Energy Proceedings
Vol 43, 2024
ISSN 2004-2965

Transformer-Based Online Battery State of Health Estimation from Electric Vehicle Driving Data

Kosaku Nakano[1], Kenji Tanaka [1*]

1 Graduate School of Engineering, The University of Tokyo, Tokyo 113-8656, Japan

(*Corresponding Author: tanaka@tmi.t.u-tokyo.ac.jp)

ABSTRACT

The surge in electric vehicle (EV) popularity necessitates innovative approaches for estimating the state of health (SOH) of EV lithium-ion batteries. This study introduces a transformer-based online SOH estimation model that leverages actual EV driving data, marking a departure from conventional methods that rely on lab-experimented battery cycle data. Our model comprises a transformer encoder and processes the raw sequences of battery voltage, current, state of charge, and vehicle speed. Despite the inherent noise in the EV battery readings while driving, the model shows high accuracy, with a mean absolute error of 1.31% and a root mean square error of 2.08%. Furthermore, this study unveils through self-attention map analysis that the model attends the stationary period of EVs to estimate the SOH. Although this study has a limitation in the dataset which lacks a wide range of driving route patterns, it still demonstrates the significant potential of transformer models in online SOH estimation for EVs while also providing valuable insights for future data collection.

Keywords: Lithium-ion batteries, Electric vehicles, State of health estimation, Deep learning, Transformer

NONMENCLATURE

Abbreviations

EV	Electric Vehicle
SOH	State of Health
LIB	Lithium-ion Battery
SOC	State of Charge
MLP	Multi-Layer Perceptron
MSE	Mean Squared Error
MAE	Mean Absolute Error
RMSE	Root Mean Squared Error

1. INTRODUCTION

The electrification of transportation is a critical factor in achieving a decarbonized society, wherein electric vehicles (EVs) play a pivotal role. A crucial technological component for expanding EVs is a battery, which is essential for storing electricity to power the motor. Lithium-ion batteries (LiBs) are commonly employed in EVs due to their high energy and power density [1]. Despite these advantages, one challenge LiBs face is their degradation in power and capacity over time through usage. This degradation is critical, especially for EV applications, because it directly relates to its performance.

The state of health (SOH) is an indicator for evaluating the degree of battery degradation. Knowing accurate SOH is essential from the perspectives of safety and performance, as well as second-life applications and recycling. However, accurately determining SOH poses a significant challenge because it cannot be directly measured and requires complex estimation techniques. In this study, we focus on the problem of online SOH estimation of EVs. Online SOH estimation is an SOH estimation that does not need to disrupt regular use, which is more suitable for EV applications.

Various methods to estimate SOH online have been explored, such as Coulomb counting methods which integrate currents and open circuit voltage-based methods [2]. In addition, recent advances in machine learning techniques have led to the exploration of many data-driven approaches. The studies of SOH estimation techniques can be categorized into two clusters based on data types: lab-experimented and EV operational data-based studies. In lab-experimented data-based studies, various methods are proposed [3,4], yet they fall short in real-world EV applications for two primary reasons. First, they rely on cell-level data to estimate SOH, but such

This is a paper for 15th International Conference on Applied Energy (ICAE2023), Dec. 3-7, 2023, Doha, Qatar.

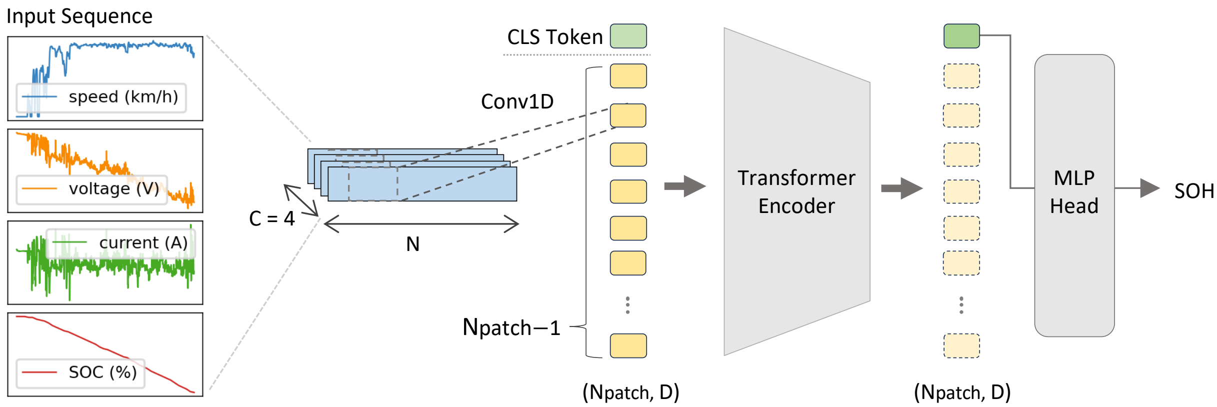

Fig. 1 The network architecture of the transformer-based online SOH estimation model.

data is often inaccessible in EV application contexts. Second, these methods usually assume consistent charging and discharging conditions in a controlled lab setting, including cycle intervals, depth of discharge, and temperature. However, these factors significantly affect battery degradation and vary widely in real-world scenarios, making such assumptions impractical. Meanwhile, studies based on real-world EV operational data have shown significant progress [5,6], yet still face challenges such as unreliable reference SOH and limited applicability in various situations.

This research introduces a transformer-based model [7] for EV online SOH estimation to tackle these challenges by leveraging complex, noisy, real-world driving data. This data is coupled with reliable reference SOH obtained from occasional full-discharge tests. While there are several definitions of SOH [8], this study focuses on energy capacity-based SOH, which is a critical factor directly related to the cruising range in EVs. In this case, SOH is defined as the ratio of current energy capacity in kWh to initial energy capacity in kWh.

The proposed model incorporates a transformer encoder to process time series data of battery voltage, current, state of charge (SOC), and vehicle speed from individual trips, enhancing its applicability in various situations. The model's design obviates the necessity for feature engineering, streamlining the training and inference process. Additionally, the transformer architecture not only enhances the model's estimation performance but also aids in interpreting the estimation process, which is attributable to the self-attention mechanism.

Our model is trained and evaluated using real-world EV driving test data amassed over about three years from three distinct EVs in Japan. The model's input data (battery voltage, current, SOC, and vehicle speed) is captured in real-time during driving. It should be noted that the dataset may not fully capture diverse real-world driving patterns because the EVs are repeatedly driven on the same route for set periods.

Despite this limitation, our research offers valuable insights into model development and data acquisition, marking a significant advancement in online SOH estimation for LIBs based on EV-driving metrics. The contribution of this study is as follows:

- We demonstrate the significant potential of using a transformer in EV online SOH estimation by developing a model that processes raw EV driving data from a single trip.
- We train and evaluate our model using real-world EV driving data with reliable reference SOH collected over three years from three vehicles.
- We investigate the nuances of data collection, which helps to further the SOH estimation studies using EV-driving data.

2. METHODS

2.1 Network architecture

To accurately deduce the SOH from EV driving data, we develop a neural network architecture that incorporates the transformer encoder. The motivation behind leveraging the transformer encoder is its competency in recognizing the dependencies inherent within sequential data. A visual representation of the proposed model architecture can be found in Fig. 1.

The network intakes time series data of battery voltage, current, SOC, and vehicle speed from a single trip and subsequently projects the estimated SOH. A single trip is defined as a period of time during which the vehicle is in continuous drive mode. For a comprehensive discussion regarding the nature, preprocessing, and source of input data, we refer the reader to Section 2.2.

The data input is structured as a multi-channel sequence, represented by the shape (C, N). Here, C refers to the number of data channels with a value of

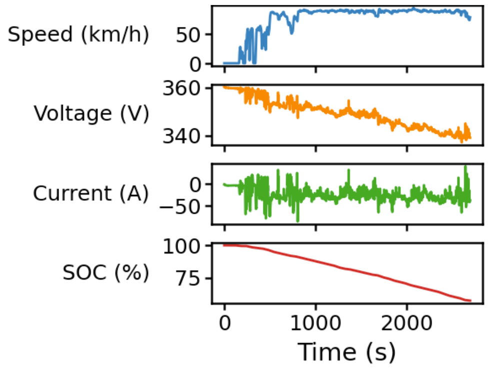

Fig. 2 The illustration of EV driving data of a single trip which the model takes as an input.

four (speed, voltage, current, and SOC), while N indicates the sequence's length. Initially, this input undergoes a transformation via a 1D convolutional layer, which maps it into a latent space. This process results in a tensor that has the dimensions of $(N_{patch} - 1, D)$. A CLS token, a special tensor used to aggregate information about the entire sequence, is added at the beginning of the embedded sequence. The transformer encoder subsequently processes this modified sequence. The latent representation corresponding to the CLS token the encoder produces is then fed into a multi-layer perceptron (MLP) head, resulting in an estimated SOH.

2.2 Dataset

The dataset utilized in this research was generated by the Japan Automobile Research Institute, collecting data from August 2011 to June 2014, with missing data from September 2012 to March 2013. This dataset encompasses readings from three EVs with an initial capacity of 16 kWh. During data collection, each vehicle was assigned a fixed driving schedule, which comprised a unique combination of predetermined and varied driving routes. These routes included some long-distance highway routes and shorter-distance city routes.

During EV driving, the speed, voltage, and current, which are used as inputs for the neural network model detailed in section 2.1, are recorded every second. In contrast, the SOC is logged every minute. This SOC data is then interpolated every second using linear interpolation. To exclude anomalous data, only records with trip lengths between 10 and 90 minutes, with a mileage of 10 km or more, and without measurement errors are used. The total number of trips after considering these conditions is 3,335. When inputting data into the model, only the first 45 minutes, or 2,700 seconds of each trip, are used. This duration is chosen for the model as it effectively captures critical features necessary to estimate the SOH while reducing the

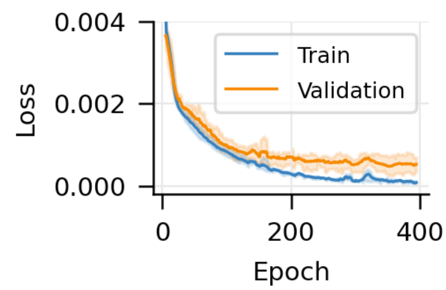

Fig. 3 The loss curve for the training and validation dataset over three random seeds.

model's size and computational demands. In the case that a trip is less than 45 minutes, padding is applied based on the length of the trip. The average mileage of the 3,335 trips is 33.9 km, and the average trip duration is 51.7 minutes. An illustration of the driving data from a single trip, utilized as input for the model, can be found in Fig. 2.

During the data collection period, a constant-current discharge test with c-rate 0.3 was conducted roughly once every six months using a chassis dynamometer to evaluate the battery degradation. The SOH is calculated based on the discharge capacity obtained from the test. It is then linearly interpolated based on the odometer reading, which reflects the vehicle's accumulated mileage. This interpolated SOH is used as the reference SOH, serving as the target value for the SOH estimation.

2.3 Model training

For the model training process, we randomly split the dataset into training, validation, and test sets, with a ratio of 8:1:1. The model is trained using the training dataset, and its performance is assessed with the validation dataset at the end of each epoch. The best model, determined by the lowest validation loss, is then evaluated on the test dataset to assess its effectiveness.

We employ the AdamW optimizer with a learning rate of 1e-4 and a batch size of 16. The loss function is the mean squared error (MSE) between the estimated SOH and the reference SOH. The model is trained for 400 epochs.

3. RESULTS AND DISCUSSION

3.1 Training and evaluation results

We evaluate the stability and performance of our model by training it using three different random seeds. Fig. 3 presents the loss curves for the training and validation datasets across these seeds. The plotted line for each dataset represents the mean loss, while the shaded region indicates the standard deviation. Learning curves are smoothed by applying a moving average with

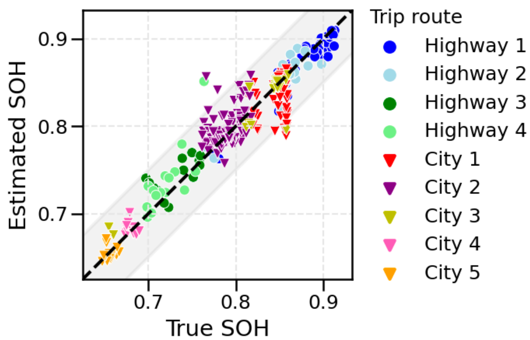

Fig. 4 Scatter plot of reference SOH vs. estimated SOH with marker pattern indicating trip routes.

a window size of 11, and the vertical axis is truncated at 0.004. The loss convergence is observed for both training and validation in Fig. 3, indicating a consistent learning process across different random seeds.

In the evaluation with the test dataset across three random seeds, the average mean absolute error (MAE) was 1.31%, with the root mean square error (RMSE) measuring 2.08%. Among the three instances of the model, each trained with different random seeds, the best-performing instance yielded an MAE of 1.22% and an RMSE of 1.83%.

Fig. 4 illustrates the SOH estimation result of the best-performing instance across various driving routes in the test dataset. The instance achieves high precision, with 97.3% of the SOH estimates falling within a 5% error margin, depicted by the gray area.

3.2 Self-attention map analysis

To elucidate the process through which the model estimates the SOH from the EV driving data, we examine the self-attention weights of the last layer assigned to the CLS token for different inputs. Fig. 5 shows examples of self-attention map visualization. These attention maps correspond to (a) highway driving input and (b) city driving input. The red heatmaps indicate the model's attention intensity.

Irrespective of the type of driving, the self-attention map analysis demonstrates that the model tends to attend the data points where the vehicle speed is recorded as 0 or where the EV comes to a complete stop, especially towards the latter segments of the trip. This observation implies that the battery readings in stationary periods of the vehicle after driving emerge as crucial features for the model to estimate SOH.

3.3 Limitation

While the proposed method attains satisfactory accuracy, a noteworthy limitation of our study hinges on

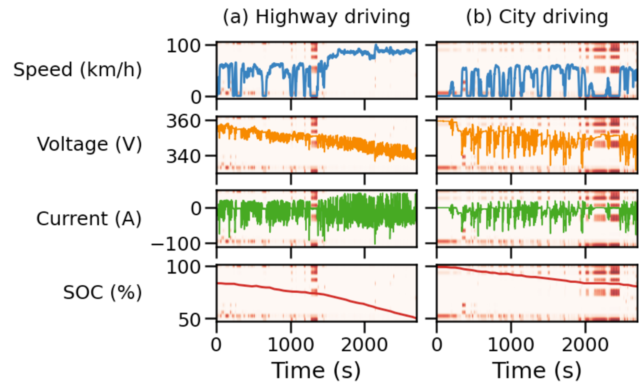

Fig.5 Self-attention of CLS token demonstrated by the red heatmap on the input sequences.

the limited and repeated driving route patterns in the dataset. The model might mistakenly correlate specific driving patterns to SOH due to the repetitiveness of the same driving route patterns. In particular, patterns associated with city driving routes 4 and 5, as seen in Figure 4, might be mistakenly correlated. Such unintended associations could prevent the model from accurately learning the relationship between EV driving data and SOH.

4. CONCLUSION

In this research, we address online SOH estimation from EV driving data. Our proposed transformer-based approach showcases the significant potential in EV online SOH estimation while achieving notable accuracy. The self-attention map analysis shows that the EVs' battery readings in their stationary period are employed as clues by the model to estimate SOH. While the model demonstrates remarkable results, it should be noted that its performance might be influenced by the repetitiveness and limited diversity of driving patterns in the dataset. For enhanced model validity and improved estimation accuracy in future iterations, it is important to incorporate a more diverse range of driving patterns.

DECLARATION OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. All authors read and approved the final manuscript.

REFERENCE

[1] Lu L, Han X, Li J, Hua J, Ouyang M. A review on the key issues for lithium-ion battery management in electric vehicles. Journal of Power Sources 2013;226:272–88.

[2] Lucian Ungurean, Cârstoiu G, Micea MV, Groza V. Battery state of health estimation: a structured review of models, methods and commercial devices. International Journal of Energy Research 2016;41(2):151–81.
[3] Gu X, Khay Wai See, Li P, Shan K, Wang Y, Zhao L, et al. A novel state-of-health estimation for the lithium-ion battery using a convolutional neural network and transformer model. Energy 2023;262:125501–1.
[4] Bockrath S, Lorentz VRH, Pruckner M. State of health estimation of lithium-ion batteries with a temporal convolutional neural network using partial load profiles. Applied Energy 2023;329:120307–7.
[5] Wen S, Lin N, Huang S, Wang Z, Zhang Z. Lithium battery health state assessment based on vehicle-to-grid (V2G) real-world data and natural gradient boosting model. Energy. 2023;284:129246–6.
[6] Liu H, Deng Z, Yang Y, Lu C, Li B, Liu C, et al. Capacity evaluation and degradation analysis of lithium-ion battery packs for on-road electric vehicles. Journal of Energy Storage. 2023;65:107270–0.
[7] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is All you Need. Advances in Neural Information Processing Systems. 2017;30.
[8] S.M. Rezvanizaniani, Liu Z, Chen Y, Lee J. Review and recent advances in battery health monitoring and prognostics technologies for electric vehicle (EV) safety and mobility. Journal of Power Sources 2014;256:110–24.