Journal of Power Sources 281 (2015) 131–137

Contents lists available at ScienceDirect

Journal of Power Sources

journal homepage: www.elsevier.com/locate/jpowsour

Electric vehicle state of charge estimation: Nonlinear correlation and fuzzy support vector machine

Hanmin Sheng*, Jian Xiao

School of Electrical Engineering, Southwest Jiaotong University, Chengdu 610031, China

HIGHLIGHTS

- We propose a measure of nonlinear correlation.
- We propose a reliability identification method for multidimensional samples.
- We propose an error interval for SVM regression.
- The daily use of electric vehicle is simulated in a filed test.

ARTICLE INFO

Article history:
Received 27 October 2014
Received in revised form
22 January 2015
Accepted 23 January 2015
Available online 28 January 2015

Keywords:
Lithium battery
Electric vehicle
State of charge
Nonlinear correlation
Support vector machine

ABSTRACT

The aim of this study is to estimate the state of charge (SOC) of the lithium iron phosphate (LiFePO4) battery pack by applying machine learning strategy. To reduce the noise sensitive issue of common machine learning strategies, a kind of SOC estimation method based on fuzzy least square support vector machine is proposed. By applying fuzzy inference and nonlinear correlation measurement, the effects of the samples with low confidence can be reduced. Further, a new approach for determining the error interval of regression results is proposed to avoid the control system malfunction. Tests are carried out on modified COMS electric vehicles, with two battery packs each consists of 24 50 Ah LiFePO4 batteries. The effectiveness of the method is proven by the test and the comparison with other popular methods.

Published by Elsevier B.V.

1. Introduction

Due to the ever increasing concern about the environmental problems and energy crisis, electric vehicles (EVs) have been regarded as the future of the automobile and are highly sought after by investors. But because the energy density of commonly used electric batteries are far lower than the fuel like gasoline or diesel, the mileage of EVs has always been a problem [1,2]. To obtain higher mileage, traction batteries must be designed with a high ampere-hour capacity. Lithium battery compared to the currently widely used lead-acid battery is with higher energy density, which means the units can be smaller and lighter [3]. Lithium batteries have higher charge and discharge rate, thus is more effective in energy feedback process. These years, the technology of lithium battery is developing rapidly, more and more new lithium batteries with higher energy density like the LiFePO4 batteries, lithium Cobalt Acid (LCO) Batteries etc. have been applied in EVs. The high energy density is followed by a series of detrimental situations [4]. Up to now, cases of lithium battery on electric vehicles spontaneous combustion accidents happened occasionally [5]. Boeing 787 dreamliner battery problems deepened the safety suspicion of lithium-ion batteries.

Subject to the influence of redox potential, the nominal voltage of lithium battery is relatively low. To improve the energy conversion efficiency, battery cells should be connected in series to obtain higher output voltage. Under the current manufacturing process, it is difficult to guarantee the uniformity of different Li-ion Batteries. The over-charge and over-discharge of battery cell caused by battery energy mismanagement is common. Which will dramatically reduce the battery cycle life. SOC is the key parameters in EVs energy management system, it is shows the ratio of remaining capacity over the nominal capacity [6]. Precise measurements of SOC are necessary to ensure the safe operation meanwhile maximizing the use of battery capacity.

* Corresponding author.
E-mail addresses: shenghanmin@hotmail.com (H. Sheng), jxiao@swjtu.edu.cn (J. Xiao).

http://dx.doi.org/10.1016/j.jpowsour.2015.01.145
0378-7753/Published by Elsevier B.V.

In the early application, people found that the battery SOC has a strong linear relationship with the open circuit voltage (OCV). To measure the accurate OCV, the battery need hours to rest. While in most conditions, accurate OCV is unlikely to obtain. Usually is substitute by mathematically transformed online data. But due to the various internal and external conditions [7], fixed mathematical transformation cannot be accurate. The coulomb counting method (CCM) is another strait forward method, has the advantage of simple computation and easy implementation. The CCM actually has been widely used in BMS systems. However, CCM is calculated by charge and discharge current time integral, unable to eliminate cumulative error and is sensitive to the initial value. If the initial SOC value is inaccurate, it will affect all estimates and the error will be accumulated during the whole estimation process. Researchers discovered that the internal resistance has a strong correlation with SOC, and accordingly proposed resistance acquaintance method (IRM) by exciting the battery using current with different frequency. Internal resistance method do not accumulate error, but the relationship between the internal resistance and SOC is affected by many factors, it is hard for us to build the model for each battery cell. Kalman filtering method has achieved great success in pattern recognition. The performance of Kalman Filter subject to the accuracy of the model and initial value.

SOC is affected by multiple factors directly or indirectly. For Li-ion battery itself, migration of the electrolyte and the activity of electrode materials is affected by the temperature, hysteresis, cell age and self-discharge rate [4]. For EVs, the complex terrain, road conditions brought the complicate power output and energy feedback has further increased the estimation difficulty. Besides, in order to reduce the cost and volume, on-board measuring equipment compared with the measurement equipment in the laboratory contains more noise. These factors result in highly nonlinear and complex coupling of SOC.

These unavoidable factors compel us to explore a new approach with no accumulate error and consider the effect of multiple factors. For the past few years, intelligent method like Artificial Neural Network (ANN) has been proved can approximate any nonlinear function [8] in arbitrary precision. But the generalization performance of ANN is unsatisfactory, susceptible to local minimum points especially in particular in small sample issue. Vapnik V. proposed Support Vector Machine (SVM) theory in 1995, it is a kind of supervised learning method based on structure risk minimization risk, has good generalization ability. Least Square Support Vector Machine (LS-SVM) replace the inequality constrain in SVM by equality constrain, avoid the continuous iterative and precision decline. But no matter ANN, SVM or LS-SVM is sensitive to noise and isolated samples. Currently common practice is to increase the number of training samples to decrease the affection of error from the validation set, but do not help to decrease the affection come from the training samples. Researchers tried to distinguish the high noise level samples with relatively accurate samples. Support Vector Data Description (SVDD) is a well-known method developed by Tax and Duin [9,10] that map the data into high dimensional feature space. Researchers found that to find a minimum-volume hypersphere in the feature space could enclose most of the data with low noise level [11]. And considered to use the distance between mapped sample in the feature space with the center of data description as the reference standard for data reliability [12]. SVDD has been widely studied in outlier detection [13], fault diagnosis [14], clustering [15], prediction [16] and other fields. But in the case that the data description is not spherical, it will filter useful data in SVM regression. Besides, the application of SVDD needs to solve quadratic programming problems, the efficiency will decline sharply with the increase of the samples. Li suppose the distribution of outlier should be a Gaussian process, and introduced a method to determine the distribution of the outlier degree by comparing the residual distribution. However, this method lacks a theoretical basis [17]. In this paper, proposed a SOC estimation method optimized by outlier identification method and fuzzy logic. This method has the advantage of wide applicability and low calculation complexity.

1.1. Weighted support vector machine

Suppose we are given a training data set,

$$\{(\mathbf{x}_1, y_1, \mu_1), (\mathbf{x}_2, y_2, \mu_2), \ldots, (\mathbf{x}_n, y_n, \mu_n)\} \tag{1.1}$$

Where n is the number of training samples,$\mathbf{x}$ is the input vector contains the voltage, current and temperature parameters, and the μ_i indicates the outlier degree of the sample. The target function is

$$f(\boldsymbol{x}) = \boldsymbol{\omega}^T \varphi(\boldsymbol{x}) + b, \quad \boldsymbol{\omega} \in \boldsymbol{R}^n \; b \in \boldsymbol{R} \tag{1.2}$$

$\varphi(\boldsymbol{x})$ is the mapping from original space to the feature space. Coefficentω and deviation$b$ can be found by solving the following optimization problem

$$\min J(\omega, e) = \frac{1}{2}\omega^2 + \gamma \frac{1}{2l} \sum_{i}^{l} \mu_i e_i^2$$

$$s.t. \;\; y_i = \boldsymbol{\omega}^T \varphi(\boldsymbol{x}_i) + b + e_i, \quad i = 1, \ldots, n \tag{1.3}$$

where$J$ is the loss function , indicates the structure risk, $e_i \in \boldsymbol{R}$ is residual. Penalty function $\gamma \in \boldsymbol{R}^+$. Usually the dimension of ω is high , makes the calculation in original space difficult. By solving the dual problem, calculation can be greatly simplified. The Lagrange function of the dual problem can be expressed as

$$\begin{aligned} &L(\boldsymbol{\omega}, b, \boldsymbol{e}, \alpha) = J(\boldsymbol{\omega}, \boldsymbol{e}) - \sum_{i=1}^{n} \alpha_i \left[\boldsymbol{\omega}^T \varphi(\boldsymbol{x}_i) + b + e_i - y_i\right] \\ &\alpha_i \geq 0, \quad i = 1, 2, \ldots, l \end{aligned} \tag{1.4}$$

α_i is the Lagrange multiplier. According to Karush–Kuhn–Tucker Theorem [18], the optimal value satisfy

$$\begin{cases} \dfrac{\partial L}{\partial \boldsymbol{\omega}} = 0 \rightarrow \boldsymbol{\omega} = \displaystyle\sum_{i=1}^{l} \alpha_i \varphi(\mathbf{x_i}) \\ \dfrac{\partial L}{\partial b} = 0 \rightarrow \displaystyle\sum_{i=1}^{l} \alpha_i = 0 \\ \dfrac{\partial L}{\partial e_i} = 0 \rightarrow \alpha_i = \dfrac{\gamma}{l} \mu_i e_i \\ \dfrac{\partial L}{\partial \alpha_i} = 0 \rightarrow \omega^T \varphi(\mathbf{x_i}) + b + e_i - y_i = 0 \end{cases} \tag{1.5}$$

The optimizedω,b,$\boldsymbol{e}$,α can be solved separately, we get

$$\begin{bmatrix} 0 & \mathbf{1}_{n\times n}^T \\ \mathbf{1}_{n\times n} & \varphi(x_i)\varphi(x_j) + n\boldsymbol{\mu}_\gamma \end{bmatrix} \begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \boldsymbol{y} \end{bmatrix}$$

$$\boldsymbol{\mu}_\gamma = \operatorname{diag}\left\{\frac{1}{\gamma\mu_1}, \ldots, \frac{1}{\gamma\mu_n}\right\} \tag{1.6}$$

By introducing the kernel function, the inner product computation in feature space can be converted into the kernel computing in original space, the nonlinear mapping help prevent the curse of dimensionality. We chose RBF (Radial Basis Function) as the kernel function, the form is