Measurement 46 (2013) 2391–2404

Contents lists available at SciVerse ScienceDirect

Measurement

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

Multiple optimized online support vector regression for adaptive time series prediction

Datong Liu *, Yu Peng, Junbao Li, Xiyuan Peng

Department of Automatic Test and Control, Harbin Institute of Technology, Harbin 150080, China

ARTICLE INFO

Article history:
Received 18 December 2012
Received in revised form 18 March 2013
Accepted 17 April 2013
Available online 9 May 2013

Keywords:
Data-driven prognostics
Online prediction
Adaptive prediction strategy
Online SVR

ABSTRACT

Data-driven prognostics based on sensor or historical test data have become appropriate prediction means in prognostics and health management (PHM) application. However, most traditional data-driven prognostics methods are off-line which would be seriously limited in many PHM systems needed on-line predicting or real-time processing. Furthermore, even in some on-line prediction algorithms such as Online Support Vector Regression (Online SVR) and Incremental learning algorithm, there are conflicts and trade-offs between prediction efficiency and accuracy. Therefore, in different PHM applications, prognostics algorithms should be on-line, flexible and adaptive to balance the prediction efficiency and accuracy. An on-line adaptive data-driven prognostics strategy is proposed with five various optimized on-line prediction algorithms based on Online SVR. These five algorithms are improved with kernel combination and sample reduction to realize higher precision and efficiency. These algorithms can achieve more accurate results by data preprocessing and model optimization, moreover, faster operating speed and lower computational complexity can be obtained by optimization of training process with on-line data reduction. With these different improved Online SVR methods, varies of prediction with different precision and efficiency demands could be fulfilled by an adaptive strategy. To evaluate the proposed prognostics strategy, we have executed simulation experiments with Tennessee Eastman (TE) process. In addition, the prediction strategies are also applied and evaluated by traffic mobile communication data from China Mobile Communications Corporation Heilongjiang Co., Ltd. Experiments and test results prove its effectiveness and confirm that the algorithms can be effectively applied to the on-line status prediction with increased performance in both precision and efficiency.

1. Introduction

The desire and need for accurate diagnostic and real-time prediction have been around since human beings began operating complex and expensive machinery. Recently, intensive research has presented on fault detection, diagnosis, prognosis and prediction of the various systems or applications [1,2]. Accurate prognostics and remaining useful life (RUL) estimation is the key technique in prognostics and health management (PHM). With the advances of computer, communication and sensor technologies, it is feasible and practical to monitor the performance and health state of a complex system at several different levels included system level, board level and even chips level, etc. Generally speaking, fault diagnosis and prognosis methods can be broadly classified into three categories, namely data-driven method, model-based method and statistical-based method. In model-based methods, models should be derived from the fundamental understanding of the mechanism of a system [3], thus it can only be applied to systems with clear mechanics

* Corresponding author. Tel.: +86 45186413533; fax: +86 45186402953.
E-mail address: liudatong@hit.edu.cn (D. Liu).

0263-2241/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.measurement.2013.04.033

description. However, it is difficult to find the accurate models in most applications of the complex systems especially for electronic complicated systems. Furthermore, statistical-based methods are based on large number of samples, while they are not suitable for the individual system or subsystem.

Because of increased automation, faster sampling rate and advances in computing, large amount of data is available on-line. Many researchers have attached great importance to the data-driven methods in diagnosis and prognosis [4,5]. Therefore, data-driven prognostics based on the sensor or historical test data, such as Artificial Neural Networks (ANNs), Support Vector Regression (SVR) and other computational intelligence methods, have become the primary prediction approaches for complex systems. In data-driven methods, state monitoring and predicting become the main factors in actual application to complex systems. Typical data-driven prognostics approaches include ANN, SVR, Fuzzy Systems and other statistical approaches.

Liu et al. [6] introduced a type of Adaptive Recurrent Neural Network (ARNN) for dynamic system state prediction and applied the proposed method into RUL prediction of lithium–ion batteries. The improved ARNN is constructed based on the adaptive/recurrent neural network architecture as well as the network weights are adaptively optimized with the recursive Levenberg–Marquardt (RLM) method. Saha et al. [7] integrated the SVR with other statistical methods to improve the ability of system prognostics. Zio and Maio [4] proposed a similarity-based approach for prognostics of the RUL of the system by fuzzy similarity analysis the evolution data to the reference trajectory patterns. Orchard and Vachtsevanos [8] presented an on-line particle-filtering-based framework for failure prognosis in nonlinear, non-Gaussian systems. Moreover, lots of researchers also did amount of work with particle filter algorithm to realize the prognostics for its uncertainty representing ability [9–11]. Similar to the SVR, the Relevance Vector Machine (RVM) [12,13] is a type of improved machine learning algorithm based on Bayesian framework. In [7], the prognostics model is constructed with the RVM algorithm to realize degradation trend estimation for the lithium–ion battery.

However, most of the data-driven approaches and models for prognostics are off-line, which could not meet the dynamic prediction demands and the unknown system features. Besides various data-driven approaches, more and more research work focused on the fault prognostics in time series prediction theory and its development [4]. SVR algorithm is a typical machine learning method based on statistical learning theory and widely applied in time series prediction. It can also be used for system state monitoring and forecasting. However, SVR cannot meet the demands of on-line and real-time application because of the time-consuming computation. In the same way, most of traditional off-line forecasting data-driven methods are facing the same challenges. Hence many training algorithms, such as incremental algorithm and decremental algorithm [13,14], have been proposed to update these methods to on-line style and further decreased the computing complexity. Incremental algorithm is the most popular on-line learning strategy for the SVR [15–21]. Kivinen et al. [15], Syed et al. [19] proposed a simple incremental algorithm that very few samples were training with the traditional quadratic programming each time. In each training process, the non-support vectors were omitted while the support vectors were kept to training with the new updated samples. As a result, the training samples decreased to accelerate the training speed. Meanwhile, the probable support vectors might be forgotten to reduce the precision of modeling and prediction. In [16], the author proposed the accurate incremental learning algorithm and applied the new algorithm in the classification of SVM. Furthermore, Ma et al. [20], Martin [21] introduced the Accurate Online SVR algorithms based on the incremental algorithm to solve the approximation and regression problems. Lots of researchers further proposed many improved Online SVR algorithms for different applications [22–25].

Although all the on-line algorithms introduced above can achieve the model dynamically learning with the sample updating, the computing complexity were very high. Moreover, in some on-line prediction methods such as Online SVR algorithm, conflicts and trade-offs between prediction efficiency and accuracy still existed.

In addition, the on-line prognostics are becoming more popular and necessary with the increasing demands for industrial applications. Jin et al. [26] presented an on-line prognostic framework consisting of the off-line degradation modeling and on-line Bayesian approach. The posterior degradation model is updated iteratively and the degradation state is estimated with the particle filter-based state and static parameter joint estimation method. Qu and Zuo [27] proposed a joint algorithm to realize condition indicators of the machine system. LSSVR is used to predict true values of condition indicators and is also optimized by Genetic Algorithm (GA). To fit the on-line changes of indicator values and noise effects, the cumulative sum technique is employed to determine the re-training of the of LSSVR parameters when current predictions are not acceptable. Liu et al. [28] presented prognostic methods based on hidden semi-Markov model (HSMM) using sequential Monte Carlo (SMC) method for equipment health prognosis. The HSMM for transition probabilities and SMC for probability relationships between health states and monitoring data are integrated. On-line multi-step-ahead RUL estimation algorithm for equipment is developed based on joint probability distribution. Though these research work aimed to realize on-line prognostics, the on-line prediction approaches above achieved the on-line prediction by iteration or re-training are not truly dynamic models. The computing complexity is generally too high, or the parameters are hard to determined for complicated algorithms.

In PHM systems, regarding the demands for fast monitoring and predicting, it is essential to develop an adaptive model for on-line prediction. Since for different applications, the needs with operation efficiency and prognostic precision are different. In some demands, the high efficiency is needed while more accurate prediction capacity is required for other requirements. Therefore, adaptive prediction strategy becomes essential for PHM systems,