Battery Fault Diagnosis for Electric Vehicles Based on Voltage Abnormality by Combining the Long Short-term Memory Neural Network and the Equivalent Circuit Model

Da Li, Zhaosheng Zhang, Peng Liu, Zhenpo Wang and Lei Zhang, *Member IEEE*

***Abstract*—Battery fault diagnosis is essential for ensuring safe and reliable operation of electric vehicles (EVs). In this study, a novel battery fault diagnosis method is presented by combining the long short-term memory recurrent neural network (LSTM) and the equivalent circuit model (ECM). The modified adaptive boosting method is utilized to improve diagnosis accuracy, and a prejudging model is employed to reduce computational time and improve diagnosis reliability. Considering the influence of the driver behavior on battery systems, the proposed scheme is able to achieve potential failure risk assessment and accordingly to issue early thermal runaway warning. A large volume of real-world operation data is acquired from the National Monitoring and Management Center for New Energy Vehicles (NMMC-NEV) in China to examine its robustness, reliability and superiority. The verification results show that the proposed method can achieve accurate fault diagnosis for potential battery cell failure and precise locating of thermal runaway cells.**

Index Terms—Electric vehicles, lithium-ion battery, fault diagnosis, equivalent circuit model, long short-term memory recurrent neural network, modified adaptive boosting.

I. Introduction

A. Motivations

In order to meet the challenges of fossil oil depletion and environmental pollution, electric vehicles (EVs) are being actively developed and extensively adopted worldwide [1]. Battery systems are an integral component of EVs and largely determine their driving performance and cost-effectiveness [2]. Multiple battery cells are often connected in series and/or parallel configurations to meet the voltage and capacity requirements [3]. Fault occurrence may precipitate to these constituent battery cells owing to battery cell degradation, electrical failure or misuse [4]. Thermal runaway may be triggered if these faults are left untended [5][6]. So accurate and timely battery fault diagnosis and safety warning issuing are essential to prevent thermal runaway occurrence and ensure safe operation of real-world EVs.

B. Literature review

Many fault diagnosis approaches have been proposed in the literature, which can be generally classified into three categories, i.e., threshold-, model- and data-driven-based methods. Threshold-based methods have been used to diagnose over- and under-voltage battery faults [7]. For instance, Duan et al. [8] calculated the standard deviation of the selected indicators based on the information entropy and set certain thresholds to assess cell inconsistency. However, the threshold-based methods cannot predict battery fault occurrence in advance, and it is difficult to determine appropriate threshold values in practice. In contrast, model-based methods can accurately describe the evolutions of battery states under normal and/or faulty conditions [9]-[12]. For example, Wang et al. [13] presented a model-based fault diagnosis method to assess the insulation state of battery packs. Similarly, Liu et al. [14] proposed a model-based sensor fault detection and isolation scheme for a series-connected battery pack. However, the models built in these studies can only be used for limited fault types. In order to reveal the nonlinear relationship between battery failures and their characteristic parameters, data-driven methods have also been widely explored [15]-[17]. In this regard, Kang et al. proposed a multi-fault diagnosis method based on an interleaved voltage measurement topology for series-connected battery packs [18]. Analogously, Shang et al. proposed a multi-fault diagnosis method for early battery failure prediction based on the modified Sample Entropy [19]. Nevertheless, data-driven methods always require a large amount of data for model training, and exhibit poor robustness to unseen datasets. In addition, these threshold-, model- and data-driven-based methods in the existing studies were only examined based on simulations or lab-experiments and seldom verified in real-world EV implementations.

As there are numerous fault types within a battery system, it is preferable to develop a comprehensive diagnosis scheme albeit there are only a limited number of measurable parameters during vehicular operation. To solve this problem, some data-

Manuscript received October 18, 2018; revised XXXXXXX; accepted XXXXXXXX. Date of publication XXXXXXXX; date of current version XXXXXXXX. This work was supported in part by the National Key R&D Program of China under Grant 2018YFB0105700, and in part by the Natural Science Foundation of China under Grants No. 61703024 and No. U1764258.

(Corresponding authors: Zhaosheng Zhang, zhangzhaosheng@bit.edu.cn; Peng Liu, roc726@163.com; and Lei Zhang, lei_zhang@bit.edu.cn.)

Da Li (li_da_bit@126.com), Zhaosheng Zhang, Peng Liu, Zhenpo Wang and Lei Zhang are with the National Engineering Laboratory for Electric Vehicles and the Collaborative Innovation Center for Electric Vehicles in Beijing, Beijing Institute of Technology, Beijing 100081, China.

driven fault diagnosis methods have been proposed based on real-world operation data [20]. For instance, Zheng et al. presented a power fault diagnosis method for a battery pack composed of 96 series-collected battery cells [21]. Similarly, Wang et al. proposed an in-situ voltage diagnosis method based on the modified Shannon entropy, which can predict the voltage fault by monitoring cell voltages during vehicular operation [22]. Similar studies have also been conducted for battery voltage fault diagnosis, in which the battery cells with poor consistency are considered faulty [23]-[25]. However, these methods failed to reveal how to determine proper thresholds for cell inconsistency evaluation. Besides, batteries may become faulty even though they exhibit good cell consistency within the battery pack.

Some untreated battery cell faults can evolve into thermal runaway of the entire battery system. Thus, it is meaningful to develop enabling methods for thermal runaway prediction. The existing thermal runaway prediction methods can be sorted into two groups, i.e., characteristic- and model-based approaches. For characteristic-based methods, voltage and temperature evolutions during thermal runaway processes are especially studied with the purpose of finding out useful information for early thermal runaway diagnosis [26]-[30]. For example, Zhao et al. analyzed the temperature distribution of the separator and the thermal runaway characteristics for NCM lithium-ion batteries under different heat dissipation conditions [31]. However, battery voltage and temperature often exhibit varied behaviors under different thermal runway conditions for disparate battery chemistries. Some scholars have attempted to develop thermal runaway prediction models based on the first principles [32]-[34]. For instance, Said et al. established a model to predict thermal runaway propagation by combining the kinetics of thermal decomposition of materials with heat transfer modelling [35]. Analogously, Qi et al. established a mathematical model to depict the voltage and temperature evolutions of lithium-ion batteries during overcharging [36]. Nevertheless, these laboratory-based methods cannot fully consider all the influencing factors during real-world vehicular operation.

The development of artificial intelligence (AI) provides enormous opportunities for uncovering the underlying mechanisms of battery failures [37][38]. Some scholars have leveraged deep-learning neural networks to achieve accurate battery fault diagnosis and prediction. For instance, Tobar et al. incorporated external variables associated with the evolution of battery voltage to improve battery fault prediction performance [39]. Zhao et al. proposed a prediction approach by combining an accurate lithium-ion battery model with a recurrent neural network (RNN) [40]. However, these methods are either model- or experiment-based; their validity in real-world EVs is doubtful since there are still some factors that are neglected. Instead, Hong et al. proposed a deep-learning-enabled method to perform accurate multi-forward-step voltage prediction for battery systems [41]. However, the presented method may fail when the battery voltage is within the preset thresholds.

Limitations of the existing studies can be summarized as follows. Firstly, the accuracy of the existing diagnosis methods needs further improvement. Secondly, most diagnosis methods are developed based on laboratory-obtained data and their performance has seldom verified in real-world EV operation. Additionally, the hyper-parameters of the developed data-driven models need to be optimized for improved performance.

C. Contributions of this study

This study can contribute to the related literature in the following aspects:

1) A novel battery fault diagnosis model by combining the long short-term memory recurrent neural network (LSTM) and the equivalent circuit model (ECM) is proposed to realize precise locating of the potential faulty battery cells. A coupling module based on the modified adaptive boosting (MAB) is applied to combining the advantages of the LSTM and the ECM to improve the diagnosis accuracy. Besides, a prejudging model is also employed to reduce computational time and improve diagnosis reliability.
2) A large volume of real-world operation data is used to train and verify the proposed method. The proposed model can be trained with the datasets collected from normal vehicles, and can be used for battery fault diagnosis in the vehicles with the same specifications.
3) An approaching optimization method (AOM) is proposed to derive the hyperparameters for optimized diagnosis performance.
4) A real-world fault diagnosis and risk assessment strategy is proposed to realize online fault diagnosis. Correspondingly, a fault alarming protocol is proposed, and its effectiveness is verified based on the real-world EV data with fault occurrence and thermal runaway accidents.

D. Organization of paper

The remainder of the paper is organized as follows: Section II gives a brief introduction of the National Monitoring and Management Center for New Energy Vehicles (NMMC-NEV) and the acquisition of real-world EV operation data. Section III introduces the voltage prediction method including the LSTM, the ECM, the MAB and the pre-judgment module. Section IV provides the voltage prediction results and discussions. Section V elaborates on the real-world fault diagnosis strategy and offers detailed discussions on its robustness, reliability and superiority, followed by the key conclusions summarized in Section VI.

II. Data Acquisition and Introduction

A. Description of the NMMC-NEV

The NMMC-NEV is responsible for the supervision of all new energy vehicles running in China. A schematic of the NMMC-NEV platform is illustrated in Fig. 1. The big data platform uses the Hadoop architecture to ensure the reliability of data collection and storage. It now has a service capability of more than 3 million units of EVs.

B. Description of data acquisition

In this study, nine electric vehicles with the same

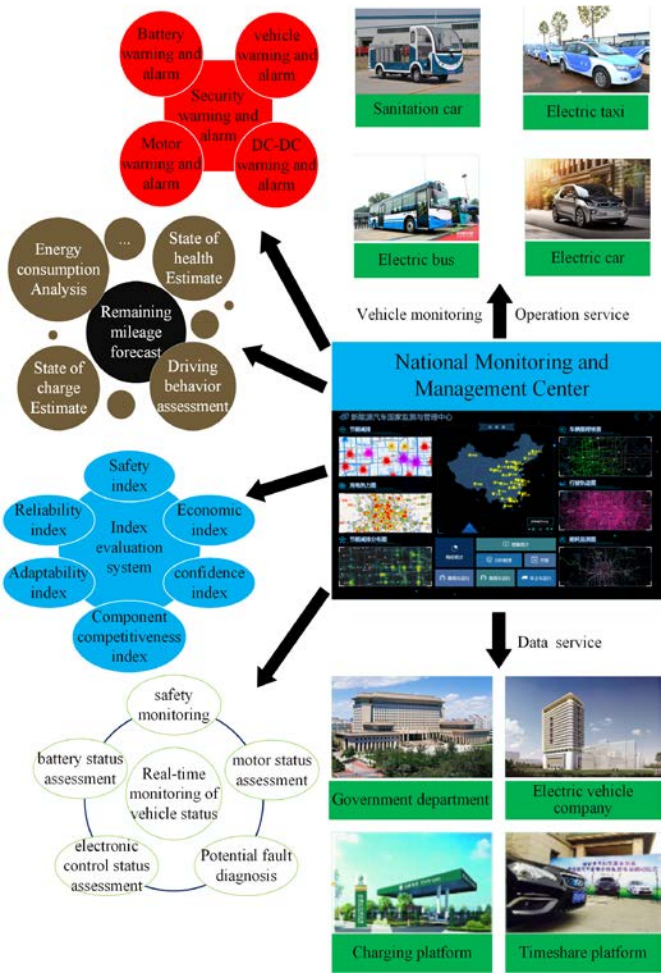


Fig. 1. The schematic of the NMMC-NEV.

specifications are used, and the data are collected with a frequency of 0.1 Hz. According to battery pack maintenance records, these vehicles are divided into three categories:

1) Normal vehicles No. 1-4: These vehicles experience no fault alarming, and no cell abnormalities have been detected in regular maintenance records.
2) Vehicles with potential failures No. 5-8: Unqualified batteries are installed in these vehicles. There were at least one fire accident occurrence during battery cell production, and the internal resistance of these battery cells are larger than that of the normal vehicles.
3) Thermal runaway vehicle No. 9: Unqualified batteries were installed in the vehicle, and there was thermal runaway occurrence during usage. The battery cell that aroused the thermal runaway accident has been identified.

The data are collected from the normal and potentially faulty vehicles in this study. The data collected from the normal vehicle No. 1 are used to train the LSTM model. The data collected from the normal vehicle No. 2 are used to verify the robustness of voltage prediction for the vehicles with the same specifications. The data collected from the normal vehicle No. 3 are used to verify the reliability of the proposed model. The data collected from the normal vehicle No. 4 are used to manifest the necessity of coupling the LSTM with the ECM. The data collected from the potentially faulty vehicle No. 5 are used to compare the effectiveness of different coupling methods.

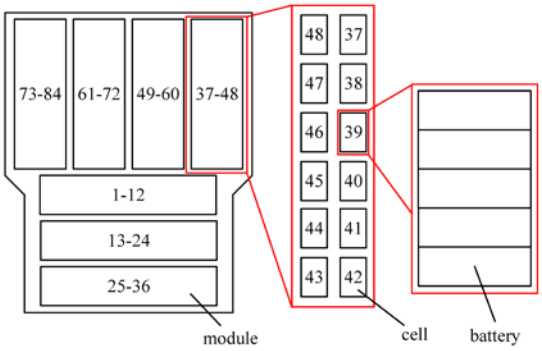


Fig. 2. Layout of the battery pack in the studied EVs.

The data collected from the potentially faulty vehicle No. 6 are used to verify the effectiveness of the model in the potential failure risk assessment. The data collected from the potentially faulty vehicle No. 7 are used to further verify the robustness of the model in the potential failure risk assessment. The data collected from the potentially faulty vehicle No. 8 are used to verify the necessity of coupling the LSTM with the ECM. The data collected from the thermal-runaway vehicle No. 9 are used to verify the effectiveness of the model for the early warning of impending thermal runaway accident.

All the batteries in the studied vehicles are ternary lithium-ion batteries. The battery pack used in the studied EVs comprises of several series-connected battery modules, and each module contains a certain number of series-connected battery cells. The layout of the installed battery pack is shown in Fig. 2. It can be seen that four modules and three modules are respectively arranged longitudinally and laterally, each containing twelve battery cells. Voltage sensors are installed for each battery cell.

III. Battery Voltage Prediction

To realize online battery fault diagnosis, it is necessary to make full use of measurable parameters with low cost sensors, which mainly includes voltage, current and temperature [42][43]. Transient currents depend on the driver behavior, and temperature measurement cannot fully reflect the electrical characteristics of battery systems. Instead, it is a natural idea to diagnose battery failure by comparing real-time cell voltages with their normal values. Therefore, it is desirable to predict the battery voltage using enabling battery models during practical vehicular operation. The existing battery models can be grouped into three categories, i.e., electrochemical models, equivalent circuit models and data-driven models [44]. For electrochemical models, it is difficult to obtain the model parameters, which hinders its application in realistic battery management systems (BMSs). For data-driven models, the model parameters can only be derived based on enough and diversified training data, and neural networks are dominant in this kind of models [45]. The ECMs have been widely used to identify internal battery parameters such as resistance and open circuit voltage (OCV) with acceptable accuracy and computational intensity.

In this study, a battery voltage fault diagnosis model combining the LSTM and the ECM is proposed. A coupling

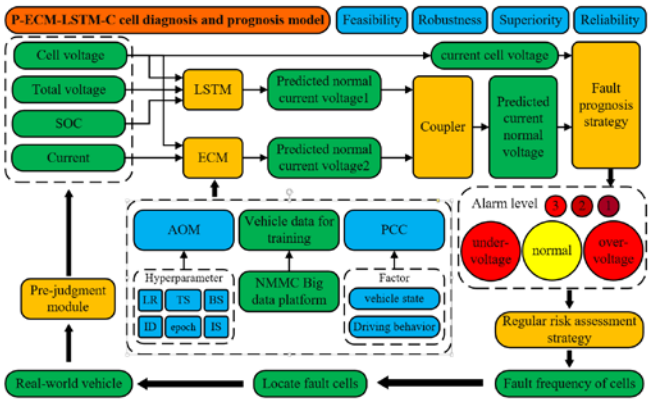


Fig. 3. The schematic of the proposed coupled LSTM and ECM model for battery cell voltage prediction.

module based on the MAB is employed to combine the voltages predicted by the LSTM and the ECM to improve diagnosis accuracy. A prejudging model is used to reduce computational intensiveness and improve diagnosis reliability. The schematic of the proposed scheme is depicted in Fig. 3, and the detailed procedures for battery voltage prediction are given as follows:

(1) Pre-process the historical data of vehicles using the pre-judgment module.

(2) Feed the processed historical data into the LSTM and the ECM to get the voltage predictions 1 and 2, respectively. The LSTM is trained offline in advance based on the data collected from normal vehicles, and the AOM and the Pearson correlation coefficient (PCC) are used to optimize hyper-parameters and to select input features, respectively.

(3) Calculate the predicted normal voltage using the coupling module based on the voltage predictions 1 and 2.

(4) Compare the predicted normal voltage with the real one, and determine the alarming levels of each battery cell based on the proposed fault prognosis strategy.

(5) Locate the potential failure and thermal runaway cells according to the alarming levels.

(6) Repeat steps (1)-(5) with new fed data during vehicular operation.

The model can be replenished in real-time based on the incoming data with EV operation.

A. Model training and voltage prediction based on the LSTM

As a recurrent neural network [46], the LSTM excels in avoiding gradient disappearance and explosion [47][48], and is thus widely used in the fields of speech emotion classification [49], financial market prediction [50], language identification [51], and smarter surveillance [52].

During practical EV operation, the voltage, temperature, and current are all time-dependent parameters, and the voltage prediction model based on the LSTM can be established to predict the voltage evolution along with time.

The history data play an important role in predicting battery states, so a 'Many to one' structure for the LSTM is chosen and the input matrix is given by

$$\boldsymbol{A}_{k\times(n+1)}=\begin{pmatrix} U_{1,c} & F_{1,1} & \cdots & F_{1,n} \\ \cdots & \cdots & F_{t,j} & \cdots \\ U_{k,c} & F_{k,1} & \cdots & F_{k,n} \end{pmatrix} \tag{1}$$

where n is the number of input features; k is the time step; $U_{t,c}$ (t=1, 2,⋯, k) is the voltage of the battery cell c at the time step t; $F_{t,j}$ (t=1, 2, ⋯ , k; j=1, 2, ⋯ , n) is the j-th input feature at the time step t.

The output matrix is given by

$$\boldsymbol{B}_{1\times1}=\left(U_{k+1,c}\right) \tag{2}$$

The battery cell voltage is influenced by various factors, and driving behaviors such as hard acceleration and deceleration would cause remarkable fluctuations of terminal voltages. Therefore, the driving behavior is considered as one of the input features.

When it comes to vehicle states, current, temperature and state of charge (SOC) are parameters that can directly affect the battery voltage. The readings of each temperature probe can be collected from the NMMC-NEV platform and the highest measured temperature in a battery pack is used to represent the overall temperature of the battery pack. In addition, the battery aging may cause the internal resistance increase and the capacity degradation, both of which would impose unavoidable impacts on the battery voltage. However, online estimations of battery state of health (SOH) and remaining useful life (RUL) are still open topics to surmount, and thus the accumulated mileage of vehicles is used in this study to approximately represent battery health levels. In addition, the insulation resistance which can reflect the insulation level of electrical systems is also selected as an input.

The driving behavior can be interpreted by the brake and acceleration pedals. Additionally, the motor parameters including rotational speed, temperature, voltage and bus current and the vehicle operation modes including driving, charging and parking are considered.

To reduce the model input number to avoid overfitting, the PCC is utilized to analyze the correlation between these parameters as well as to extract highly-relevant parameters [53][54]. The PCCs between the voltage and different input features are shown in Fig. 4.

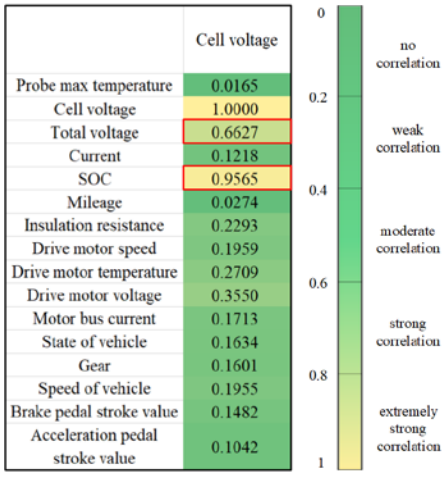


Fig. 4. The PCC values of different input features.

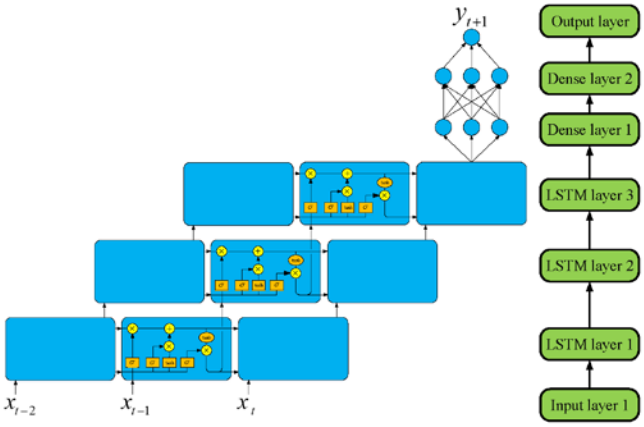


Fig. 5. The structure of the proposed LSTM model.

The correlation between two parameters can be evaluated via absolute PCC values with the ranges of 0.8-1.0, 0.6-0.8, 0.4-0.6, 0.2-0.4 and 0-0.2 meaning extremely strong, strong, moderate, weak and very weak correlations. It can be seen that the parameters most relevant with the cell voltage are the total voltage and SOC of battery packs. In this study, the cell voltage and the total voltage and SOC of battery pack are finally chosen as the inputs of the LSTM.

Cell voltage should be predicted timely with newly fed operation data. In this study, a sliding window is used to extract sub-calculation intervals and to feed the features of sub-calculation intervals into the trained LSTM for voltage prediction.

The number of layers and the activation function of the model are crucial to the voltage prediction performance. The LSTM network is constructed with a cascaded sequence of seven layers, i.e., one input layer, three LSTM hidden layers, two dense layers with the *tanh* activation function, and one output layer. The structure of the proposed neural network is illustrated in Fig. 5 for demonstration.

The widely used optimizers are the SGD, Adagrad, Adam, Adamelta, RMSprop, Nadam, and Adamax [55]. The Adam is selected as the optimizer to update the weights and bias based on the gradient of the loss function [56]. The MSE is chosen as the loss function, which is given by

$$MSE(y, y') = \frac{\sum_{i=1}^{n} (y_i - y_i')^2}{n} \tag{3}$$

where y_i' and y_i are the prediction and the real voltage at the time step i; n is the time step.

A bunch of hyper-parameters need to be determined and optimized in the LSTM, including the learning rate (LR), time step (TS), batch size (BS), input dimension (ID), output dimension (OD), neurons number of dense layer (NNDL), neurons number of LSTM layer (NNLL), regularization factor (RF), dropout factor (DF), input samples and epochs. Here, the length of sub-calculation interval (LSCI) is also an important hyper-parameter. There are four major methods for determining the hyper-parameters, i.e., the babysitting, grid, random grid and Bayes SMBO [57]. With the increasing hyper-parameter dimensions, the time complexities of the babysitting and grid

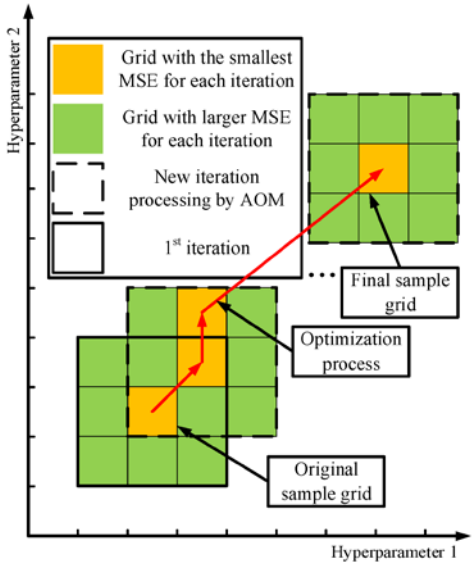


Fig. 6. The flowchart of the proposed approaching optimization method.

increase exponentially, preventing their use in the case of high hyper-parameter dimensions. The time complexity can be reduced through the random grid and the Bayes SMBO only if the sampling set is big enough. In this study, the AOM is employed with the procedure with two hyper-parameter dimensions illustrated in Fig. 6.

The specific steps are as follows:

1) Set the possible values of hyper-parameters respectively and form the grid space.
2) Determine a set of hyper-parameters through the trial-and-error method as an original sample grid, and train the LSTM model and calculate the MSE of the original sample grid.
3) Calculate the MSE under the hyper-parameters corresponding to all grids near the sample grid and select the one whose MSE is smaller than that of other grids as a new sample grid.
4) Repeat steps 1)-3) until the MSE corresponding to all adjacent grids is greater than that of the center sample grid. Select the hyper-parameters of the recent sample grid as the final hyper-parameters.

Since there are many hyper-parameters in the LSTM, some need to be determined beforehand to reduce the grid space dimensions. According to the inputs and output analyzed by the PCC, we set: ID=3, OD=1, and NNDL=NNLL=100. The LSTM is trained offline, and thus the LR can be set small to get a better LSTM model. Due to the large amounts of data and the offline trained algorithm, 1368920 samples of input data are used, which is one year's operation data of the vehicle No. 1. Then, the AOM is used to determine the other hyper-parameters. A 3-D grid space is established by the BS, epoch and LSCI. The initial values of the hyper-parameters are set as: BS=16, epoch=20, and LSCI=300. The hyper-parameters after optimization are obtained as: LR=0.001, TS=LSCI=360, BS=32, epoch =50, ID=3, OD=1, and NNDL=NNLL=100. No over-fitting was observed during the training process, and thus the RF and DF are not set.

B. *Voltage prediction based on the equivalent circuit model*

Various battery equivalent circuit models such as the Rint model, the Thevenin model and the PNGV model have been widely used for SOC estimation and open circuit voltage identification in the literature [58]. In particular, the Thevenin model outstands because of high modeling accuracy and low computational complexity, and is adopted here to delineate battery dynamics [59]. It can be reasonably assumed that battery SOC, temperature, and aging level remain unchanged during three successive time steps, and thereby the voltage can be predicted by

$$\begin{cases} \boldsymbol{K}_{Ls,t} = \boldsymbol{P}_{Ls,t-1}\boldsymbol{\Phi}_t^T[\boldsymbol{\Phi}_t\boldsymbol{P}_{Ls,t-1}\boldsymbol{\Phi}_t^T + \mu]-1 \\ \hat{\boldsymbol{\theta}}_t = \hat{\boldsymbol{\theta}}_{t-1} + \boldsymbol{K}_{Ls,t}[U_{t,c} - \boldsymbol{\Phi}_t\hat{\boldsymbol{\theta}}_{t-1}] \\ \boldsymbol{P}_{Ls,t} = \dfrac{1}{\mu}[\boldsymbol{I} - \boldsymbol{K}_{Ls,t}\boldsymbol{\Phi}_t]\boldsymbol{P}_{Ls,t-1} \end{cases} \tag{4}$$

where $\boldsymbol{K}_{Ls,t}$ is the gain matrix; $\boldsymbol{P}_{Ls,t\text{-}1}$ and $\boldsymbol{P}_{Ls,t}$ are the error covariance matrices of states estimate at the time steps t-1 and t; $\boldsymbol{\Phi}_t$ is the data matrix with $\boldsymbol{\Phi}_t$=[1 $U_{t\text{-}1}$ I_t $I_{t\text{-}1}$]$^{\mathrm{T}}$, where I_t (t=1, 2, ⋯ , k) is the current at the time step t; $\hat{\boldsymbol{\theta}}_t$ is the parameter matrix with $\hat{\boldsymbol{\theta}}_t$ =[(1-a_1)$U_{oc,t,c}$ a_1 a_2 a_3]$^{\mathrm{T}}$, where $U_{oc,t,c}$ (t=1, 2, ⋯ , k) is the OCV at the time step t of the battery cell c; a_1, a_2 and a_3 are the ECM parameters; μ is the forgetting factor.

The voltage prediction can be given by

$$U_{t+1,c} = \boldsymbol{\Phi}_{t+1}\hat{\boldsymbol{\theta}}_{t+1} = \boldsymbol{\Phi}_{t+1}\hat{\boldsymbol{\theta}}_t \tag{5}$$

C. *The coupling module*

The Adaptive boosting (AdaBoost) is an algorithm that can combine multiple weak classifiers to form a strong one. The improved AdaBoost can be used to solve the regression problem [60]. The weighted summation is adopted to combine the voltages predicted by the two models.

Predict the voltages using the LSTM and the ECM to get the prediction matrix $\boldsymbol{P}$ as

$$\boldsymbol{P} = \begin{pmatrix} \hat{U}_{1,1} & \cdots & \hat{U}_{1,b} \\ \cdots & \hat{U}_{t,c} & \cdots \\ \hat{U}_{a,1} & \cdots & \hat{U}_{a,b} \end{pmatrix} \tag{6}$$

where $\hat{U}_{t,c}$ (t=1, 2, ⋯ , a; c=1, 2, ⋯ , b) is the prediction voltage at time step t of the battery cell c.

Calculate the regression error rates of the LSTM and the ECM by

$$e = \frac{1}{m}\sum_{t=1}^{a}\sum_{c=1}^{b}\frac{(\hat{U}_{t,c} - U_{t,c})^2}{E^2} \tag{7}$$

where $U_{t,c}$ (t=1, 2, ⋯ , a; c=1, 2, ⋯ , b) is the real voltage at the time step t of the battery cell c; E is the maximum error in the training set with E=max|$\hat{U}_{t,c}$-$U_{t,c}$|; m is the number of samples with m=a×b.

Calculate the weight α for the LSTM and the ECM by

$$\alpha = \ln\left(\frac{1-e}{e}\right) \tag{8}$$

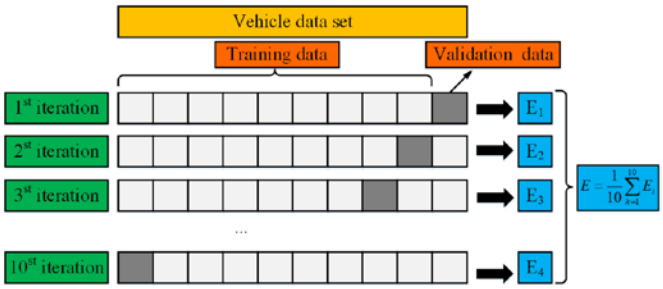


Fig. 7. The 10-fold cross-validation.

It is worth noting that α is calculated based the last ten time steps of data for online update. α_{LSTM} for the LSTM and α_{ECM} for the ECM are obtained based on Eq. (8). Calculate the final weights of the LSTM and the ECM, i.e., α'_{LSTM} and α'_{ECM}, to ensure the sum of the two weights equaling 1 by

$$\alpha'_{LSTM} = \frac{e^{\alpha_{LSTM}}}{e^{\alpha_{LSTM}} + e^{\alpha_{ECM}}} \tag{9}$$

$$\alpha'_{ECM} = \frac{e^{\alpha_{ECM}}}{e^{\alpha_{LSTM}} + e^{\alpha_{ECM}}} \tag{10}$$

Calculate the final prediction voltage $\hat{U}$ by

$$\hat{U} = \alpha'_{ECM} * \hat{U}_{ECM} + \alpha'_{LSTM} * \hat{U}_{LSTM} \tag{11}$$

where $\hat{U}_{LSTM}$ is the voltage predicted by the LSTM and $\hat{U}_{ECM}$ is the voltage predicted by the ECM.

Another way to obtain the prediction voltage is to switch the use of the two models based on an appropriate mechanism. At each time step, a sequence of the last ten-time-step data is used by the two models to generate their respective predictions, and the bigger one is adopted as the final prediction.

D. *The pre-judgment module*

It is essential to reduce computational time to achieve online voltage prediction and fault diagnosis. Here, the pre-judgment module is employed to select the battery cells which can be used to represent the performance of the entire battery pack. Two pre-judging rules are set to improve the accuracy while reducing computational time. They are:

1) Select the median value among all the battery cell voltages at the time step t-1 to get $\hat{U}_t$ which serves as the base for fault diagnosis.
2) If the median voltage $U_{t,median}$ at a certain moment t is predicted to be $\hat{U}_{t,median}$ and is judged to be faulty, $\hat{U}_{t,median}$ is used as the substitute of $U_{t,median}$ for use in the following time steps.

IV. Voltage Prediction Results and Discussions

A. *LSTM-based voltage prediction results*

In order to verify the accuracy of the LSTM for voltage prediction, 10-fold cross-validation is performed as shown in Fig. 7. The data collected from the vehicle No. 2 are evenly sorted into 10 datasets. For each interaction, one dataset is used as the validation dataset, and the others are used as the training datasets. The accuracy of the LSTM is then evaluated by the average MSE of the 10 validation datasets.

The data of the vehicle No. 2 are retrieved from the big data platform to serve as an example, and the data period covers

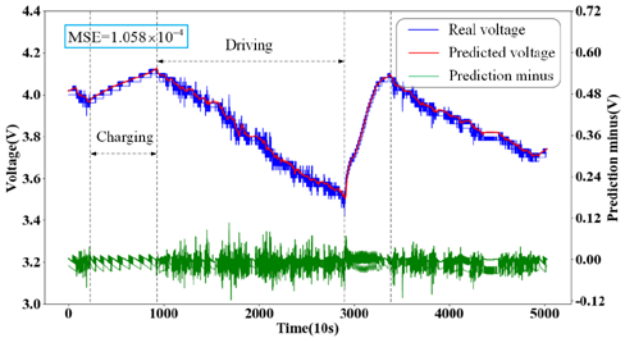


Fig. 8. Battery voltage prediction results based on the LSTM.

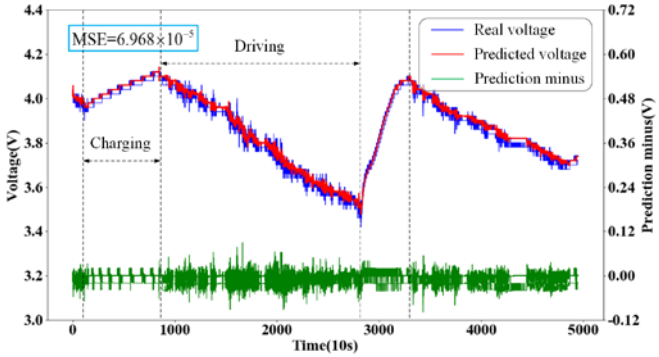


Fig. 9. Battery voltage prediction results based on the ECM.

from 10 June 2016 00:00:04 to 10 June 2016 23:59:57 with a sampling frequency of 0.1 Hz. The battery cell prediction results are illustrated in Fig. 8. It can be seen that the LSTM model exhibits good prediction performance and the predicted voltage can closely follow the real voltage with an MRE of $1.058*10^{-4}$.

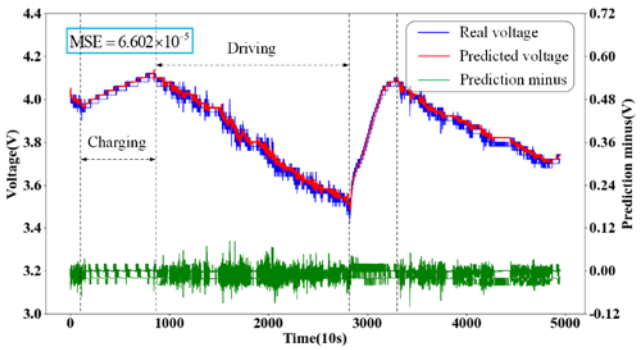
B. ECM-based voltage prediction results

The same data are used for battery voltage prediction based on the ECM. A sliding window is also used to extract sub-calculation interval for accurate battery voltage prediction. Based on Eqs. (4) and (5), the voltage and current at the time steps *t*-1 and *t* are used to calculate $\hat{\theta}_t$, while the current at the time step *t*+1 and $\hat{\theta}_t$ are used to predict the voltage at the time step *t*+1. The prediction results are shown in Fig. 9. It can be seen that the ECM model also has good prediction accuracy. The MSE is used to evaluate the errors between the prediction and the real values, which registers a reading of $6.968*10^{-5}$.

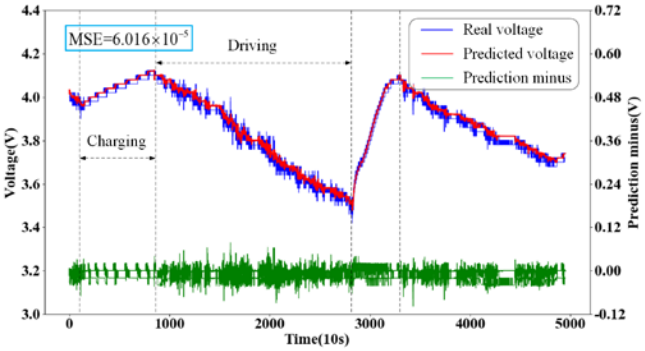
C. Discussions about the coupling module

Both the LSTM and the ECM perform well in battery voltage prediction. Compared with the LSTM, the prediction accuracy of the ECM is better, but the ECM cannot be used directly to distinguish the faulty battery cells from the normal ones. In contrast, the LSTM is able to better capture the voltage evolution trend, and is suitable for fault diagnosis based on the data collected from normal vehicles. Thus, a coupling module is utilized to combine the battery voltages predicted by the LSTM and the ECM to reduce the prediction error and improve the accuracy of fault diagnosis.

The results of model combining and switching are shown in Fig. 10. Their respective prediction MSEs are $6.602*10^{-5}$ and

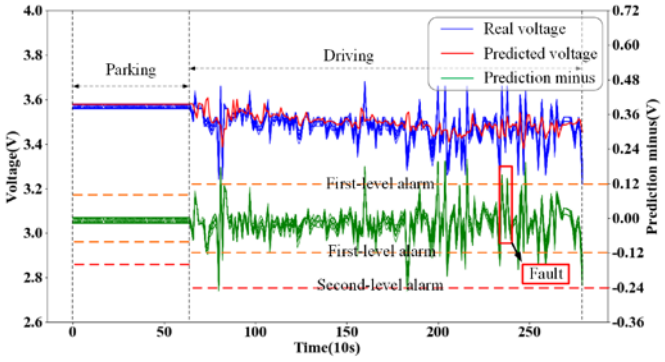


(a)

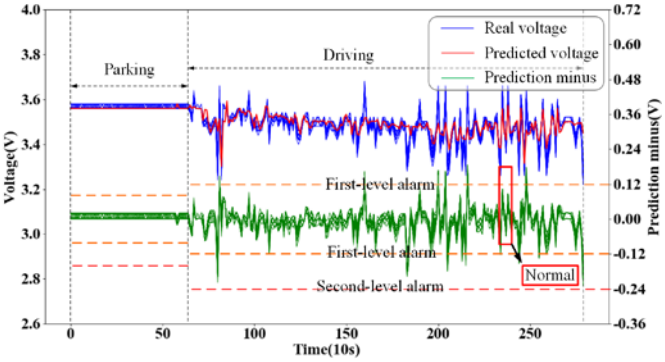


(b)

Fig. 10. Battery voltage prediction results based on: (a) the model combining method; (b) the model switching method.



(a)



(b)

Fig. 11. Battery fault diagnosis results based on: (a) the model combining method; (b) the model switching method.

$6.016*10^{-5}$, which means that the model switching method outperforms the model combining method. However, as the voltage predicted by the ECM is close to the real voltage, and the ECM is selected most of the time when the model switching method is used. This renders poorer fault diagnosis effect than

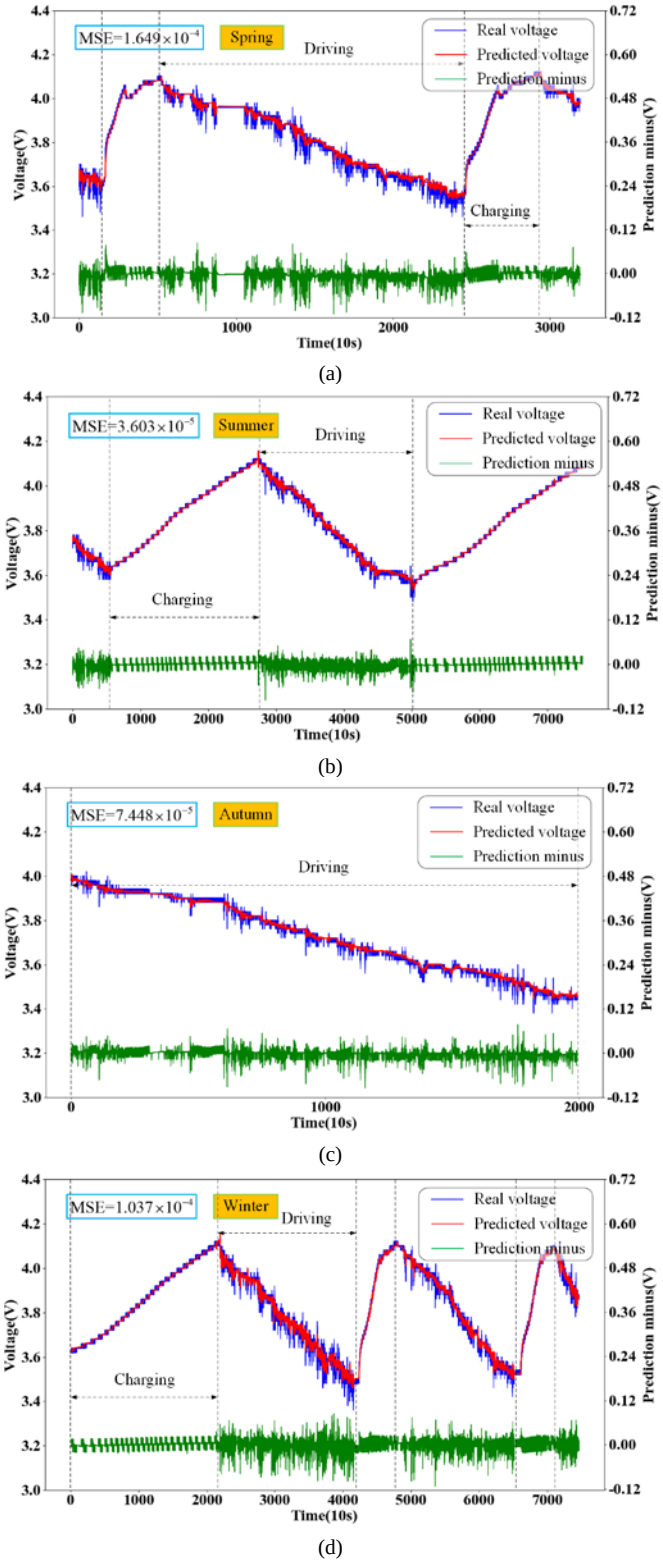


Fig. 12. The voltage prediction results for the vehicle No. 2 in different seasons: (a) Spring; (b) Summer; (c) Autumn; (d) Winter.

that using the model combining method.

The data of the vehicle No. 5 are also collected from the center, and the data period covers from 10 May 2016 15:32:04 to 10 May 2016 16:17:14. The results of the model combining and model switching methods in battery fault diagnosis are shown in Fig. 11. It can be seen that the batteries exhibit high

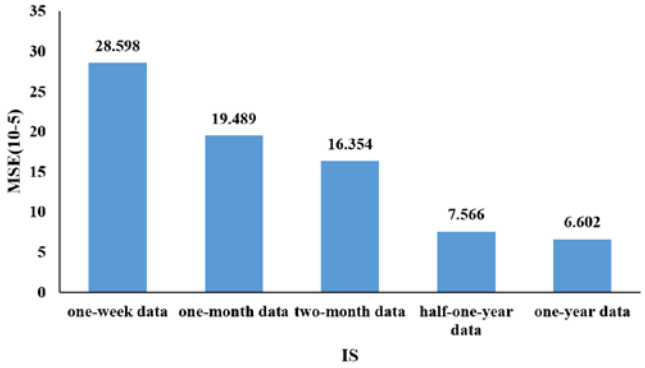


Fig. 13. The MSEs for different numbers of training samples.

fault tendency, and abnormal voltage fluctuations occur during hard accelerations and decelerations. As shown in Fig. 11(b), the model switching method can be misled by the faulty battery cell, thereby failing fault diagnosis in certain situations.

In general, the LSTM is based on large amounts of real-world operation data collected from normal vehicles, which can predict the battery voltage more accurately, thus improving fault diagnosis accuracy. The ECM considers the influence of current in the process of voltage prediction, and can have voltage prediction accuracy during practical vehicular operation. The model combining method can fully capitalize on the advantages of the two models for better diagnosis performance and is thereby adopted in this study.

D. Verification of the proposed voltage prediction method based on real-world operation data

In order to verify the robustness of the proposed voltage prediction method, the data of the vehicle No. 2 are collected from the center, and the operation data on 1st February 2016, 1st May 2016, 1st August 2016 and 1st November 2016 are selected to represent the operating conditions of the Spring, the Summer, the Autumn and the Winter. The voltage prediction results are shown in Fig. 12. The MSEs of the four validation datasets are $1.649*10^{-4}$, $3.603*10^{-5}$, $7.448*10^{-5}$, and $1.037*10^{-4}$, respectively. It is evident that the MSE for the Summer is the smallest, which means the best prediction accuracy. The MSEs in the Spring and the Winter are slightly higher, indicating that the accuracy of voltage prediction can be affected by seasons and temperature. In general, the MSEs of four seasons are low, which verifies the effectiveness of the proposed voltage prediction method.

E. Discussions about training samples

In order to assess the influence of the number of training samples on voltage prediction accuracy, the MSEs with different numbers of input samples are calculated as shown in Fig.13. There is a decreasing trend for the MSEs with the increasing number of training samples. The MSE varies according to the number of training samples, which can be set beforehand in practice.

TABLE I
ALARMING LEVELS AND THRESHOLDS TO PERFORM REAL-WORLD FAULT DIAGNOSIS

Threshold (discharge/charge or park)	Fault type	Alarm level
$U_t-\hat{U}_t \le$-0.36 V / $U_t-\hat{U}_t \le$-0.24 V	Under-voltage	3
-0.36 $V \le U_t-\hat{U}_t \le$-0.24 V / -0.24 $V \le U_t-\hat{U}_t \le$-0.16 V	Under-voltage	2
-0.24 $V \le U_t-\hat{U}_t \le$-0.12 V / -0.16 $V \le U_t-\hat{U}_t \le$-0.08 V	Under-voltage	1
-0.12 $V \le U_t-\hat{U}_t \le$0.12 V / -0.08 $V \le U_t-\hat{U}_t \le$0.08 V	Normal	
0.12 $V \le U_t-\hat{U}_t \le$0.24 V / 0.08 $V \le U_t-\hat{U}_t \le$0.16 V	Over-voltage	1
0.24 $V \le U_t-\hat{U}_t \le$0.36 V / 0.16 $V \le U_t-\hat{U}_t \le$0.24 V	Over-voltage	2
$U_t-\hat{U}_t \ge$0.36 V / $U_t-\hat{U}_t \ge$0.24 V	Over-voltage	3

V. BATTERY FAILURE DIAGNOSIS

A. Battery fault diagnosis

In order to achieve fault diagnosis in real-world EV operation, the data of the vehicle No. 1 are used to train the LSTM model, and different fault levels with their respective voltage thresholds are set to perform in-site fault diagnosis as shown in Table I. As sensor and prediction errors may have an impact on the settings of the thresholds, the statistical method is used to calculate the maximum prediction error of the vehicle No. 1 which is employed as the threshold of the first-level warning. It is worth mentioning that the proposed fault diagnosis method is also feasible for other battery systems based on proper threshold settings.

As seen from Table I, the voltage abnormities are divided into over-voltage and under-voltage, each of which is further divided into three levels. The first-level warning is a relatively safe state, but it is necessary to timely intervene for avoiding the potential failures. The second-level warning means the relevant batteries are in a dangerous state, and the vehicle is required to be stopped for a close and carful check. When the third-level warning is triggered, the driver needs to stop the vehicle and immediately get out of it. The first-level and second-level warnings are used to assess the potential failure risk while the third-level warning is used as the early alarming of thermal runaway occurrence.

B. Potential failure risk assessment

In order to achieve online fault diagnosis, the fault matrix $F_{k\times n}$ is defined to record the warning levels of each battery cell at each time step, which is given by

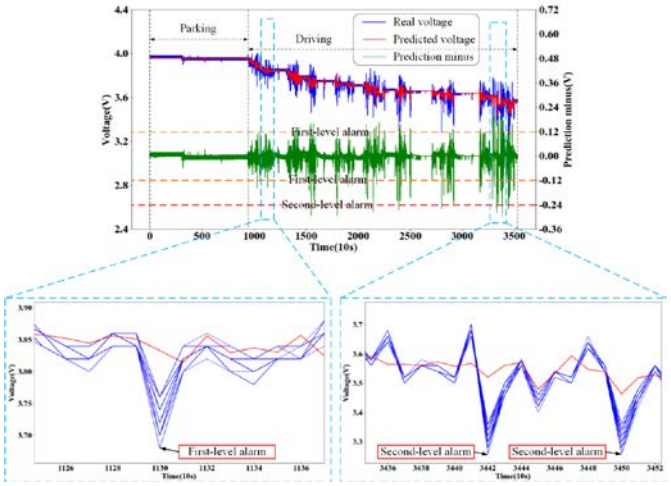


Fig. 14. Potential battery failure risk assessment results based on the vehicle No. 6.

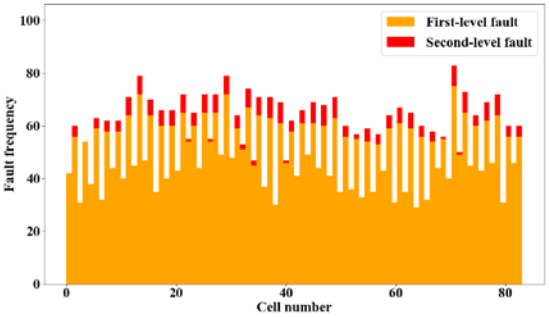


Fig. 15. Fault frequencies of the battery cells based on the vehicle No. 6.

$$\boldsymbol{F}_{k\times n}=\begin{pmatrix} f_{1,1} & \cdots & f_{1,n} \\ \cdots & f_{t,j} & \cdots \\ f_{k,1} & \cdots & f_{k,n} \end{pmatrix} \tag{12}$$

where $f_{t,j}$ (t=1, 2, ⋯ , k; j=1, 2, ⋯ , n) is the warning level of the battery cell j at the time step t. At each time step, the vehicle generates a new data sequence, and the warning level is calculated according to Table I.

Define the fault frequency matrix to quantitatively describe the failure risk for a period of time with

$$\boldsymbol{R}_{k\times n}=\begin{pmatrix} r_1 & \cdots & r_n \end{pmatrix} \tag{13}$$

where r_j (j=1, 2, ⋯ , n) is the fault occurrence frequency of the battery cell j from the time step x to y, with each warming level calculated with $r_j=\sum_{t=x}^{y} f_{t,j}$; x and y are the starting and final time steps of the selected time interval.

The data of the vehicle No. 6 are retrieved from the center, and the data period covers from 15 February 2016 17:46:33 to 16 February 2016 18:14:50. As shown in Fig. 14, abnormal voltage fluctuations occur to some battery cells during driving, which are accurately detected by the proposed method. As shown in Fig. 10, the voltage can also be accurately predicted when there is normal voltage fluctuation caused by large currents due to the use of the ECM. The battery cells can be diagnosed as faulty ones only when the error between the real and the predicted voltage is large enough. The fault frequencies of the battery cells calculated by the regular risk assessment

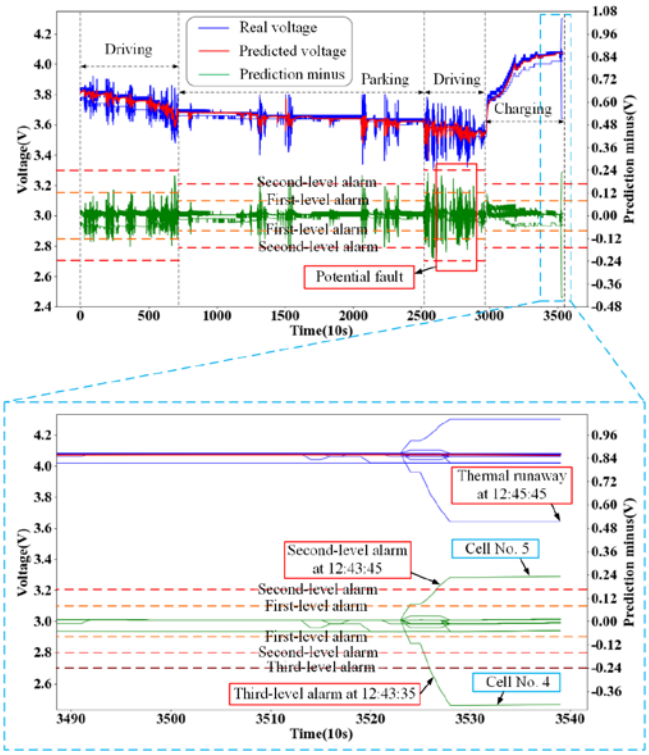


Fig. 16. Thermal runaway early warning results based on the vehicle No. 9.

strategy are high as shown in Fig. 15, which verifies the effectiveness of the proposed fault diagnosis method.

C. Thermal runaway early warning

The data of the vehicle No. 9 are retrieved from the center, and the data period covers from 18 March 2016 1:23:55 to 18 March 2016 12:45:45. Thermal runaway occurred to the battery cells No. 4 and No. 5 at 12:45:45, and their predicted voltages are depicted in Fig. 16. It can be seen that the first-level fault warnings are triggered multiple times in the last driving segment before thermal runaway occurrence, and the potential faulty cells can be precisely located by the proposed model. In the last charging segment before thermal runaway occurrence, the first-level and the third warnings are triggered for the battery cell No.4 150 seconds and 130 seconds before thermal runaway occurrence, respectively. Similarly, the first-level and the second-level warnings are also issued for the battery cell No.4 150 seconds and 120 seconds before thermal runaway occurrence. This verifies that the proposed model can not only detect potential cell failures in battery packs, but also provide the early warning of thermal runaway occurrence.

D. Verification of robustness, reliability and superiority

An effective fault diagnosis model should have the following three properties. They are:

1) Robustness: The model can accurately diagnose faults for other vehicles.
2) Reliability: The model should not send out false warnings to normal vehicles.
3) Superiority: The model has higher prediction accuracy compared with other methods.

The data of the vehicle No. 7 are retrieved from the center, and the data period covers from 27 February 2016 16:43:42 to

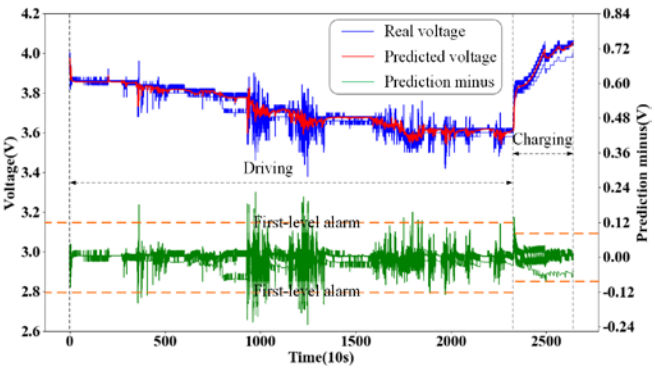


Fig. 17. Potential battery failure risk assessment results based on the vehicle No. 7.

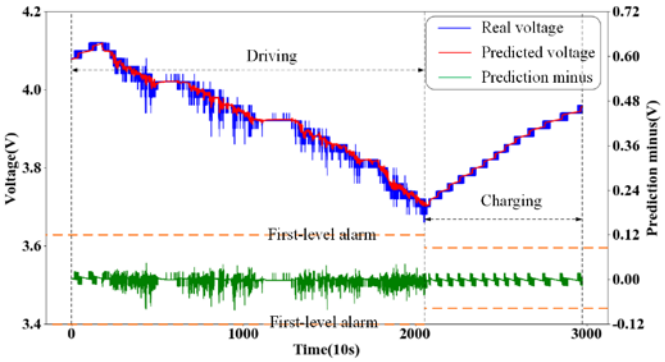


Fig. 18. Battery voltage prediction results based on the vehicle No. 3.

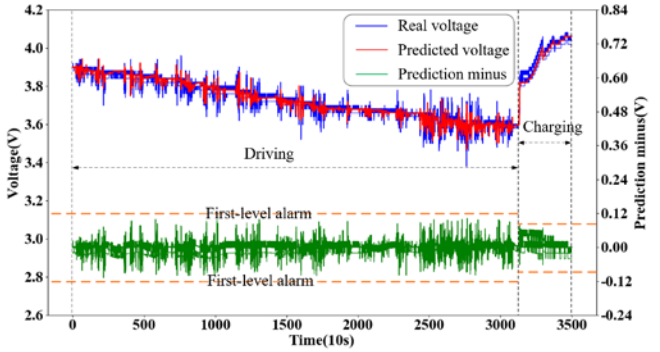
2 March 2016 16:46:02. The real and predicted voltages are shown in Fig. 17. At the time steps 987, 1257 and 1286, there are abnormal voltage fluctuations for most battery cells and the first-level alarming is triggered. All the abnormal battery cells can be detected by the proposed method, which verifies the robustness of the method.

The data of the vehicle No. 3 are retrieved from the center, and the data period covers from 28 April 2016 00:00:05 to 28 April 2016 23:59:54. The real and predicted voltages are depicted in Fig. 18. There are no abnormities in battery cells and all battery cells are considered in normal operation by the proposed diagnosis method. This serves as an example to illustrate that the proposed fault diagnosis method will not falsely send out warnings, which verifies its reliability.

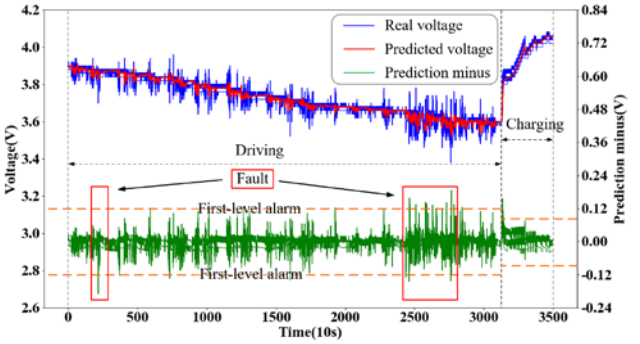
To verify the superiority of the method, the MSE of voltage prediction based on the method presented in Ref. [41] is used for comparison. As the voltages can be predicted in the six following time steps by the method, and only one time step prediction is used for fair comparison. The MSE of the comparison method is $7.042*10^{-3}$ while the proposed method has a MSE of $6.602*10^{-5}$, which verifies the superiority of the proposed method. This may be ascribed to the fact that the comparison method fails to consider the influence of current for voltage prediction and strongly relies on battery data including fault occurrence for battery fault diagnosis. Besides, it cannot detect the faulty battery cells whose voltage and voltage change rate are within their respective thresholds.

E. Further discussions about necessity

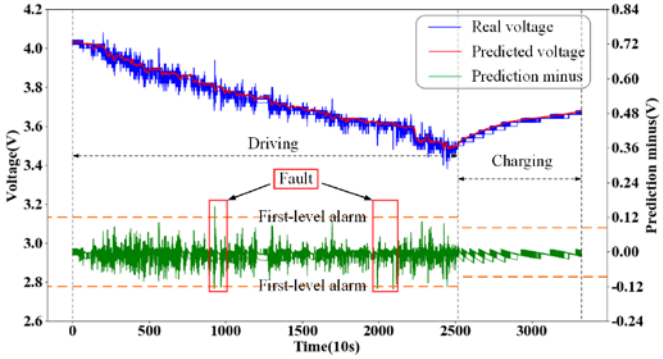
The data of the vehicles No. 4 and No. 8 are retrieved from the center. The data period of the vehicle No. 4 covers from 30



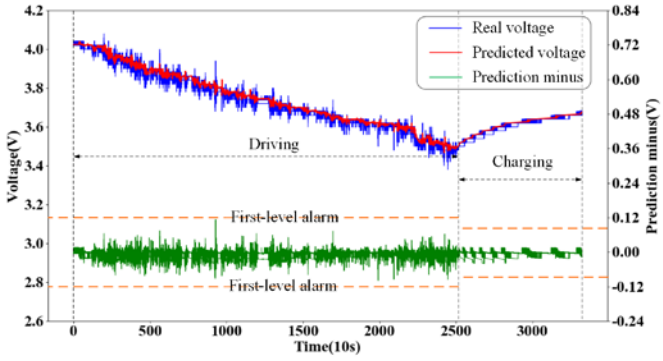
(a)



(b)



(c)



(d)

Fig. 19. Battery voltage prediction results. The real and predicted voltages: (a) by the ECM based on the vehicle No. 8; (b) by the model combining method based on the vehicle No. 8; (c) by the LSTM based on the vehicle No. 4; (d) by the model combining method based on the vehicle No. 4.

July 2016 03:45:53 to 30 July 2016 22:10:43, and the data period of the vehicle No. 8 covers from 26 June 2016 08:52:51 to 26 June 2016 22:25:46. The real and predicted voltages are shown in Fig. 19. As shown in Fig. 19(a), there are abnormal voltage fluctuations for some battery cells in the vehicle No. 8. However, the ECM alone cannot diagnose the faulty battery cells. The ECM can estimate the internal parameters with the first two time steps of data, based on which the terminal voltage can be predicted at the third time step. The primary goal of the ECM is to minimize the difference between the predicted and the actual voltages, and this results in good voltage prediction performance yet is unable to distinguish the faulty battery cells from the normal ones. It necessitates the combination with the LSTM for accurate diagnosis as shown in Fig. 19(b). Battery cells in the vehicle No. 4 are normal without any fault occurrence. As shown in Fig. 19(c), during regenerative braking, some battery cells may be falsely diagnosed as faulty battery cells when only employing the LSTM for battery fault diagnosis. This shows that the LSTM also needs to be combined with the ECM to achieve better diagnosis performance, which indicates the necessity of combining the two models.

VI. Conclusions

This paper presents a battery fault diagnosis method by combining the LSTM and the ECM model. A coupling module based on the modified adaptive boosting is adopted to make full use of the respective advantages of the LSTM and ECM models. In order to improve computational efficiency, a pre-judgment module is also used. The LSTM model can be offline trained and online implemented, which facilitates fast diagnosis for battery systems. The proposed method can utilize large quantities of data collected from normal EVs for the LSTM model training. Moreover, a real-world fault diagnosis strategy is proposed to realize online fault diagnosis. Finally, the robustness, reliability, superiority and necessity of the proposed method are verified based on massive real-world operation data.

References

[1] J. Zhang, L. Zhang, F. Sun, and Z. Wang, “An overview on thermal safety issues of lithium-ion batteries for electric vehicle application,” IEEE Access, vol. 6, pp. 23848-23863, 2018.

[2] J. Shen, J. Shen, Y. He, and Z. Ma, “Accurate state of charge estimation with model mismatch for Li-ion batteries: A Joint Moving Horizon Estimation Approach,” *IEEE Trans. Power Electron.*, vol. 34, no. 5, pp. 4329-4342, May 2019.

[3] Z. Song, X. Wu, X. Li, J. Sun, H. Hofmann, and J. Hou, “Current profile optimization for combined state of charge and state of health estimation of lithium ion battery based on Cramer-Rao bound analysis,” *IEEE Trans. Power Electron.*, vol. 34, no. 7, pp. 7067-7078, Jul. 2019.

[4] A. El Mejdoubi, H. Chaoui, H. Gualous, P. Van den Bossche, N. Omar, and J. Van Mierlo, “Lithium-ion batteries health prognosis considering aging conditions,” IEEE Trans. Power Electron., vol. 34, no. 7, pp. 6834-6844, Jul. 2019.

[5] X. Feng, M. Ouyang, X. Liu, L. Lu, Y. Xia, X. He, “Thermal runaway mechanism of lithium ion battery for electric vehicles: A review,” *Energy Storage Mater*, vol. 10, pp. 246-267, 2018.

[6] Q. Wang, B. Mao, S. Stoliarov, and J. Sun, “A review of lithium ion battery failure mechanisms and fire prevention strategies,” *Prog. Energy Combust. Sci.*, vol. 73, pp. 95-131, Jul. 2019.

[7] R. Zhao, J. Liu and J. Gu, “Simulation and experimental study on lithium ion battery short circuit,” *Appl. Energy*, vol. 173, pp. 29-39, 2016.

[8] B. Duan, Z. Li, P. Gu, Z. Zhou, C. Zhang, “Evaluation of battery inconsistency based on information entropy,” *J. Energy Storage*, vol. 16, pp. 160-166, 2018.

[9] Z. Chen, R. Xiong, J. Tian, X. Shang, and J. Lu, “Model-based fault diagnosis approach on external short circuit of lithium-ion battery used in electric vehicles,” *Appl. Energy*, vol. 184, pp. 365-374, 2016.

[10] J. Son and Y. Du, “Model-Based Stochastic Fault Detection and Diagnosis of Lithium-Ion Batteries,” *Processes*, vol. 7, no. 1, 2019.

[11] X. Feng, Y. Pan, X. He, L. Wang, and M. Ouyang, "Detecting the internal short circuit in large-format lithium-ion battery using model-based fault-diagnosis algorithm," *J. Energy Storage*, vol. 18, pp. 26-39, Aug. 2018.

[12] R. Yang, R. Xiong, H. He, and Z. Chen, "A fractional-order model-based battery external short circuit fault diagnosis approach for all-climate electric vehicles application,", *J. Clean Prod.*, vol. 187, pp. 950-959, Jun. 2018.

[13] Y. Wang, J. Tian, Z. Chen, and X. Liu, "Model based insulation fault diagnosis for lithium-ion battery pack in electric vehicles," *Measurement*, vol. 131, pp. 443-451, 2019.

[14] Z. Liu, and H. He, "Sensor fault detection and isolation for a lithium-ion battery pack in electric vehicles using adaptive extended Kalman filter," *Appl. Energy*, vol. 185, pp. 2033-2044, 2017.

[15] B. Xia, Y. Shang, T. Nguyen, C. Mi A correlation based fault detection method for short circuits in battery packs," *J. Power Sources*, vol. 337, pp. 1-10, 2017.

[16] X. Gong, R. Xiong, and C. Mi, "A data-driven bias-correction-method-based lithium-ion battery modeling approach for electric vehicle applications," *IEEE Trans. Ind. Appl.*, vol. 52, no. 2, pp. 1759-1765, 2016.

[17] X. Hu, J. Jiang, D. Cao, and B. Egardt, "Battery health prognosis for electric vehicles using sample entropy and sparse bayesian predictive modeling," *IEEE Trans. Ind. Electron.*, vol. 63, pp. 1-11, 2015

[18] Y. Kang, B. Duan, Z. Zhou, Y. Shang, and C. Zhang, "A multi-fault diagnostic method based on an interleaved voltage measurement topology for series connected battery packs," *J. Power Sources*, vol. 417 , pp. 132-144, 2019.

[19] Y. Shang, G. Lu, Y. Kang, Z. Zhou, B. Duan, and C. Zhang, "A multi-fault diagnosis method based on modified Sample Entropy for lithium-ion battery strings," *J. Power Sources*, vol. 446, no. 227275, 2020.

[20] Y. Zhao, P. Liu, Z. Wang, L. Zhang, and J. Hong, "Fault and defect diagnosis of battery for electric vehicles based on big data analysis methods," *Appl. Energy*, vol. 207, pp. 354-362, 2017.

[21] Y. Zheng, X. Han, L. Lu, J. Li, and M. Ouyang, "Lithium ion battery pack power fade fault identification based on Shannon entropy in electric vehicles," *J. Power Sources*, vol. 223, pp. 136-146, 2013.

[22] Z. Wang, J. Hong, P. Liu, and L. Zhang, "Voltage fault diagnosis and prognosis of battery systems based on entropy and Z -score for electric vehicles," *Appl. Energy*, vol. 196, pp. 289-302, 2017.

[23] D. Li, Z. Zhang, P. Liu, and Z. Wang, "DBSCAN-Based Thermal Runaway Diagnosis of Battery Systems for Electric Vehicles," *Energies*, vol. 12, no. 15, 2019.

[24] P. Liu, Z. Sun, Z. Wang, and J. Zhang, "Entropy-based voltage fault diagnosis of battery systems for electric vehicles," *Energies*, vol. 11, no. 1, 2018.

[25] F. Feng, X. Hu, L. Hu, F. Hu, Y. Li, L. Zhang, "Propagation mechanisms and diagnosis of parameter inconsistency within Li-Ion battery packs," *Renewable Sustainable Energy Rev.*, vol. 112, pp. 102-113, 2019.

[26] M. Chen, J. Liu, D. Ouyang, and J. Wang, "Experimental investigation on the effect of ambient pressure on thermal runaway and fire behaviors of lithium-ion batteries," *Int. J. Energy Res.*, vol. 43, pp. 4898-4911, Jul. 2019.

[27] S. Abada, M. Petit, A. Lecocq, G. Marlair, V. Sauvant, and F. Huet, "Combined experimental and modeling approaches of the thermal runaway of fresh and aged lithium-ion batteries," *J. Power Sources*, vol. 399, pp. 264-273, 2018.

[28] X. Cheng, T. Li, X. Ruan, and Z. Wang, "Thermal Runaway Characteristics of a Large Format Lithium-Ion Battery Module," *Energies*, vol. 12, no. 16, Aug. 2019.

[29] H. Li, Q. Duan, C. Zhao, Z. Huang, and Q. Wang, "Experimental investigation on the thermal runaway and its propagation in the large format battery module with Li (Ni1/3Co1/3Mn1/3) O-2 as cathode,", *J. Hazard. Mater.*, vol. 375, pp. 241-254, Aug 2019.

[30] D. Ouyang, Y. He, M. Chen, J. Liu, and J. Wang, "Experimental study on the thermal behaviors of lithium-ion batteries under discharge and overcharge conditions," *J. Therm. Anal. Calorim.*, vol. 132, pp. 65-75, 2017.

[31] L. Zhao, M. Zhu, X. Xu, and J. Gao, "Thermal runaway characteristics on NCM lithium-ion batteries triggered by local heating under different heat dissipation conditions," *Appl. Therm. Eng.*, vol. 159, Aug. 2019.

[32] Z. An, K. Shah, L. Jia, and Y. Ma, "Modeling and analysis of thermal runaway in Li-ion cell," *Appl. Therm. Eng.*, vol. 160, Sep 2019.

[33] T. Cai, A. Stefanopoulou, and J. Siegel, "Modeling Li-Ion Battery Temperature and Expansion Force during the Early Stages of Thermal Runaway Triggered by Internal Shorts," *J. Electrochem. Soc.*, vol. 166, no. 12, pp. A2431-A2443, Jul. 2019.

[34] D. Ren, X. Feng, L. Lu, M. Ouyang, S. Zheng, J. Li, and X. He, "An electrochemical-thermal coupled overcharge-to-thermal-runaway model for lithium ion battery," *J. Power Sources*, vol. 364, pp. 328-340, 2017.

[35] M. Said and M. Tohir, "Prediction of Lithium-ion Battery Thermal Runaway Propagation for Large Scale Applications Fire Hazard Quantification," *Processes*, vol. 7, no. 10, Oct. 2019.

[36] C. Qi, Y. Zhu, F. Gao, K. Yang, and Q. Jiao, "Mathematical model for thermal behavior of lithium ion battery pack under overcharge," *Int. J. Heat Mass Transf.*, vol. 124, pp. 552-563, 2018.

[37] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436-444, May 2015.

[38] F. Pedregosa, "Scikit-learn: Machine Learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825-2830, Oct. 2011.

[39] F. Tobar, I. Castro, J. Silva, and M. Orchard, "Improving battery voltage prediction in an electric bicycle using altitude measurements and kernel adaptive filters," *Pattern Recognit. Lett.*, vol. 105, pp. 200-206, 2017.

[40] R. Zhao, P. J. Kollmeyer, R. D. Lorenz, and T. M. Jahns. "A compact unified methodology via a recurrent neural network for accurate modeling of lithium-ion battery voltage and state-of-charge," in *IEEE Energy Conversion Congress and Exposition Proceedings*, 2017, pp. 5234-5241.

[41] J. Hong, Z. Wang, Y. Yao, "Fault prognosis of battery system based on accurate voltage abnormity prognosis using long short-term memory neural networks," *Appl. Energy*, vol. 51, 2019.

[42] M. Hannan, M. S. H. Lipu, A. Hussain, and A. Mohamed, "A review of lithium-ion battery state of charge estimation and management system in electric vehicle applications: Challenges and recommendations," *Renew. Sust. Energ. Rev.*, vol. 78, pp. 834-854, Oct. 2017.

[43] R. Xiong, J. Cao, Q. Yu, H. He, and F. Sun, "Critical review on the battery state of charge estimation methods for electric vehicles," *IEEE Access*, vol. 6, pp. 1832-1843, 2018.

[44] R. Xiong, L. Li, and J. Tian, "Towards a smarter battery management system: A critical review on battery state of health monitoring methods," *J. Power Sources*, vol. 405, pp. 18-29, Nov. 2018.

[45] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85-117, Jan. 2015.

[46] L. Zhou and Z. Zhao, "Exponential synchronization and polynomial synchronization of recurrent neural networks with and without proportional delays," *Neurocomputing*, vol. 372, pp. 109-116, 2020..

[47] S. Hochreiter and J. Schmidhuber, ""Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735-1780, Nov. 1997.

[48] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with LSTM," *Neural Comput.*, vol. 12, no. 10, pp. 2451-2471, Oct. 2000.

[49] Y. Xie, R. Liang, Z. Liang, C. Huang, C. Zou, and B. Schuller, "Speech emotion classification using attention-based LSTM," *IEEE-ACM Trans. Audio Speech Lang.*, vol. 27, no. 11, pp. 1675-1685, Nov. 2019.

[50] T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," *Eur. J. Oper. Res.*, vol. 270, no. 2, pp. 654-669, Oct. 2018.

[51] R. Zazo, A. Lozano-Diez, J. Gonzalez-Dominguez, D. T. Toledano, and J. Gonzalez-Rodriguez, "Language identification in short utterances using long short-term memory (LSTM) recurrent neural networks," *PLoS One*, vol. 11, no. 1, Jan. 2016.

[52] A. Ullah, K. Muhammad, J. Del Ser, S. W. Baik, and V. H. C. de Albuquerque, "Activity recognition using temporal optical flow convolutional features and multilayer LSTM," *IEEE Trans. Ind. Electron.*, vol. 66, no. 12, pp. 9692-9702, Dec. 2019.

[53] F. Galton, "Section h; anthropology; opening address," *Nature*, vol. 32, pp. 507-510, 1885.

[54] K. Pearson, "Note on regression and inheritance in the case of two parents," *Proc. Roy. Soc. London* vol. 58, pp. 240-242, 1895.

[55] Y. Zhang, R. Xiong, H. He, and M. Pecht. "Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries," *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 5695-5705, 2018.

[56] D. P. Kingma, and J. Ba, "Adam: A method for stochastic optimization," *Comput. Sci.*, 2014.

[57] G. Ian, B. Yoshua, and C. Aaron, "Deep Learning: Adaptive Computation and Machine Learning series," *People Post Press*, 2018.

[58] X. Hu, S. Li, and H. Peng, "A comparative study of equivalent circuit models for Li-ion batteries," *J. Power Sources*, vol. 198, pp. 359-367, 2012.

[59] H. He, R. Xiong, and J. Fan, "Evaluation of lithium-ion battery equivalent circuit models for state of charge estimation by an Experimental Approach," *Energies*, vol. 4, no. 4, pp. 582-598, Apr. 2011.

[60] J. Zhu, H. Zou, S. Rosset, and T. Hastie, "Multi-class AdaBoost," *Stat. Interface*, vol. 2, no. 3, pp. 349-360, 2009.

Da Li received the B.S. degree in Automotive Engineering from the Harbin Institute of Technology, China, in 2018, and he is currently pursuing the Ph.D. degree in mechanical engineering with the National Engineering Laboratory for Electric Vehicles, Beijing Institute of Technology, Beijing, China. His research interests mainly include battery fault diagnosis and deep learning algorithms development.

Zhaosheng Zhang received the Ph.D. degree in Automotive Engineering from the Tsinghua University, Beijing, China, in 2013. He is currently a Lecturer with the School of Mechanical Engineering, Beijing Institute of Technology. He has published 2 monographs and over 20 technical papers. He also holds over 10 patents. His current research interests include intelligent transportation and big data analysis.

Peng Liu received the Ph.D. degree in Mechanical Engineering from the Beijing Institute of Technology, Beijing, China, in 2011. He is currently an Associate Professor with the School of Mechanical Engineering, Beijing Institute of Technology. He has published 2 monographs and 10 technical papers. He also holds 17 patents. His current research interests include battery fault diagnosis, intelligent transportation and big data analysis. He was a recipient of numerous awards including the third Prize of Beijing Science and Technology Award and the First Prize of China Automotive Industry Science and Technology Award.

Zhenpo Wang (M'11) received the Ph.D. degree in Automotive Engineering from Beijing Institute of Technology, Beijing, China, in 2005.

He is currently a Professor with the Beijing Institute of Technology, and the Director of National Engineering Laboratory for Electric Vehicles. His current research interests include pure electric vehicle integration, packaging and energy management of battery systems and charging station design.

Prof. Zhenopo Wang has been the recipient of numerous awards including the second National Prize for Progress in Science and Technology and the first Prize for Progress in Science and Technology from the Ministry of Education, China and the second Prize for Progress in Science and Technology from Beijing Municipal, China. He has published 4 monographs and translated books as well as more than 80 technical papers. He also holds more than 60 patents.

Lei Zhang (S'12-M'16) received the Ph.D. degree in Mechanical Engineering from Beijing Institute of Technology, Beijing, China, and the Ph.D. degree in Electrical Engineering from University of Technology, Sydney, Australia, in 2016. He is now an Associate Professor with the School of Mechanical Engineering, Beijing Institute of Technology.

His research interest includes management techniques for energy storage systems, and vehicle dynamics and advanced control for electric vehicles.