

A Novel State of Charge Estimation Method of Batteries Using Recurrent Neural Networks

Anyu Cheng^{a)}, Yao Wang^{b)}

School of Chongqing University of Posts and Telecommunications, Chongqing 400065, China.

^{a)} caycat@163.com, ^{b)} s150331046@stu.cqupt.edu.cn

Abstract. This paper, an improved recurrent neural network, long and short time memory model (LSTM) is used to estimate the SOC estimation of vehicle lithium ion battery, and the SOC estimation model of the battery based on LSTM network is established. Based on the electrochemical reaction of lithium ion batteries and the complex operating conditions of the electric vehicle, a battery model was established, and the experimental verification was carried out. The results showed that the accuracy of the SOC estimation model could meet the requirements of the SOC estimation application.

Key words: power battery, SOC, LSTM network.

INTRODUCTION

It is becoming increasingly difficult to ignore the energy crisis and environmental pollution. Because electric vehicles (EV) have the advantages of high energy utilization rate, no emission, low noise and diversification of energy sources, it has become a critical research field of automobile industry. And the BMS (battery management system) is one of the three core technologies of EV. Currently, one of the most significant discussions in battery management system is accurate of the SOC estimation. It is effectively to re-duce the cost of the battery that enhances the estimation of SOC accuracy.

The focus of the current research on battery SOC is to improve the accuracy of SOC estimation. Some re-searchers use the linear relationships between battery open-circuit voltage and SOC to estimate the SOC. This method has achieved good experimental results, especially in early stage of charge and late stage of discharge. When the battery is powered off for an hour to stabilize its internal state, an accurate open circuit voltage value is obtained. This limitation also makes this method difficult to achieve in practical applications. Ah counting method is the most commonly used SOC estimation method, also known as the Coulomb counting method. From the perspective of electrical energy, the SOC value is obtained by monitoring the battery current value and integrating it. This method has a measurement error when the charge and dis-charge current fluctuate in a large range, and it is easy to form SOC accumulation error.

At present, the emerging estimation methods include impedance measurements, neural networks, and Kalman filter, etc. It is difficult to test the impedance spectrum in the actual operation of an electric vehicle, and the impedance spectrum test is also affected by the temperature. Kalman filter method can quickly converge to the battery SOC value near, and it reduced the requirement for the accuracy of the initial SOC value. However, the drawback of this method is over-reliance on the choice of battery equivalent circuit model and the computing power of BMS. The neural network can avoid the complexity of the internal battery but requires a lot of data to train the model. The estimation error of the model is greatly affected by the training data set and training method. By comprehensive comparison, neural network as SOC estimation method is selected.

The rest of this paper is structured as follows: Section II discusses the definition of battery SOC and the choice of neural networks. Section III analyzes the effect of external parameters on SOC in conjunction with the battery

equivalent circuit model. Section IV goes on to show how to use LSTM method to establish SOC estimation model and corresponding simulation results and analysis.

THEORETICAL FRAMEWORK

State of Charge.

From the point of view of electricity, the conventional definition of the SOC is given as follows:

$$\text{SOC} = \frac{Q_c}{Q_I} \quad (1)$$

The above formula is proposed and widely used by USABC in the "Electric Car Battery Test Manual". It defines SOC as the percentage of rated capacity of the battery at a specific discharge rate. In Eq. (1), Q_c is rated capacity, Q_I is the discharge capacity of the battery when the discharge current is I. When the battery is fully charged, SOC=1; when the battery is fully discharged, SOC=0. According to the above equation, the actual amount of electricity discharged from the battery will be vary depending on the charge and discharge rate, and the Q_I obtained from Eq. (1) also will change. Therefore, the nominal capacity of the battery can be used, instead of the variable battery capacity Q_I at different discharge rates.

$$\text{SOC} = \frac{Q_c}{Q_N} \quad (2)$$

Considering that the power battery application field and the power battery complex operating conditions, from the perspective of electrical energy, can more accurately reflect the changes in the state of charge and discharge of the battery. So, we can get it:

$$\text{SOC} = \frac{Q_0 - \int_0^t \eta I dt}{Q_N} \quad (3)$$

The Q_0 in the above equation is the battery initial value of the nominal condition.

Through the previous study, we can see that the factors that affect the battery capacity, including the battery charge and discharge current, voltage, state of health (SOH), temperature, self-discharge factor, charge or discharge efficiency. Therefore, the battery SOC should be strictly defined above Parameters of the function, SOC is obtained from the following equations:

$$\text{SOC} = f(U, I, T, \text{SOH}, K_{dis}, \eta) \quad (4)$$

where U and I are voltage and current, respectively. K_{dis} and η represent the self-discharge efficiency and charge-discharge efficiency of the battery, respectively. The researchers' goal is to determine the relationship between the SOC and the corresponding parameter values.

The previous section describes some of the SOC estimation method and their advantages and disadvantages. In order to overcome these shortcomings, the researchers proposed a combination of different methods to overcome their own problems.

Now the artificial neural network method is often combined with fuzzy logic, while other types of neural networks are less used, taking into account the recurrent neural network (RNN) in the processing sequence on the excellent performance and SOC value is gradually changing, the paper try to combine RNN to solve the problem of accurate estimation of SOC value.

Long Short-Term Memory.

In the traditional neural network model, the layer and the layer are fully connected between the nodes between each layer is no connection. But this kind of ordinary neural network is helpless for many problems, especially for

some of the time sequence of the sample appears more important situation. In order to meet this demand, there has been such a neural network structure - the recurrent neural network (RNN). In the RNN, the output of the neuron can be directly applied to itself at the next moment. In RNN, the nodes between the hidden layers are no longer connected but connected, and the input of the hidden layer includes not only the output of the input layer but also includes the output of the hidden layer at the last moment. This actually contains the core concepts of programmable neural networks. At this point, RNN is Turing complete.

Similarly, in the RNN, the gradient information is propagated in time will appear decay phenomenon, that is, the gradient disappearance of many neural networks. In order to improve the common RNN model of the various shortcomings, machine learning areas to develop a variety of complex RNNs, the following will be a brief overview:

Based on the traditional RNN, long short-term memory is proposed. It is a new structure, through the door switch to achieve the memory function on time and to prevent the gradient disappeared. The re-researcher introduces the cell definition to hold the state before the current input, or you can add a restriction to determine which cells suppress which excitement. In this model, the introduction of cells in the structure, that is, add a hidden variable as a memory unit, and then a layer of the neural network before the increase of three, namely, input gate, output door, forgotten door, respectively. It has been shown that the network structure is very effective for long sequence dependency problems.

SYSTEM DESCRIPTION

The lithium-ion battery is a chargeable and discharge-able high-energy battery, it works is lithium ions between the positive and negative. When it is in the charge and discharge state, the battery will appear in the external characteristics of some resistance and capacitance characteristics. According to these characteristics, we can use the battery model to simulate the dynamic characteristics of the battery. The equivalent circuit model is such a battery model that conforms to the above description. So, we will combine one type equivalent circuit model to analyze the factors that affect the battery SOC.

The PNGV model is a low-order model. But it is more accurate than the simple first-order circuit model, Thevenin battery model, RC model. So, then we will combine PNGV model to analyze the impact of various factors on SOC. The PNGV model is shown below.

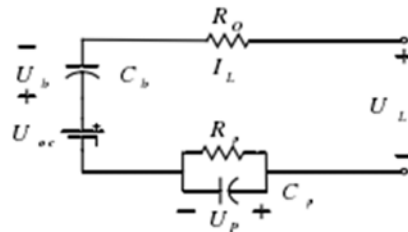


FIG 1. PNGV model

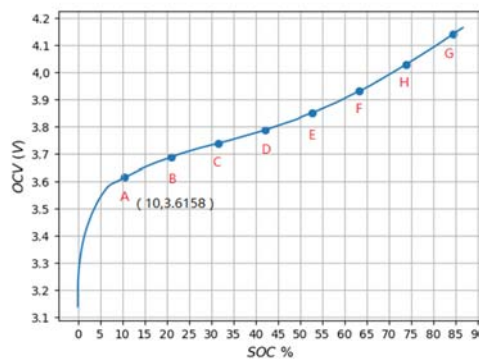


FIG 2. the discharge voltage and SOC obtained under specific conditions

Figure 2 shows the relationship between the discharge voltage and SOC obtained under specific conditions. From the previous analysis, we can see that the discharge voltage is similar to the open circuit voltage, which can be used to predict the SOC of the battery by the relationship curve of the battery SOC. But it is different from the open circuit voltage, the value of the open circuit voltage need to stand for a period of time to get, it can be measured in the battery discharge process in real time to get. We can see through the analysis of the curve in the battery discharge process, the initial discharge voltage reduction rate was a gradually reduced trend, to the late discharge, the voltage reduction rate began to gradually become larger. When the battery discharging, the voltage reducing rate is a trend of decrease in the beginning, but at the later, it is a trend of increasing. But the influence of the size of the discharge current and discharge temperature in actual use cannot be ignored, otherwise, it will cause a lot of errors. Also, the relationship between the discharge voltage and the SOC will change due to other factors, that is, the same voltage corresponding to the SOC will be different in different circumstances.

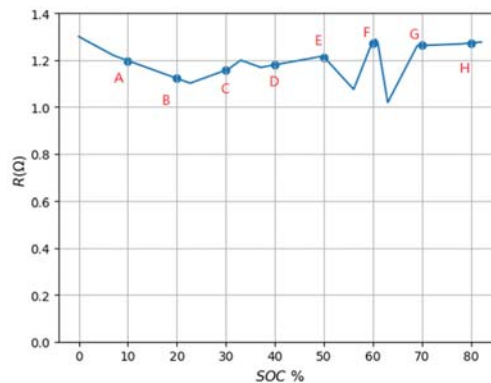


FIG 3. Battery resistance and SOC

Figure 3 for a certain condition under the relationship between resistance and SOC. It can be seen from the figure, in the initial discharge, the battery resistance does not change with the SOC changes in the period, such as the use of internal resistance SOC prediction cannot achieve the desired results. In the early stage of charging, there is a close correlation between the internal resistance of the battery and SOC. Until the last remaining 50% of the electricity, the change in battery internal resistance shows greater irregularity. In addition, the internal resistance of the battery will be affected by the discharge current and temperature, which will be described in terms of temperature and discharge current analysis. Therefore, we cannot directly use the battery internal resistance to obtain the SOC in real-time.

According to lithium ion battery characteristics and battery SOC definition, this paper analyzes the charge-discharge characteristics of the lithium-ion battery and the corresponding relationship between battery voltage, internal resistance, temperature and charge-discharge efficiency and SOC, and the relationship between these factors. In the next part, the external parameters of the battery include current, voltage, internal resistance, and temperature, which together serve as the input to build the SOC estimation model.

EXPERIMENTAL ANALYSIS

Before training the model, we need to take the step forward is data preprocessing. Because our car terminal to set the frequency of data obtained for five seconds, and the battery charge and discharge are required a certain time, so in a relatively short period of time SOC has not changed. So, we will be in the collection of data to see multiple sets of duplicate data, which is redundant data. In order to improve the prediction accuracy of the system and reduce its computational load, we will preprocess the data to eliminate the redundant data and improve the accuracy of the data and improve the system accuracy. Because the prediction of this paper is dependent on historical data, if the historical information can be memorized, the better the prediction is, so we will use 25% as the test set, and all the rest of the data is used to predict the training of the model.

The first step is to consolidate the date-time information into a single date-time so that we can use it as an index in Pandas. There are also a few scattered "NA" values later in the dataset; we can mark them with 0 values for now.

The second step is to prepare the dataset for the LSTM. This involves framing the dataset as a supervised learning problem and normalizing the input variables. We will frame the supervised learning problem as predicting the SOC

at the current hour (t) given the external parameters of the battery at the prior time step. Next, all features are normalized, then the dataset is transformed into a supervised learning problem.

In the third step, we will fit an LSTM on the multivariate input data. First, we must split the prepared dataset into train and test sets. To speed up the training of the model for this demonstration, we will only fit the model on the 75% of data, then evaluate it on the remaining 25% of data. We will define the LSTM with 50 neurons in the first hidden layer and 1 neuron in the output layer for predicting pollution. The input shape will be 1-time step with 5 features. We will use the Mean Squared Error (MSE) loss function and the efficient Adam version of stochastic gradient descent. The model will be fit for 50 training epochs with a batch size of 128.

After the model is fit, we can forecast for the entire test dataset. We combine the forecast with the test dataset and invert the scaling. We also invert scaling on the test dataset with the expected SOC. With forecasts and actual values in their original scale, we can then calculate an error score for the model. In this case, we calculate the Mean Absolute Percentage Error (MAPE) that gives an error in the same units as the variable itself.

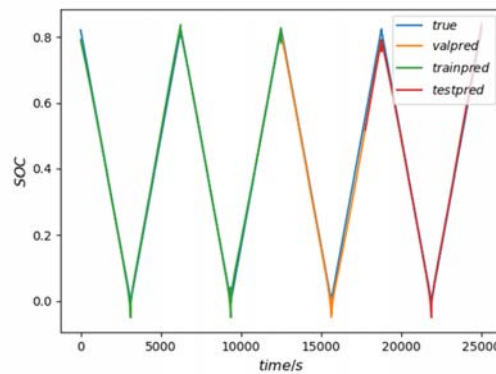


FIG 4. Comparison between real and predicted values

Figure 4 shows the comparison between the predicted value and the actual value after the model training is completed. The blue line is the lithium-ion battery SOC true value, the green line is the training set predicted value, the orange line is the verification set predicted value, and the red line is the test set predicted value. From Figure 3 we can see that the model performs well in predicting SOC and the generalization ability has also been verified.

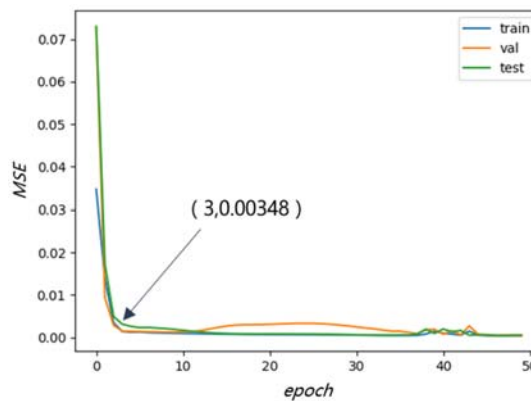


FIG 5. Model loss analysis

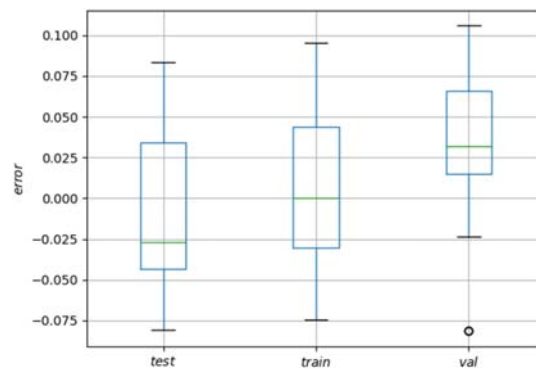


FIG 6. Model error rate distribution

Figure 5 shows the estimated SOC model loss analysis. The blue line is the training set error trend, the orange line is the validation set error trend, and the green line is the test set error trend. From Figure 5 we can see that the error value of the LSTM-based SOC estimation model has stabilized at the iteration number of 3. It proving that the model converges faster. At this time, the error of the test set was 0.00348, which verifies the accuracy of the model. The training error is lower than the test error, which proves that the model does not overfit the training data and the overall performance is good.

Figure 6 is a box plot of the predicted and true error rates on training set, verification set, and test set. A boxplot is a method that uses statistics defined by data to describe the distribution of the dataset. The main purpose is to reflect symmetry or dispersion of the data. From Fig.6, it can be concluded that SOC estimation error rate of the test set is mainly distributed in the range of -0.041 to 0.035, and stays in a lower range, which verifies the accuracy of the estimated SOC value of the model.

CONCLUSION

This paper describes in detail the time-varying dis-charge characteristics of the power battery in the electric vehicle. It introduces the improved RNN algorithm, which is the long-term memory network LSTM and establishes the LSTM-based SOC estimation model. This section describes in detail the optimization setting and processing of the LSTM network structure to deal with timing issues. Next, introducing the sliding window algorithm to handle the battery experimental data obtained in the previous chapter is more conducive to the training of the SOC estimation model. The experimental data was simulated in the following section to verify the accuracy of the SOC-based battery SOC estimation model established in this paper.

REFERENCES

1. Zhou Y, Li X. Overview of lithium-ion battery SOC estimation[C]// IEEE International Conference on Information and Automation. IEEE, 2015:2454-2459.
2. Zhang Xiao qiang, Wang Xiao cheng, Zhang Wei ping, et al. A Simplified Li-ion Battery SOC Estimating Method[J]. Transactions on Electrical & Electronic Materials, 2016, 17(1):13-17.
3. Baoen L I, Xianwei L I, Bingxin W U, et al. Re-search on energy optimal dispatching strategy for mi-crogrid based on battery SOC[J]. Power System Protection & Control, 2017, 45(11):108-114.
4. Liu C, Liu W, Wang L, et al. A new method of modeling and state of charge estimation of the battery[J]. Journal of Power Sources, 2016, 320:1-12.
5. Ranjbar A H, Banaei A, Khoobroo A, et al. Online Estimation of State of Charge in Li-Ion Batteries Using Impulse Response Concept[J]. IEEE Transactions on Smart Grid, 2012, 3(1):360-367.
6. Wei Z, Lim T M, Skyllas-Kazacos M, et al. Online state of charge and model parameter co-estimation based on a novel multi-timescale estimator for vana-dium redox flow battery[J]. Applied Energy, 2016, 172:169-179.

7. Baccouche I, Jemmali S, Manai B, et al. Improved OCV Model of a Li-Ion NMC Battery for Online SOC Estimation Using the Extended Kalman Filter[J]. *Energies*, 2017, 10(6):764.
8. Zhang X, Wang X, Zhang W, et al. A simplified Li-ion battery SOC estimating method[J]. *Transactions on Electrical & Electronic Materials*, 2016, 17(1):13-17.
9. Zahid T, Qin F, Li W, et al. Sequential Monte Carlo based technique for SOC estimation of LiFePO₄ battery pack for electric vehicles[C]// *IEEE International Conference on Information and Automation*. IEEE, 2017:1308-1312.