Journal of Power Sources 239 (2013) 680–688

Contents lists available at SciVerse ScienceDirect

Journal of Power Sources

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

Health diagnosis and remaining useful life prognostics of lithium-ion batteries using data-driven methods☆

Adnan Nuhic [a,*], Tarik Terzimehic [a], Thomas Soczka-Guth [a], Michael Buchholz [b], Klaus Dietmayer [b]

[a] Deutsche ACCUmotive GmbH & Co. KG, Neue Str. 95, 73230 Kirchheim u. Teck (Nabern), Germany
[b] Institute of Measurement, Control, and Microtechnology, Ulm University, 89081 Ulm, Germany

HIGHLIGHTS

► SOH and RUL estimation of a lithium-ion battery with a support vector machine.
► New method for training data processing by load collectives.
► Estimation accuracy over lifetime approved on real driving profiles.
► Application on-board a battery management system conceivable.

ARTICLE INFO

Article history:
Received 17 September 2012
Received in revised form
26 November 2012
Accepted 27 November 2012
Available online 13 December 2012

Keywords:
Lithium-ion battery
Battery management system
State of health
Remaining useful life
Support vector machine

ABSTRACT

The accurate estimation of state of health (SOH) and a reliable prediction of the remaining useful life (RUL) of Lithium-ion (Li-ion) batteries in hybrid and electrical vehicles are indispensable for safe and lifetime-optimized operation. The SOH is indicated by internal battery parameters like the actual capacity value. Furthermore, this value changes within the battery lifetime, so it has to be monitored on-board the vehicle. In this contribution, a new data-driven approach for embedding diagnosis and prognostics of battery health in alternative power trains is proposed. For the estimation of SOH and RUL, the support vector machine (SVM) as a well-known machine learning method is used. As the estimation of SOH and RUL is highly influenced by environmental and load conditions, the SVM is combined with a new method for training and testing data processing based on load collectives. For this approach, an intensive measurement investigation was carried out on Li-ion power-cells aged to different degrees ensuring a large amount of data.

1. Introduction

The most limiting factor of electric and hybrid vehicles popularization in means of transport is currently the vehicle's battery. The battery increases the price of the vehicle, thus, it becomes more expensive compared to conventional vehicles [1,2]. Because of many advantages, Lithium-ion (Li-ion) batteries are the most used battery type in hybrid and electric vehicles, nowadays [3]. Since this technology is present on the market for only a relatively short period, not all its characteristics are well-known. Gaining more knowledge about battery lifetime behavior would eventually result in the development of cost-effective and long lasting batteries.

However, independent from battery design, environmental impacts and dynamical cycling will always push the battery aging and thereby impede the battery in its maximum performance over lifetime. Therefore, it is always desirable to monitor the underlying degradation to be able to track the actual performance and take countermeasures if developing faults occur. This task is called health diagnosis. A recent summary on methods for Li-ion battery diagnosis can be found in Ref. [4]. Prognostics for batteries, on the other hand, predict the remaining useful life (RUL), i.e. how soon a battery pack component (e.g. cells) will fail or reach a level that cannot guarantee satisfactory performance. Diagnosis and prognostics, therefore, are two integral parts in realizing a battery health monitoring system.

Health monitoring embedding diagnosis and prognostics for machinery has gained much attention in the research community in recent years [5,6]. However, an electro-chemical system is fundamentally different from a mechanical system in various aspects. The electro-chemical reactions inside a Li-ion battery pack

☆ This contribution was originally presented as a poster at UECT 2012.

* Corresponding author. Tel.: +49 (0)15158624085; fax: +49 (0)7113052131657.
E-mail address: adnan.nuhic@daimler.com (A. Nuhic).

0378-7753/$ – see front matter
http://dx.doi.org/10.1016/j.jpowsour.2012.11.146

are almost inaccessible by using common sensor technologies. Therefore, the most available monitoring data collected from Li-ion batteries are from terminal behavior such as voltage, current, and temperature. Finally, compared to mechanical systems, the operation profiles of Li-ion batteries show much more dynamics. A good example for that is a hybrid vehicle, where the Li-ion battery condition is affected by the driver's behavior and environment. Factors affecting the performance and health of Li-ion batteries include aging-dependent capacity loss, capacity imbalance among battery cells, self-discharge, etc. Therefore, the development of appropriate methodologies and algorithms for monitoring these values have to take into account the uniqueness of Li-ion battery system [7].

The permanent reliable operation of the battery requires these monitoring algorithms to be implemented on-board the vehicle within the battery management system (BMS) [8]. For choosing appropriate algorithms, compromise needs to be made between their complexity and their diagnosis and prognostics accuracy/capability.

Several approaches for on-board suitable algorithms exist today. The usage of model-based tracking methods is a common way to achieve desired results [4]. The usage of Kalman filtering with electro-chemical or electrical equivalent-circuit models for monitoring was reported in a lot of works, e.g. Refs. [9,10]. But multiple sources of errors like sensor offsets, degrading sensor fidelity, or the quality of measured data impede this estimation, especially when it is used on-board a vehicle with reduced computation capabilities. Automated reasoning schemes based on neuro-fuzzy and decision theoretic methods, like Autoregressive Integrated Moving Average (ARIMA), have been investigated for both diagnostics and prognostics tasks [11]. Not disclaiming the work done before, it still remains difficult to accurately monitor the battery health or predict the remaining useful life under arbitrary environmental and load conditions.

At this point, the usage of data-driven methods is convenient due to their ability to transform high-dimensional and noisy environmental data into lower-dimensional information for diagnostics and, especially, for prognostics tasks [12]. In this contribution, a new data-driven approach is developed for embedding diagnosis and prognostics of battery health in automotive applications. For the estimation of SOH and RUL, one of today's most powerful and popular machine learning algorithms, the support vector machine (SVM), is combined with a completely new method for data processing. The input and output vectors of the required SVM learning data set are generated by processing the measured data through load collectives. As the estimation of SOH and RUL is strongly influenced by environmental, ambient, and load conditions, this method processes the data in respect to these dependencies, including even the operation history. Furthermore, to ensure a large amount of training and testing data, an intensive measurement investigation was carried out on automotive Lithium-ion power-cells aged to different degrees.

The following sections will expand more on the chosen algorithm in Section 2, our implementation approach in Section 3, the experimental setup and corresponding results in Section 4, and concludes with a summary in Section 5.

2. Intelligent battery health monitoring

2.1. Defining battery health

Usually, the term state of health (SOH) is used to characterize the battery health status. The SOH describes the physical condition of the battery, which is commonly characterized by the loss of rated capacity:

$$\mathrm{SOH} = \frac{C_{\mathrm{act}} - C_{\mathrm{EOL}}}{C_{\mathrm{nom}} - C_{\mathrm{EOL}}} \cdot 100\% \quad C_{\mathrm{act}} \geq C_{\mathrm{EOL}}. \tag{1}$$

Here, C_{act} is the actual capacity of the battery and C_{nom} represents the nominal capacity of a brand-new battery. For the Eq. (1), an end of life (EOL) capacity C_{EOL} at SOH = 0% has to be defined, too. In the battery manufacturing industry, this value is often reached if the actual capacity drops below 80% of its initial value

$$C_{\mathrm{EOL}} = 0.8 \cdot C_{\mathrm{nom}}. \tag{2}$$

However, the SOH value declines as a function of time through battery usage and aging from 100% to 0%. The number of charge–discharge cycles related to the specific performance (until i.e. 80% of the nominal capacity is reached) is the remaining useful life (RUL) of the battery. In this work, the degradation trend of the time-varying capacity is tracked, and the number of cycles to SOH = 0% is estimated to realize the proposed approach.

2.2. Support vector regression

The support-vector-machines (SVMs) have been applied for classification problems in various domains of pattern recognition. A comprehensive introduction can be found e.g. in Refs. [13,14]. However, the SVM can also be applied to regression problems, although regression is inherently more difficult than classification. The SVM used for regression as a non-linear estimator is more robust than a least-squares estimator because it is insensitive to small changes [15].

Let the training data set be given with $(x_1, y_1),\ldots,(x_l, y_l) \subset X \times R$, where X denotes the space of input patterns (e.g. q-dimensional space $X = R^q$). The goal of ε support vector regression (ε-SVR) is to find a function $f(x)$that has a ε deviation at the most from the target patterns y_i for all training data, while at the same time the function is as flat as possible. In other words, attention is not paid to the errors as long as they are smaller than ε, but also deviations bigger then ε are not accepted. However, sometimes it is not possible to find such a function, or it is desirable to allow some errors. For this purpose, the so-called slack variables ξ_i, ξ_i^* are introduced in order to cope with otherwise unsolvable optimization problem constraints. In the case that the demanded function is linear,

$$f(x) = \langle \omega, x \rangle + b, \omega \in X, b \in R, \tag{3}$$

the optimization problem has the form

$$\begin{aligned} &\min_{\omega} \frac{1}{2}\left\| \omega^2 \right\| + C \sum_{i=1}^{l} \left(\xi_i + \xi_i^* \right) \\ &\text{subject to} \begin{cases} y_i - \|\omega, x_i\| - b \leq \varepsilon + \xi_i \\ \|\omega, x_i\| + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0. \end{cases} \end{aligned} \tag{4}$$

The parameter $C > 0$ determines the trade-off between flatness of the function $f(x)$ (i.e. simplicity of the function) and the amount of deviations higher than ε that is tolerated. Tolerating deviations can be represented by the ε-insensitive loss function $|\xi|_\varepsilon$:

$$|\xi|_\varepsilon = \begin{cases} 0 & \text{if} |\xi| \leq \varepsilon \\ |\xi| - \varepsilon & \text{otherwise}. \end{cases} \tag{5}$$

Graphically, this is shown in Fig. 1. Only points outside the shaded areas increase the amount of deviation.

Eq. (4) represents a dual optimization problem which is much easier for solving, and, more importantly, makes it possible to apply