Engineering Applications of Artificial Intelligence 25 (2012) 814–823

Contents lists available at SciVerse ScienceDirect

Engineering Applications of Artificial Intelligence

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

A data-model-fusion prognostic framework for dynamic system state forecasting

J. Liu [a,*], W. Wang [b], F. Ma [c], Y.B. Yang [d], C.S. Yang [e]

[a] Department of Mechanical and Aerospace Engineering, Carleton University, Ottawa, ON, Canada K1S 5B6
[b] Department of Mechanical Engineering, Lakehead University, Thunder Bay, ON, Canada P7B 5E1
[c] Department of Mechanical Engineering, University of California, Berkeley, CA 94720, USA
[d] State Key Laboratory for Novel Software Technology, Department of Computer Science, Nanjing University, Nanjing 210093, China
[e] National Research Council, Ottawa, ON, Canada K1A 0R6

A R T I C L E I N F O

Article history:
Received 1 February 2011
Received in revised form
20 February 2012
Accepted 20 February 2012

Keywords:
Nonlinear prediction
Fault diagnosis
Failure prognostics
Neural networks
Neural fuzzy systems
Remaining useful life prediction

A B S T R A C T

A novel data-model-fusion prognostic framework is developed in this paper to improve the accuracy of system state long-horizon forecasting. This framework strategically integrates the strengths of the data-driven prognostic method and the model-based particle filtering approach in system state prediction while alleviating their limitations. In the proposed methodology, particle filtering is applied for system state estimation in parallel with parameter identification of the prediction model (with unknown parameters) based on Bayesian learning. Simultaneously, a data-driven predictor is employed to learn the system degradation pattern from history data so as to predict system evolution (or future measurements). An innovative feature of the proposed fusion prognostic framework is that the predicted measurements (with uncertainties) from the data-driven predictor will be properly managed and utilized by the particle filtering to further update the prediction model parameters, thereby enabling markedly better prognosis as well as improved forecasting transparency. As an application example, the developed fusion prognostic framework is employed to predict the remaining useful life of lithium ion batteries through electrochemical impedance spectroscopy tests. The investigation results demonstrate that the proposed fusion prognostic framework is an effective forecasting tool that can integrate the strengths of both the data-driven method and the particle filtering approach to achieve more accurate state forecasting.

1. Introduction

Condition-based maintenance is a program that recommends maintenance decisions based on the information collected through system condition monitoring (or system state estimation) and equipment failure prognostics (or system state forecasting), in which prognostics still remains as the least mature element in both research and real-world applications (Jardine et al., 2006). Prognostics entails the use of the current and previous system states (or observations) to predict the future states of a dynamic system. Reliable forecast information can be used to schedule repairs and maintenance in advance and provide an alarm before faults reach critical levels so as to prevent machinery performance degradation, malfunction, or even catastrophic failures (Liu et al., 2009).

In general, prognostics can be conducted using either data-driven methods or model-based approaches. Data-driven methods use pattern recognition and machine learning to detect changes in system states (Yagiz et al., 2009; Gupta and Ray, 2007). The classical data-driven methods for nonlinear system prediction include the use of stochastic models such as the autoregressive (AR) model, the threshold AR model (Tong and Lim, 1980), the bilinear model (Subba, 1981), the projection pursuit (Friedman and Stuetzle, 1981), the multivariate adaptive regression splines (Friedman, 1991), and the Volterra series expansion (Brillinger, 1970). Since the last decade, more research interests in data-driven system state forecasting have been focused on the use of flexible models such as various types of neural networks (NNs) (Atiya et al., 1999; Liang and Liang, 2006) and neural fuzzy (NF) systems (Husmeier, 1999; Korbicz, 2004; Jang, 1993). Data-driven methods rely on past patterns of the degradation of similar systems to project future system states; their forecasting accuracy depends on not only the quantity but also the quality of system history data, which could be a challenging task in many real applications (Liu et al., 2009;

* Corresponding author. Tel.: +1 613 5202600x8257; fax: +1 613 5205715.
E-mail address: jliu@mae.carleton.ca (J. Liu).

0952-1976/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.engappai.2012.02.015

Wang and Vrbanek, 2008). Another principal disadvantage of data-driven methods is that the prognostic reasoning process is usually opaque to users (Tse and Atherton, 1999); consequently, they are not suitable for some advanced applications where forecast reasoning transparency is required (e.g., credit card cheating, earthquake and stock market prediction).

Model-based approaches typically involve building models (or mathematical functions) to describe the physics of the system states and failure modes; they incorporate physical understanding of the system into the estimation of system state and/or remaining useful life (RUL) (Adams, 2002; Luo et al., 2003; Chelidze and Cusumano, 2004). Model-based approaches, however, may not be suitable for many industrial applications where the physical parameters and fault modes may vary under different operation conditions (Pecht and Jaai, 2010). On one hand, it is usually difficult to tune the derived models in situ to accommodate time-varying system dynamics. On the other hand, model-based approaches cannot be used for complex systems whose internal state variables are inaccessible (or hard) to direct measurement using general sensors. In this case, inference has to be made from indirect measurements using techniques such as particle filtering (PF). The PF-based approaches have been used for prognostic applications (Saha et al., 2009), in which the PF is employed to update the nonlinear prediction model and the identified model is applied for system state forecasting. However, a limitation associated with the classical PF-based predictors is that the prediction model parameters cannot be updated during the prognostic period since no new measurements are available. The prediction accuracy could be low in many applications because the identified model during the state estimation period may not be accurate and robust.

To address the aforementioned challenges, a data-model-fusion framework is proposed in this work for system state prognostics. The developed framework aims to integrate the strengths of the data-driven prognostic method and the model-based PF approach for a more reliable system state forecasting. The proposed fusion framework is new in the following aspects: (1) the prediction uncertainties from the data-driven predictor can be properly managed and utilized through the fusion framework so as to further update the prediction model parameters; (2) the fusion prognostic framework can overcome the aforementioned limitations of both the data-driven method and the model-based PF approach so as to make prediction models more interpretable and transparent; (3) as an application example, the developed fusion prognostic framework is implemented for the RUL prediction of lithium-ion batteries.

This paper is organized as follows. The proposed fusion prognostic framework is described in Section 2. The effectiveness of this fusion framework is demonstrated in Section 3 via an application in battery RUL prediction. A summary of important observations and conclusive remarks are given in Section 4.

2. The fusion prognostic framework for dynamic system state forecasting

In this section, we first briefly discuss two principal components of the proposed fusion prognostic framework: the data-driven prognostic method and the PF-based prognostic approach. The limitations of each component will be examined, which, in turn, motivates the advanced research of this work. The fusion prognostic framework will then be described. This framework aims to integrate the advantages of both the data-driven predictor and the PF approach while alleviating their respective limitations, so as to develop a more reliable system state forecasting paradigm.

2.1. The data-driven prognostic method

Data-driven predictors employ pattern recognition and machine learning to forecast changes in system states (Yagiz et al., 2009; Gupta and Ray, 2007). Since the last decade, more research interests in data-driven system state forecasting have shifted to the use of flexible models such as NNs (Atiya et al., 1999; Husmeier, 1999), NF systems (Jang, 1993), and recurrent neural fuzzy (RNF) systems (Liu et al., 2009). The authors' research group has also developed several data-driven predictors for machinery applications (Liu et al., 2009; Wang and Vrbanek, 2008), and the investigation results have shown that if an NF predictor is properly trained, it performs better than both the feedforward and the recurrent NN forecasting schemes. The prediction output of a data-driven predictor can be generally described as

$$Y_k = g(C_{1:q}, Y_{1:k-1}) + u_k, \tag{1}$$

where Y_k is the predicted measurement at step k, $Y_{1:k-1}$ is the system's historical measurements up to time step $k-1$, $C_{1:q}$ are the system inputs (or system operational conditions), $g(\cdot)$ denotes the nonlinear prediction reasoning, and u_k is a random noise that represents the prediction uncertainty. The uncertainty term u_k generally pertains to the specific data-driven prognostic scheme (i.e., the structure and training algorithm) as well as the quality and quantity of training data, which can be estimated through a large number of simulations (Tiwari and Chatterjee, 2010).

Although data-driven prognostic methods have demonstrated some superior properties to other classical forecasting tools, they still have some limitations in industrial applications (Walter and Pronzato, 1997): (1) the forecasting accuracy strictly depends on if the training data are adequate and representative of all the possible application conditions. Such a requirement is usually difficult to achieve in real-world applications because, on one hand, running a system to failure could be a lengthy and rather costly process and the training data are usually inadequate in industrial applications; on the other hand, most machines/systems operate in noisy and/or uncertain environments and machinery dynamic characteristics may change suddenly (e.g., just after repairs or regular maintenance), thus the training data cannot cover all the possible operational conditions. (2) For NN/NF-based predictors, the forecasting reasoning structures are usually difficult to be understood by users. This limits their applications in which reasoning transparency (or understandability) is required. (3) The prediction uncertainty u_k usually increases dramatically as the prediction step becomes larger; as a result, an appropriate filtering process is required to further improve the forecasting accuracy. The aforementioned limitations associated with data-driven prognostic methods can be properly alleviated through the proposed data-model-fusion framework, which will be discussed in Section 2.3.

2.2. The particle filtering-based prognostic approach

For complex systems whose internal state variables are inaccessible (or hard) to direct measurement using general sensors, inference has to be made from indirect measurements, for which Bayesian learning provides a rigorous framework. Given a general discrete-time state estimation problem, the unobservable state vector $X_k \in R^n$ evolves according to the following system model

$$X_k = f(X_{k-1}) + w_k, \tag{2}$$

where $f{:}R^n \to R^n$ is the system state transition function and $w_k \in R^n$ is a noise whose known distribution is independent of time. At each discrete time instant, an observation (or measurement) $Y_k \in R^p$ becomes available. This observation is related to the