1. INTRODUCTION

Equipment failure will bring great security risks, the RUL (Remaining Useful Life) expectancy of the device is very necessary. At present, there are mainly three types of prediction models for RUL, one is based on empirical knowledge, the other is based on data-driven approach and the third is based on the mechanism model approach. Knowledge-based approaches include expert system approach and fuzzy-based approach. Because these methods are too dependent on the prior knowledge of experts, they lead to poor generalization ability. The methods based on mechanism model mainly include Markov-based method, Kalman-based method, and particle filter-based method. Due to the complexity of current devices, it is difficult to establish an accurate physical model. The data-driven method is currently the hotspot of research, including statistical analysis methods and artificial intelligence methods. Statistical learning methods mainly include trend extrapolation and AR methods. Artificial intelligence methods mainly include ANN, SVM and deep learning methods. The method of statistical learning recursively after multi-step to predict the remaining life, it can not guarantee real-time, while the shallow artificial intelligence method has the problem of feature extraction is not accurate, so the deep learning based method is one of the most popular method about RUL prediction.

The traditional deep learning methods include stacking auto-encoder method, which can extract more accurate features. However, since the feature of RUL prediction is obviously time-dependent, the AE method can not extract the time-dependent features of time series data , So people put forward the recurrent neural network model (RNN), RNN can establish the time-dependent relationship between data, but with the development of time, the latter nodes perceived ability of the previous node decreased, can not remember the information for a long time, This is the problem of the gradient disappearing. In view of the above problems, long and short term memory neural network (LSTM) redesigns computing nodes based on keeping the structure of RNN so that it can remember the information before long time, which makes it have a significant effect in speech recognition and other fields. However, the application of LSTM in equipment RUL prediction is very rare, especially the research of RUL prediction on equipment that plays a very crucial role in important fields is even less. For example, the key components of the spacecraft, due to the difficulty of sampling to the full life cycle of the data, so how to use these precious non-life-cycle data, and how to use LSTM advantage of such data to establish an effective RUL prediction model is very Important issues.

In this paper, we aim at optimizing the network parameters, and propose a real-time updating LSTM prediction model based on minimized cost parameters. The RUL predictive model can be established without full life-cycle data of the equipment. With the increase of equipment service time, the established model parameters are gradually optimized by using the collected data. By comparing with many typical time series prediction models, it is verified that the proposed real-time updated LSTM model has strong applicability and higher accuracy in non-full life-cycle time series prediction.

The remaining part of this paper is organized as follows: Section 2 introduces LSTM theories and techniques; Section 3 introduces the proposed LSTM prediction model and its parameters real-time update method; Section 4 presents the experimental results and analysis of their results; Sections 5 conclude and look forward to future work.

The remaining part of this paper is organized as follows: Section 2 introduces LSTM theories and techniques; Section 3 introduces the proposed LSTM prediction model and its parameters real-time update method; Section 4 presents the experimental results and analysis of their results; Sections 5 conclude and look forward to future work.

**LSTM Theory and Prediction Methods**

General neural network can not solve the time-dependent problem, and RNN is an improved multi-layer neural network, which consists of input layer, hidden layer and output layer. The RNN can be thought of as multiple replications of the same neural network, with each neural network module passing the message to the next. When it is expanded in time, as shown in Fig. 1, it can be seen that RNN is a chained structure, and chained features indicate that it is related to sequences and lists. The output of the current moment is related not only to the input at the current time but also to the output at the previous moment, which allows the previous information to be passed back, which is why RNN can solve the time-related problem. The training of parameter of RNN is realized by BP Natural Netwrok(BPNN). However, one of the most obvious problems of RNN is that the gradient of training needs to be back propagated with time. When the input sequence is relatively long, there will be gradient disappearance and gradient explosion problem.



Fig.1 Unfolding of single RNN hidden layer unit

LSTM can solve the above problems in RNN by introducing a group of memory cells, allowing the network to learn when to forget the historical information and when to update the memory cells with new information. LSTM and RNN have the same chained structure, except that the hidden layer in the standard RNN has only a simple tanh layer, but in LSTM it is four interactive layers, and the structure is as shown in Fig. 2.



Fig.2 Structure of LSTM unit

First, the state of the cell is the most critical component of the LSTM. The state of the cells resembles the carousel, and information can circulate throughout the chain. Second, LSTM has the ability to remove or increase information to the state of the cell due to its well-designed gate structure. The door is a way of selectively passing information. The key to the gate structure is a sigmoid neural network layer and a pointwise multiply operation. The output of the sigmoid layer is between 0 and 1, 0 means all do not pass and 1 means all pass. LSTM has three doors, namely forgotten door, input door, output door, use these three doors to protect and control the cell status.

LSTM network training process is described as follows:

**Step 1**, training data preprocessing (normalization and other operations).

**Step 2**, initialize the weights and offsets of the network, and the hyperparagraphs of the network.

**Step 3**: Calculate the LSTM forward operation process shown in formula 5 to obtain the LSTM prediction value.

**Step 4**, find the error between the predicted value and the actual value.

**Step 5**, to determine whether to reach the error threshold or the maximum number of iterations, if reached, exit the training, or use BPTT algorithm to update the network parameters, and then go to step three cycles, until the conditions are met, exit the loop. LSTM workflow shown in Figure 3.



Fig.3 The training process of LSTM

**LSTM prediction model with real-time updating of parameters**

Based on practical problems, this paper gives a method to establish an effective RUL prediction model without full life-cycle samples. First of all, the LSTM prediction model should be established by making use of the known historical data reasonably. Secondly, when the actual online data comes, the predicted value can be obtained by using the established prediction model. The new data of the next time is used as the true value of the previous time. The error between the predicted value and the real value is added to the overall sample error. Then, the model parameters are iteratively updated by using the error minimization method. By updating the parameters in this way, the model can be more and more accurate over time. This modeling idea is more conducive to the actual system of online monitoring, as well as failure warning. When the predicted value reaches the point of failure to issue a fault warning signal.

The first step in LSTM is to decide what to discard from the state of the cell. This is done by forgetting the gate, which reads theand and then outputs a value between 0 and 1 after sigmoid, each of the numbers in the cell state is multiplied by this number so that the purpose of selectively discarding information from the cell state is achieved. The second step is to determine what information can be stored in the cell state, including two parts here. The first is the sigmoid entry gate that determines what value we are going to update. The second is to use the tanh layer to create a new Candidate value. The process is to confirm the updated information. The third step is to update the old cell state, that is,  is updated to, first of all, the old stateis multiplied by the output of the forgotten gate, discarding the unwanted information, then add the new candidate value. The final step is to determine what value to output, first through a sigmoid layer (output gate) to determine which parts of the cell state will be output. Then, the cell state is processed by tanh and multiplied by the output of the output gate to determine which parts are output. In general, LSTM is updated as shown in formula1 at time t.

 (1)

Assuming the actual time series is, the improved LSTM steps are as follows:

**Step1:**Expand  to  ,Where n is the time series length, k is the sample dimension, the number of samples is n-k+1, the training data is labeled, and X normalized, E.g: .

**Step2:** Initialize network parameters, for example, the weight and biasof forget door. Similarly, there are input doors, output doors and so on. And set the error threshold Error\_Cost, the maximum number of iterations Max\_iter and other parameters.

**Step3:** Calculate what information needs to be forgotten about the state of the cell. First calculate the output of the forgotten door, Then multiply the forgotten door's output by the cell state of the previous moment.

**Step4:** Calculate which input information can stay in the cell state at time t. This includes two parts, one is the output of the input door that determines which values we will update. The second is to use the tanh layer to create a new candidate phasorthen Multiplies the output of the input gate by the candidate phasor.

**Step5:** Calculate the cell state ,The state of the cell is the result of a combination of the forgotten and the input gates on the state of the cell

**Step6:** Calculate the network output at time t. First calculate the output of the output gate The output of the current moment is obtained by multiplying the output of the output gate by the cell state of the current moment.is the current moment of the predictive value. Repeat steps 3 to 6 to calculate the predicted values for all training samples.

**Step7:** Calculate the error between the predicted and the true values of all samples, If error <Error\_Cost or the current number of iterations iter> Max\_iter, then the end of training, Otherwise use BPTT algorithm to update the network parameters, the number of iterations plus one, and then go to step three cycles until it reaches the error threshold or the maximum number of iterations, exit the loop. Save the trained network parameters

**Step8:** New samplewith Perform the forward operation of LSTM shown in step 3-6,and obtain the output predicted value of the new sample, When collected next data   is the true value of predicted value . Calculate the total error. Then the BPTT algorithm is used to update the model parameters to . A is the learning rate,  and  are the gradient matrixes and vectors of the neuron's weights and offsets respectively. Since the parameter initialization is the global optimal solution of the historical sample when adding a new sample, the global optimal solution under the new sample can be achieved with a few simple steps.

**Step9:** When the predicted value at the next moment reaches the point of failure, the system will initiate a fault warning and react promptly to prevent any further loss or even a safety accident.

Improved LSTM flowchart shown in Fig. 4.



Fig.4 The training process of improved LSTM

**Experiment Analysis**

In this paper, partial ground experimental data of NASA lithium-ion battery is used, and different prediction algorithms and LSTM algorithm are adopted to predict the cycle of charge-discharge cycle when the capacity of lithium-ion battery declines to 70%, in order to achieve the life expectancy of lithium-ion battery.

LSTM cost function value with the number of iterations as shown in Figure 4, the left is the cost function curve with a fixed learning rate of 0.01. the middle one is the cost function curve of time-varying learning rate, the right is cost function of the method actually used in this article.



Fig.5 Training error of various methods

The experimental results are shown in Figure 5. The solid line is the actual SOH curve, the red dashed line is the SOH curve predicted by AR, the blue dashed line is predicted by the neural network, the yellow dashed line is predicted by LSTM, and the black dashed line is predicted by the improved LSTM .forecasts show that all the forecasting methods work well with historical data, but the predictions of several traditional methods are not good when they do not have historical data, but this article The proposed method can be a good solution to this situation, the experimental results shown in Figure 5, the black dotted line. It can be clearly seen that all kinds of prediction methods work well with historical data, but the prediction performance of several traditional methods is not good when there is no historical data, but the method proposed in this paper works well. The experimental results shown in Figure 6,and mean square error shown in Figure 7.



Fig.6 Battery health prognostics based on various methods model



Fig.7 MSE of Various Methods

**Conclusion**

Aiming at the problem that the traditional LSTM model can not effectively use the non-life-cycle data to establish the excellent RUL prediction model and the inability to utilize the online data reasonably, this paper proposes a real-time updating method based on minimizing the error to solve the problem of online modeling and model correction of small sample data. Taking NASA lithium-ion battery data as an example, the applicability of the proposed LSTM model parameters real-time updating method in the field of reliability prediction is verified. At the same time, it is proved that this method can be effectively used in lithium ion battery life prediction and has high practical value in engineering application. Based on the current work, follow-up we can conduct further research to find more effective LSTM parameter optimization methods. In addition, in the future, non-numerical data such as image data can also be fused together to make the extracted features more comprehensive and accurate to establish the prediction model.