The State of Charge Estimation of Li-Ion Batteries in Electric Vehicle Based on RBF Neural Network Optimized by Genetic Algorithm

Jun Bi, Qiuping Xu, Kai Wang, Dong Zhang

School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China, MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology, Beijing Jiaotong University, Beijing, China,

Abstract

In order to guarantee security and stable operation of electric vehicle, it is necessary to on-line estimation for the state of charge (SOC) of batteries. The power battery is a complex nonlinear system, and Radial Basis Function Neural Network (RBF NN) has advantages in solving nonlinear problems, so the model of on-line SOC estimation based on RBF NN is proposed. In order to improve the prediction accuracy of SOC, genetic algorithm is used to optimize the model, which can make global optimization search for the center and spread of each neuron in hidden layer of the RBF NN to get the most optimal value. The experiments are based on the battery data achieved from the pure electric buses with LiFePO4 Li-ion batteries running during the period of 2010 Shanghai World Expo. The results show that compared with RBF NN, RBF NN optimized by genetic algorithm significantly can improve the prediction accuracy of SOC.

Keywords: SOC Estimation, Battery, Electric Vehicle, RBF Neural Network, Genetic Algorithm

1. Introduction

Now many countries are increasingly concerned with the oil energy conservation and environmental protection. The power source of the pure electric vehicle are the batteries instead of the fuel oil, so the electric vehicle has the advantages of zero gas emission, less environmental pollution and energy saving. The electric vehicle with Li-ion batteries is being widely researched. The state of charge (SOC) is the most important parameter of batteries, which can be used to estimate the cell’s performance, avoid over-charge or over-discharge and predict the driving distance of electric vehicles. So it is significant to estimate SOC when the vehicle is in operation [1][2].

At present, the conventional methods for SOC estimation are chemistry-dependent methods, electrochemical modeling [3], circuit models [4], open-circuit voltage (OCV) measurements [5], impedance spectroscopy [6], Kalman filter algorithm and artificial neural network [7].Comparatively, because the nonlinear characteristics are existed in the batteries, while RBF NN has the advantage of describing the nonlinear characteristics, RBF can be used to predict SOC.

In this paper, an on-line practical estimation method for SOC of lithium batteries in electric vehicle based on RBF NN is proposed. Firstly, the model of on-line SOC estimation with RBF NN is set. Secondly, in order to improve the prediction accuracy of SOC, genetic algorithm is used to optimize the model, which can make global optimization search for the center and spread of each neuron in hidden layer to get the most optimal value. Finally, the pure electric buses with LiFePO4 Li-ion batteries running during the period of 2010 Shanghai World Expo are considered as the experimental object. All experimental data of batteries are achieved from the practical process of electric bus in operation instead of laboratory test.

2. Background of battery data acquisition

During the period of 2010 Shanghai World Expo exhibition, 120 pure electric buses were granted to carry the visitors by Chinese government. Each bus owns 104 LiFePO4 Li-ion batteries which are existed in the 10 battery packs in series. In order to ensure the safety of electric bus in operation, it is necessary to monitor the status of all batteries in the remote monitoring center. The equipment installed in the electric bus can sample the battery data such as voltage, current, temperature and transmit the...
data to the monitoring center by GPRS wireless technology according to the transmission period of 20s. So a large number of battery data from each electric bus in operation are stored in the monitoring center’s server computer, which provides the data basis for research of SOC estimation.

3. SOC estimation model based on RBF neural network

3.1. Structure of RBF neural network

Radial Basis Function Neural Network (RBF NN) [8] is a typical feed-forward neural network, which has many merits, such as simple structure, training fast, nonlinear mapping characteristic, self-organized study ability, capable of converging to global optimization and approaching function best. At the same time, it can be widely used in pattern recognition and other areas of nonlinear function approximation. The structure of the RBF NN is composed of three layers such as one input layer, one hidden layer and one output layer. The RBF neural network designed for SOC estimation is shown in Fig.1.

![Fig.1. Structure of RBF NN for SOC estimation](image)

3.2 Design of RBF neural network input layer

The network input signals $x_i (i=1,2,\ldots,n)$ indicate the influence factors of SOC, where $n$ is the number of input neurons. There are many factors influencing the SOC. By using contribution analysis method [9], 4 factors are determined as the input variables, such as SOC at the last time, total current, voltage and average temperature. The network input variables may be derived from a set with 4 factors as follows:

$$X = (x_{LastSoc}, TotalI, TotalV, AvgT)$$

Where $LastSoc$ is the SOC of all battery packs in series at the last time. $TotalI$ is the total current of all battery packs in series at the current time. $TotalV$ is the total terminal voltage of all battery packs in series at the current time. $AvgT$ is the average temperature of all battery packs at the current time.

3.3 Design of RBF neural network hidden layer

The number of hidden layer is $m$, and the output $h_j(X)$ of the $j$th neuron in the hidden layer is as follows:

$$h_j(X) = \varphi \left( \|X - C_j\| \right)$$

(1)

$\varphi(\cdot)$ is adopted to be the Gaussian function. $h_j(X)$ is described by Eq.(2):
where $X$ is the input matrix of network, $\varphi(\cdot)$ is the radial basis function, $C_j (C_j \in R^n)$ is the center of the $j$th hidden neuron, $\| \cdot \|$ is the Euclidean norm, and $\sigma_j$ is the spread of the $j$th neuron for the $i$th input signal.

### 3.4 Design of RBF neural network output layer

The output layer only has one node $Y(X)$ which indicates the SOC estimation. $Y(X)$ is the linearly combined signal of the outputs $h_j(X)$ from the hidden layer with the synaptic weights $w_j$, as follows:

$$Y(X) = \sum_{j=1}^{m} w_j h_j(X)$$

where $m$ is the number of hidden neurons.

### 4. SOC estimation algorithm based on RBF neural network optimized by genetic algorithm

The K-means clustering algorithm of RBF NN is sensitive to the initial parameters, that the different initial value may cause different clustering results and local optimal solution. The genetic algorithm is not sensitive to the initial parameters and cannot easily fall into the local minimum, so it is appropriate to find the optimum solution in the complex network structure [10]. Based on the above analysis, it is efficient to make the value of error function about the center, spread and weights tend to be a minimum by using genetic algorithm to optimize the RBF NN and select the center and spread of each neuron in hidden layer to get the optimal value [11][12]. The improved model algorithm is described as follows:

**Step1**: Normalize and standardize the samples.

**Step2**: Determine the input variables. The contribution analysis method is studied to determine the important input variables according to contribution factor of each input variable to the output variable of the RBF network. Four factors are determined as the neural network input variables, such as SOC at the last time, total current, voltage and average temperature.

**Step3**: Encode the chromosome. Calculate the encoding length based on the range and computational accuracy of the center and spread of the RBF NN, and encode the center and spread by using of binary numbers. Where, the encoding length of the central value is $\text{HiddenNum} \times \text{InputNum}$, and the encoding length of the spread is $\text{HiddenNum}$, and their values are all real number string ranging from 0.1 to 1. $\text{HiddenNum}$ indicates the number of neuron in hidden layer, and $\text{InputNum}$ indicates the number of input variables.

**Step4**: Initialize population. Generate the initial population with random method and set the evolution algebra of the population.

**Step5**: Decode the chromosomes. Calculate the center and spread of each neuron in hidden layer, and figure out the output of hidden layer by Gaussian function. The weights between hidden layer and output layer are adjusted by the least square method.

**Step6**: Calculate the error of each individual. The performance indicator of the training sample data
in the structure of the RBF NN is expressed:

\[
E(w) = \frac{1}{2} \sum_{i=1}^{N} \left( y_i - f(x_i) \right)^2
\]  

(4)

The purpose of RBF NN training is to make error E tend to the minimum, depending on each set of samples. If E is less than the specified threshold value, the algorithm is stopped, otherwise go to the next step.

Step7: Select the fitness function. It is relatively easy for the individual with high adaptability inherited to the next generation. The reciprocal of the square of the error value between the network output value and the expected value is used as the fitness function, taking into account the accuracy of the neural network. Calculate and sort the fitness value of each individual in the population. The fitness function of the ith individual is expressed in Eq.(5).

\[
f(i) = \frac{1}{e^2(i)} = \frac{1}{\sum_{i=1}^{N} (y_i^* - y_i)^2}
\]

(5)

where, N indicates the number of neuron in hidden layer, \( y_i^* \) is the expected output of the ith input sample, \( y_i \) is the network output of the ith input sample.

When the evolution algebra of the population is over, or there is no improvement after several times, the algorithm is stopped, otherwise goes to the next step.

Step8: Genetic operation

(1) Selecting operation. This algorithm uses the roulette selection method, in which the probability of the individual with high fitness selected into the next generation will be bigger. Assume \( f(X_i) \) is the fitness of \( X_i \), the probability selected into the next generation is,

\[
P = \frac{f(X_i)}{\sum_{j=1}^{n} f(X_j)}
\]

(6)

where, \( \sum_{j=1}^{n} f(X_j) \) is the sum of individual fitness of the population.

(2) Crossover operation. Genetic algorithm improves its search capabilities by cross-operating. In this paper, we adopt two-point crossover, which can be sketched as follows:

\[
\begin{align*}
X_1 : & \text{aaa|aaa|aaa} \quad X_1 : \text{aaa|b|b|b|a|a} \\
X_2 : & \text{b|b|b|b|b|b} \quad X_2 : \text{b|b|a|a|a|b|b}
\end{align*}
\]

\[
\text{junction1} \quad \text{junction2}
\]

Crossover probability should not be too small, otherwise it will lead to the stagnation of search process. Generally, the range is between [0.5, 1].

(3) Mutation operation. The mutation operator can change all individual genetic bits randomly, which can guarantee the genetic algorithm with global convergence, increase the diversity of the population, and avoid training process into the local minimum value. When mutation probability is
over large, the genetic algorithm will be a random search, so in general the value is between [0.01, 0.2].

The next generation of population is generated by above three basic operations of genetic algorithm, after that go to Step5.

5. Experimental validation

5.1. Parameters identification

The number of 4300 batteries data sampled from No.1 pure electric bus running during the period of 10th to 20th July 2010 are used to train RBF NN. The mean square error (MSE) expressed in Eq.(7) is used as the target error of network training.

\[ \text{MSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2} \]

(7)

where \( \hat{Y}_i \) is the true value, \( Y_i \) is the estimation value of RBF NN, \( n \) is the number of training times. When MSE is below 0.001, the network training is stopped.

After repeated experiment, in the RBF NN model optimized by genetic algorithm, the size of population is 30, the evolution algebra is 200, the computational accuracy is 0.0001, the range of the center and spread of each neuron in hidden layer is [0.1, 1], the crossover probability is 0.7, the mutation probability is 0.005, the binary encoding length of each parameter is 15.

5.2. Experimental results and analysis

The batteries data of No.1 pure electric bus respectively running on 7th and 17th August 2010 are used to verify SOC estimation of RBF NN by the K-means clustering algorithm and the algorithm optimized by genetic algorithm. The number of hidden layer neuron is set respectively to be 30, 50, 70. The MSE of both learning algorithm is shown in Table 1 and Table 2.

| Table 1. Comparison of the MSE of both learning algorithm on 7th August 2010 |
|--------------------------|---|---|---|
| The number of hidden layer neuron | 30 | 50 | 70 |
| Learning algorithm | K-means clustering algorithm | 0.0053 | 0.0036 | 0.0027 |
| | Algorithm optimized by genetic algorithm | 0.0036 | 0.0023 | 0.0020 |

| Table 2. Comparison of the MSE of both learning algorithm on 17th August 2010 |
|--------------------------|---|---|---|
| The number of hidden layer neuron | 30 | 50 | 70 |
| Learning algorithm | K-means clustering algorithm | 0.0067 | 0.0041 | 0.0034 |
| | Algorithm optimized by genetic algorithm | 0.0038 | 0.0024 | 0.0019 |

The results show that the more hidden layer neuron, the smaller the MSE predicting the SOC of batteries, and the calculation speed is significantly slower. So we set the number of hidden layer neuron as 70 to estimate the SOC with both learning algorithm. The estimation results are shown in Fig.2-Fig.5.

The prediction accuracy of the algorithm optimized by genetic algorithm is significantly better than
that of K-means clustering algorithm, and the error of SOC estimation is below 0.01, which shows the optimized algorithm has good SOC estimation performance.

Fig.2. SOC estimation of batteries on 7th August 2010

Fig.3. The error of SOC estimation on 7th August 2010

Fig.4. SOC estimation of batteries on 17th August 2010
6. Conclusion

In this paper, the model of on-line SOC estimation based on RBF NN is proposed and optimized by genetic algorithm, which can make global optimization search for the center and spread of each neuron in hidden layer. The network training and SOC estimation are conducted by using the data of LiFePO4 Li-ion batteries in the electric bus in operation during the period of 2010 Shanghai World Expo. The practical results show the optimized algorithm has good SOC estimation performance. However, in this paper, we only establish the model based on the battery parameters which can be collected, without considering the degree of aging and ambient temperature of batteries. At the further research, we should consider the impact of these factors on the SOC, after the long-term operation of batteries.

7. Acknowledgment

This work was supported by “the Fundamental Research Funds for the Central Universities” and Beijing Municipal Science & Technology Project (No. Z111109073511001)

8. References

The State of Charge Estimation of Li-Ion Batteries in Electric Vehicle Based on RBF Neural Network Optimized by Genetic Algorithm
Jun Bi, Qiuping Xu, Kai Wang, Dong Zhang


