Data-model alliance network for the online multi-step thermal warning of energy storage system based on surface temperature diffusion

Highlights
- A data-model alliance module combining electrothermal model and LSTM is established
- An improved adaptive boosting algorithm is employed
- The data-model alliance network realizes the multi-step-ahead thermal warning

Authors
Marui Li, Chaoyu Dong, Yunfei Mu, Xiaodan Yu, Qian Xiao, Hongjie Jia

Correspondence
cydong@tju.edu.cn

In brief
The temperature of a lithium-ion battery energy storage system has a great impact on its safety and performance. Combined with the thermal model of lithium-ion battery and the method of long short-term memory network, a data-model alliance network is established to realize the multi-step-ahead thermal warning of a lithium-ion battery energy storage system, and the accuracy is more than 97%.
Data-model alliance network for the online multi-step thermal warning of energy storage system based on surface temperature diffusion

Marui Li,1 Chaoyu Dong,1,2,* Yunfei Mu,1 Xiaodan Yu,1 Qian Xiao,1 and Hongjie Jia1

1School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China
2Lead contact
*Correspondence: cydong@tju.edu.cn
https://doi.org/10.1016/j.patter.2021.100432

SUMMARY

As an important type of energy storage, battery energy storage systems have been widely used. However, there are frequent cases of battery explosion due to high temperature. To address this issue, researches have been carried out either in the model-driven or the data-driven aspects to predict the temperature of the battery. In this paper, a two-node electrothermal model and a multi-scale long short-term memory network are established formulating a data-model alliance network (DMAN) for surface temperature diffusion. An improved adaptive boosting algorithm is then employed to enhance the bridge of the two models. Integrating a data-model alliance module (DMAM) and multi-step-ahead thermal warning network (MATWN), this DMAN provides an advanced online multi-step-ahead thermal warning structure to achieve early warning of temperature crossing. Experimental results verify the progressiveness of the proposed technique.

INTRODUCTION

Energy storage systems are an important part of energy systems. They are becoming increasingly important with the development of distributed energy sources and electric vehicles, etc. With the deterioration of the environment and the depletion of non-renewable resources, lithium batteries have become one of the most promising energy storage systems.1 This is because lithium batteries have the advantages of high energy density, high power density, and long life.2 However, the application of battery energy storage systems is seriously hindered by inadequate methods and technologies to address battery safety issues.3

The safety of battery energy storage systems includes many aspects, of which temperature is a major issue affecting battery reliability and safety.4 The low temperature will slow down the migration and diffusion rate of ions in the electrolyte. Lithium precipitation may even occur on the surface of the negative electrode, which may lead to diaphragm puncture and other failures.5 High temperature will accelerate the internal side reactions of the battery, such as the decomposition of solid electrolyte interphase (SEI) film and the accelerated rate of diffusion.
irreversible loss of active material. When the temperature rises further, a series of side reactions such as electrolyte decomposition and decomposition of cathode material may occur, even leading to serious faults such as internal short circuit and thermal runaway. Therefore, it is important to monitor and predict the temperature of battery energy storage systems.

At present, the fault diagnosis methods of battery energy storage systems are mainly divided into battery model methods and data-driven methods. The method based on battery model achieves fault diagnosis by comparing the predicted value of the model with the actual measured value of the battery. The premise is to establish a reliable and accurate battery model. Xie et al. proposed an improved three-dimensional thermal model for a pouch battery. Murashko et al. also described a three-dimensional thermal model of lithium batteries.

Although these models can simulate the temperature distribution of the battery, they are too computationally intensive and thus not conducive to online applications. To simplify the model, a multi-node model has been developed. A simple two-node thermal model has been widely used.

A battery is a dynamic and time-varying electrochemical system with nonlinear behavior and complicated internal mechanisms. Due to the complexity and computational difficulty of electrothermal models and with the development of deep learning, many scholars have tried to use data-driven black-box models to predict the temperature of batteries. Among them, Ojo et al. proposed a method to accurately estimate the cell surface temperature, which relies on a long short-term memory neural network (LSTM) in conjunction with an alteration to the walk-forward technique. LSTM is an excellent kind of network that can deal with time series and has been widely used in battery state estimation and fault diagnosis. Hong et al. used LSTM to accurately predict the multi-step voltage of a battery system and further evaluated battery safety. Wang et al. proposed a power consumption predicting and anomaly detection approach based on LSTM. However, few scholars have combined electrothermal models and neural networks.

An advanced cyber hierarchy and interactional network (CHAIN) framework for energy system design and management has been proposed. CHAIN is based on artificial intelligence technology and cloud computing, and it connects the cyber end and vehicle end through wireless communication technology. The structure of the CHAIN framework decomposes a complex system into hierarchical interdependent layers with various functions for full lifetime management of battery systems. This method reduces the requirement of computing power of the vehicle end and is applicable to a wider range of users and scenarios. Similarly, the battery digital twin system plays a revolutionary role in the multi-scale design and intelligent management system of a battery system, but it also faces higher requirements for the multi-physical model, communication network, data security, and so on.

The main work of this paper is as follows. A data-model alliance network (DMAN) is established. First, a data-model alliance module (DMAM) is developed to couple the prediction results of the electrothermal model and multi-scale LSTM network model. The parameters of the electrothermal model are identified and updated by the recursive least-square method. The multi-scale LSTM prediction network has two inputs and two LSTM networks. The electrothermal model and multi-scale LSTM network model are used to predict the surface temperature diffusion of the battery energy storage system at the next two sampling times, and then modified by the improved adaptive boosting method. Interfacing with the DMAM, an online multi-step-ahead thermal warning network (MATWN) for battery energy storage system is established. The online MATWN takes the output of the DMAM as the input to judge whether the temperature exceeds the boundary at the next two sampling times and decides whether to give an early warning signal. The two parts of DMAM and MATWN constitute the DMAN. In this paper, the DMAM is used to improve the prediction accuracy of the surface temperature diffusion of a
battery energy storage system. Then a classification neural network model is established, and the prediction temperature with high accuracy is used as the input to further improve the accuracy of online temperature warning. Finally, the advantages of the proposed model are verified by experiments. We define nomenclature, Greek symbols, subscripts, superscripts, and acronyms separately in Box 1.

Method

Surface temperature prediction for battery energy storage systems

Surface temperature prediction based on electrothermal model. First, the two-node electrothermal model of the batteries is established. Similar to the equivalent circuit model, the battery thermal model consists of a heat source, two thermal resistances, and two thermal capacities, and it concentrates the temperature distribution into three points of temperature, i.e., core temperature, surface temperature, and ambient temperature. The thermal resistance causes a temperature difference at three points. The battery thermal model is shown in Figure 1. Here, Q is the heat generation rate of lithium-ion batteries, R1, R2 and the heat capacity C1, C2 constitute the heat transfer part of the thermal model. Tin, TS, and Tamb represent three concentrated temperature points.

Figure 1. Battery thermal model

The heating rate Q represents the heat generation part of the thermal model of the battery, and the thermal resistance R1, R2 and the heat capacity C1, C2 constitute the heat transfer part of the thermal mode. Tin, TS, and Tamb represent three concentrated temperature points.

Heat Q is generally composed of irreversible heat caused by electrode polarization, reversible heat caused by entropy change, heat generated by concentration gradient, and phase change reaction. Since the input of our model is a large load current, and the main source of heat generation is ohmic heat under a large load current, the simplified heat generation equation is given by

\[ Q = R \cdot i^2 \]  

(Equation 4)

where \( i \) is the battery terminal current, and \( R \) denotes the internal resistance of the battery. \( R \) is assumed to be a function of the battery’s core temperature \( T_{in} \) only. \( R \) is a variable that varies with temperature and state of charge (SOC). Although the effect of SOC on the internal resistance of the battery at low temperatures cannot be ignored, the internal resistance of the battery is almost independent of SOC in the range of 20°C–50°C. And normally the battery is only cycled in a limited range, e.g., 30%–70% SOC or 20%–80% SOC, in which the internal resistance of the battery is not much affected by the change of SOC. The simplified R equation is a function of temperature.

\[ R = f(T_{in}) \]  

(Equation 5)

The correlation curve of internal resistance \( R \) and temperature can be obtained by offline identification method. And the value of \( R \) at the current time can be obtained by the interpolation method when predicting the temperature of the battery energy storage system.

Applying the following discretization. \( T_s \) is the sampling time.

\[ \frac{dT(k)}{dt} = z - 1 \cdot T(k) \]  

(Equation 6)

\[ zT(k) = T(k + 1) \]  

(Equation 7)

and setting \( \Delta t = 1 \), Equations 2 and 3 reduce to

\[ C_1 \cdot \left( T_{in}(k) - T_{in}(k - 1) \right) = Q(k - 1) - K_1 \cdot \left( T_{in}(k - 1) - T_s(k - 1) \right) \]  

(Equation 8)

\[ C_2 \cdot \left( T_s(k) - T_s(k - 1) \right) = K_1 \cdot \left( T_{in}(k - 1) - T_s(k - 1) \right) - K_2 \cdot \left( T_s(k - 1) - T_{amb}(k - 1) \right) \]  

(Equation 9)

We assume that the ambient temperature is constant. Then Equations 8 and 9 are combined to eliminate \( T_{in} \) to obtain the equation of \( T_s \) at the current time.

\[ T_s(k) = aT_s(k - 1) + bT_s(k - 2) + cQ(k - 2) + dT_{amb} \]  

(Equation 10)

\[ a = 1 - \frac{K_1 \cdot C_2}{C_1} + \frac{K_1 - K_2}{C_2} \]  

(Equation 11)

\[ b = \frac{K_2^2}{C_1 C_2} - 1 + \frac{K_1}{C_1} \]  

(Equation 12)
The values of $a$, $b$, $c$, and $d$ can be obtained from Equations 15 and 16. When the model parameters are identified, we can use Equation 10 to predict the surface temperature of the battery at sampling time $k + 1$. Then, using the predicted surface temperature at $k + 1$ and applying Equation 10 again, the surface temperature at sampling time $k + 2$ can be predicted. In practice, parameter identification and temperature prediction are carried out simultaneously. When the parameters of Equation 10 are updated using the RLS method, the surface temperature for the next two consecutive sampling times is immediately derived using Equation 10. Here, the signal we import into the model is the measurement signal with noise. We use the electrothermal model to suppress the noise and accurately predict the real surface temperature of the battery energy storage system.

Surface temperature prediction based on multi-scale LSTM. A recurrent neural network (RNN) is a kind of network that can deal with time series. This is because the standard RNN structure has a chain form of recurrent neural network module. Thus, the output of the network layer at the previous time can be fed back to the input of the same network layer the next time. However, when the gradients are multiplied multiple times in multiple periods, the gradients tend to zero gradually. As a result, RNN stops learning and cannot learn long-term dependence. Schmidhuber et al. proposed a recursive neural network with LSTM in 1997. LSTM is a special recurrent neural network that not only solves the problem of gradient explosion and disappearance during the back-propagation process of RNN, but it also overcomes the long-term dependencies of RNN. The key to overcoming the long-term dependence of LSTM lies in the addition of cell state. The structure of a recurrent LSTM cell is shown in Figure 2. $x_t$, $h_t$, and $C_t$ represent input, output, and cell state respectively. The cell state $C_t$ represents long-term memory. The cell state $C_t$ and output $h_t$ are transitive overtime in an LSTM network. Each LSTM cell is made of input, forget, and output gates. The gate consists of a sigmoid neural network layer and a bitwise multiplication operation. The sigmoid neural network layer can convert the input signal to a value between 0 and 1 to determine how many input signals can pass through. 0 means no signal is allowed to pass, and 1 means all signals are allowed to pass. The first stage is the forgetting gate, which determines the information in the cell state at the last moment to be forgotten. The next stage is the input gate, which determines the new information stored in the cell state.
Finally, the output gate is determined as the output value. LSTM cells update the cell state and generate the input of the next time through the output of the previous time, the cell state, and the current input. The calculation process is shown in Equation 17. W and b represent weight and deviation respectively.

\[
\begin{align*}
    f_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\
    i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\
    C_t &= \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \\
    C_t &= f_t \cdot C_{t-1} + i_t \cdot C_t \\
    \tanh(x) &= \frac{e^x - e^{-x}}{e^x + e^{-x}} \\
    o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\
    h_t &= o_t \cdot \tanh(C_t)
\end{align*}
\] (Equation 17)

First, the forgetting gate works. The output \(h_{t-1}\) of the previous time and the input \(x_t\) of the current time are accepted, and the signal \(f_t\) is output through the sigmoid layer. Then \(f_t\) is a value from 0 to 1, which is multiplied by \(C_{t-1}\) to determine the information retained in \(C_{t-1}\). In the input gate, \(x_t\) and \(h_{t-1}\) are input into the sigmoid layer, and a value \(i_t\) between 0 and 1 is output. At the same time, \(x_t\) and \(h_{t-1}\) create a new state candidate vector \(\overline{C}_t\) with values between \(-1\) and \(1\) through a tanh neural network layer. Then \(i_t\) and \(\overline{C}_t\) are multiplied to determine which information is added to the cell state. Therefore, the forgetting gate and the input gate retain the previous important information and add some new information in the cell state, to get a new cell state \(C_t\). In the output gate, \(x_t\) and \(h_{t-1}\) are input into the sigmoid layer, and a value \(o_t\) between 0 and 1 is output. Then the updated cell state \(C_t\) is converted to a function between \(-1\) and \(1\) by a tanh function. The new output is obtained by multiplying \(o_t\) and \(C_t\).

Due to the different characteristics of multiple inputs and their effects on the output of the neural network, we divide the unique input of the neural network into two parts by adjusting the weight, deviation, and learning rate of the full connection layer. The multi-scale LSTM network for temperature prediction established in this paper is shown in Figure 3. One input layer connects two fully connected layers. The input is divided into two parts by the fully connected layers with learning rate of 0. In this way, the weight and deviation of the fully connected layer will not change. Setting the weight to 1 and the deviation to 0 can ensure that the input will not change through the fully connected layer. The two inputs are then fed into two different integrated LSTM networks. Each integrated LSTM layer consists of two LSTM layers and two dropout layers. Then the two integrated LSTM layers are connected by a concatenation layer, and then the two fully connected layers are connected. The activation function of the fully connected layer containing 32 neurons is tanh, and the activation function of the fully connected layer with four neurons is rule. The features of samples are extracted and learned through the integrated LSTM layer, and then mapped to the output through the fully connected layer. Since deep learning networks often suffer from overfitting,
reduces the generalization ability of the network, a dropout layer is added after the LSTM layer to prevent overfitting.  

The number of neurons in each layer is indicated in brackets in Figure 3. Because the characteristics of battery temperature, heating, and SOC may affect each other and their waveforms are different, these measurements are divided into two inputs in this paper. One input is the surface temperature $T_s$ and ambient temperature $T_{\text{amb}}$, the other input is the $Q$ and SOC. In this paper, the noise measurement signal of 20 consecutive sampling points is used as the input. The outputs of the neural network are the real surface temperature $T_s(k+1)$ and $T_s(k+2)$ of the battery energy storage system in the next two consecutive sampling times.

**Surface temperature prediction based on data-model alliance module.** Adaptive boosting (AdaBoost) is an iterative algorithm. Its core idea is to train different weak classifiers for the same training set, and then combine these weak classifiers to form a strong classifier. AdaBoost can deal with classification and regression problems. The algorithm is realized by changing the data distribution. It determines the weight of each sample according to whether the classification of each sample in each training set is correct and the accuracy of the last overall classification. The new data set with modified weights is sent to the lower classifier for training. Finally, the classifier obtained from each training is fused as the final decision classifier. The improved AdaBoost can be used to solve the regression problem, and the weighted summation is adopted to combine the surface temperature predicted by the two models. Then the DMAM is established.

First, the regression error rate is calculated by Equation 18.

$$e = \frac{1}{n} \sum_{t=1}^{n} \frac{(T_{s,t} - \hat{T}_{s,t})^2}{E^2}$$  

(Equation 18)

Here, $\hat{T}_{s,t}$ is the predicted surface temperature of the model at the time $t$. $T_{s,t}$ is the real surface temperature of the battery energy storage system at the time $t$. $n$ is the number of samples needed for each update of weight. $E$ is the maximum error in the training set.

$$E = \max |T_{s,t} - \hat{T}_{s,t}|$$  

(Equation 19)

The weight is then calculated by

$$a = \ln \left( \frac{1 - e}{e} \right)$$  

(Equation 20)

Then the weights of the model prediction $a_{\text{MOD}}$ and the multi-scale LSTM network prediction $a_{\text{LSTM}}$ are calculated through Equations 18 and 20 respectively. Then, the weights of model prediction and LSTM prediction are modified by Equations 21 and 22, so that the sum of the two weights is 1.

$$a_{\text{MOD}} = \frac{e_{\text{MOD}}}{e_{\text{MOD}} + e_{\text{LSTM}}}$$  

(Equation 21)

$$a_{\text{LSTM}} = \frac{e_{\text{LSTM}}}{e_{\text{MOD}} + e_{\text{LSTM}}}$$  

(Equation 22)

Finally, the coupled surface temperature $T_s$ is predicted by

$$T_s = a_{\text{MOD}} \cdot T_{s,\text{MOD}} + a_{\text{LSTM}} \cdot T_{s,\text{LSTM}}$$  

(Equation 23)

Here, $T_{s,\text{MOD}}$ and $T_{s,\text{LSTM}}$ are the surface temperature predicted by the model and LSTM neural network respectively.

This paper uses the data of ten sampling points to calculate the weight, that is, $a = 10$. It is assumed that the prediction weights of the two models remain unchanged in the next two sampling times. Then, the predicted value of the surface temperature $T_s(k+1)$ and $T_s(k+2)$ in the next two consecutive sampling times is corrected by Equation 23. The weights of model prediction and multi-scale LSTM prediction are continuously updated by the sliding window method, and the surface temperature is predicted while updating the weights.

**Online thermal diagnosis and multi-step-ahead thermal warning based on data-model alliance network**

Because high temperature will seriously affect the operation safety of a battery energy storage system, it is necessary to diagnose and warn in advance whether the temperature exceeds the limit. More accurate prediction of temperature will reduce misjudgment. However, there are deviations between even the most accurate prediction and the real temperature. To further enhance the early warning capability, a multi-step-ahead thermal warning model (MATWM) based on DMAM is developed in this paper as a discriminator to realize the diagnosis and early warning of the surface temperature diffusion of the battery storage system.

Because the multi-scale LSTM network model established in the previous section has a good prediction, the online MATWM developed in this article is basically the same as the model used in the previous paper for surface temperature prediction. The differences are as follows. First, the regression network is transformed into a classification network with cross-entropy as the loss function. The cross-entropy is represented by Equation 24. The probability distribution $p$ is the expected output, and the probability distribution $q$ is the actual output. Cross-entropy represents the distance between the actual output probability and the expected output probability. Second, the input to the

---

Table 1. Comparison of different networks

| Method | MSE of 1 s | MSE of 2 s | Method | Accuracy |
|--------|-----------|-----------|--------|---------|
| The single input temperature prediction network with one LSTM layer | 0.0150 | 0.0119 | The single channel thermal warning network with one LSTM layer | 0.9682 |
| The single input temperature prediction network with two LSTM layers | 0.0208 | 0.0155 | The single channel thermal warning network with two LSTM layers | 0.9692 |
| Multi-scale LSTM prediction network | 0.0105 | 0.0099 | Multi-step-ahead thermal warning network | 0.9712 |

---
network becomes the output of the DMAM developed in the previous section, i.e., the surface temperature at the next two thermal states. Third, the output of the network changes to 0 and 1. A value of 0 means that the surface temperature of the battery storage system does not cross the boundary in the next two thermal states, and 1 means that the temperature may cross the boundary and needs advance warning.

Table 2. MSE of prediction results of different prediction models under 3C current

| MSE                  | Model-based method | Data-based method | DMAN model |
|----------------------|--------------------|-------------------|------------|
| The next thermal state | 0.0062             | 0.0105            | 0.0058     |
| The thermal diffusion of two-step ahead | 0.0135 | 0.0099 | 0.0078 |



\[ H(\rho, q) = - \sum \rho_i \log q_i \]  

(Equation 24)

The MATWM established in this paper is shown in Figure 4. First, the input features are extracted and mapped through the LSTM layer and full connection layer. Then the output of the fully connected layer refracts to the probability between 0 and 1 through the softmax layer. Then the output of softmax and the real output calculate the network loss through cross-entropy and calculate the gradient of loss to update the weight and deviation of each hidden layer.

Table 3. The calculation time of different prediction models

| Method | Model-based method | Data-based method | DMAN model |
|--------|--------------------|-------------------|------------|
| Time/s | 5.62e-6            | 7.46e-5           | 8.40e-5    |
The complete DMAN process for online diagnosis and early warning of temperature overrun is shown in Figure 5. First, the parameters of the electrothermal model are identified by RLS method, and then the surface temperature of the next two sampling times is predicted by the electrothermal model and the previous real-time measurement. At the same time, the multi-scale LSTM network prediction model is used to get the predicted temperature. The results of the two prediction models are modified by DMAM, and a better-coupled prediction temperature is obtained. Finally, the predicted temperature of the DMAM is input into the MATWM to judge whether the temperature will exceed the boundary. Early warning is given if the temperature may exceed the limit.

RESULTS

We use the same hidden layer structure in the multi-scale LSTM prediction model and multi-step-ahead thermal warning network, which is obtained through experiments. The experimental results are shown in Table 1.

It can be seen from Table 1 that the multi-scale LSTM prediction network and multi-step-ahead thermal warning network designed with our multi-scale LSTM network structure have improved prediction ability compared with other networks.

Results of multi-step prediction of surface temperature of a battery energy storage system

For the electrothermal model, the RLS method is used to update the parameters to get $q(k)$, and then the surface temperatures $T_s(k+1)$ and $T_s(k+2)$ are predicted by the model. For the multi-scale LSTM prediction model, there are many hyperparameters. This paper uses MATLAB to train the network. We choose the training parameters to make the network prediction best. This paper sets InitialLearnRate = 0.001, LearnRateDropPeriod = 400, LearnRateDropFactor = 0.2, MaxEpochs = 550, and uses the Adam as the optimizer to train the network. Finally, a multi-scale LSTM network with two input layers is formulated to predict the change of surface temperature obtained. Then, the measured data $T_s$, $Q$, SOC, and $T_{amb}$ of 20 s are input into the trained network to obtain the predicted surface temperature of the next two steps. Finally, the prediction results are corrected by combining the two models through the DMAM. We have carried out the experiment of 3C current. The prediction results and errors for the three models are shown in Figure 6. The mean square error (MSE) of the three methods is calculated, and the results are shown in Table 2. The calculation time of the three models is shown in Table 3.

Figure 6 shows that the DMAM effectively combines the advantages of the electrothermal model and multi-scale LSTM prediction model. The modified prediction temperature not only has a good waveform, but it also reduces the prediction error. It can be seen from Table 1 that the prediction of the electrothermal model for the first step is more accurate than the multi-scale LSTM prediction model. However, for the second step prediction, the accuracy of the multi-scale LSTM prediction model is basically unchanged, whereas the accuracy of the electrothermal model decreases significantly. From the perspective of MSE, the accuracy of the two-step prediction results modified by the DMAM is significantly improved.

And to expand the temperature range, we added a large rate current experiment. In the process of the 10,000 s test, we used the alternating current of 4C square wave and 2C square wave with a period of 200 s. The experimental results are shown in Figure 7 and Table 4. It can be seen from Figure 7 and Table 3 that the proposed DMAM is still effective.

| Table 4. MSE of prediction results of different prediction models under 4C current |
|-----------------|-----------------|-----------------|
| MSE             | Model-based method | Data-based method | DMAM model |
| The next thermal state | 0.0342 | 0.0422 | 0.0240 |
| The thermal diffusion of two-step ahead | 0.0748 | 0.0423 | 0.0328 |
Results of multi-step-ahead thermal warning

The online MATWM built in this paper adopts a hidden layer structure that is basically the same as the multi-scale LSTM prediction model. The input of the network is the result $T_s(k+1)$ and $T_s(k+2)$ predicted by the DMAM. The output of the network is 0 and 1, representing whether the temperature is out of the limit. The MATWM network is also built using MATLAB and trained using the Adam algorithm. This article changes training parameters and the probability of dropout. Then we test different networks. The results of the tests are shown in Figure 8.

Figure 8A shows that the number of iterations and the initial learning rate of the network are interrelated. When both of them are too small, network training is insufficient. But when both of them are too large, there will be overfitting, which will reduce the generalization ability of the network. Only with a reasonable number of iterations and learning rate can we get a network with high accuracy. In Figure 8B, if the dropout probability is too small, the ability to prevent overfitting is weak, and the accuracy of the training network is low. When the dropout probability exceeds the optimal value, there is too much information lost, so the network training is not sufficient and the accuracy is not high.

We choose the training parameters to make the network prediction best. Finally, this article sets InitialLearnRate = 0.0001, LearnRateDropPeriod = 50, LearnRateDropFactor = 0.2, MaxEpochs = 200, MiniBatchSize = 512, and Dropout = 0.2. The training process is shown in Figure 9.

The network established in this article is evaluated by four indexes described in Equations 25, 26, 27, and 28.

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad \text{(Equation 25)}
\]

\[
\text{Precision} = \frac{TP}{TP + FP} \quad \text{(Equation 26)}
\]

\[
\text{Recall} = \frac{TP}{TP + FN} \quad \text{(Equation 27)}
\]

\[
F1 = \frac{2 \cdot \text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} = \frac{2TP}{2TP + FP + FN} \quad \text{(Equation 28)}
\]

TP represents the number of samples with the real value of 1 and the predicted value of 1.

TN represents the number of samples with the real value of 0 and the predicted value of 0.

FP represents the number of samples with the real value of 0 and the predicted value of 1.

FN represents the number of samples with the real value of 1 and the predicted value of 0.

Table 5 illustrates the comparison between the network model established in this paper and other networks. Table 6 compares the real-time computing time of these networks.

From Table 5, we can see that the precision of the simple network is higher than that of the complex network; i.e., the accuracy of the simple network in discriminating positive examples is higher. However, its recall is lower, i.e., the ability to discriminate positive cases is poor. The $F1$ index integrates precision and recall, which can better illustrate the effectiveness of the test method. Table 5 shows that the accuracy and $F1$ score of the proposed MATWM are improved. In the 4C current experiment, the accuracy of the proposed MATWN can even reach 99.97%.

**Figure 9. Training process**

The blue curve shows the change of accuracy during training. The red curve shows the change of loss during training.
Network 2 represents a single channel network with two LSTM layers in series.

From Tables 3 and 6, we can see that the time is mainly spent on training the network, and the more complex the network is, the longer will be the training time. But a test takes only about one thousandth of a second. So the real-time thermal warning basically does not take time.

**DISCUSSION**

In this article, a data-model alliance network (DMAN) is established. DMAN consists of a data-model alliance module (DMAM) and an online multi-step-ahead thermal warning network (MATWN). A DMAM is developed by integrating the thermal model and data-driven algorithm. With the convenient interface, the deep neural network with cross-entropy loss serves as the MATWN. From the perspective of MSE, during the prediction of the next thermal state, the experiments show that the MSE of the DMAM is 6.8% less than that of the electro-thermal model, and 44% less than that of the traditional LSTM approach. For the thermal diffusion of two-step ahead, the MSE of the proposed DMAM reduces by 42% and 21% respectively. On the basis of DMAM, an online MATWN also achieves good performance. This model employs the output of the boosted temperatures as the input to assess the diffusion of the battery energy storage system. The experiments show that the proposed model is better than both traditional LSTM network and thermal model in accuracy and F1 score. The accuracy and F1 of the proposed model achieve an improvement of 0.3% and 0.2%, respectively.

**EXPERIMENTAL PROCEDURES**

**Resource availability**

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Chaoyu Dong (cydong@tju.edu.cn).

**Materials availability**

This study did not generate new unique reagents.

**Data and code availability**

The codes generated during this study are available at Github: https://github.com/15098781837/Thermal-Warning.git. The datasets generated during this study are available at Mendeley: https://data.mendeley.com/datasets/xd8bn5tktw/draft?a=4038b51c-fa1f-4660-a30a-874f46f57a824.

**Table 5. Comparison of prediction ability of different networks**

| Index          | Accuracy | Precision | Recall | F1   |
|---------------|----------|-----------|--------|------|
| Proposed Network | 0.9712   | 0.9838    | 0.9796 | 0.9817 |
| Network 1      | 0.9682   | 0.9845    | 0.9749 | 0.9797 |
| Network 2      | 0.9692   | 0.9841    | 0.9766 | 0.9803 |

Network 1 represents a single channel network with one LSTM layer. Network 2 represents a single channel network with two LSTM layers in series.

**Table 6. The calculation time of different networks**

| Method        | Network 1 | Network 2 | Proposed network |
|---------------|-----------|-----------|------------------|
| Time/s        | 9.36e-5   | 7.79e-5   | 1.040e-4         |

Network 1 represents a single channel network with one LSTM layer. Network 2 represents a single channel network with two LSTM layers in series.

Acknowledgments

This work was funded by China Postdoctoral Science Foundation (grant no. 2020M680880) and National Natural Science Foundation of China (no. 5206163S103, 51625702).

**AUTHOR CONTRIBUTIONS**

M.L., C.D., and Q.X. derived the data-model alliance network for energy storage system based on the surface temperature diffusion and original draft preparation. Y.M., X.Y., and H.J. contributed to review and editing and provided computational resources and supervision. All authors have read and agreed to the revised version of the manuscript.

**DECLARATION OF INTERESTS**

The authors declare no competing interests.

Received: April 28, 2021
Revised: June 23, 2021
Accepted: December 21, 2021
Published: January 26, 2022

**REFERENCES**

1. Hu, X., Zhang, K., Liu, K., Lin, X., Dey, S., and Onori, S. (2020). Advanced fault diagnosis for lithium-ion battery systems: a review of fault mechanisms, fault features, and diagnosis procedures. IEEE Ind. Electron. Mag. 14, 65–91.
2. Qian, Z., Li, Y., and Rao, Z. (2016). Thermal performance of lithium-ion battery thermal management system by using mini-channel cooling. Energy Convers. Management 126, 622–631.
3. Liu, P., Sun, Z., Wang, Z., and Zhang, J. (2018). Entropy-based voltage fault diagnosis of battery systems for electric vehicles. Energies 11, 136.
4. Tang, X., Wang, Y., Zou, C., Yao, K., Xia, Y., and Gao, F. (2019). A novel framework for Lithium-ion battery modeling considering uncertainties of temperature and aging. Energy Convers. Manage. 170, 162–170.
5. Burow, D., Sergeeva, K., Calles, S., Schorb, K., Börger, A., Roth, C., and Heitjans, P. (2016). Inhomogeneous degradation of graphite anodes in automotive lithium ion batteries under low-temperature pulse cycling conditions. J. Power Sourc. 307, 806–814.
6. Feng, X., Fang, M., He, X., Ouyang, M., Lu, L., Wang, H., and Zhang, M. (2014). Thermal runaway features of large format prismatic lithium ion battery using extended volume accelerating rate calorimetry. J. Power Sourc. 255, 294–301.
7. Finegan, D.P., Zhu, J., Feng, X., Keyser, M., Ulmefors, M., Li, W., Bazant, M.Z., and Cooper, S.J. (2021). The application of data-driven methods and physics-based learning for improving battery safety. Joule 5, 316–329.
8. Finegan, D.P., and Cooper, S.J. (2019). Battery safety: data-driven prediction of failure. Joule 3, 2599–2601.
9. Xie, Y., He, X., Hu, X., Li, W., Zhang, Y., Liu, B., and Sun, Y. (2020). An improved resistance-based thermal model for a pouch lithium-ion battery considering heat generation of posts. Appl. Therm. Eng. 164, 114455.
10. Murashko, K., Pyrhonen, J., and Laurila, L. (2013). Three-dimensional thermal model of a lithium ion battery for hybrid mobile working machines: determination of the model parameters in a pouch cell. IEEE Trans. Energy Convers. 28, 9.
11. Pan, Y., Hua, Y., Zhou, S., He, R., Zhang, Y., Yang, S., Liu, X., Lian, Y., Yan, X., and Wu, B. (2020). A computational multi-node electro-thermal model for large prismatic lithium-ion batteries. J. Power Sourc. 459, 228070.
12. Ruan, H., Jiang, J., Xu, Q., Sun, B., and Cheng, G. (2017). A reduced wide-temperature-range electro-thermal model and thermal parameters determination for lithium-ion batteries. Energy Proced. 105, 805–810.
13. Lin, X. (2014). A lumped-parameter electro-thermal model for cylindrical batteries. J. Power Sourc. 257, 1–11.
14. Zhang, C., Li, K., and Deng, J. (2016). Real-time estimation of battery internal temperature based on a simplified thermoelectric model. J. Power Sourc. 302, 146–154.

15. Chen, L., Hu, M., Cao, K., Li, S., Su, Z., Jin, G., and Fu, C. (2020). Core temperature estimation based on electro-thermal model of lithium-ion batteries. Int. J. Energy Res 44, 5320–5333.

16. Hu, X., Xu, L., Lin, X., and Pecht, M. (2020). Battery lifetime prognostics. Joule 4, 310–346.

17. Hasan, M.M., Ali Pourmousavi, S., Jahanbani Ardakani, A., and Saha, T.K. (2020). A prediction model of vanadium redox flow battery based on multiple innovation recursive least squares. In 2020 Asia Energy and Electrical Engineering Symposium (AEEES) (IEEE), pp. 758–762.

18. Hussein, A.A., and Chehade, A.A. (2020). Robust artificial neural network-based models for accurate surface temperature estimation of batteries. Int. J. Energy Res. 44, 101879.

19. Fang, K., Mu, D., Chen, S., Wu, B., and Wu, F. (2012). A prediction model based on artificial neural network for surface temperature simulation of nickel–metal hydride battery during charging. J. Power Sourc. 208, 378–382.

20. Ojo, O., Lang, H., Kim, Y., Hu, X., Mu, B., and Lin, X. (2021). A neural network based method for thermal fault detection in lithium-ion batteries. IEEE Trans. Ind. Electron 68, 4068–4078.

21. Hong, J., Wang, Z., and Yao, Y. (2019). Fault prognosis of battery system based on accurate voltage abnormality prognosis using long short-term memory neural networks. Appl. Energy 251, 113381.

22. Wang, X., Zhao, T., Liu, H., and He, R. (2019). Power consumption predicting and anomaly detection based on long short-term memory neural network. In 2019 IEEE 4th International Conference on Cloud Computing and Big Data Analysis (ICCCBDA) (IEEE), pp. 487–491.

23. Yang, S., He, R., Zhang, Z., Cao, Y., Gao, X., and Liu, X. (2020). CHAIN: cyber hierarchy and interactional network enabling digital solution for battery full-lifespan management. Mater. 3, 27–41.

24. Wu, B., Widanage, W.D., Yang, S., and Liu, X. (2020). Battery digital twins: perspectives on the fusion of models, data and artificial intelligence for smart battery management systems. Energy and AI 1, 100016.

25. Huo, Y., Hu, W., Li, Z., and Rao, Z. (2020). Research on parameter identification and state of charge estimation of improved equivalent circuit model of Li-ion battery based on temperature effects for battery thermal management. Int. J. Energy Res. 44, 11583–11596.

26. Liao, Z., Xia, B., Wang, W., Sun, W., Bai, Y., and Wang, M. (2018). A novel method for lithium-ion battery online parameter identification based on variable forgetting factor recursive least squares. Energies 11, 1358.

27. Sun, X., Ji, J., Ren, B., Xie, C., and Yan, D. (2019). Adaptive forgetting factor recursive least square algorithm for online identification of equivalent circuit model parameters of a lithium-ion battery. Energies 12, 2242.

28. Chang, Y., Sun, M., Jia, W., and Liu, Z. (2020). Online model identification method of vanadium redox flow battery based on multiple innovation recursive least squares. In 2020 Asia Energy and Electrical Engineering Symposium (AEEES) (IEEE), pp. 758–762.

29. Hong, J., Wang, Z., Chen, W., and Yao, Y. (2019). Synchronous multi-parameter prediction of battery systems on electric vehicles using long short-term memory networks. Appl. Energy 254, 113648.

30. Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. Neur. Comput. 9, 1735–1780.

31. Zhang, Y., Xiong, R., He, H., and Pecht, M.G. (2018). Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries. IEEE Trans. Veh. Technol. 67, 5695–5705.

32. Tan, Y., and Zhao, G. (2020). Transfer learning with long short-term memory network for state-of-health prediction of lithium-ion batteries. IEEE Trans. Ind. Electron 67, 8723–8731.

33. Li, W., Sengupta, N., Dechent, P., Howey, D., Annaswamy, A., and Sauer, D.U. (2021). Online capacity estimation of lithium-ion batteries with deep long short-term memory networks. J. Power Sourc. 482, 228863.

34. Chemali, E., Kollmeyer, P.J., Preindl, M., Ahmed, R., and Emadi, A. (2018). Long short-term memory networks for accurate state-of-charge estimation of li-ion batteries. IEEE Trans. Ind. Electron 65, 6730–6739.

35. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 15, 1929–1958.

36. Das, R., Lopez, P.S., Ahmed Khan, M., Iversen, H.K., and Puthusserypady, S. (2020). FBCSP and adaptive boosting for multiclass motor imagery BCI data classification: a machine learning approach. In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (IEEE), pp. 1275–1279.

37. Li, D., Zhang, Z., Liu, P., Wang, Z., and Zhang, L. (2021). 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. 36, 1300–1315.