Engineering Applications of Artificial Intelligence 67 (2018) 1–13

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

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

Data-driven fault prediction and anomaly measurement for complex systems using support vector probability density estimation

Hong-Qiao Wang *, Yan-Ning Cai, Guang-Yuan Fu, Ming Wu, Zhen-Hua Wei

Xi'an Research Institute of Hi-Tech, Xi'an 710025, China

A R T I C L E I N F O

Keywords:
Data-driven
Fault prediction
Anomaly degree measurement
Support vector machine
Probability density estimation

A B S T R A C T

To quantitatively monitor the state of complex system, a data-driven fault prediction and anomaly degree measurement method based on probability density estimation is studied in this paper. First, an anomaly index is introduced and defined to measure the anomaly degree of samples. Then By improving the form of constraint condition, a single slack factor multiple kernel support vector machine probability density estimation model is presented. As a result, the scale of object function and the solution number are all reduced, and the computational efficiency of the presented model is greatly enhanced. On the other hand, as the introduction of multiple kernel functions, a multiple kernel matrix with better data mapping performance is obtained, which can well solve the composite probability density estimation for uncoupled data. The simulation test shows that the presented model has higher estimation precision and speed. The experiments on complex system fault prediction also show that the system's anomaly degree can be quantitatively and accurately measured by the anomaly index gained from the prediction results, which can effectively improve the fault prediction precision and increase the prediction advances.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

With the increasing needs for system reliability, it is hoped that not only the fault's detection and isolation can be provided when it occurs, but also the fault can be forecasted before it occurs. It also means that the fault can be discovered, be located and be eliminated in the early period, when it has not caused serious damage to the whole system. In this way, enough time will be obtained to prevent the emerging of fault by taking necessary measures, which can avoid unnecessary loss and is important to system. Especially for the systems requiring high reliability, such as aerospace and nuclear energy, fault prediction has been a very important problem presented in recent years (Zhou and Xu, 2009; Dai and Gao, 2013). In fault prediction field, the system commonly has the fault state and the failure state. The fault state means that an anomaly of the system index occurs, but the system can still in a normal working process. Correspondingly, the failure state means the system index exceeds some threshold, in this case, the system will cannot work.

Different systems have different demand levels for reliability, so it would be best if the anomaly index measuring the system's anomaly degree can be calculated from algorithms. As for whether the fault should be predicted, the operator can decide according to the practical security requirement. In the domain of data-driven fault prediction, one of the methods can be used is probability density estimation to samples. On this basis, we can seek for an evaluation index characterizing the system's anomaly degree can be found and utilized.

Probability density estimation from the observed dataset is a basic problem of machine learning. There are two types of probability density estimation methods at present, one is the parameter estimation, the other is the non-parameter estimation. Maximum likelihood method is one of the representative parameter estimation measures, but this method has some limitations, for example, it cannot be used to estimate the probability density of the function compounded with several normal distributions. By contrast, the non-parameter estimation methods have been more widely used. The Parzen window density estimation (Parzen, 1962; Jenssen et al., 2006; Mohamed et al., 2004) is the most representative non-parameter estimation method, which is also a classical kernel density estimator. But the Parzen window method has a disadvantage that it does not have sparseness. When the probability densities of new samples are estimated, all the samples of the dataset are concerned and the computational complexity will become huge. Therefore, researchers have expected for a long time to find a probability density estimation method, which only uses some training samples having great influence

* Corresponding author.
E-mail address: whq05@mails.tsinghua.edu.cn (H.-Q. Wang).

https://doi.org/10.1016/j.engappai.2017.09.008
Received 27 November 2015; Received in revised form 19 December 2016; Accepted 7 September 2017
0952-1976/© 2017 Elsevier Ltd. All rights reserved.

on density estimation, instead of all the training samples. The essence of this method is to seek a sparse solution, so as to reduce the computation cost and improve the applicability. The support vector machine (SVM) provides a good approach for obtaining sparse solutions (Vapnik and Mukherjee, 2000), as the solution of SVM is only concerned with the support vectors in training samples. So the SVM method can be used to estimate probability density, and the operational steps are as follows: firstly, start from the definition of probability density and estimate an approximate distribution function from the empirical cumulative distribution function values. Secondly, get the density function by differential computing. In fact, the linear operator equation solutions are computed by SVM in the above-mentioned method, as a result, a sparse probability density estimation which is similar to the Parzen Window in form is obtained. By improving the form of constraint condition of SVM probability density estimation model, a single slack factor SVM probability density estimation model is presented in this paper. On this basis, the measurement of system's anomaly degree is achieved.

In the remainder of this paper, we go along through different sections which are organized as follows: in Section 2, we summarize the data-driven fault prediction methods for complex systems, and introduce the quantitative measurement of system anomaly based on anomaly index. The principle of probability density model based on single slack factor SVM is introduced detailedly in Section 3, the corresponding algorithm's complexity is also analyzed. In Section 4, several experiments are carried out to testify the effectiveness of the proposed method. Finally, a conclusion is drawn and the future work is also planned in Section 5.

2. Fault prediction and anomaly degree measurement for complex systems

2.1. Data-driven fault prediction for complex systems

In the operating process of some practical industry systems, the fault prediction and reliability evaluation technologies can be used to reduce the cost of system's maintenance (Wang et al., 2008; Ding et al., 2014; Alghazzawi and Lennox, 2009). The technologies also can provide reliable evidence for system's repairing opportunity determination, under this circumstance, the blindness of device maintenance can be reduced, and the effective time of system running can be greatly increased. Fault prediction and reliability evaluation, which are important measures guaranteeing the reliability of system and have received more attention in recent ten years, are key technologies for complex engineering systems' predictive maintenance. According to the difference of the known condition forms for specific problems, the fault prediction and reliability evaluation methods fall into three categories: model-driven method, data-driven method and qualitative knowledge-driven method. The precondition of model driven method (Isermann, 2005; Si et al., 2011) is that the mathematical model characterizing the physical laws of a system has been known. In data-driven method (Jiang et al., 2014; Wang and Yin, 2014; Mahadevan and Shah, 2009), the analytical model of system is not required to be known, but the quantitative data samples characterizing the state of system must be known. The qualitative knowledge driven method (Hu et al., 2011) allows the analytical model to be unknown, but the qualitative knowledge characterizing the features of system must be gained. From the characteristics of the above methods, the data driven method, which takes the gathered data as basis and is independent of the object's prior knowledge, is a more useful approach for fault prediction and reliability evaluation (Hsu et al., 2010; Chirico and Kolodziej, 2014; Svärd et al., 2014; Si et al., 2012).

As an important direction, the time series analysis and prediction using some learning algorithms in data-driven fault prediction and reliability evaluation methods has received widespread attention. On this basis, some good research results appeared, which mainly focused on the time series prediction and detection analysis using the intelligent algorithms such as neural networks and support vector machine (SVM) (Dash et al., 2007; Daewon and Jaewook, 2007; EI-Koujok et al., 2014). But in practical problems, the state of system is usually decided by multiple correlative factors, so the system being observed is often characterized by multiple correlated variables, the time series observed is commonly called multivariate correlated time series. For this reason, the characters of data must be fully considered in some procedures including the modeling based on system data, monitoring of system state and evaluation of system reliability. While for the above mentioned multivariate time series prediction, many problems can be classified as the modeling category on multi-input multi-output (MIMO) samples, such as fault prediction for MIMO systems, fault forecasting based on multi-step prediction of time series, and so on. The essence of all these problems is seeking for a mapping relationship between the multi-input samples and the multi-output samples. In recent years, the black-box modeling based on input and output data is also a research focus, the relevant techniques gain wide attention and some research findings have emerged. Some effective modeling methods, such as neural network based modeling, wavelet network based modeling, gain much popularity since then. Due to the perfect nonlinear mapping performance and generalization ability, the SVM based on statistical learning theory is introduced to the black-box modeling field and has acquired good effect.

2.2. Anomaly degree measurement for complex systems

Consider the following problem: set a system's original state is the normal operation state, and the sample set characterizing the normal operation state has been known with the form as $\{v_i\}, i = 1, \dots, l$. When a new sample v' is obtained, it is necessary to judge whether v' drifts away from the normal state and to determine the drift degree. The character of this problem is that the samples in system anomaly state or in fault state are unknown.

Then the above mentioned problem can be analyzed from the pattern recognition perspective. Let class ω_1 be the normal state of system, and class ω_2 be the system state with anomaly or fault. The prior probabilities of every class are $P(\omega_1)$ and $P(\omega_2)$, their probability density functions can be expressed as $p(v|\omega_1)$ and $p(v|\omega_2)$ respectively. According to Bayes decision theory, we can draw a conclusion that $v \in \omega_1$ if and only if

$$P(\omega_1|v) > P(\omega_2|v), \tag{1}$$

or

$$p(v|\omega_1)P(\omega_1) > p(v|\omega_2)P(\omega_2). \tag{2}$$

As we have not any information on the anomaly or fault state class ω_2, the prior probabilities $P(\omega_1)$ and $P(\omega_2)$, the probability density functions $p(v|\omega_2)$ are all unknown. So only $p(v|\omega_1)$ can be calculated using non-parametric probability density estimation method.

Based on the hypothesis testing method adopted by Yeung and Chow (2002), we know that the probability density estimation model can be constructed by the sample sets having been known, then the model can be used to calculate the probability density value of a new sample, the greater the value is, the higher the level of similarity between the new sample and the known sample set becomes. In the fault prediction methods, the training sample sets are commonly composed of the samples characterizing the system's normal operation state. So the greater the probability density value of the new sample is, the more likely the sample belongs to the normal samples. And the smaller the probability density value is, the more likely the sample drifts away from the system's normal state. To further measure the anomaly degree of a sample, we can define the anomaly index (AI).

Definition 1. Set the sample set $\{v_i,\ i = 1, \dots, l\}$ characterizing the normal system state ω_1 is known, but the samples in anomaly state or in fault state are unknown. The probability density of v_i is $p(v_i|\omega_1)$. Let $\rho_1 = \min(p(v_i|\omega_1))$, $\rho_2 = \max(p(v_i|\omega_1))$, $\rho = p(v'|\omega_1)$, then the anomaly

Download English Version:

https://daneshyari.com/en/article/4942562

Download Persian Version:

https://daneshyari.com/article/4942562

Daneshyari.com