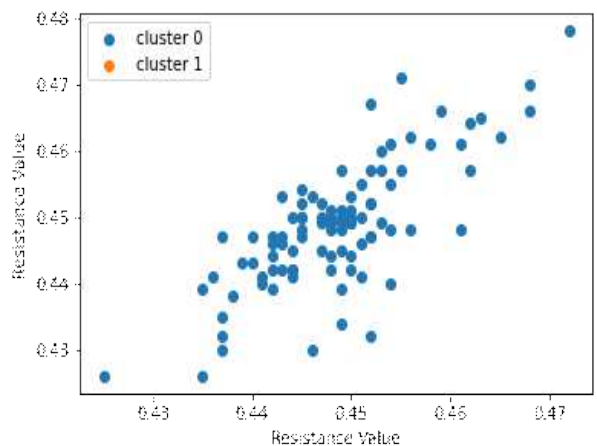


그림 19. 첫 번째 달 데이터 셋에 대한 클러스터링 결과
 Figure 19. The result of clustering for the 1st month dataset

그림 19은 데이터 셋 04의 첫 번째 달 데이터 셋을

DBSCAN으로 클러스터링 했을 때 결과를 나타낸 것이다. 첫 번째 달 데이터 셋의 모든 셀들은 하나의 클러스터로 클러스터링 되었기 때문에 이상이 없다고 판단한다.

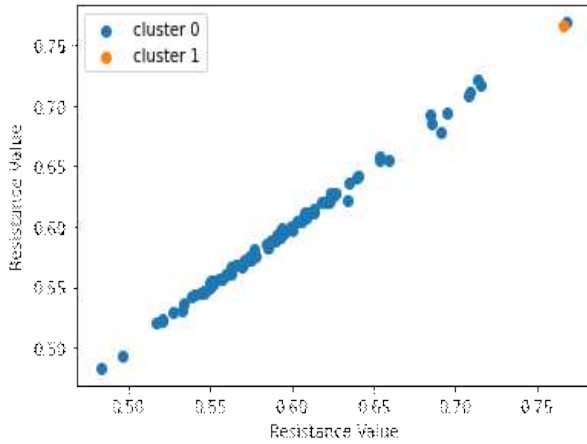


그림 20. 24 번째 달 데이터 셋에 대한 클러스터링 결과
Figure 20. The result of clustering for the 24th month

그림 20은 데이터 셋 04의 24 번째 달 데이터 셋을 DBSCAN으로 클러스터링한 결과이다. 첫 번째 달부터 23번째 달까지 클러스터의 변화가 없었지만 24번째 달부터 48번 셀이 다른 클러스터로 클러스터링 된 것을 확인할 수 있고 해당 셀에 대하여 열폭주 위험 셀이 맞는지 열폭주 위험 셀 검증 알고리즘을 실행시킨다.

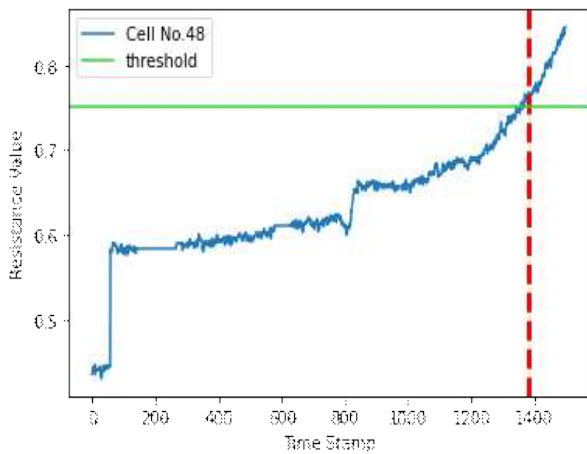


그림 21. 48번 셀의 저항 값, 임계값 그리고 탐지된 시점
Figure 21. Resistance values, threshold and time of detection of Cell 48

그림 21은 48번 셀의 내부저항 값, 임계값 그리고 탐지된 시점을 나타내고 있다. 48번 셀의 평균 내부저항 값은 임계값보다 높기 때문에 열폭주 위험 셀이 맞다고

판단하였고 실제 데이터 셋 04의 열폭주 위험 셀은 48번 셀이었다. 따라서 두 가지 알고리즘에 의하여 열폭주 위험 셀을 잘 탐지할 수 있었다.

5. Dataset 05

데이터 셋 05는 실제 BMS에서 약 3년간 측정하였으며 본 논문에서 제안한 두 가지 알고리즘에 의해 열폭주 위험 셀 탐지 실험을 진행한다.

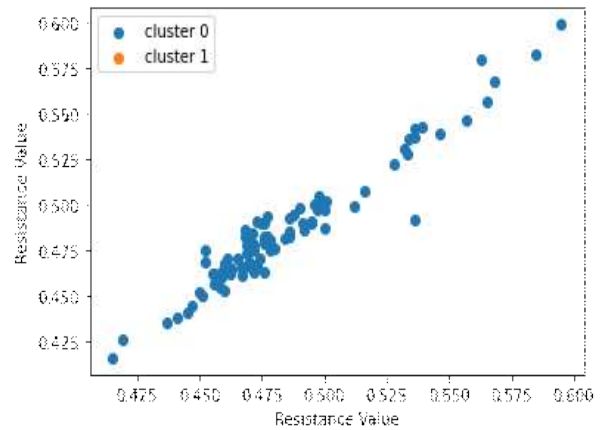


그림 22. 첫 번째 달 데이터 셋에 대한 클러스터링 결과
Figure 22. The result of clustering for the 1st month dataset

그림 22는 데이터 셋 05의 첫 번째 데이터 셋을 DBSCAN으로 클러스터링한 결과이다. 모든 셀이 같은 클러스터로 클러스터링되었기 때문에 열폭주 위험 셀이 없다고 판단한다.

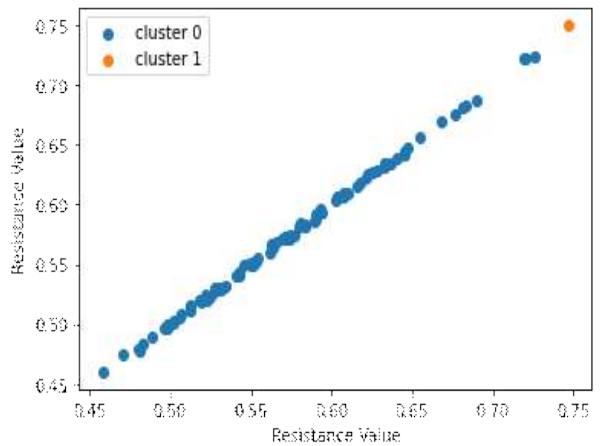


그림 23. 35 번째 달 데이터 셋에 대한 클러스터링 결과
Figure 23. The result of clustering for the 35th month dataset

그림 23은 데이터 셋 05의 35번째 달 데이터 셋을

클러스터링한 결과이다. 첫 번째 달부터 34번째 달까지 하나의 클러스터로 클러스터링 되었지만 35번째 달부터 20번 셀이 새로운 클러스터에 클러스터링 되었다. 열폭주 위험 셀 탐지 알고리즘에 의해 20번 셀이 열폭주 위험 셀로 탐지되었고 해당 셀에 대하여 열폭주 위험 셀이 맞는지 열폭주 위험 셀 검증 알고리즘을 사용한다.

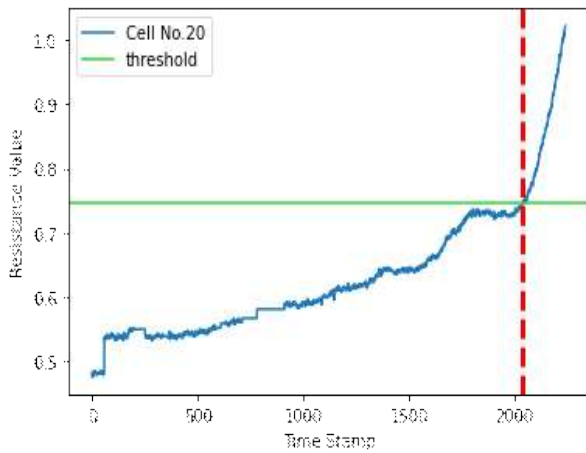


그림 24. 20번 셀의 저항 값, 임계값 그리고 탐지된 시점
 Figure 24. Resistance values, threshold and time of detection of Cell 20

그림 24를 통해 20번 셀의 내부저항 값, 임계점 그리고 열폭주 위험 셀로 탐지된 시점을 확인할 수 있다. 20번 셀의 경우 월 평균 내부저항 값이 임계값보다 높기 때문에 열폭주 위험 셀로 분류하였다. 데이터 셋 05 중 20번 셀이 전문가가 진단한 열폭주 위험 셀이었고 본 논문에서 제안한 두 알고리즘은 열폭주 위험 셀을 성공적으로 검출하였다.

6. Dataset 06

데이터 셋 06은 인위적으로 만들었으며 하나의 열폭주 위험 셀을 포함한 데이터 셋이다. 인위적으로 만든 데이터도 본 논문에서 제안하는 두 가지 알고리즘을 통해 열폭주 위험 셀을 탐지하는지 실험한다.

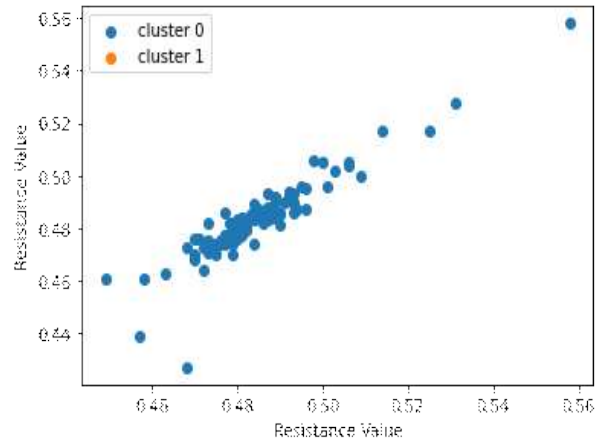


그림 25. 첫 번째 달 데이터 셋에 대한 클러스터링 결과
 Figure 25. The result of clustering for the 1st month dataset

그림25는 데이터 셋 06의 첫 번째 데이터 셋을 클러스터링한 결과이다. 모든 셀이 하나의 클러스터로 클러스터링 되었기 때문에 열폭주 위험 셀이 없다.

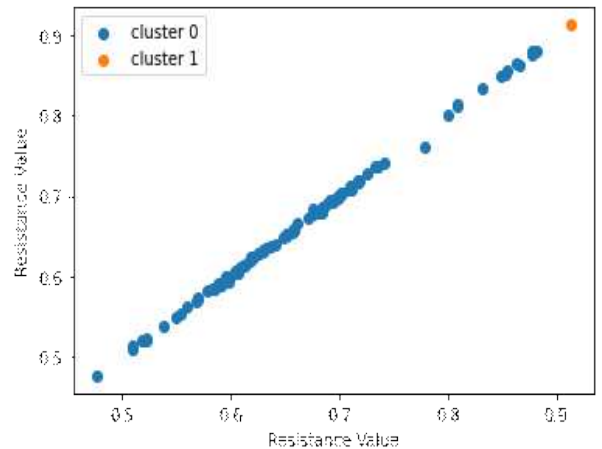


그림 26. 25 번째 달 데이터 셋에 대한 클러스터링 결과
 Figure 26. The result of clustering for the 25th month dataset

첫 번째 달부터 24번째 달까지 하나의 클러스터로 클러스터링 되었지만 그림26과 같이 25번째 달부터 두 개의 클러스터로 클러스터링 되었다. 따라서 새로운 클러스터로 클러스터링 된 35번 셀은 열폭주 위험이 있다고 의심된다. 해당 셀을 열폭주 위험 셀 검증 알고리즘을 사용하여 실제 열폭주 위험이 있는지에 대해 검증한다.

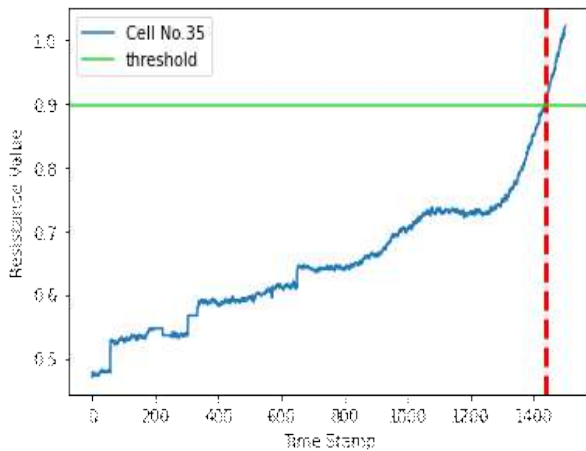


그림 27. 35번 셀의 저항 값, 임계값 그리고 탐지된 시점
Figure 27. Resistance values, threshold and time of detection of Cell 35

그림27는 35번 셀의 내부저항과 임계값 그리고 열폭주 위험 셀 탐지 시점을 나타낸다. 열폭주 위험 셀 탐지 알고리즘에 의해 탐지된 시점의 월평균 내부저항 값이 임계값보다 높기 때문에 35번 셀은 열폭주 위험 셀이 맞다고 분류한다. 데이터 셋 06은 인위적으로 생성한 데이터로써 35번 셀을 열폭주 위험 셀로 만들었다. 따라서 본 논문에서 제안한 두 알고리즘은 데이터 셋 06의 열폭주 위험 셀을 성공적으로 분류했다.

V. 결 론

배터리 관리 시스템을 통해 열폭주 위험이 있는 셀을 탐지하는 방법으로 본 논문에서는 DBSCAN을 사용한 열폭주 위험 셀 탐지 알고리즘과 통계적 방법인 3 시그마 규칙을 활용한 검증 알고리즘으로 열폭주 위험 셀 분류하는 것을 제안한다. 실제 BMS에서 측정된 납축전지 배터리의 저항, 전압 그리고 온도 데이터가 포함된 데이터 셋 5개와 인위적으로 생성한 데이터 셋 1개를 사용하였다. 본 논문에서 제안한 열폭주 위험 셀 탐지 알고리즘, 열폭주 위험 셀 검증 알고리즘을 사용하여 6개의 데이터 셋의 열폭주 위험 셀을 분류하는 것에 성공했다. 그러나 서론에 언급한 논문[14]의 방식으로는 본 논문에서 사용한 6개 데이터 셋 중 오직 3개의 데이터 셋에서 열폭주 위험 셀을 분류하는 것에 성공하였다. 또한 그리드 서치를

활용하여 적절한 DBSCAN의 파라미터 값을 찾음으로써 열폭주 위험 셀 탐지 알고리즘이 빠르게 실행되도록 했다. 이러한 결과를 통해 본 논문에서 제안한 방법이 효율적이고 성능이 개선되었다는 것을 확인할 수 있었다.

향후 연구에서는 이상이 있는 셀을 검출하는 것뿐만 아니라 이상이 생길 것 같은 셀을 예측하는 연구가 필요하다고 제안한다.

References

- [1] Wang, Z. H., Hendrick, Horng, G. J., Wu, H. T., & Jong, G. J., "A prediction method for voltage and lifetime of lead-acid battery by using machine learning," *Energy Exploration & Exploitation*, Vol. 38, No.1, pp.310-329, 2020. DOI :10.1177/01445987198811223
- [2] Voronov, S., Frisk, E., Krysander, M., "Lead-acid battery maintenance using multilayer perceptron models," *proc. of the 2018 IEEE International Conference on Prognostics and Health Management (ICPHM)*, pp. 1-8, 2018. DOI : 10.1109/ICPHM.2018.8448472
- [3] Weighall, M. J., "Techniques for jar formation of valve-regulated lead-acid batteries," *Journal of power sources*, Vol.116, No.1-2, pp. 219-231, 2003. DOI : 10.1016/S0378-7753(02)00706-1
- [4] Matrakova, M., Pavlov, D., "Thermal analysis of lead-acid battery pastes and active materials," *Journal of power sources*, Vol.158, No.2, pp. 1004-1011, 2006. DOI : 10.1016/j.jpowsour.2005.11.007
- [5] Liu, X., & Wang, W., "VRLA battery system reliability and proactive maintenance," *proc. of the Intelec 2010*, pp. 1-7, 2010. DOI :10.1109/INTLEC.2010.5525726
- [6] Rand, D. A. J., Holden, L. S., May, G. J., Newnham, R. H., Peters, K., "Valve-regulated lead/acid batteries," *Journal of Power Sources*, Vol.59, No.1-2, pp. 191-197, 1996. DOI : 10.1016/0378-7753(96)02322-1
- [7] Culpin, B., "Thermal runaway in valve-regulated lead-acid cells and the effect of separator structure," *Journal of Power Sources*, Vol.133, No.1, pp. 79-86, 2004. DOI : 10.1016/j.jpowsour.2003.09.078
- [8] McMenamin, D., "Thermal Runaway-A system

- solution for a system problem,” proc. of the Fourteenth International Telecommunications Energy Conference-INTELEC’92, pp. 18–21, 1992. DOI : 10.1109/INTLEC.1992.268469
- [9] Rutledge, W. T., Bowers, R. J., “Electrical energy distributions in VR battery applications that can trigger thermal runaway,” proc. of the Intelec 94 , pp. 168–171, 1994. DOI : 10.1109/INTLEC.1994.396659
- [10]Culpin, B., & Rand, D. A. J., “Failure modes of lead/acid batteries,” Journal of power sources, Vol.36, No.4, pp. 415–438, 1991. DOI : 10.1016/0378-7753(91)80069-A
- [11]Diao, W., Naqvi, I. H., Pecht, M., “Early detection of anomalous degradation behavior in lithium-ion batteries,” Journal of Energy Storage, Vol. 32 pp. 101710. DOI : 10.1016/j.est.2020.101710
- [12]Ossai, C. I., Egwutuoha, I. P., “Anomaly Detection and Extra Tree Regression for Assessment of the Remaining Useful Life of Lithium-Ion Battery,” Proc. of the International Conference on Advanced Information Networking and Applications, pp. 1474–1488, 2020.
- [13]Samanta, A., Chowdhuri, S., Williamson, S. S., “Machine learning-based data-driven fault detection/diagnosis of lithium-ion battery: A critical review,” Electronics, Vol. 10, No.11, pp. 1309, 2021. DOI : 10.3390/electronics10111309
- [14]Çelik, M., Dadaşer-Çelik, F., Dokuz, A. Ş., “Anomaly detection in temperature data using DBSCAN algorithm,” proc. of the 2011 international symposium on innovations in intelligent systems and applications, pp. 91–95, 2011.
- [15]Chesnokov, M. Y., “Time series anomaly searching based on DBSCAN ensembles,” Scientific and Technical Information Processing, Vol. 46, No.5, pp.299–305, 2019.
- [16]Li, D., Zhang, Z., Liu, P., Wang, Z., “DBSCAN-based thermal runaway diagnosis of battery systems for electric vehicles,” Energies, Vol. 12, No. 15, pp. 2977, 2019. DOI : 10.3390/en12152977
- [17]Feng, X., Zhang, X., Xiang, Y., “An inconsistency assessment method for backup battery packs based on time-series clustering,” Journal of Energy Storage, Vol.31 pp. 101666, 2020. DOI : 10.1016/j.est.2020.101666
- [18]Li, D., Zhang, Z., Liu, P., Wang, Z., & Zhang, L. “Battery fault diagnosis for electric vehicles based on voltage abnormality by combining the long short-term memory neural network and the equivalent circuit model,” IEEE Transactions on Power Electronics, Vol. 36 No. 2, pp. 1303–1315. DOI : 10.1109/TPE.2021.3117841
- [19]Haider, S. N., Zhao, Q., Li, X., “Data driven battery anomaly detection based on shape based clustering for the data centers class,” Journal of Energy Storage, Vol. 29, pp. 101479, 2020. DOI : 10.1016/j.est.2020.101479
- [20]Choi, H. S., Choi, J. W., Whangbo, T. K., “Design and Development of a Battery State of Health Estimation Model for Efficient Battery Monitoring Systems,” Sensors, Vol. 22, No. 12, pp. 444, 2022. DOI : 10.3390/s22124444
- [21]Ceraolo, M., “New dynamical models of lead-acid batteries,” IEEE transactions on Power Systems, Vol. 15, No. 4, pp. 1184–1190, 2000. DOI : 10.1109/59.898088
- [22]Moubayed, N., Kouta, J., El-Ali, A., Dernayka, H., Outbib, R., “Parameter identification of the lead-acid battery model,” proc. of the 2008 33rd IEEE Photovoltaic Specialists Conference, pp. 1–6, 2008.
- [23]Hyperphysics. [Online]. <http://hyperphysics.phy-astr.gsu.edu/hbase/electric/leadacid.html>, (download 2021, December 17)
- [24]Samuel, A. L., “Some studies in machine learning using the game of checkers II—Recent progress,” IBM Journal of research and development, Vol. 11, No. 6, pp. 601–617, 1967. DOI : 10.1147/rd.116.0601
- [25]El Naqa, I., Murphy, M.J.. Machine Learning in Radiation Oncology, pp. 3–11, Newyork, 2015.
- [26]Caruana, R., Niculescu-Mizil, A., “An empirical comparison of supervised learning algorithms,” proc. of the 23rd international conference on Machine learning, pp. 161–168, 2006.
- [27]Van Der Maaten, L., Postma, E., Van den Herik, J., “Dimensionality reduction: a comparative,” J Mach Learn Res, Vol. 10, No. 66–71, pp. 13, 2009.
- [28]Kumbhare, T. A., Chobe, S. V., “An overview of association rule mining algorithms,” International Journal of Computer Science and Information Technologies, Vol. 5, No. 1, pp. 927–930, 2014.
- [29]Ester, M., Kriegel, H. P., Sander, J., Xu, X., “A density-based algorithm for discovering clusters in large spatial databases with noise,” proc. of the kdd, pp. 226–231, 1996.
- [30]Khan, K., Rehman, S. U., Aziz, K., Fong, S., Sarasvady, S., “DBSCAN: Past, present and future,” proc. of the fifth international conference

- on the applications of digital information and web technologies (ICADIWT 2014), pp. 232-238, 2014. DOI : 10.1109/ICADIWT.2014.6814687
- [31]Duan, L., Xu, L., Guo, F., Lee, J., Yan, B., “A local-density based spatial clustering algorithm with noise,” Information systems.
- [32]Ankerst, M., Breunig, M. M., Kriegel, H. P., Sander, J., “OPTICS: Ordering points to identify the clustering structure,” ACM Sigmod record, Vol. 28, No. 2, pp. 49-60, 1999.
- [33]Paparrizos, J., Gravano, L., “k-shape: Efficient and accurate clustering of time series,” proc. of the 2015 ACM SIGMOD international conference on management of data, pp. 1855-1870, 2015.
- [34]Lee, Y. H., Wei, C. P., Cheng, T. H., & Yang, C. T., “Nearest-neighbor-based approach to time-series classification,” Decision Support Systems, Vol. 53, No. 1, pp. 207-217, 2012. DOI : 10.1016/j.dss.2011.12.014
- [35]Niennattrakul, V., Ratanamahatana, C. A., “On clustering multimedia time series data using k-means and dynamic time warping,” proc. of the 2007 International Conference on Multimedia and Ubiquitous Engineering (MUE'07), pp. 733-738, 2007.
- [36]He, L., Agard, B., Trépanier, M., “A classification of public transit users with smart card data based on time series distance metrics and a hierarchical clustering method,” Transportmetrica A: Transport Science, Vol. 16, No. 1, pp. 56-75, 2020. DOI : 10.1080/23249935.2018.1479722
- [37]Guan, Y., Ghorbani, A. A., Belacel, N., “Y-means: a clustering method for intrusion detection,” proc. of the CCECE 2003-Canadian Conference on Electrical and Computer Engineering, pp. 1083-1086, 2003. DOI : 10.1109/CCECE.2003.1226084

<p>※ 이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No.2022R1F1A1066017)</p>
