Applied Energy 118 (2014) 114–123

Contents lists available at ScienceDirect

Applied Energy

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

A naive Bayes model for robust remaining useful life prediction of lithium-ion battery

Selina S.Y. Ng *, Yinjiao Xing, Kwok L. Tsui

Department of Systems Engineering and Engineering Management, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong

HIGHLIGHTS

- Robustness of RUL predictions for lithium-ion batteries is analyzed quantitatively.
- RUL predictions of the same battery over cycle life are evaluated.
- RUL predictions of batteries over different operating conditions are evaluated.
- Naive Bayes (NB) is proposed for predictions under constant discharge environments.
- Its robustness and accuracy are compared with that of support vector machine (SVM).

ARTICLE INFO

Article history:
Received 12 September 2013
Received in revised form 29 November 2013
Accepted 16 December 2013
Available online 8 January 2014

Keywords:
Lithium-ion battery
Prognostic
State-of-health
Remaining useful life
Prediction
Naive Bayes

ABSTRACT

Online state-of-health (SoH) estimation and remaining useful life (RUL) prediction is a critical problem in battery health management. This paper studies the modeling of battery degradation under different usage conditions and ambient temperatures, which is seldom considered in the literature. Li-ion battery RUL prediction under constant operating conditions at different values of ambient temperature and discharge current are considered. A naive Bayes (NB) model is proposed for RUL prediction of batteries under different operating conditions. It is shown in this analysis that under constant discharge environments, the RUL of Li-ion batteries can be predicted with the NB method, irrespective of the exact values of the operating conditions. The case study shows that the NB generates stable and competitive prediction performance over that of the support vector machine (SVM). This also suggests that, while it is well known that the environmental conditions have big impact on the degradation trend, it is the changes in operating conditions of a Li-ion battery over cycle life that makes the Li-ion battery degradation and RUL prediction even more difficult.

1. Introduction

Battery-powered systems constitute an essential part of our everyday life. The consequences of battery failure can range from inconvenience, production downtime, to catastrophic failure. The health of a battery may refer to a number of parameters that indicate the battery condition, such as the cycle number, the full charge capacity (FCC), and the state-of-charge (SoC). Three critical parameters for battery health management are the real-time estimation of the remaining charge within one discharge cycle, or SoC estimation; the real-time estimation of the battery's capability to deliver its specified output, or state-of-health (SoH) estimation; and the prediction of remaining useful life (RUL). The ultimate goal of SoC estimation to ensure optimum control of the charging/discharging process, thereby maximize the run-time per discharge cycle [1]. SoH estimation and RUL prediction, on the other hand, aim at determining when the battery can no longer hold a useful amount of energy and needs to be replaced, thereby maximize the number of cycles attainable for the life of the battery. These two problems are modeled with battery data on different time-scales. Some recent papers on SoC and capacity estimation include [2–6]. This paper focuses on the RUL to end-of-life (EOL) prediction in SoH estimation. A more comprehensive introduction to battery health management and Li-ion batteries is in [7].

There are fewer papers on battery SoH estimation and RUL prediction than in SoC estimation. One challenge in battery RUL prediction is the controversy over the definition of SoH. For example, capacity indicates how much electricity can be stored, and power indicates how fast electricity can be given out. They represent different aspects of battery health. Both capacity and power fade have been used by Pattipati et al. [1] as SoH indicators. Recently, Liu et al. [8] proposed a novel health indicator with observable parameters instead of parameters from battery properties,

* Corresponding author. Tel.: +852 3442 2584; fax: +852 3442 0173.
E-mail address: selina.ng@my.cityu.edu.hk (S.S.Y. Ng).

0306-2619/$ - see front matter
http://dx.doi.org/10.1016/j.apenergy.2013.12.020

although more evidence of its effectiveness is required. The end-of-life (EOL) is usually defined as the point at which the full charge capacity (FCC) drops by a certain percentage from its initial rated capacity, typically 20–30% [9,10]. Predicting the RUL from the SoH usually involves finding the k-step-ahead projection of SoH, then transforming the capacity predictions to time-to-EOL [9]. In this paper, the most popular health indicator, the capacity, is used and the EOL is defined at 70% of rated capacity as in [10].

Techniques for prognostics and health management (PHM) [11] of systems and components can be roughly classified into the Physics-of-failure-based approach (PoF-based approach), the data-driven approach, and the fusion approach. A short review on various methods for Li-ion battery prognostics and health monitoring can be found in [12]. A discussion of SoH estimation methods categorized under these approaches can be found in [13]. PoF-based approach takes into account the physics-of-failure of a component or system. Ramadass et al., Ning et al., and Zhang and White [14–16] proposed different PoF-based techniques for modeling capacity fade and cycle life of Li-ion batteries. However, physical models are usually presented in the form of partial differential equations with many unknown parameters. They are accurate but not desirable in online applications due to a high requirement for memory and computation.

In contrast, data-driven approach does not require product-specific knowledge. Statistical techniques and/or machine learning algorithms are used to detect parameter changes, isolate faults and estimate RUL for failure diagnosis and prognosis [17]. He et al. [4] proposed a SoH and RUL estimation method using Dempster–Shafer theory and Bayesian update with Monte Carlo approach. Pattipati et al. [1] proposed a data-driven battery management system (BMS) that used support vector machine (SVM) for both SoH and SoC estimation; and support vector regression (SVR) for RUL prediction. Xing et al. [18] predicted RUL with a combination of exponential and polynomial regression followed by particle filtering (PF). Data-driven techniques are flexible and can be applied to problems with a similar format, even when the underlying physics are different. In recent years, techniques that utilize PoF-based models with data-driven methods for parameter estimation are sometimes referred to as the fusion approach. The relevance vector machine with particle filtering (RVM–PF) framework for battery RUL prediction proposed by Saha et al. [19] falls into this category.

A common problem in the existing data-driven models is that modeling battery degradation was based on the capacity data, which was collected at a constant discharge rate and a constant ambient temperature. In practice, the battery is operated in a varying operational condition. Uncertainties will make these models underperform. Although the PF-based models have paid more attention to address the uncertainties, the degradation data have still been employed under a constant environmental condition, such as Saha et al. [19] and Xing et al. [18] model. The fact that more uncertainties are introduced can reduce both the performance and the computational efficiency of the PF-based methods. Therefore, it is important to model the battery degradation taking into account different operational conditions including different usage conditions and ambient temperatures. Goebel et al. [10] compared the capacity estimation and RUL prediction over time of three methods qualitatively. Xing et al. [20] studied the influences of ambient temperature on SoC estimation. The authors are not aware of any studies on the robustness of RUL prediction over different operations conditions of different batteries. This paper is a quantitative study on the robustness of RUL prediction over time for the same battery and for different batteries under different constant operations conditions. The purpose procedure is to develop an effective battery degradation model considering both various operational conditions and the feasibility for real application. It will help a better planning of battery replacement schedule, a proper maintenance strategy, and an early precaution to protect against battery failure.

In this paper, the problem of Li-ion battery SoH estimation is formulated such that a data-driven classification method, the naive Bayes (NB), can be applied directly for RUL prediction. Their RUL prediction accuracy and robustness against constant discharging at different operating conditions are evaluated. Section 2 is a brief background on naive Bayes. Section 3 introduces the proposed prognostics procedure and Section 4 shows the empirical analysis with publicly available Li-ion battery data. This is followed by a discussion of the results in Section 4 and the conclusion in Section 5.

2. Modeling the capacity depletion

The NB used in the proposed algorithm belongs to the data-driven approach. NB is a popular data mining algorithm [21] known for its empirical performance.

NB model is the simplest form of a Bayesian network [22,23]. It calculates the probabilities of an observation belonging to a particular class according to the Bayes' theorem, by assuming that each predictor is conditionally independent of every other predictor [24]. An NB normally assigns an observation to the class with the largest probability, but it has also been applied to regression problems [25] in which the probabilities of each class can be combined differently. It is well-known for its competitive performance in real-world applications, even though the conditional independent assumption is rarely valid. Its strengths include its simplicity, efficiency, and robustness to noise and missing data. It also permits the use of more than two classes.

2.1. A short review on the naive Bayes literature

The naive Bayes has a long history in the literature, sometimes under various names in earlier articles [26]. A recent summary of the technique can be found in [21].

It is well documented in the literature that the NB often performs surprisingly well and outperforms more sophisticated algorithms in classification, even at times where the features are clearly not independent [27]. This has motivated three kinds of research [28]: (1) improve the algorithm by relaxing the independence assumption; (2) modify the feature set to comply with the independence assumption as much as possible; and (3) explain why the independence assumption is not necessary. Quantifying the degree of dependence which the NB can tolerate was started by Ref. [29]. Identifying the NB's true region of optimal performance was started by Ref. [27]. A flexible tree-augmented naive Bayes using kernel density estimation was proposed by Ref. [23]. More recently, a 'non-parametric' naive Bayes was proposed by Ref. [30], motivated by surprisingly well performance of the NB on a medical data set.

There has been continued interest in explaining and quantifying the conditions for good NB performance for practical application [22,26,27,31]. One obvious reason for good classification performance is the insensitivity of zero-one loss. This refers to a model's ability to assign the correct class to a test observation as long as the probability of belonging to this class is greater than that of other class(es), even though the probability estimates may not be accurate. However, this property applies to all classification algorithms in general and does not explain why the NB performs better than more complicated classification algorithms such as C4.5 decision tree, SVM, and neural networks on real data sets.

Another reason for good performance unique to the NB is that the conditional independence is only a sufficient but not a