ARTICLE IN PRESS

ISA Transactions xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

ISA Transactions

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

Research article

An improved incipient fault detection method based on Kullback-Leibler divergence

Hongtian Chen, Bin Jiang*, Ningyun Lu

College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, 169 Shengtai West Road, Jiang Ning District, Nanjing, 211106, China

ARTICLE INFO

Keywords:
Incipient fault detection
Kullback-Leibler (KL) divergence
Principal component analysis (PCA)
Electrical drive systems

ABSTRACT

This paper presents an improved incipient fault detection method based on Kullback-Leibler (KL) divergence under multivariate statistical analysis frame. Different from the traditional multivariate fault detection methods, this methodology can detect slight anomalous behaviors by comparing the online probability density function (PDF) online with the reference PDF obtained from large scale off-line data set. In the principal and residual subspaces obtained via PCA, a symmetric evaluation function is defined for both single variate and multivariate cases. The uniform form of probability distribution and fault detection thresholds associated with all eigenvalues are given. In addition, the robust performance is analyzed with respect to a wide range of Signal to Noise Ratio (SNR). Case studies are conducted with three types of incipient faults on a numerical example; combining with two nonlinear projections, the proposed scheme is successfully used for incipient fault detection in non-Gaussian electrical drive system. The results can demonstrate the superiority of the proposed method than several other methods.

1. Introduction

With the increasing demands on reliability and safety of complicated modern systems with high degree of automation, fault detection and diagnosis (FDD) has become a critical issue in engineering and research domains [1–13]. Model-based FDD methods are usually implemented under some rigorous assumptions on system dynamics and operating circumstances. They can achieve satisfactory results only when well-established models are available [14]. For the knowledge-based methods [15], the advantages of FDD can be brought out if human's expertise is sufficient and complete. But this situation may be not common when confronting with complicated systems or various operational cases. Serving as an alternative methodology, data-driven FDD methods have been receiving considerably increasing attentions during the past two decades. In general, they are easy for applications comparing with the model-based or knowledge-based FDD methods [11].

PCA is an effective multivariate statistical analysis technique, based on which has been widely used for FDD in the past two decades. It can deal with high-dimensionally correlated data by projecting the original data set into a low dimensional principal component subspace [10,16] which contains most of the original system information. Primary work about PCA and lots of improved methods can be found in Refs. [3,5–8,10,11,14,17].

Among those PCA-based FDD methods, two test statistics, T^2 and SPE, are usually used for fault detection. For Hotelling's T^2, its confident region is determined by Mahalanobis distance in the principal subspace (Span$\{P\}$) which is spanned by the principal loading matrix; for SPE, it uses Euclidean distance to define the corresponding normal region in the residual subspace (Span$\{P\}^{\perp}$) which is spanned by the residual loading matrix [17]. As pointed in Ref. [18], the Hotelling's T^2 measures the deviation of the measurement sample inside the PCA model, and SPE indicates the deviations away from the normal PCA model.

It should be pointed that some slight changes caused by faults may mainly occur either in Span$\{P\}$ or in Span$\{P\}^{\perp}$, because the nature of singular value decomposition (SVD) makes the two subspaces be orthogonal. In this case, this type of faults by projections mainly affecting one subspace will be not detected via the other test statistic, especially for incipient faults. Incipient faults mean those small changes with respect to the normal condition, which may be unnoticed at the early stage but will evolve to be serious faults and increase the risk of possible hazards. The incipient faults are easily masked by normal variations in system [7]. Their amplitudes are small compared to system signals, usually from 1% to 10% [19–21]. They can be modeled as deviations of the mean and variance of the normal signals. In a short time window, the amplitude of an incipient fault can be treated as a constant value [22].

Generally, from the detectabilities to incipient faults, T^2 statistic

* Corresponding author.
E-mail addresses: binjiang@nuaa.edu.cn, ebjiang@yahoo.com (B. Jiang).

https://doi.org/10.1016/j.isatra.2018.05.007
Received 18 April 2017; Received in revised form 20 March 2018; Accepted 10 May 2018

Please cite this article as: Chen, H., ISA Transactions (2018), https://doi.org/10.1016/j.isatra.2018.05.007

used in Span$\{P\}$ shows less efficiency than it used in Span$\{P\}^{\perp}$ when no information about incipient faults is available. The main reason is that the direction in Span$\{P\}^{\perp}$ corresponding to the smaller singular value delivers a stronger weighting. Accordingly, if there is no a priori about fault information, detection of incipient faults will be more effective in the Span$\{P\}^{\perp}$ than the Span$\{P\}$. For the decision scheme in the online detection phase, it was usually suggested that the two statistics using T^2 statistic should be monitored simultaneously in both of subspaces [10]. Besides that, some combined indexes [17] or improved indexes [18] were investigated to ensure better fault detection performance for complex industrial systems. However, efficient FDD methodology for tiny abnormalities caused by incipient faults is still an unsettled problem, especially for the low signal to noise ratio (SNR) cases.

Recently, a new probabilistic measure, called Kullback-Leibler (KL) divergence, was introduced from information theory to deal with fault detection problem [19]. It can measure the difference between two statistical distributions [20]. Under the PCA framework, the existing methods using KL divergence were proposed, which clearly presented how to optimally compute the KL divergence of score vectors between offline samplings and online readings [21–24]. Following this idea, a multivariate KL divergence-based method was then proposed in Ref. [19] because the authors were confident that PCA is not necessary to decompose the original data into two orthogonal subsets. All the above mentioned studies provided an alternative solution of the detection of incipient faults.

In fact, these methods seem imperfect because of some improper assumptions used. For example [21,23], declared that the means of score vectors were unchanged before and after an incipient faults appearing; likewise, the assumption on variances were used in Ref. [24]. Although these assumptions can simply the theoretical analysis of detection of incipient faults, it is obvious that these improper conditions may be not always satisfied in practical applications. These observations motivate us, based on the existing methods, to develop an improved method to achieve satisfied performance on incipient fault detectability.

In this paper, on the basis of the existing KL divergence and PCA-based studies, we propose an improved version where the theoretical analysis and some practical guides are presented, and its application is also extended to the electrical drive system. The main contributions of our work are summarized as follows.

1) No assumptions on unchanged mean or variance are needed in the proposed method. It is sensitive to incipient faults, and is robust to Gaussian noises as well.
2) Precise limiting distribution of the evaluation function is given by theoretical derivation, which can help determine rejection region in principal and residual subspaces accurately.
3) In order to achieve a remarkable computational cost of probability density estimation, the KL divergence between reference probability density function (PDF) and online PDF are simply to calculate mean and variance in principal and residual subspaces.
4) The proposed method can be applied to non-Gaussian electrical traction systems by introducing two coordinate transformations skillfully, and a practical algorithm is given for this purpose.

This paper is organized as follows. Section 2 gives a brief introduction about PCA and KL divergence. In Section 3, the evaluation function regrading as a fault indicator is presented in detail. Then, an online monitoring strategy is given in Section 4. The proposed method is illustrated through a numerical example and electrical traction system in Section 5. Finally, conclusions are given in Section 6.

2. Related work

2.1. PCA-based FDD

Let N, m be the numbers of samples and variables, and l be the number of principal components, then a data matrix $X \in R^{N\times m}$ can be decomposed as

$$X = \hat{X} + \breve{X} = TP^T + \breve{T}\breve{P}^T. \tag{1}$$

where $\hat{X}$ and $\breve{X}$ are the modeled and residual part of the data matrix X. $T \in R^{N\times l}$ is the score matrix and $P \in R^{m\times l}$ is the loading matrix in the principal subspace. In the residual subspace, $\breve{T} \in R^{N\times(m-l)}$ and $\breve{P} \in R^{m\times(m-l)}$ are the score matrix and the loading matrix, respectively. Eq. (1) can be further rewritten as

$$X = TP^T + E \tag{2}$$

The loading matrices P and $\breve{P}$ can be determined by SVD as

$$S = \frac{1}{N-1}X^TX = [P \quad \breve{P}]\begin{bmatrix} \Lambda_{pc} & 0 \\ 0 & \Lambda_{res} \end{bmatrix}[P \quad \breve{P}]^T \tag{3}$$

where $\Lambda_{pc} = \text{diag}(\lambda_1,\cdots,\lambda_l)$, $\Lambda_{res} = \text{diag}(\lambda_{l+1},\cdots,\lambda_m)$, and $\lambda_1,\cdots,\lambda_m$ are the eigenvalues of the covariance matrix S. The modeled components can capture most of the original variation information. The number of principal components can be chosen by the cumulative percent variance method [25].

Hotelling's T^2 test statistic, with l and $N-l$ degrees of freedom, is defined as $T^2 = \| \Lambda_{pc}^{-\frac{1}{2}}P^TX^T \|^2$. Its corresponding confident region can be found in Ref. [11]. In addition, SPE is defined as $SPE = X(I_{m\times m} - PP^T)X^T$, where $I_{m\times m}$ is the identity matrix with the rank m. Its control limit can be determined by the method in Ref. [10]. As described in Refs. [14,18], there may be advantageous to use the two statistics or some combined indexes to detect abnormalities when there is a change in the original data set X.

2.2. Definition of KL divergence

The KL divergence is a popular tool to measure the difference between two probability density functions (PDFs). Its wide applications include model selection [26], speech recognition [27], estimating divergence [28], etc. The KL divergence of two continuous PDFs f_1 and f_2 is [29]:

$$I(f_1 \parallel f_2) = \int f_1(x)\log\frac{f_1(x)}{f_2(x)}dx\mathbf{f} \tag{4}$$

where $\log(\cdot)$ is the natural logarithm. Generally, the KL divergence $I(f_1 \parallel f_2)$ is not equal to $I(f_2 \parallel f_1)$ unless the non-singular condition is satisfied [29]. Furthermore, a symmetric version [22,23] has been presented as

$$K(f_1, f_2) = I(f_1 \parallel f_2) + I(f_2 \parallel f_1) \tag{5}$$

The value of $K(f_1, f_2)$ is nonnegative, $K(f_1, f_2) = 0$ if and only if $f_1 = f_2$.

Lemma 1. [24] *Assume that there are two density functions for Gaussian distribution signals $f_1 = \mathcal{N}(\mu_1, \delta_1^2)$ and $f_2 = \mathcal{N}(\mu_2, \delta_2^2)$, the KL divergence of f_2 with respect to f_1 can be expressed as*

$$I(f_1 \parallel f_2) = \frac{1}{2}\left[\log\left(\frac{\delta_2^2}{\delta_1^2}\right) + \frac{\delta_1^2}{\delta_2^2} + \frac{(\mu_1-\mu_2)^2}{\delta_2^2} - 1\right] \tag{6}$$

Computation of KL divergence can be simplified to estimate the means and deviations of two PDFs from Lemma 1. Then $K(f_1, f_2)$ can be calculated by