An online data driven fault diagnosis and thermal runaway early warning for electric vehicle batteries

Sun, Zhenyu; Wang, Zhenpo; Liu, Peng; Qin, Zian; Chen, Yong; Han, Yang; Wang, Peng; Bauer, Pavol

**Citation (APA)**

**Important note**
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

# An Online Data-Driven Fault Diagnosis and Thermal Runaway Early Warning for Electric Vehicle Batteries

Zhenyu Sun , Zhenpo Wang , *Senior Member, IEEE*, Peng Liu, Zian Qin , *Senior Member, IEEE*,
Yong Chen , Yang Han, Peng Wang, and Pavol Bauer , *Senior Member, IEEE*

*Abstract*—**Battery fault diagnosis is crucial for stable, reliable, and safe operation of electric vehicles, especially the thermal runaway early warning. Developing methods for early failure detection and reducing safety risks from failing high energy lithium-ion batteries has become a major challenge for industry. In this article, a real-time early fault diagnosis scheme for lithium-ion batteries is proposed. By applying both the discrete Fréchet distance and local outlier factor to the voltage and temperature data of the battery cell/module that measured in real time, the battery cell that will have thermal runaway is detected before thermal runaway happens. Compared with the widely used single parameter based diagnosis approach, the proposed one considerably improve the reliability of the fault diagnosis and reduce the false diagnosis rate. The effectiveness of the proposed method is validated with the operational data from electric vehicles with/without thermal runaway in daily use.**

*Index Terms*—**Discrete Fréchet distance (DFD), fault diagnosis, lithium-ion battery (LIB), local outlier factor (LOF).**

## I. INTRODUCTION

WITH the ambition to achieve carbon neutrality in 2050, vehicle electrification has been attracting a significant amount of attention [1]. Lithium-ion batteries (LIBs) are widely used for applications on electric vehicles (EVs) due to their relatively low self-discharge rates, high energy density, high power density, long cycle life compared with the counterparts, such as lead–acid batteries, nickel–cadmium batteries, nickel–hydrogen batteries, etc. [2]. To meet the voltage, power, and energy requirements of the applications, a typical LIB system contains many battery cells connected in series and parallel [3], [4]. It is of high importance and also quite challenging to maintain safety of battery packs under complex operating conditions [5]. Recently, the increasing number of reported EV fire accidents has drawn much attention, most of which are due to battery system catching fire [6], [7]. As a result, the fault diagnosis and early detection of thermal runaway of EVs become essential.

However, the failure mechanism of LIB system is extremely complicated because it is a nonlinear time-varying system with a mass of dynamic electrochemical and mechanical phenomena [8]–[10]. Faults of the LIB system can be categorized into internal and external ones [11]. The former include overcharge, overdischarge, internal short circuit, external short circuit, and overheat. The latter include fault of sensors, connection, and cooling system. Both of the two types of faults can lead to abnormal voltages, temperatures, and pressure in the battery pack [12], [13]. Nonetheless, there is a higher possibility that the internal faults cause a fire accident or explosion rather than the external faults.

The fault diagnosis approaches of LIB systems can be classified into three types: rule-based, model-based, and data-driven methods. For rule-based methods, if the value of critical parameters, such as voltage and temperature, exceed the specified threshold, an alarm will be tripped. For example, Pesaran *et al.* [14] suggested the maximum temperature difference between modules should be less than 5 °C. To diagnose external short circuit fault, Xia *et al.* [15] proposed three diagnostic criteria utilizing measurement parameters, such as the voltage, current, and temperature increase. However, the appropriate threshold is hard to define, due to the complex operating circumstances of the EV. If the threshold is too low, the alarm can be too sensitive and easily faulty tripped; on the other hand, if the threshold is too high, when the alarm is tripped, severe faults may already happen and make the fault detection meaningless.

For model-based methods, analytical models are used to generate residuals, which can be monitored and analyzed to detect faults. With simulation on a three-cell battery string model, Chen *et al.* [16] proposed a bank of reduced-order Luenberger observers and the learning observers for fault isolation and estimation. Liu and He [17] developed adaptive extended Kalman filter to estimate output voltage of equivalent circuit

models (ECMs) and detect sensor fault by residual evaluation. Dey *et al.* presented a thermal model to diagnose thermal faults of LIBs by Luenberger observer [18], nonlinear observer [19], and partial differential equation [20]. Feng *et al.* [21], [22] raised an electrochemical–thermal coupled model to internal short circuit. In general, the model-based methods are more accurate and reliable than threshold-based methods. Nevertheless, the computational cost is much larger, and some sophistic and highly coupled electrochemical models can hardly be implemented for online implementation.

In recent years, data-driven fault diagnosis methods are drawing more and more attention, including statistical methods [23], signal-based [10], machine learning [24], etc. Xia *et al.* [25] applied the correlation coefficient between cell voltages to identify the short circuit fault. On top of this, by incorporating also the measured battery cell data from the laboratory setup, Kang *et al.* [26] implemented an improved correlation coefficient method to diagnose several types of faults, such as short circuits, sensor faults, and connection faults. To detect connection failure, Yao built general regression neural networks and grid search support vector machine [27], [28]. Yao *et al.* [29] proposed local outlier factor (LOF) and Grubbs outlier filter to identify the battery with abnormal internal resistance. The aforementioned methods however have not been verified by operational data from vehicles in daily use. The fault diagnosis approaches by using cloud data platform, which is in real-time gathering the operational data from vehicles in daily use, are thus studied. Based on statistics, Zhao *et al.* [30] proposed a $3\sigma$ multilevel screening strategy to find the faulty cell. Based on this, a combined method of k-mean clustering, z-score, and $3\sigma$ screening approach is used to locate abnormal cells by Xue *et al.* [31]. Li and Wang [32] utilized the interclass correlation coefficient method to detect voltage faults. Hong *et al.* [33] and Wang *et al.* [34] established a fault diagnosis mechanism for voltage and temperature faults using the modified Shannon entropy approach, respectively. In order to overcome the difficulty of selecting the calculation window, Hong *et al.* [35] proposed the modified multiscale sample entropy to predict the thermal runaway of LIB. Li *et al.* [36] combine long short-term memory neural network and ECM to diagnose overvoltage and undervoltage. These methods focus more on the diagnosis of single-parameter faults, such as voltage or temperature, rather than multiple parameters coupled with each other in actual battery operation. Thus, they also overlook the information regarding fault reflected by other parameters. In this article, a multiparameter-based battery early fault diagnosis approach is proposed. The operational data of EV batteries are obtained from the National Big Data Alliance of New Energy Vehicles (NDANEV) in Beijing, including the voltages and temperatures of the battery cells. As shown in Fig. 1, the fault diagnosis approach has two layers. The first layer is for severe fault detection, in which the battery cell voltage and temperature are compared with predefined upper and lower limits (for temperature only upper limit). Once the voltage or temperature is beyond the limit, alarm is tripped. Since severe faults already happen in this scenario, the battery management system will stop the operation of the EV ASAP. As the first layer is nothing new, it will not be further elaborated. The focus of this article and the major challenge are the second layer, where severe faults

have not happened yet, but some battery cells already show abnormities. If they are detected, damage of the battery will be avoided. Moreover, it will also buy enough time to remind the driver of the EV to go for maintenance. In the end, the cost of the fault will be very much reduced.

However, identifying the abnormal voltages or temperatures is complicated as they marginally deviate from the normal value, thus false alarm often happens, especially only single parameter (voltage or temperature) is used for diagnosis. To improve the reliability of the fault diagnosis, the correlation of the voltage and temperature is considered in the second layer by combing the discrete Fréchet distance (DFD) with the LOF. More details are elaborated in Section II. The content of this article is structured as follows: Section II explains the data source and studied battery systems. In Section III, diagnostic strategies based on voltage and temperature data are described. Results and discussions are presented in Section IV. Finally, Section V concludes this article.

## II. DATA SOURCE AND ONLINE DIAGNOSIS PROCESS

### A. National Big Data Alliance of New Energy Vehicles

As depicted in Fig. 2, the NDANEV is built to provide data resources of BEV for researchers and pursue a safe operation of EVs. Using real-operating data of battery systems from NDANEV can solve the problem Of insufficient BMS calculation capabilities. According to the protocol named technical specifications of remote service and management system for EVs, vehicle status data and power battery data are collected from real running vehicles. To be more specific, power battery data include voltage and current of battery pack, cell voltage, and temperature. Vehicle status data consist of speed, mileage, and positional information.

### B. Object of Study

Two accident EVs (vehicle 1# and vehicle 3#) and one normal EV (vehicle 2#) with a data sampling frequency of 0.1 Hz are considered in this study. Some more information of the EVs is shown in Table I and II. Vehicle 1# and vehicle 2# are of the same model, whereas vehicle 3# is a different model. Vehicle 1# is used as the training detection approach. Vehicles 2# and 3# are utilized to verify it. The battery pack of vehicles 1# and 2# consists 17 modules (95 cells in total) connected in series, and has 95 voltage sensors and 34 temperature sensors, as indicated in Fig. 3 and Table I. The LIB system of the vehicle 3# is a bit different, and it has 84 cells connected in series and 28 temperature sensors. In this article, the algorithm code was compiled based on Python 3.7.5 and implemented on a PC (processor AMD Ryzen 7 4800H with Radeon Graphics CPU @2.9 GHz).

Taking vehicle 1# as an example, the voltage and temperature of the accident vehicle 1# near the thermal runaway are shown in Fig. 4. Note that, since the sampling rate is 0.1 Hz, there is one frame per 10 s. However, as a rule of the *NDANEV*, the data sampling only happens when the vehicle is running. Thus, the voltage and temperature shown in Fig. 4 are the same as the voltage and temperature in time series but with the data during parking (except charging) of the vehicle cut out. Nonetheless, it
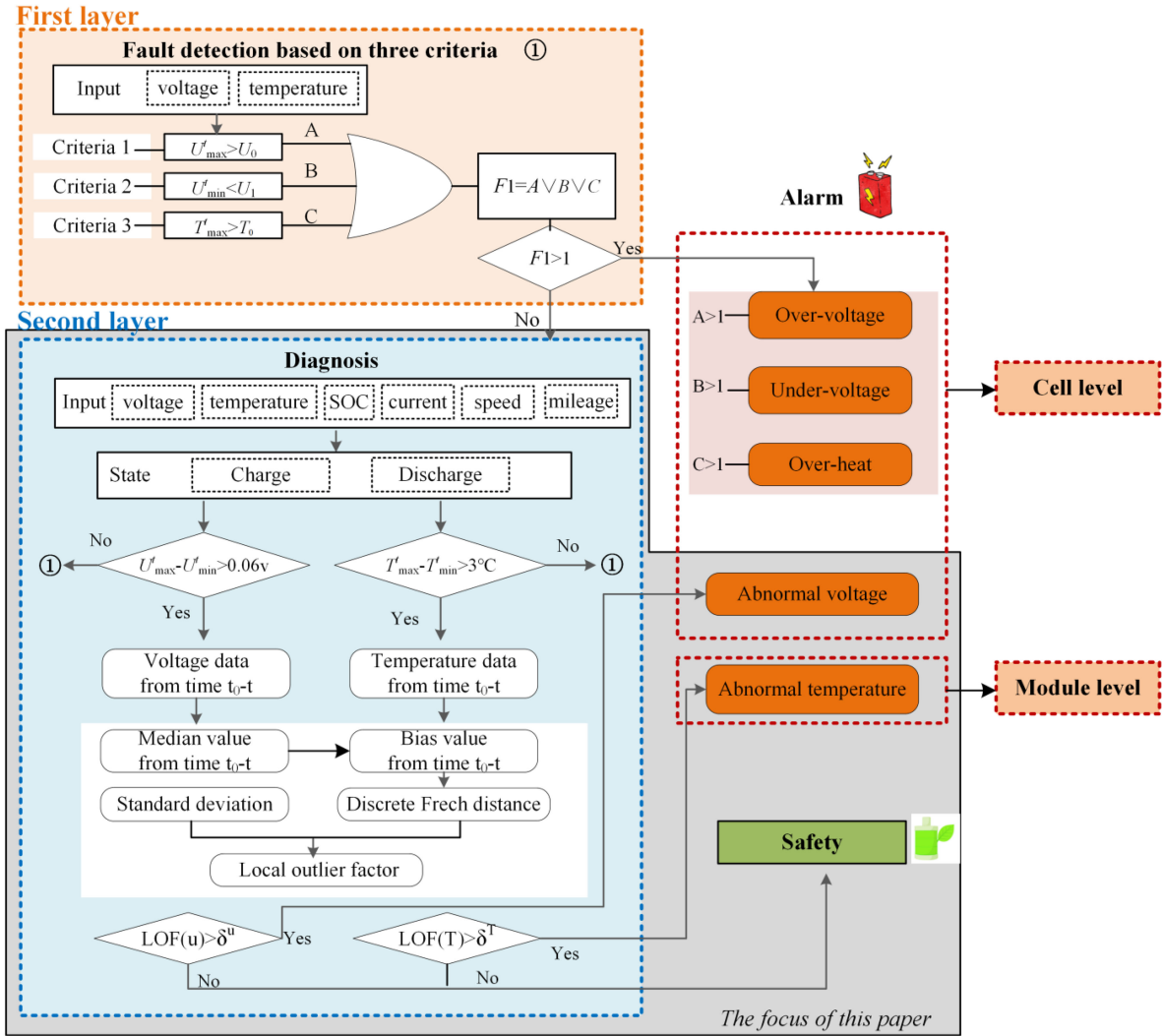
Fig. 1.    Schematic of the proposed method.
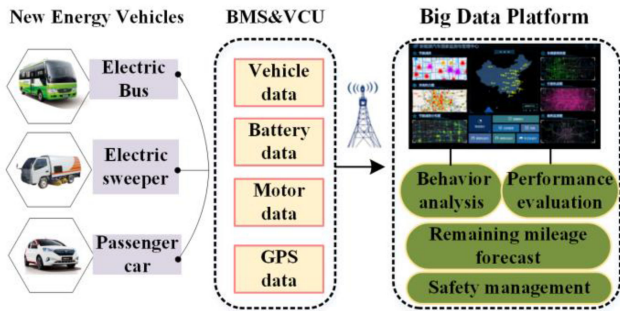


Fig. 2.    Data collection schematic.

would not affect the fault detection, since faults rarely happen during parking while not charging. As seen in the figure, the battery pack encountered a thermal runaway during charging process. The No.17 cell of module 4# is the source that triggered the thermal runaway. An overdischarge warning (under 2.8 V) occurred in the No.17 cell whose voltage was 0.03 V at the

5241-the frame (2019-07-05 14:27:32) in Fig. 4(a). But the maximum temperature of this LIB system was 50 °C (below 55 °C), as depicted in Fig. 4(b), which proves the ineffectiveness of the fault detection approaches that rely on one parameter. After the accident, the battery system structure has been severely damaged, as seen in Fig. 4(c).

### C. Online Diagnosis Process

The first layer algorithm is applied every time point. The voltage and temperature of one battery primarily measured at one time point is compared with two thresholds, upper and lower. If the voltage/temperature at that time point is beyond the upper threshold, the battery is considered as overvoltage/overheat; if the voltage is below the lower threshold, the battery is considered as undervoltage. It is usually used for detection of extreme overvoltage, undervoltage/overheat. If no alarm occurred during the first level detection process, the second layer is activated. First, abnormal behavior has not been detected by the first layer. When maximum voltage range or temperature range in charging/

TABLE I
DETAILED INFORMATION OF ALL VEHICLES

| The number of vehicle | Cathode Chemistry | The number of voltage sensors [a] | The number of temperature sensors | The resolution of the voltage sensor (V) | The resolution of the temperature sensor °C | Vehicle state |
|---|---|---|---|---|---|---|
| Vehicle 1# | Li(NiMnCo)O2 | 95 | 34 | 0.01 | 1 | Thermal runaway |
| Vehicle 2# | Li(NiMnCo)O2 | 95 | 34 | 0.01 | 1 | Normal |
| Vehicle 3# | Li(NiMnCo)O2 | 84 | 28 | 0.02 | 1 | Thermal runaway |

TABLE II
BATTERY PACK PROPERTIES OF NO. 1 VEHICLE

| Items | Parameter |
|---|---|
| Nominal capacity (Ah) | 126 |
| Total voltage range (V) | 266–408.5 |
| Total energy (kWh) | 43.5 |
| Battery pack size (length×width×height, mm) | 2040×1096×291 |
| Cell nominal voltage (V) | 3.65 |
| Operating temperature (°C) | -30−55 |



Fig. 3. Structure of the battery pack. (a) Battery cell. (b) One battery module. (c) Connection structure of the modules.

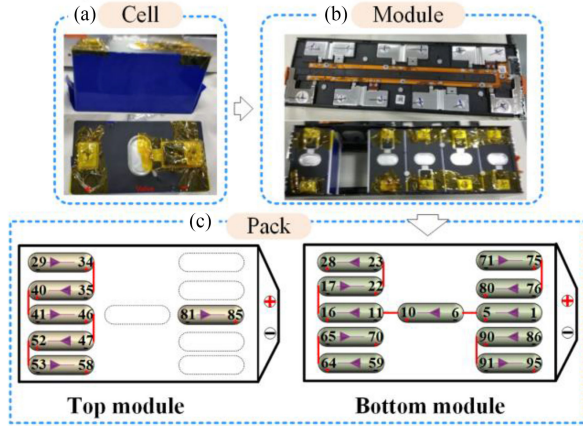discharging cycles is more than 0.06 V or 3 °C at time $t_e$, all voltage/temperature data from the start of cycle (charging/discharging) to time $t_e$ will be used as input for the DFD and LOF. If the LOF values of voltage/temperature are less than the given threshold, no alarm will be triggered. On the contrary, the faulty cell/module can be diagnosed and effective measures should be taken to prevent further damage from happening.

## III. EARLY STAGE ABNORMAL OPERATION DETECTION OF BATTERY

Thermal runaway of battery system is always triggered by various kinds of abuse, including mechanical abuse, electrical abuse, and thermal abuse. The mechanical abuse condition includes crash, penetration, and bend. Overheating and fire exposure belong to the thermal abuse. Overcharge, overdischarge, and short circuit are electrical abuse. In the case of electrical abuse or faults, multiple failure modes can be coupled to result in complicated mechanism. Without suitable diagnostics and fault

handling, a minor fault of battery system could develop into thermal runaway [37]. Voltage and temperature are the two factors controlling the battery reactions. To prevent thermal runaway and other severe damage of the battery pack, the abnormal cell or module should be identified before the severe damage happens. In order to achieve this, the LOF of the battery cell voltages can be calculated at each time being. The cell that has the largest LOF is then considered as the abnormal cell. However, this approach does not consider the correlation of the voltages of one cell at different time beings. In fact, in the early stage of abnormal operation, the cell voltage/temperature will gradually deviate from the healthy cells. This gradual change in the cell voltage/temperature can be very useful for detecting the abnormal operation. Therefore, the DFD between the voltage/temperature curve of a cell and the median voltage/temperature curve is calculated first, where Savitzky–Golay filter is employed to remove the noise from the voltage and temperature data beforehand. Then LOF is calculated based on DFD and standard deviation of each cell voltage and module temperature in time series. The LOF value is then used to judge if the cell is in early stage of a fault or abnormal operation. More details are as follows.

### A. Discrete Fréchet Distance

Owing to the discrete acquisition data, the DFD with origin from Fréchet distance [38] needed to be used [39]. The DFD was defined by Eiter and Mannila in 1994 and was further expanded by Mosig and Clausen in 2005 [40], [41].

As mentioned earlier, the cell voltage and module temperature are monitored and recorded with 0.1-Hz sampling frequency. The time series voltage and temperature are chopped into pieces for analysis. The starting and ending point of each piece is actually the starting of a charging event and the next charging event, respectively. As such, an abnormal operation detection of the battery is executed after each charging/discharging cycle, and no data point is missing in the analysis. Taking the piece between $t_0 - t_e$ as an example to elaborate the detection approach, the cell voltages are expressed as

$$U^{t_0-t_e} = \begin{bmatrix} U_1^{t_0-t_e} \\ \vdots \\ U_i^{t_0-t_e} \\ \vdots \\ U_n^{t_0-t_e} \end{bmatrix} \qquad (1)$$

$$U_i^{t_0-t_e} = [u_i^{t_0}, u_i^{t_1}, \cdots, u_i^{t_e}] \qquad (2)$$
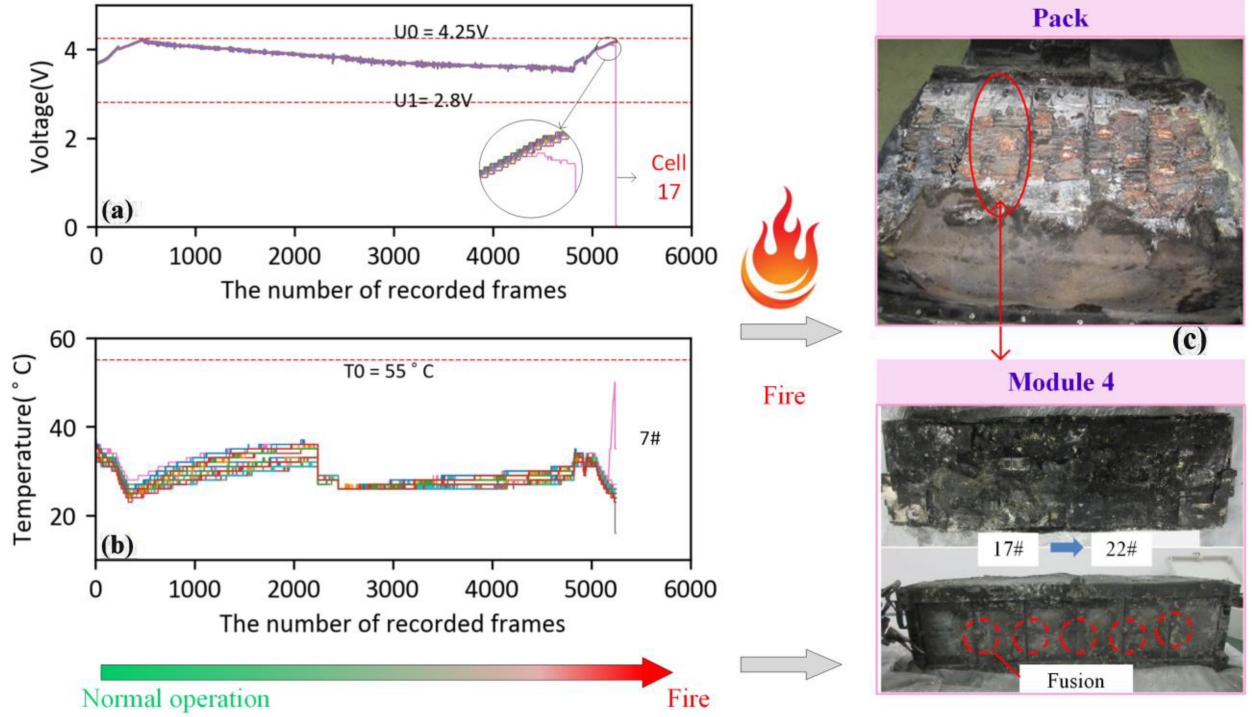
Fig. 4.    Vehicle 1#: (a) voltages (cell 1 to 95); (b) temperature (sensor 1 to 34); (c) battery pack and module 4 after fire.
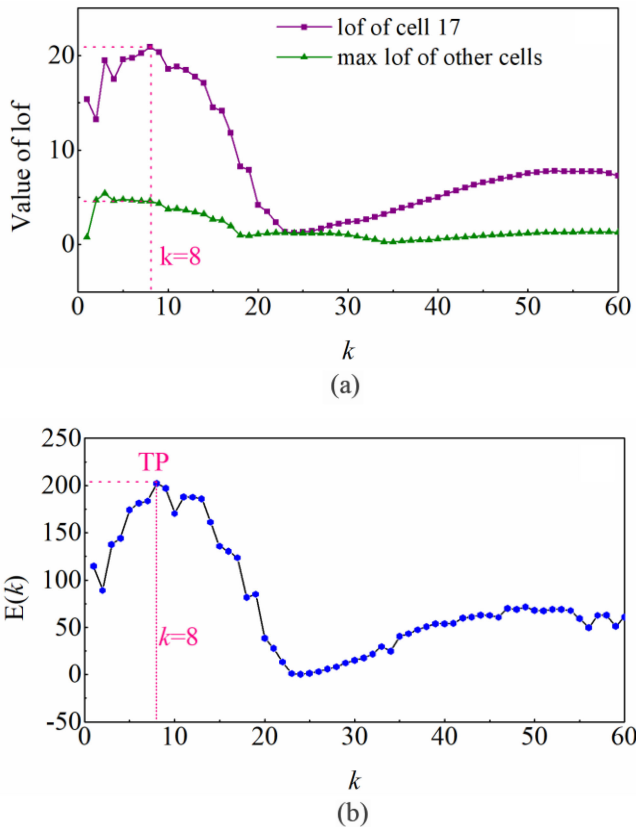


Fig. 5.    Variation of values at different $k$. (a) LOF value of cell 17 and remaining cells' maximums. (b) Cost function.

where $n$ is the total number of cells, so $0 \leq i \leq n$. The median voltage vector between time $t_0$ and time $t_e$ is indicated as

$$\overline{U}^{t_0-t_e} = \left[\overline{u}^{t_0}, \overline{u}^{t_1}, \cdots, \overline{u}^{t_e}\right]$$
$$= \left[\frac{1}{n}\sum_{i=1}^{n} u_i^{t_0}, \frac{1}{n}\sum_{i=1}^{n} u_i^{t_1}, \cdots, \frac{1}{n}\sum_{i=1}^{n} u_i^{t_e}\right]. \quad (3)$$

In order to eliminate the influence of cell voltage bias error on the DFD, the voltages of each cell are deducted by a bias voltage, so that the initial voltages of all the cells will equal to the median value of the initial voltages of all cells. The bias voltage deduction for the $i$th cell is defined as

$$u_{bud,i} = u_i^{t_0} - \overline{u}^{t_0}. \quad (4)$$

The calibrated cell voltages then become

$$U_{cali}^{t_0-t_e} = \begin{bmatrix} U_{cali,1}^{t_0-t_e} \\ \vdots \\ U_{cali,i}^{t_0-t_e} \\ \vdots \\ U_{cali,n}^{t_0-t_e} \end{bmatrix} = \begin{bmatrix} U_1^{t_0-t_e} - u_{bud,1} \\ \vdots \\ U_i^{t_0-t_e} - u_{bud,i} \\ \vdots \\ U_n^{t_0-t_e} - u_{bud,n} \end{bmatrix} \quad (5)$$

where $U_{cali,i}^{t_0-t_e} = [u_{cali,i}^{t_0}, u_{cali,i}^{t_1}, \cdots, u_{cali,i}^{t_e}]$.

The DFD between the calibrated voltages of the $i$- and the median voltage vector $\overline{U}^{t_0-t_e}$ is then calculated and shown as follows:

$$DFD_i = \inf_{\alpha,\beta \in [t_0, t_e]} \max |u_{cali,i}^{\alpha} - \overline{u}^{\beta}|. \quad (6)$$

Fig. 6. LOF value for vehicle 1# at different cycle. (a) Charging voltage data. (b) Discharging voltage data. (c) Charging temperature data. (d) Discharging temperature data.

## B. Local Outlier Factor

In the end, each cell should be represented by two parameters as input of LOF calculation. Beside DFD, the standard deviation of each cell's voltage between $t_0 - t_e$ is used as the other parameter, and it is calculated as

$$\sigma_i = \sqrt{\frac{1}{n} \sum_{t=t_0}^{t_e} (u_i^t - \mu)^2} \qquad (7)$$

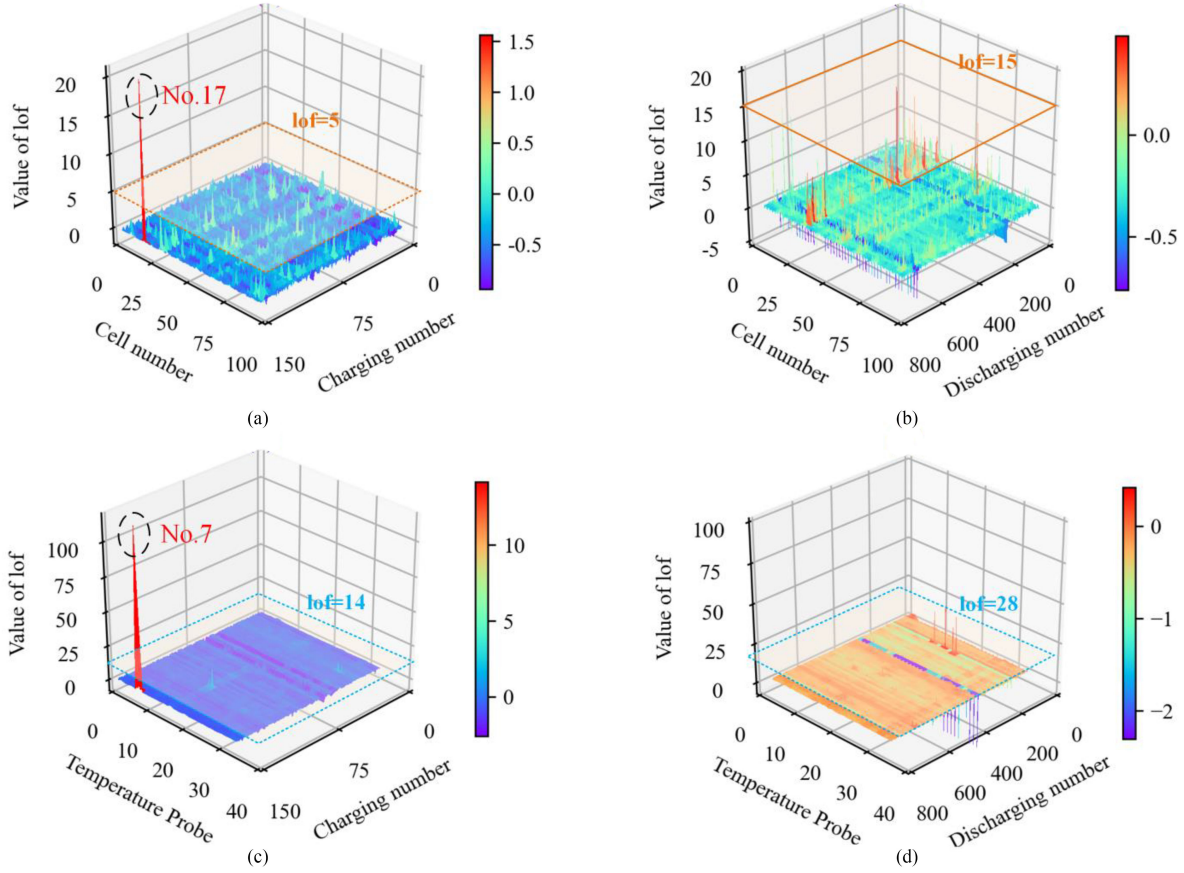where $\mu = \frac{1}{n}\sum_{t=t_0}^{t_e} u_i^t$. So, $i$th cell now is expressed as $(DFD_i, \sigma_i)$. LOF of $i$th cell's voltage is depicted as $LOF_k(u_i)$. More details of the LOF calculation can be found in [27].

Similarly, LOF of $i$th module's temperature can be calculated, and it is illustrated as $LOF_k(T_i)$. Note that since the temperature sensor resolution is 1 °C, the LOF values of temperatures are reset to 0 if the temperature variation is less than 3 °C. A cell or module will be labeled as abnormal, if $LOF_k(u_i)$ or $LOF_k(T_i)$ is beyond the threshold. The flag $F_u^i$ or $F_T^i$ is then set to 1, as follows:

$$F_u^i = \begin{cases} 1, & \text{if } LOF_k(u_i) \geq \delta^u \\ 0, & \text{if } LOF_k(u_i) < \delta^u \end{cases} \qquad (8)$$

$$F_T^i = \begin{cases} 1, & \text{if } LOF_k(T_i) \geq \delta^T \\ 0, & \text{if } LOF_k(T_i) < \delta^T \end{cases} \qquad (9)$$

where $\delta^u$ and $\delta^T$ are the LOF threshold of voltage and temperature, respectively. Apparently, these two thresholds affect the sensitivity of the detection. Moreover, the LOF value relates to $k$. The selection of $k$, $\delta^u$, and $\delta^T$ is then elaborated as follows.

## C. Selection of k and Threshold δ

The goal of setting $k$, $\delta^u$, and $\delta^T$ is that abnormal cells and modules can be screened out as outliers and require a low calculation time. The evaluation function is expressed as

$$E(k) = \frac{\Delta LOF(k)}{t_{\cos t}(k)} \qquad (10)$$

where $t_{\cos t}(k)$ is calculation time, and $\Delta LOF(k)$ is the maximum LOF variation at $k$. Thus, larger value of $E(k)$ indicates that the abnormal cells are more differentiated from the normal cells, and thereby the detection becomes easier. The idea here is to use the data from an accident vehicle to extract a proper value of $k$, $\delta^u$, and $\delta^T$. Fig. 5(a) shows the voltage LOF value of No.17 cell and other cells at different $k$ of the accident vehicle 1#, where the time piece includes the time being of the fault shown in Fig. 4. Fig. 5(b) shows $t_{\cos t,(k)}$ of No.17 cell as a function of $k$. As seen, E(k) has the maximum value at $k = 8$. Therefore, $k$ in this article is taken as 8.

The charging data are relatively stable, whereas the data during the discharging process are more complex. In this article,
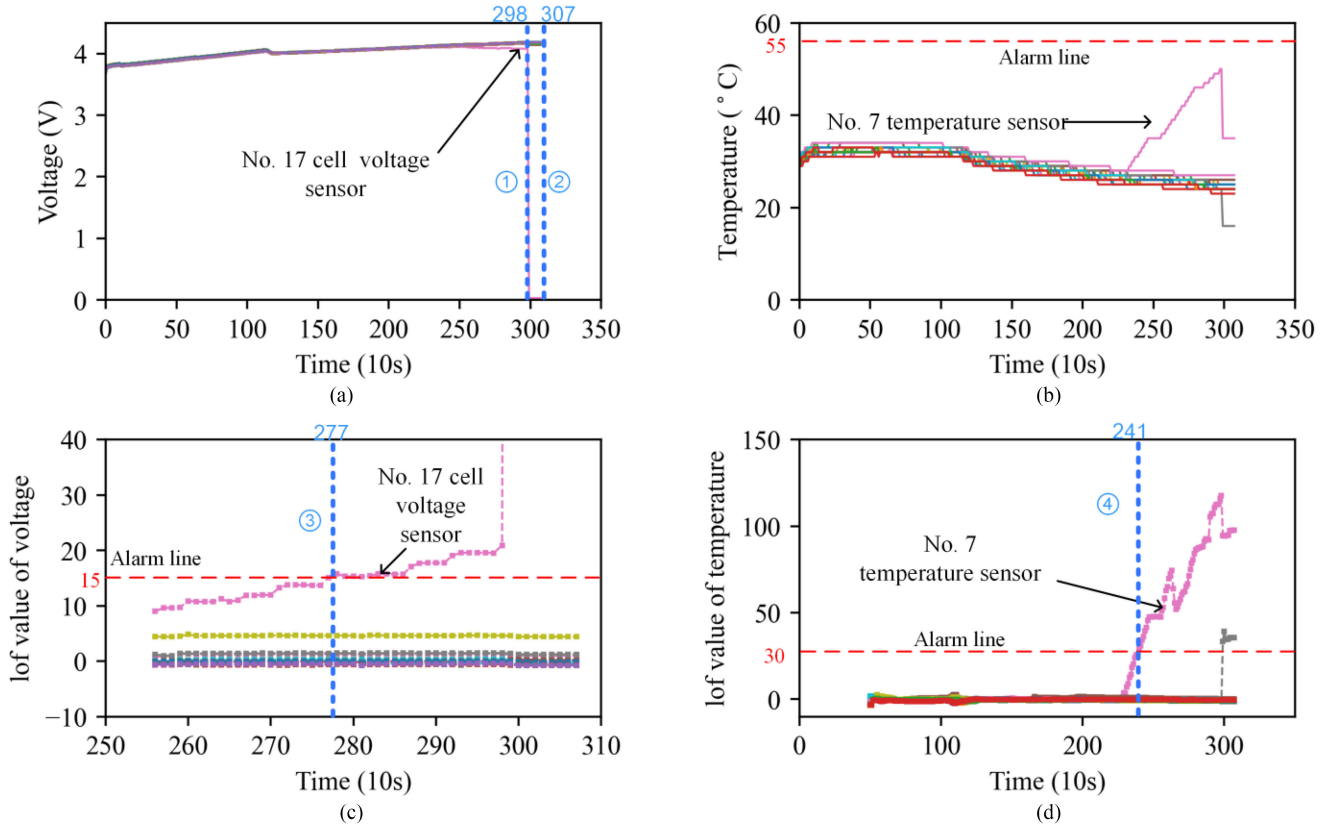
Fig. 7. Last charging process of vehicle 1#. (a) Voltage curve (cell 1 to 95). (b) LOF value of voltage. (c) Temperature curve (sensor 1 to 34). (d) LOF value of temperature.

the charging and discharging conditions are studied separately. Charging and discharging processes with less than 50 frames are excluded from the preprocessing. As shown in Fig. 4, thermal runaway happens to the cell No.17. To prevent thermal runaway, the fault needs to be detected in its early stage. The data of vehicle 1# 2 min before the thermal runaway (14:27:22, 2019-07-05) is then used to train the detection algorithm, where both the voltage and temperature are within the safe range, and the layer 1 is not tripped yet. The LOF of the voltage and temperature are then calculated and they are shown in Fig. 6. The voltage LOF of No.17 cell is greater than 20 during the 148th charging cycle. Before it, the voltage LOF during charging is not higher than 5. The voltage LOF does not exceed 15 during discharging cycle. Fig. 6(c) shows that the temperature LOF of No.7 sensor (which is in the same module as the No. 17 cell) within module 4# is much larger than 30, whereas others are less than 14 during charging. Fig. 6(d) shows that the temperature LOF of all sensors during discharging is less than 28. Therefore, the voltage diagnostic factor $\delta^V$ is set to 15, and the temperature diagnostic factor $\delta^T$ is set to 30.

### D. Verifying the Approach

Taking the voltage and temperature data of 148th charging process for vehicle 1# as an example, all data are shown in Fig. 7 and the vehicle caught fire at the 307th frame [14:28:52, 2019-07-05, event ② in Fig. 7(a)]. As shown in Fig. 7(a), the cell No. 17 was detected undervoltage (first layer) at event ①, which

is 8 frames earlier than thermal runaway (298th frame, 14:27:22, 2019-07-05). While by using LOF voltage-based detection (second layer), the faulty cell No. 17 can be detected at frame 277, which is 21 frames earlier than the first layer detection. It might be noted that in Fig. 7(c), the LOF is null before frame 256. That is because, LOF calculation of voltage is only activated when the voltage difference is larger than 0.06 V, which is only the case science frame 256.

Fig. 7(b) and (d) shows the comparison between the first and second layer detection in terms of temperature. Specifically, the first layer detection is not tripped because the temperature is always below the threshold [typically 55 °C for EV batteries, red line in Fig 7(b)]. However, with the second layer detection relying on LOF of temperature, the faulty cell is detected at frame 241 (event ④), which is 66 frames (660 s) earlier than thermal runaway. Note that, each temperature sensor is shared two or three cells, and the temperature sensor No. 7 is the one most close to the cell No. 17.

### IV. RESULT AND DISCUSSION

### A. Validity Analysis

Vehicle 2# (same model as vehicle 1#, but with no fault happened) is selected to verify if the proposed detection approach can have faulty trip. Fig. 8(a) and (b) showed the voltage and temperature of the cells during a whole charging and discharging cycle. In the two processes, the maximum voltage and minimum voltage in the battery system are, respectively, 4.24 and 3.63 V,
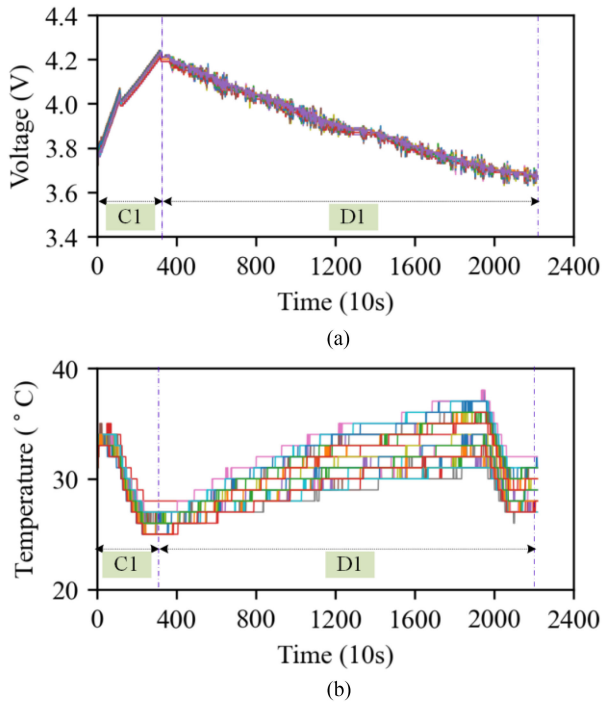
Fig. 8. Parametric curves of vehicle 2#. (a) Voltages values (cell 1 to 95). (b) Temperature values (sensor 1 to 34).

and the maximum temperature is 38 °C. These values are in the normal range, so no alarm can be tripped in the first layer.

The voltage and temperature LOF are shown in Fig. 9. As seen, the voltage LOF is always less than 3, no matter in charging (C1) or discharging (D1), and it is far lower than the threshold 15. The temperature LOF is always less than 1. Therefore, it can be concluded that battery that in normal condition will not trip the proposed detection approach in the second layer.

### B. Universality Analysis

Vehicle 3#, which is a different model from vehicle 1# but also encountered thermal runaway, is selected to verify if the proposed detection approach is effective in a generic way. The battery system of vehicle 3# caught fire during discharging on March 14, 2019. The ignition source was No. 47 cell of module 4# according to the incident investigation. From 20:48:19, on March 13, 2019, to 12:22:59, on March 14, 2019, voltage data of the last discharging cycle of vehicle 3# were shown in Fig. 10(a). Before the 928th frame (09:59:03, March 14, 2019), voltage values of all cells were within normal range. After the 928th frame, voltage data of all cells, including No. 47, were 1.44 V, which can trigger an undervoltage alarm on the first layer because the voltage is lower than the cutoff voltage in ① of Fig. 10(a). At the 1064th frame (12:22:59, March 14, 2019), the vehicle 3# has experienced thermal runaway and data cannot be obtained in ② of Fig. 10(a). The original temperature curve during the 3rd discharging process was shown in Fig. 10(c). At the 855th frame (08:25:34, March 14, 2019), the temperature of No.13 and No.14 sensors from module 4# reached 60 °C and an overheat alarm of the first layer was triggered in ④ of Fig. 10(c).
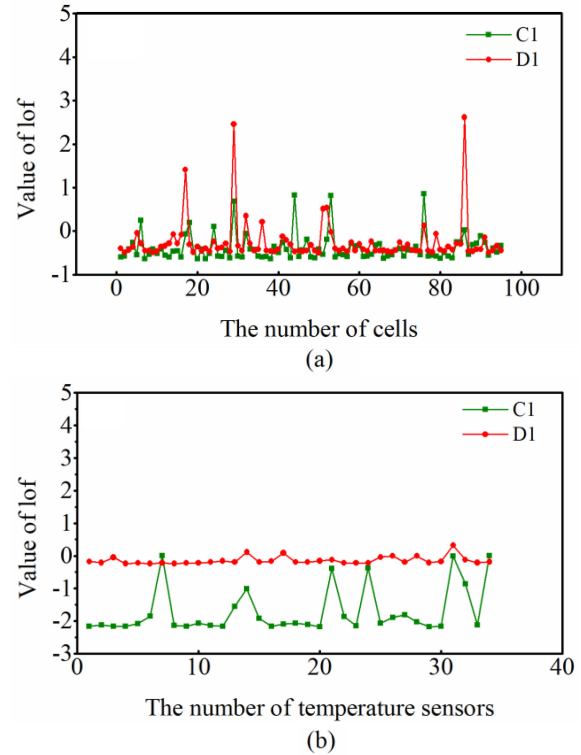


Fig. 9. LOF value of (a) voltage data and (b) temperature data of vehicle 2# during charging and discharging processes.

The voltage and temperature LOF values of vehicle 3# for fault detection were shown in Fig. 10(b) and (d). At the 829th frame, (08:25:08, March 14, 2019) in ③ of Fig. 10(b), the No. 47 cell is detected by the second layer, which is 1000 s earlier than the over low voltage detection in the first layer. As shown in ⑤ of Fig. 10(d), the temperature of No.13 sensor can be detected at 145th frame (04:57:45, March 14, 2019). This abnormal temperature alarm lasted for two frames. Then, at 867th frame (08:25:46, March 14, 2019), abnormal temperature alarm is activated again in ⑥ of Fig. 10(d). Because No. 47 voltage sensor is in the same module 4# as No. 13 temperature sensor, it proved that the fault of module 4# can be accurately detected by the proposed approach and it has considerable leading time than the thermal runaway.

### C. Different Methods Comparison

In order to evaluate the performance of proposed method, other three methods, including correlation coefficient [25], $3\sigma$ multilevel screening strategy [30], and Shannon entropy [34], were used to compare. The data of the first 930 frames of the last process of vehicle 3# were selected. The length of each window in [25] and [34] was 30 and 100 sample points.

The correlation coefficient of the selected voltage data was shown in Fig. 11(a). The calculation time of each window of this method is 0.03 s. However, false alarms are given frequently when the correlation coefficient value was below 0.9. It was noted that this method had the short calculation time and low accuracy.
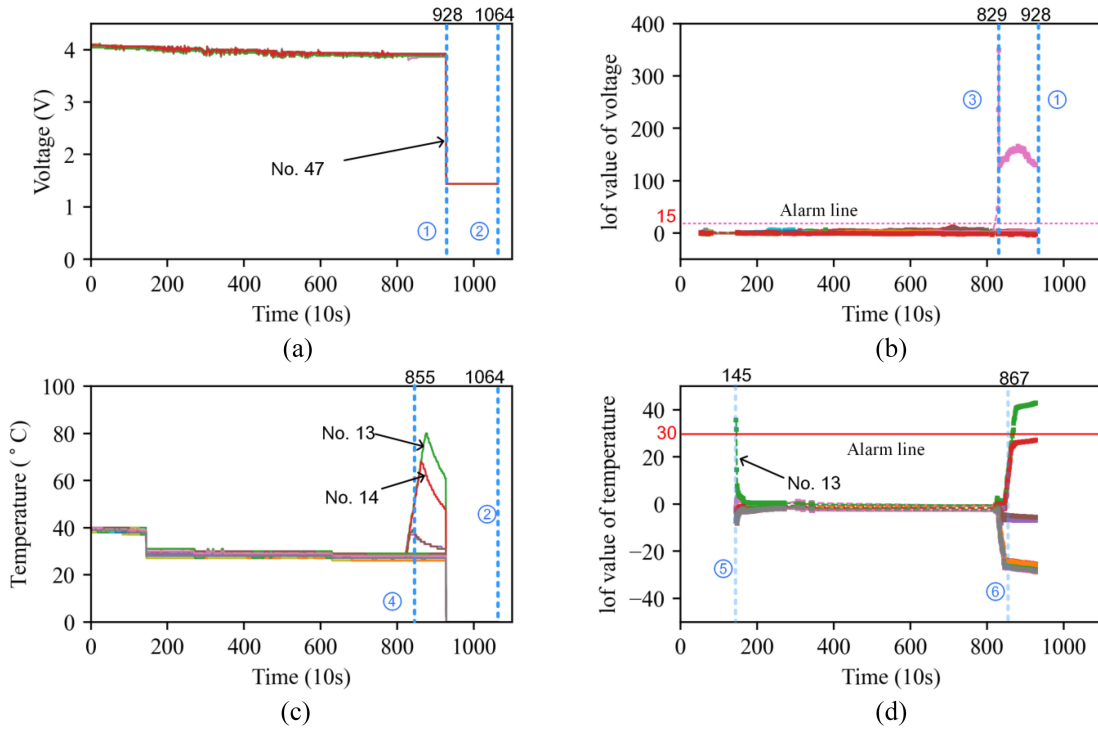
Fig. 10. Last discharging process of vehicle 3#. (a) Voltage curve (cell 1 to 84). (b) LOF value of voltage. (c) Temperature curve (sensor 1 to 28). (d) LOF value of temperature.
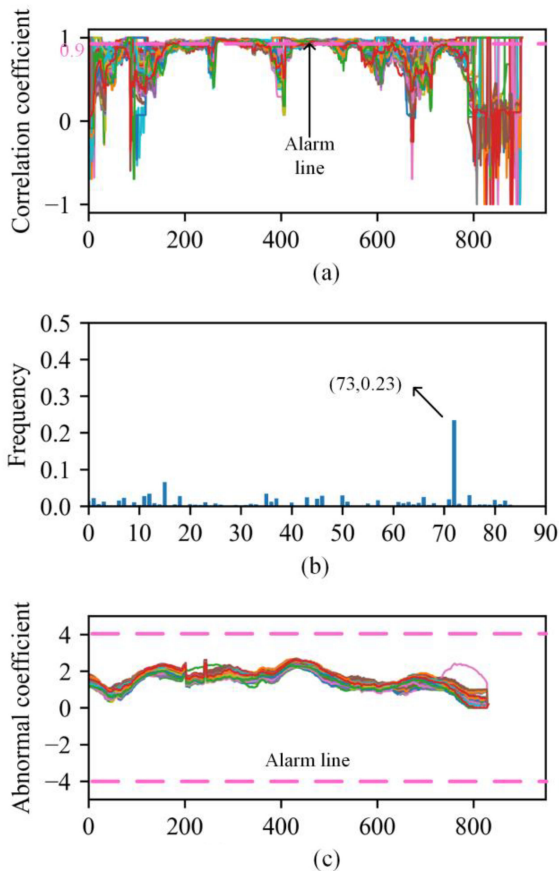


Fig. 11. (a) Correlation coefficient, (b) $3\sigma$ multilevel screening strategy, and (c) abnormal coefficient of Shannon entropy of vehicle 3#.

Fig. 11(b) showed the cell faulty frequency by $3\sigma$ multilevel screening strategy. The calculation time of all data is 0.86 s. Cell 73 was considered the most likely faulty cell in battery pack because its frequency was 0.23. But the faulty source was No. 47 cell, which indicated that this method cannot give an accurate fault cell of vehicle 3#.

Fig. 11(c) showed the abnormal coefficient based on Shannon entropy. The calculation time of each window was 0.1 s. The threshold of the abnormal coefficient was $\pm 4$ in [34]. The abnormal coefficient of all cells was [–44], so no abnormal cell will be given. This indicated that this method did not accurately identify abnormal cell.

The calculation time of the proposed method is 0.09 s, and can meet the practical application. And faulty battery can be diagnosed by the proposed approach

## V. CONCLUSION

An online data-driven EV battery fault diagnosis scheme using both the cell voltage and temperature is proposed. Both DFD method and LOF are used to quantify the correlation between the different battery cells' voltage and temperature. In this way, the faulty cell can be detected considerably earlier than thermal runaway happens, which gives sufficient time to start maintenance or protection scheme, to very much reduce the cost of fault. The data from vehicles with thermal runaway are used to train the detection approach. The trained detection approach is then applied to vehicles with or without thermal runaway to check its sensitivity and it works perfectly. All the data are obtained from NDANEV, which is a national cloud

platform that gathering operational data from vehicles in daily use with 0.1-Hz sampling rate, rather than laboratory data that obtained in an artificial scenario. Therefore, the proposed early fault detection approach is very practical and promising for real application.

## REFERENCES

[1] D. Chen, J. Deng, B. Zhang, Z. Wang, and S. Wang, "A dual-transformer based hybrid semidual active bridge converter for wide voltage range applications utilizing simple segmented control," *IEEE Trans. Power Electron.*, vol. 37, no. 2, pp. 1435–1446, Feb. 2022.

[2] Z. S. Gan, Z. Zhang, S. Xu, P. Liu, and Z. Qin, "Data-driven fault diagnosis of lithium-ion battery overdischarge in electric vehicles," *IEEE Trans. Power Electron.*, vol. 37, no. 4, pp. 4575–4588, Apr. 2022.

[3] Q. Wang, Z. Wang, L. Zhang, P. Liu, and Z. Zhang, "A novel consistency evaluation method for series-connected battery systems based on real-world operation data," *IEEE Trans. Transp. Electrific.*, vol. 7, no. 2, pp. 437–451, Jun. 2021.

[4] Z. Sun *et al.*, "Detection of voltage fault in the battery system of electric vehicles using statistical analysis," *Appl. Energy*, vol. 307, 2022, Art. no. 118172.

[5] X. Hu, H. Jiang, F. Feng, and B. Liu, "An enhanced multi-state estimation hierarchy for advanced lithium-ion battery management," *Appl. Energy*, vol. 257, 2020, Art. no. 114019.

[6] R. Xiong, W. Sun, Q. Yu, and F. Sun, "Research progress, challenges and prospects of fault diagnosis on battery system of electric vehicles," *Appl. Energy*, vol. 279, 2020, Art. no. 115855.

[7] Z. Liao, S. Zhang, K. Li, G. Zhang, and T. G. Habetler, "A survey of methods for monitoring and detecting thermal runaway of lithium-ion batteries," *J. Power Sources*, vol. 436, 2019, Art. no. 226879.

[8] D. P. Finegan *et al.*, "The application of data-driven methods and physics-based learning for improving battery safety," *Joule*, vol. 5, no. 2, pp. 316–329, 2021.

[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] X. Hu, K. Zhang, K. Liu, X. Lin, S. Dey, and S. Onori, "Advanced fault diagnosis for lithium-ion battery systems: A review of fault mechanisms, fault features, and diagnosis procedures," *IEEE Ind. Electron. Mag.*, vol. 14, no. 3, pp. 65–91, Sep. 2020.

[11] M.-K. Tran and M. Fowler, "A review of lithium-ion battery fault diagnostic algorithms: Current progress and future challenges," *Algorithms*, vol. 13, no. 3, 2020, Art. no. 62.

[12] L. Liu *et al.*, "Comparative study on substitute triggering approaches for internal short circuit in lithium-ion batteries," *Appl. Energy*, vol. 259, 2019, Art. no. 114143.

[13] B. Liu *et al.*, "Safety issues caused by internal short circuits in lithium-ion batteries," *J. Mater. Chem. A*, vol. 6, no. 43, pp. 21475–21484, 2018.

[14] A. Pesaran, S. Santhanagopalan, and G. H. Kim, "Addressing the impact of temperature extremes on large format Li-ion batteries for vehicle applications (Presentation)," Office of Entific & Technical Information Technical Reports, Golden, CO (United States), No. NREL/PR-5400-58145, 2013.

[15] B. Xia, Z. Chen, and M. Chris, "External short circuit fault diagnosis for lithium-ion batteries," in *Proc. IEEE Transp. Electrific. Conf. Expo*, Dearborn, MI, USA, 2014, pp. 1–7.

[16] W. Chen, W.-T. Chen, M. Saif, and M.-F. Li, "Simultaneous fault isolation and estimation of lithium-ion batteries via synthesized design of luenberger and learning observers," *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 1, pp. 290–298, Jan. 2014.

[17] 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.

[18] H. E. P. Dey and S. J. Moura, "Model-based battery thermal fault diagnostics: Algorithms, analysis, and experiments," *IEEE Trans. Control Syst. Technol.*, vol. 27, no. 2, pp. 576–587, Mar. 2019.

[19] S. Dey *et al.*, "Model-based real-time thermal fault diagnosis of lithium-ion batteries," *Control Eng. Pract.*, vol. 56, pp. 37–48, 2016.

[20] S. Dey, H. E. Perez, and S. J. Moura, "Model-based battery thermal fault diagnostics: Algorithms, analysis, and experiments," *IEEE Trans. Control Syst. Technol.*, vol. 27, no. 2, pp. 576–587, Mar. 2019.

[21] X. Feng, X. He, L. Lu, and M. Ouyang, "Analysis on the fault features for internal short circuit detection using an electrochemical-thermal coupled model," *J. Electrochem. Soc.*, vol. 165, no. 2, pp. A155–A167, 2018.

[22] 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, 2018.

[23] M. Cheliotis, I. Lazakis, and G. Theotokatos, "Machine learning and data-driven fault detection for ship systems operations," *Ocean Eng.*, vol. 216, 2020, Art. no. 107968.

[24] R. Xiong, S. Ma, H. Li, F. Sun, and J. Li, "Toward a safer battery management system: A critical review on diagnosis and prognosis of battery short circuit," *iScience*, vol. 23, no. 4, Apr. 2020, Art. no. 101010.

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

[26] Y. Kang, B. Duan, Z. Zhou, Y. Shang, and C. Zhang, "Online multi-fault detection and diagnosis for battery packs in electric vehicles," *Appl. Energy*, vol. 259, 2020, Art. no. 114170.

[27] K. X. Chen, J. Wei, and G. Dong, "Voltage fault detection for lithium-ion battery pack using local outlier factor," *Measurement*, vol. 146, pp. 544–556, 2019.

[28] L. Yao, Z. Fang, Y. Xiao, J. Hou, and Z. Fu, "An intelligent fault diagnosis method for lithium battery systems based on grid search support vector machine," *Energy*, vol. 214, 2021, Art. no. 118866.

[29] L. Yao, Y. Xiao, X. Gong, J. Hou, and X. Chen, "A novel intelligent method for fault diagnosis of electric vehicle battery system based on wavelet neural network," *J. Power Sources*, vol. 453, 2020, Art. no. 227870.

[30] 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.

[31] Q. Xue, G. Li, Y. Zhang, S. Shen, Z. Chen, and Y. Liu, "Fault diagnosis and abnormality detection of lithium-ion battery packs based on statistical distribution," *J. Power Sources*, vol. 482, 2021, Art. no. 228964.

[32] X. Li and Z. Wang, "A novel fault diagnosis method for lithium-ion battery packs of electric vehicles," *Measurement*, vol. 116, pp. 402–411, 2018.

[33] J. Hong, Z. Wang, and P. Liu, "Big-data-based thermal runaway prognosis of battery systems for electric vehicles," *Energies*, vol. 10, no. 7, 2017, Art. no. 919.

[34] 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.

[35] J. Hong *et al.*, "Thermal runaway prognosis of battery systems using the modified multiscale entropy in real-world electric vehicles," *IEEE Trans. Transp. Electrific.*, vol. 7, no. 4, pp. 2269–2278, Dec. 2021.

[36] D. Li, Z. Zhang, P. Liu, Z. Wang, and L. Zhang, "Battery fault diagnosis for electric vehicles based on voltage abnormality by combining the long short-term memory neural network and the equivalent circuit model," *IEEE Trans. Power Electron.*, vol. 36, no. 2, pp. 1303–1315, Feb. 2021.

[37] L. Jiang *et al.*, "Data-driven fault diagnosis and thermal runaway warning for battery packs using real-world vehicle data," *Energy*, vol. 234, 2021, Art. no. 121266.

[38] M. Frechet, "Sur quelques points du calcul fonctionnel," *Rendiconti del Circolo Matematico di Palermo,* vol. 22, no. 1, pp. 1–72, 1906.

[39] T. Eiter and H. Mannila, "Computing discrete Fréchet distance," Christian Doppler Lab. Expert Syst., TU Vienna, Vienna, Austria, Tech. Rep. CD-TR 94/64, pp. 636–637, 1994.

[40] A. Mosig and M. Clausen, "Approximately matching polygonal curves with respect to the Fréchet distance," *Comput. Geometry*, vol. 30, no. 2, pp. 113–127, 2005.

[41] T. Wylie and B. Zhu, "Following a curve with the discrete Fréchet distance," *Theor. Comput. Sci.*, vol. 556, pp. 34–44, 2014.

**Zhenyu Sun** received the B.Eng. degree in vehicle engineering from Beijing Forestry University, Beijing, China, in 2015, and the M.Eng. degree in mechanical engineering in 2018 from the Beijing Institute of Technology, Beijing, China, where he is currently working toward the Ph.D. degree in mechanical engineering.

He is currently a Visiting Student with the Delft University of Technology, Delft, The Netherlands, since 2021. His research interests include electric vehicle big data analysis, fault diagnosis, and safety management of lithium-ion battery system.

**Zhenpo Wang** (Senior Member, IEEE) received the B.Eng. degree in automotive engineering from Tongji University, Shanghai, China, in 2000, and the Ph.D. degree in automotive engineering from the Beijing Institute of Technology, Beijing, China, in 2005.

He is currently a Professor with the Beijing Institute of Technology and the Director of the National Engineering Laboratory for Electric Vehicles. He has authored or coauthored four monographs and translated books and more than 80 technical papers. He also holds more than 60 patents. His current research interests include pure electric vehicle integration, packaging and energy management of battery systems, and charging station design.

Prof. Wang was 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.

**Peng Liu** received the B.S. and M.S. degrees from Chang'an University, Xi'an, China, in 2005 and 2008, respectively, and the Ph.D. degree 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. His research interests include big data analysis and safety management of electric vehicles.

**Zian Qin** (Senior Member, IEEE) received the B.Eng. degree in automation from Beihang University, Beijing, China, in 2009, the M.Eng. degree in control science and engineering from the Beijing Institute of Technology, Beijing, China, in 2012, and the Ph.D. degree from Aalborg University, Aalborg, Denmark, in 2015.

He is currently an Assistant Professor with the Delft University of Technology, Delft, The Netherlands. In 2014, he was a Visiting Scientist with Aachen University, Aachen, Germany. From 2015 to 2017, he was a Postdoctoral Research Fellow with Aalborg University. His research interests include wide bandgap devices, power electronics based grid, and Power2X.

Dr. Qin served as the Technical Program Chair of IEEE-ISIE-2020, the Technical Program Co-Chair of IEEE-COMPEL 2020, and the industrial session Co-Chair of ECCE-Asia 2020.

**Yong Chen** received the Ph.D. degree in vehicle engineering from the Beijing Institute of Technology, Beijing, China, in 2002.

He is currently a Professor with the School of Mechanical and Electrical Engineering, Beijing Information Science and Technology University, Beijing, China. He is also a Researcher with the Collaborative Innovation Center of Electric Vehicles, Beijing, China. His research interests include modeling, simulation and control of new energy source vehicles, and vehicle system dynamics.

**Yang Han** received the Ph.D. degree in statistics from the University of Southampton, Southampton, U.K., 2014.

She is currently a Senior Lecturer (Associate Professor) with the Department of Mathematics, The University of Manchester, Manchester, U.K. Her main research interests include simultaneous inference and multiple comparison procedures.

Dr. Han was the recipient of a Distinguished Achievement Award for Teacher of the Year 2019 at The University of Manchester.

**Peng Wang** received the B.S. degree in measuring and control technology and instrumentations from Zhengzhou University, Zhengzhou, China, in 2009, and the M.S. degree in engineering management from the East China University of Science and Technology, Shanghai, China, in 2018.

At present, he is responsible for the development of the battery, motor, and electric control system of electric vehicles in Zeekr Automobile (Ningbo Hangzhou Bay New Zone) Company, Ltd., and holds more than five patents.

**Pavol Bauer** (Senior Member, IEEE) received the master's degree in electrical engineering from the Technical University of Kosice, Košice, Slovakia, in 1985, and the Ph.D. degree from the Delft University of Technology, Delft, The Netherlands, in 1995.

He is currently a Full Professor with the Department of Electrical Sustainable Energy, Delft University of Technology and the Head of DC Systems, Energy Conversion and Storage group. He has authored or coauthored more than 72 journal and almost 300 conference papers in many field (with H factor Google scholar 43, Web of science 20), he is an author or coauthor of eight books, holds four international patents, and organized several tutorials at the international conferences. He has worked on many projects for industry concerning wind and wave energy, power electronic applications for power systems, such as Smarttrafo; HVdc systems, projects for smart cities, such as PV charging of electric vehicles, PV and storage integration, contactless charging; and he participated in several Leonardo da Vinci and H2020 EU projects as Project Partner (ELINA, INETELE, E-Pragmatic) and Coordinator (PEMCWebLab.com-Edipe, SustEner, Eranet DCMICRO).

Prof. Bauer was the recipient of the title "Prof." from the President of Czech Republic at the Brno University of Technology, Brno, Czechia, in 2008 and Delft University of Technology in 2016.