



An optimal fault detection threshold for early detection using Kullback–Leibler Divergence for unknown distribution data

Abdulrahman Youssef^{a,b}, Claude Delpha^{a,*}, Demba Diallo^b

^a Laboratoire des Signaux et Systèmes (L2S), Univ. Paris Sud, CNRS, CentraleSupélec, 91192 Gif Sur Yvette, France

^b Laboratoire Génie Electrique et Electronique de Paris (GeePs), CNRS, CentraleSupélec, Univ. Paris Sud, UPMC, 91192 Gif Sur Yvette, France

ARTICLE INFO

Article history:

Received 1 March 2015

Received in revised form

31 August 2015

Accepted 8 September 2015

Available online 25 September 2015

Keywords:

Incipient fault

Detection and diagnosis

Kullback–Leibler divergence

Performance modelling

Optimisation

ABSTRACT

The incipient fault detection in industrial processes with unknown distribution of measurements signals and unknown changed parameters is an important problem which has received much attention these last decades. However most of the detection methods (online and offline) need *a priori* knowledge on the signal distribution, changed parameters, and the change amplitude (Likelihood ratio test, Cusum, etc.). In this paper, an incipient fault detection method that does not need any *a priori* knowledge on the signals distribution or the changed parameters is proposed. This method is based on the analysis of the Kullback–Leibler Divergence (KLD) of probability distribution functions. However, the performance of the technique is highly dependent on the setting of a detection threshold and the environment noise level described through Signal to Noise Ratio (SNR) and Fault to Noise Ratio (FNR). In this paper, we develop an analytical model of the fault detection performances (False Alarm Probability and Missed Detection Probability). Thanks to this model, an optimisation procedure is applied to optimally set the fault detection threshold depending on the SNR and the fault severity. Compared to the usual settings, through simulation results and experimental data, the optimised threshold leads to higher efficiency for incipient fault detection in noisy environment.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

1.1. State-of-the-art

Fault detection plays a key role in enhancing today's technological systems high demands for performance, productivity and security. The sensitivity of the fault detection methods depends on the application main goals. When productivity is the main goal, the fault detection sensitivity required is weak: only large defects should be detected. However, when security is the main goal, unde-

tected faults even with very small severity assessment may result in catastrophic growing failures. Therefore, there is a need for fault detection and diagnosis (FDD) methods, with a high sensitivity to small fault but insensitive to the environment perturbations (noise, temperature, etc.) and to input changes [1].

In the literature, a fault is defined as “a non-allowed and unpredictable deviation of at least one characteristic property or variable of the system” [2]. For industrial process monitoring, when safety is the main priority, it is crucial to be able to detect very slight faults (namely incipient modification) at their earliest stage. Indeed, early detection may provide invaluable warning on emerging problems, and appropriate actions may allow to avoid serious process upset. However, the accurate detection of incipient faults, is a challenge as it requires distinguishing

* Corresponding author.

E-mail addresses: abdulrahman.youssef@supelec.fr (A. Youssef), claude.delpha@l2s.centralesupelec.fr (C. Delpha), demba.diallo@geeps.centralesupelec.fr (D. Diallo).

the fault itself from nuisance parameters like noise or environmental unpredictable changes. There exist many sources of noise in industrial processes depending on the applications. As examples we can cite vibrations, electric power fluctuations, stray radiation from nearby electrical equipment, static electricity, turbulence in the flow of gases or liquids, background radiation from natural radioactive elements, etc. [3]. In fact, in real application processes, every kind of slight disturbance can be considered as a nuisance parameter designated here by the general term “noise”. Indeed, this noise can affect the fault detection method performance in terms of false alarm probability (reliability) and missed detection probability (sensitivity).

Moreover, the faults in industrial process may manifest in different forms on the measured signals. For example, some faults change the statistics properties of the signal (Mean, Variance, Skewness, and Kurtosis), other change the spectral properties, and other manifest as a noise added to the signals. Therefore, the fault detection methods should be able to cope all these types of fault signature.

Various methods of fault detection and isolation have been proposed in different industrial contexts. They are generally classified as: model-based and data-driven-based methods. In the model-based methods, fault detection is based on the comparison of the system's measured variables with the estimated ones obtained from a mathematical model of the process. These methods include statistical hypothesis testing approach applied on the residuals (e.g., Bayesian, likelihood, and minimax) [4,5], observer-based approach, interval approach, and parity-space approach [6,7].

The model-based approach is efficient when an accurate model is available. However poor fault detection performances are obtained due to the model uncertainties, modelling errors and e.g., tricky tuning of the observers.

In contrast to the model-based approaches, where *a priori* knowledge on the process is needed [6,7] data-based methods require the availability of a sufficient amount of historical process data to perfectly describe the process behaviour using well chosen descriptive features [8]. These approaches include the latent variable methods, e.g., partial least square (PLS) regression, principle component analysis (PCA), canonical variate analysis (CVA), independent component analysis (ICA), neural network, fuzzy systems and pattern recognition methods [9].

In this paper a data-driven approach is considered using several descriptive features in the Principle Component Analysis (PCA) frame combined with multivariate statistical techniques to develop an efficient fault detection and diagnosis method.

PCA-based monitoring methods can easily handle high dimensional, noisy and highly correlated data generated from industrial processes, and provide superior performance compared to univariate methods [10]. In addition, these process monitoring methods are attractive for industrial practical processes because they only require a good historical data set of healthy operation, which are easily obtainable for computer-controlled industrial processes. The PCA can be used to reduce the m dimensional

space of process variables to a lower l -dimensional subspace termed the principal subspace while keeping a maximum information in the new space. The remaining information is in the $(m-l)$ -dimensional subspace named the residual subspace [10–14].

PCA-based monitoring methods and their extensions have been successfully applied in a wide range of applications and industries, such as in chemical processes, air quality, water treatment, aerospace, agriculture, automotive, electronics, energy, manufacturing, medical devices, and many others [15].

The most common procedure of process monitoring with PCA consists in using some metrics (known as detection indices) to identify faults. Several detection indices have been used with this multivariate technique, which include Hotelling's T^2 statistic [16] and the squared prediction error SPE [17]. The T^2 measures the variations of the principle components at different time samples, while the SPE measures the variations of the residuals. As mentioned in [18], the performances of the T^2 and SPE in terms of false alarm and miss detection probability are not so satisfactory. Also, the T^2 and SPE are sensitive to modelling errors [19]. Moreover, the control limits of T^2 and SPE are based on the assumption that the latent variables follow a multivariate Gaussian distribution. Therefore when the latent variables are non-Gaussian distributed, using T^2 and SPE may be misleading [20].

Therefore to overcome the T^2 and the SPE performances shortcomings, an approach was proposed in [15], in which the generalised likelihood ratio GLR test is used with the PCA. The GLR test is a hypothesis testing method that has been successfully used in model-based fault detection and was superior to T^2 and SPE [15].

To distinguish between the two hypotheses (faulty and healthy), each of which has known parameters, the use of the likelihood ratio test can be justified by the Neyman–Pearson lemma [21], which proves that such a test has the best performances of all competitors.

However, the GLR is a parametric hypothesis test method, so it needs *a priori* knowledge on the signal distribution type (Gaussian, gamma, etc.) and the parameters affected by the fault (mean, variance, kurtosis, etc.). If this information is not available, or if the signals have a non-usual distribution, the GLR cannot be applied fruitfully. In addition, even if these conditions are fulfilled, if the change amplitude is unknown, the GLR is not optimal.

In the literature, the most popular non-parametric hypothesis test method is the Wilcoxon Rank Sum test (WRST). However, we will show in this paper that this method is not able to be applied in the industrial processes because of its high sensitivity to the noise presence and the random type of the signals.

1.2. Paper contribution

In this paper, we develop an approach that can be applied without any *a priori* information either on the distribution type and whatever the type and without knowledge on the parameters affected by the fault, neither the fault amplitude (non-parametric method). This method is based on the Kullback–Leibler Divergence (KLD)

Download English Version:

<https://daneshyari.com/en/article/6958602>

Download Persian Version:

<https://daneshyari.com/article/6958602>

[Daneshyari.com](https://daneshyari.com)