

# 엔트로피 지수를 이용한 기계학습 기반의 배터리의 건강 상태 예측 알고리즘

## Machine Learning Based State of Health Prediction Algorithm for Batteries Using Entropy Index

김 상 진\*, 임 현 근\*, 장 병 훈\*, 우 성 민\*\*

Sangjin Kim\*, Hyun-Keun Lim\*, Byunghoon Chang\*, Sung-Min Woo\*\*

### Abstract

In order to efficiently manage a battery, it is important to accurately estimate and manage the SOH(State of Health) and RUL(Remaining Useful Life) of the batteries. Even if the batteries are of the same type, the characteristics such as facility capacity and voltage are different, and when the battery for the training model and the battery for prediction through the model are different, there is a limit to measuring the accuracy. In this paper, We proposed the entropy index using voltage distribution and discharge time is generalized, and four batteries are defined as a training set and a test set alternately one by one to predict the health status of batteries through linear regression analysis of machine learning. The proposed method showed a high accuracy of more than 95% using the MAPE(Mean Absolute Percentage Error).

### 요 약

배터리를 효율적으로 관리하기 위해서는 배터리의 건강 상태와 잔여 수명을 정확하게 추정하고 관리하는 것이 중요하다. 배터리는 같은 종류여도 설비용량 및 전압 등의 특성이 다르며 학습용 모델을 위한 배터리와 모델을 통한 예측을 위한 배터리가 서로 다를 경우에는 정확도 측정에 한계가 있다. 본 논문에서는 전압의 분포와 방전 시간을 이용한 엔트로피 지수를 일반화하고 4개의 배터리를 각각 1개씩 교차적으로 훈련 집합과 테스트 집합으로 정의하여 기계학습의 선형회귀 분석을 통하여 배터리의 건강 상태를 예측하는 방법을 제안하였다. 제안된 방법은 평균 절대값 퍼센트 오차를 이용하여 95% 이상의 높은 정확도를 나타내었다.

*Key words : Entropy, Battery, SOH, Voltage Distribution, Machine Learning*

---

\* Hankook Electric Power Information Co.

\*\* Chungbuk Technopark

★ Corresponding author

E-mail : happy@hepi.co.kr, Tel : +82-42-861-2400

※ Acknowledgment

This research was the result of being supported by Ministry of Trade, Industry and Energy(MOTIE) in 2022 (No.20215910100030).

This research was financially supported by the Ministry of Trade, Industry and Energy, Korea, under the "Regional Innovation Cluster Development Program(R&D, P0016222)" supervised by the Korea Institute for Advancement of Technology(KIAT).

Manuscript received Oct. 14, 2022; revised Nov. 10, 2022; accepted Nov. 11, 2022.

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

## I. 서론

리튬 이온 배터리는 1990년대부터 가볍고 긴 수명, 내구성 및 안정성 및 고용량, 고효율 등의 장점으로 인하여 대표적인 2차 전지로 널리 사용되고 있다. 그러나 배터리의 가장 큰 단점 중 하나는 노화이며 배터리의 충방전이 반복될수록 배터리의 용량과 전력 감소를 초래한다. 배터리의 노화 현상은 배터리 내부의 물리·화학적 변화에 의해 발생하며 전지의 내부저항과 최대 충방전 용량을 변화시킨다. 특히 전기자동차의 배터리의 경우는 방전이나 수명이 다되면 치명적인 사고가 일어 날 수 있기 때문에 주의가 많이 필요하다.

배터리를 효율적으로 관리하기 위해서는 배터리의 SOH(State of Health)와 RUL(Remaining Useful Life)을 정확하게 추정하고 관리하는 것이 중요하다. 최근에는 데이터 기반 접근 방식을 통하여 배터리 측정 데이터가 미리 수행된 학습 모델을 사용하는데 NN(Neural Network) 구조를 기반으로 한 SOH 추정 및 RUL 예측 연구는 학습 능력이 우수하기 때문에 많이 사용되고 있다. 특히 LSTM(Long Short Term Memory)은 노화추정모델로 널리 사용되며 LSTM을 기반으로 SOH를 추정하는 다양한 방법이 소개되었으며 5% 이하의 평균 오차율을 보였다[1, 2].

엔트로피는 전압분포를 통하여 전압 고장진단 및 예측을 하는 방법이 소개되었으며[3], 엔트로피는 시간의 흐름에 따른 복잡도로 정의할 수 있다.

배터리의 특성상 충방전을 반복할수록 시간의 단축과 전압 분포의 변화로 인하여 복잡도가 증가하는 특징이 있다. 이를 토대로 배터리의 노화는 충방전 사이클을 통한 전압 및 전류가 흐르는 시간이 단축되고 분산된다는 사실을 이용하여 시간 보상 엔트로피라고 하는 노화 지수를 정의하여 배터리의 수명을 예측하는 방법이 제안되었다[4].

따라서 본 논문에서 정의한 배터리 노화 예측에 사용할 엔트로피 지수가 시계열의 특성을 반영하고 있기 때문에 LSTM 알고리즘을 사용하지 않고 상대적으로 학습 속도와 예측이 빠르며 모델구현도 쉬운 선형회귀(Linear Regression) 알고리즘을 사용하였다. 선형회귀는 일반적으로 저차원의 데이터셋에서는 다른 회귀 예측 알고리즘 보다는 일반화 성능이 더 좋은 편이기 때문에 엔트로피 지수만을 통한 배터리 수명 예측 모델로 사용하기에 적합하다.

본 논문에서는 NASA에서 제공하는 배터리 데이터를 토대로 전압분포를 통한 엔트로피 지수를 산출하여 기계

학습의 선형회귀 예측 알고리즘을 통하여 SOH 예측 성능을 확인하였다.

## II. 본론

### 1. NASA 배터리 데이터 분석

#### 가. State of Health(SoH)

리튬이온 배터리와 관련된 연구 중에서 가장 필수적인 분야 중 하나는 배터리의 상태 또는 건전성 예측이다. 배터리의 SOH는 열화에 기인한 영구적 용량손실을 고려하는 개념이기에 설계치 또는 공장에서 제조된 바로 직후 건전한 배터리만 100%이며 시간이 지날수록 최대치가 낮아지게 된다. 따라서 SOH는 노화로 인한 성능 감소의 지표로 초기 용량 대비 감소된 용량 또는 저항을 백분율로 나타낸다.

$$SOH(\%) = \frac{C_n}{C_0} \quad (1)$$

여기서  $C_n$ 은  $n$ 번의 충방전 후의 배터리의 설계 용량을 뜻하며,  $C_0$ 는 초기 상태의 용량 수치를 의미한다. 배터리의 충전과 방전을 반복하면서 용량이 점차 저하되기 때문에  $C_n$ 은 점차 감소하지만  $C_0$ 는 일정하다.

#### 나. NASA 배터리 데이터

배터리 수명예측에 사용한 리튬이온 배터리 실험데이터는 NASA Ames Prognostics Center에서 공개한 자료를 활용하였다[5]. 리튬이온 배터리 데이터는 B0005, B0006, B0007, B0018로 구성되며 실온에서 3가지(충전, 방전, 임피던스)를 통하여 실행이 되었다. 표 1은 각 배터리 데이터의 초기용량과 방전시의 전류와 전압을 나타낸다. 각 데이터에 대해 표 1에 표기된 값 까지 떨어질 때까지 2A로 방전을 수행하였고 배터리가 정격 용량이 30% 감소한 수명 종료(EOL) 기준에 도달할 때까지 실험을 수행하였다.

Table 1. Specification of 4 Battery Data.

표 1. 배터리 4개의 4 종류의 기본 특성

Battery No	Initial Capacity(Ah)	Discharge Current(A)	Discharge Cut-off Voltage(V)
B0005	1.86	2.0	2.7
B0006	2.04	2.0	2.5
B0007	1.89	2.0	2.2
B0018	1.86	2.0	2.5

배터리 데이터셋은 사이클이 반복될수록 용량이 감소되지만 일정하게 감소하는 것이 아닌 용량이 상승하는 구간이 존재한다. 지속적인 배터리 사용은 저항이 증가하고 셀 전압이 감소하지만 휴지 시간(sleep time)을 가지게 되면 배터리 내부 화학 반응이 안정화되며 이는 출력 전압을 상승시켜 배터리 보존 용량(capacity retention)이 발생한다. 그림 1은 방전 사이클 횟수에 따른 SOH의 현황으로 횟수가 증가할수록 SOH의 수명이 감소하는 추세이지만 순간적으로 증가하는 구간이 존재하는 것을 알 수 있다. 그림 1을 살펴보면 배터리마다 용량이 줄어드는 속도가 다르며 B0006은 다른 배터리에 비해서 급격하게 줄어드는 것을 확인할 수 있다.

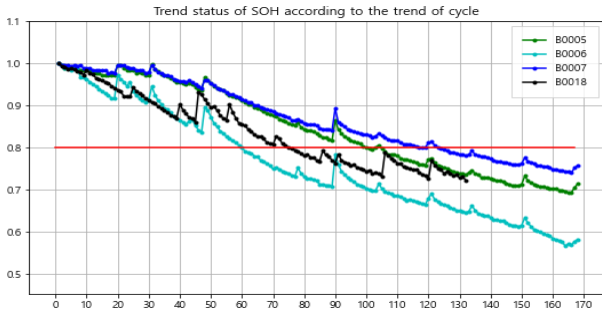


Fig. 1. Trend status of SOH according to the cycle and battery.

그림 1. 사이클과 배터리별 SOH 추세 현황

다. SOH 예측에 사용된 엔트로피 소개

엔트로피는 1948년에 Shannon이 정보를 측정하기 위해 도입한 이론의 개념으로 정보가 무질서하고 불규칙할수록 시간의 흐름에 따라 정보의 양이 증가한다. 아래의 식(2)은 정보 엔트로피의 기본 방정식으로 다음과 같다.

$$H = - \sum_{i=1}^N p(x_i) \log_b p(x_i) \quad (2)$$

$p(x_i)$ 는  $x_i$ 가 발생할 확률을 나타내며 정보 엔트로피는 시스템의 불확실성, 무질서를 나타내는데 사용할 수 있다. 배터리가 오래되면 내부 특성의 변화로 인해 배터리 작동 시스템이 무질서 할 수 밖에 없다. 정보 엔트로피는 전압분포를 통하여 생긴 변화를 나타내는데 사용되며 전압분포는 히스토그램을 통하여 아래의 식으로 표시할 수 있다.

$$H_k = - \sum_{i=1}^M p(x^{k_i}) \log_{10} p(x^{k_i}) \quad (3)$$

히스토그램의 빈의 개수를  $M$ 개로 가정하였을 때  $H_k$ 는  $k$ 번째 배터리 사이클의 정보 엔트로피이고  $p(x_i^k)$ 는

히스토그램의 빈  $x_i^k$ 에 대한 확률을 나타낸다.  $k$ 번째 배터리 사이클에서의 전압분포 및 배터리의 작동시간은 작동환경에 따라 달라지며 정보 엔트로피의 편차가 발생할 수 있다. 그리고 정보 엔트로피는 작동시간이 짧아질수록 감소하는 경향이 있기 때문에 이를 보완하는 방법으로 T. Bak[4]이 제안을 토대로 아래와 같이 정의하였다.

$$Ent_k = - \frac{1}{D_k} \sum_{i=1}^M p(x_i^k) \log_{10} p(x_i^k) \quad (4)$$

식 (4)에서  $Ent_k$ 는 예측변수로 사용될 새로 정의된 엔트로피의 지수이며  $D_k$ 는  $k$ 번째 사이클의 방전 기간으로 정의한다. 하지만 엔트로피 지수는 배터리의 종류에 따라 다르고 초기용량의 크기에 따라서 수치가 다르기 때문에 식 (4)번을 아래와 같이 일반화를 시킬 필요가 있다.

$$E_k = - \frac{1}{D_k Ent_1} \sum_{i=1}^M p(x_i^k) \log_{10} p(x_i^k) \quad (5)$$

Table 2. Correlation between SOH and entropy by the number of sections of voltage distribution.

표 2. 전압분포의 구간 개수 별 엔트로피와 SOH의 상관도

Number	B0005	B0006	B0007	B0018	Mean
10	-0.9944	-0.9877	-0.9950	-0.9889	-0.9915
20	-0.9940	-0.9928	-0.9975	-0.9888	-0.9932
30	-0.9921	-0.9932	-0.9987	-0.9893	-0.9933
50	-0.9915	-0.9928	-0.9991	-0.9890	-0.9931
100	-0.9892	-0.9925	-0.9984	-0.9886	-0.9921
200	-0.9847	-0.9921	-0.9960	-0.9889	-0.9904

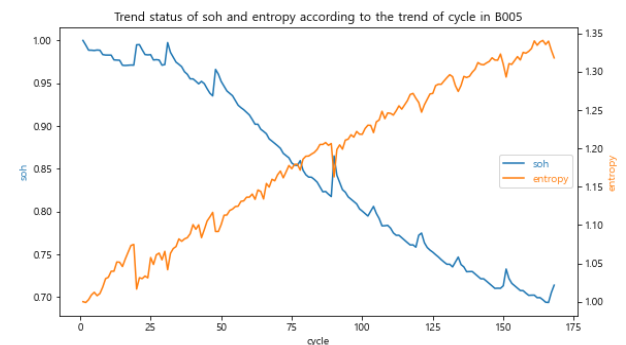


Fig. 2. Trend status of SOH and entropy according by cycle in B0005.

그림 2. B0005의 사이클에 따른 SOH와 엔트로피의 추세 현황

식 (5)에서  $E_k$ 는 식(4)에서 보정된 엔트로피의 지수이며  $Ent_1$ 는 식(4)에서 정의된 초기 방전시의 엔트로피 지수이다. 표 2는 각 배터리 별로 전압의 분포의 구간을 설

정하여 엔트로피와 SOH의 상관도를 분석한 결과이며 그림 2는 표 2의 전압분포의 구간의 개수중에서 평균 상관도가 제일 높은 30으로 하였을 때 사이클에 따른 SOH와 엔트로피의 추세 현황을 나타내었다.

그림 3의 경우에는 사이클 추세에 따른 엔트로피 값의 추세 현황을 나타낸 것으로 사이클의 증가에 따라서 엔트로피의 값이 증가하는 추세 현황을 보이는 것으로 확인 할 수 있다.

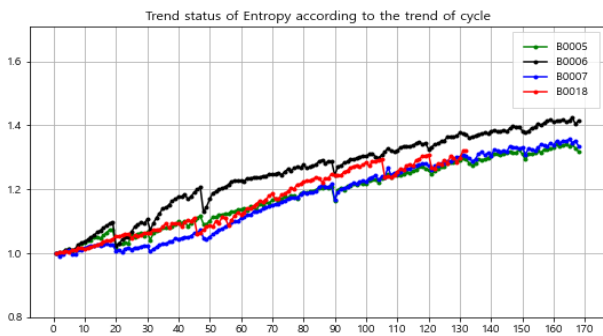


Fig. 3. Trend status of entropy according to the cycle and battery.

그림 3. 사이클과 배터리별 엔트로피 추세 현황

라. 엔트로피를 이용한 SOH 예측 결과

배터리의 수명 예측은 엔트로피와 SOH의 상관도가 0.9 이상이 되는 분석 결과를 토대로 기계학습의 선형회귀(Linear Regression, LR) 모델을 사용하였다. 훈련집합과 테스트집합을 설정하여 모델을 적용하였을 때의 정확도는 아래의 표와 같다. 모델 간의 성능 비교를 위한 평가 기준은 MAE(Mean Absolute Error)와 MAPE(Mean Absolute Percentage Error)를 사용하였다.

Table 3. Table of prediction accuracy of SOH of battery type.

표 3. 배터리 종류 별로 SOH의 예측 정확도 테이블

Train set	Test set	MAE	MAPE
B0005	B0006	0.0201	97.2
	B0007	0.0297	96.3
	B0018	0.0298	96.5
B0006	B0005	0.0201	97.6
	B0007	0.0388	95.1
	B0018	0.0218	97.4
B0007	B0005	0.0314	96.1
	B0006	0.0560	91.6
	B0018	0.0408	94.9
B0018	B0005	0.0271	97.1
	B0006	0.0243	96.6
	B0007	0.0425	94.9

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (5)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \quad (6)$$

$n$ ,  $y_i$ ,  $\hat{y}_i$ 는 각 데이터의 개수, 실제 SOH와 예측 SOH를 나타내며 오차의 단위는 %이다. MAE는 오차의 절대 값 자체를 볼 수 있어 계산과 이해가 쉽고 MAPE는 퍼센트 값을 통하여 성능을 평가하기 때문에 성능 비교 해석이 가능하며 둘 다 0에 가까울수록 모델의 성능이 높다는 것을 나타낸다.

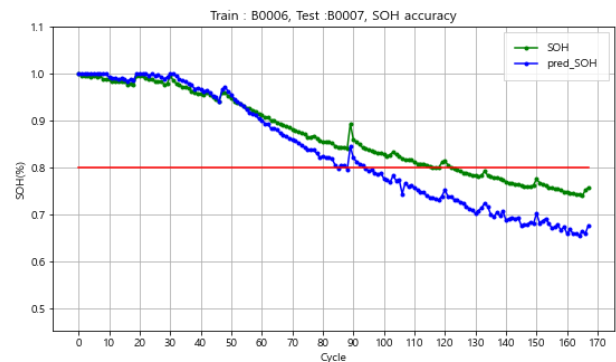


Fig. 4. Visualize actual and predicted trends of SOH.

그림 4. SOH의 실제값과 예측값의 추세 시각화

일반적으로 배터리의 수명 종료(End-of-Life, EOL)는 초기 용량의 80%로 설정을 하므로 실제 SOH가 75% 이

Table 4. Table of prediction accuracy of SOH more than 75% of battery type.

표 4. 배터리 종류 별로 75% 이상의 SOH에 대한 예측 정확도 테이블

Train set	Test set	MAE	MAPE
B0005	B0006	0.0154	98.2
	B0007	0.0283	96.5
	B0018	0.0348	96.1
B0006	B0005	0.0209	97.6
	B0007	0.0369	95.4
	B0018	0.0224	97.4
B0007	B0005	0.0232	97.2
	B0006	0.0164	97.9
	B0018	0.0391	95.2
B0018	B0005	0.0351	96.1
	B0006	0.0233	97.3
	B0007	0.041	95.04

상인 구간에 대해서 정확도를 다시 산출한 결과를 위의 표 4로 나타내었다.

교차실험을 통한 예측 정확도는 평균 MAE는 0.028, MAPE는 96.6%의 성능을 나타내었다. 제안된 모델은 기존 LSTM을 이용한 예측 사례와 비교 시 정확도는 비슷하지만 연산 속도에서 월등히 빠른 장점이 있다.

### III. 결론

본 논문에서는 NASA Ames Prognostics Center에서 공개하고 있는 데이터를 이용하여 전압 분포와 방전 시간을 이용하여 계산한 엔트로피 지수를 이용하여 SOH 예측 방법을 제안하였다. 배터리별로 노화가 촉진되는 시간이 달라서 정확도의 차이가 있지만 평균 95.9% 이상의 정확도를 보였으며 실제 수명이 75% 이상인 범위에서는 평균 96.6% 이상의 정확도를 확인하였다. 배터리의 수명에 영향을 미치는 전류, 온도 등의 특성의 반영과 엔트로피 지수 산출의 일반화를 시킬수 있으면 특성이 다른 배터리에 대해서도 높은 수명예측 정확도를 보일 수 있을 것임을 예상한다.

배터리의 수명예측을 위한 딥러닝 기반의 알고리즘 연구 결과가 많이 진행이 되고 있으며 최근에는 리튬이온 전지의 용량·수명을 더욱 정확하게 측정할 수 있는 물리 지식 기반의 딥러닝기술이 소개되었다[6].

향후에는 엔트로피 지수와 딥러닝을 이용하여 배터리 수명 예측 정확도를 더 높이는 방안을 연구해 나갈 예정이다.

### References

- [1] S. J. Jung, J. W. Hur, "Deep Learning Approaches to RUL Prediction of Lithium-ion Batteries," *Journal of the Korean Society of Manufacturing Process Engineers*, Vol.19, No.12, pp.21-27, 2020. DOI: 10.14775/ksmpe.2020.19.12.021
- [2] S. R. Hong, M. Kang, H. G. Jeong, J. B. Baek, J. H. Kim, "State of Health Estimation for Lithium-Ion Batteries Using Long-term Recurrent Convolutional Network," *The transaction of the Korean Institute of Power Electronics*, Vol.26, No.3, pp.183-191, 2021. DOI: 10.1109/IECON43393.2020.9254275
- [3] P. Liu, Z. Sun, Z. Wang and J. Zhang,

"Entropy-Based Voltage Fault Diagnosis of Battery Systems for Electric Vehicles," *Energies* 2018, Vol.11, No.1, p.136. DOI: 10.3390/en11010136

[4] T. Bak, S. Lee, "Accurate Estimation of Battery SOH and RUL Based on a Progressive LSTM with a Time Compensated Entropy Index," *Proceedings of the Annual Conference of the PHM Society 2019*, Vol.11, No.1, 2019.

DOI: 10.36001/phmconf.2019.v11i1.833

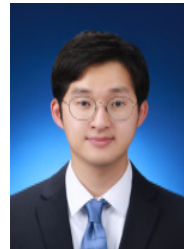
[5] B. Saha, K. Goebel, "Battery data set," *NASA AMES Prognostics Data Repository*, 2007.

[6] S. W. Kim, K. Y. Oh, S. C. Lee, "Novel informed deep learning-based prognostics framework for on-board health monitoring of lithium-ion batteries," *Applied Energy, Elsevier*, Vol.315(C), 2022.

DOI: 10.1016/j.apenergy.2022.119011

### BIOGRAPHY

#### Sangjin Kim (Member)



2012 : BS degree in mathematics, Chungnam National University.

2018 : Combined MS/PhD degree in Pure Mathematics, Chungnam National University.

2019~ : Senior Researcher, Hankook Electric Power Information Co.

#### Hyun-Keun Lim (Member)



2002 : BS degree in Computer Science, Semyung University.

2019 : MS degree in Computer Science, Paichai University.

2022 : PhD degree in Computer Science, Paichai University.

2021~ : Head of Research Institute, Hankook Electric Power Information Co.

**Byunghoon Chang** (Member)

1989 : BS degree in Electrical Engineering, Korea University.  
1992 : MS degree in Electrical Engineering, Korea University.  
1999 : PhD degree in Electrical Engineering, Korea University.  
2017~ : CEO in Hankook Electric Power Information Co.

**Sung-Min Woo** (Member)

2004 : BS degree in Electrical Engineering, Incheon National University.  
2006 : MS degree in Electrical Engineering, Incheon National University.

2013 : PhD degree in Electrical Engineering, Incheon National University.

2014~ : Senior Researcher, Chungbuk Techno park