Journal of Process Control 44 (2016) 224–235

Contents lists available at ScienceDirect

Journal of Process Control

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

Supervised linear dynamic system model for quality related fault detection in dynamic processes

Zhiqiang Ge*, Xinru Chen

State Key Laboratory of Industrial Control Technology, Institute of Industrial Process Control, College of Control Science and Engineering, Zhejiang University, Hangzhou 310027, PR China

ARTICLE INFO

Article history:
Received 15 December 2015
Received in revised form 2 April 2016
Accepted 5 June 2016
Available online 20 June 2016

Keywords:
Dynamic fault detection
Quality related
Linear dynamic systems
Data uncertainty
Supervised modeling

ABSTRACT

Dynamic and uncertainty are two main features of industrial processes data which should be paid attentions when carrying out process monitoring and fault diagnosis. As a typical dynamic Bayesian network model, linear dynamic system (LDS) can efficiently deal with both dynamic and uncertain features of the process data. However, the quality information has been ignored by the LDS model, which could serve as a supervised term for information extraction and fault detection. In this paper, a supervised form of the LDS model is developed, which can successfully incorporate the information of quality variables. With this additional data information, the new supervised LDS model can provide a quality related fault detection scheme for dynamic processes. A detailed industrial case study on the Tennessee Eastman benchmark process is carried out for performance evaluation of the developed method.

1. Introduction

To ensure process safety and improve production efficiency, fault detection has become popular in industrial processes for many years. Compared to traditional model-based fault detection methods which may highly rely on the process knowledge or experience from process engineers, the data-based modeling method is much more flexible, thus has been widely used for fault detection in recent years. This is mainly due to the wide utilization of the distributed control system (DCS) in industrial processes, an enormous number of process data have been collected, which contain useful information for fault detection [1]. Along the past years, a number of data-based fault detection methods have been developed, such as Principal Component Analysis (PCA), Partial Least Squares (PLS), Fisher Discriminant Analysis (FDA), Support Vector Machines (SVM), etc [2–6].

For the purposes of control and optimization in the industrial process, the dynamic relationships among process data should be paid particular attentions. Actually, the dynamic model can provide a powerful tool when the phenomenon of feedback, missing or multiple measurements and noise appears. For fault detection, a data-based dynamic model is also required to capture the serial correlations among the process data. In order to handle the dynamic relationship among process data, several fault detection methods have already been developed [7–14]. For example, a dynamic generalization of the conventional Principal Component Analysis (PCA) method has been developed, which defines an augmented vector that contains time-lagged process measurements [7]. Therefore, the dynamic information can be captured in the subsequent PCA model. To further reduce the false alarms in DPCA, univariate ARMA filters were proposed to remove autocorrelation from the DPCA score variables [8]. Another new dynamic latent variable model has also been developed to improve the monitoring efficiency of the traditional DPCA method [9]. Besides, the state space model has been introduced for dynamical fault detection and diagnosis, most of which are related to the subspace identification methods [10–14].

However, most existing dynamical fault detection methods have not well considered the data uncertainty. Actually, most process variables collected from noisy environments are contaminated by random noises. It is more reasonable to consider that those process variables are inherently random variables. To address both dynamical and uncertain data characteristics, dynamical Bayesian networks (DBNs) have provided an effective modeling framework [15]. There are two well-known modeling tools in DBNs: linear dynamic systems (LDS) for modeling of continuous state sequences and Hidden Markov Model (HMM) for classification of discrete state sequences. For the

* Corresponding author.
E-mail addresses: gezhiqiang@zju.edu.cn, gezhiqiang@gmail.com (Z. Ge).

http://dx.doi.org/10.1016/j.jprocont.2016.06.003

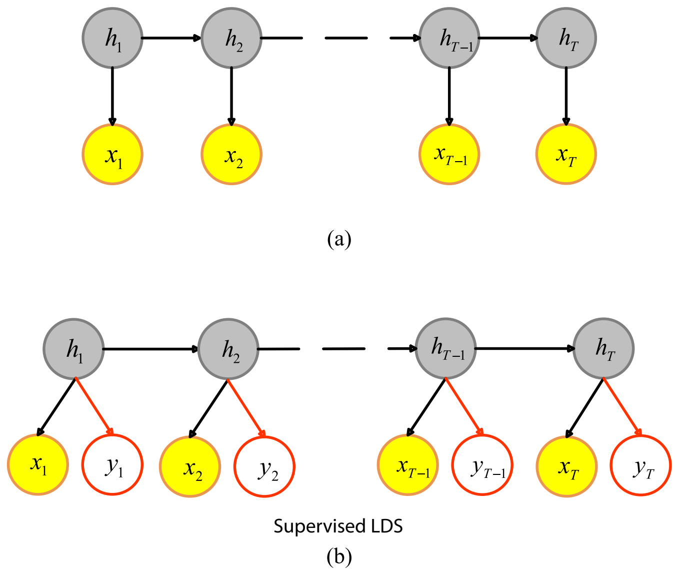

Fig. 1. Graphical representation of the LDS and SLDS models, (a) LDS model; (b) SLDS model.

fault detection purpose, the LDS model has already been introduced. For example, a data-based linear Gaussian state-space model was constructed upon the framework of LDS for dynamic process monitoring [16], a switching LDS-based approach has been proposed for fault detection and classification [17].

Unfortunately, the LDS model can only provide an unsupervised modeling framework for fault detection, which means that only the information of process variables can be incorporated into the model. Here, the unsupervised model means that the model only considers one type of variables, while the supervised model tries to build relationships between two types of variables. More illustrations about unsupervised and supervised learnings can be found in the machine learning community. In practice, data information from quality variables is important and sometimes may play as an essential role in fault detection and diagnosis. To this end, the quality relevant or related process monitoring methods have been developed in recent years. For example, a quality relevant T-PLS modeling and monitoring method has been developed for dynamic processes [18], a concurrent projection to latent structures has been proposed for quality relevant and process relevant fault monitoring [19], a new data-driven process monitoring scheme related to key performance indictors has been developed and applied to hot strip mill process [20], and so on [21–25].

In this paper, the quality data information is incorporated into the traditional LDS model, thus a supervised LDS model will be developed and used for fault detection. With the introduction of the quality data information, the supervised LDS model could be more related to the quality information. Compared to the conventional LDS model based fault detection method, the supervised LDS model can provide more relationships between process variables and the quality variables. The information of quality variables serve as supervised items for the modeling and parameter estimation process of the LDS model. Therefore, the fault detection performance is expected to be improved.

The remainder of this paper is structured as follows. In Section 2, detailed description of the supervised LDS model is presented, including the model structure, learning algorithm, and online inference for the new data. The fault detection scheme is given in Section 3. An industrial case study is carried out in Section 4 to demonstrate the performance of the proposed method. Finally, conclusions are made.

2. Supervised linear dynamic systems (SLDS)

2.1. Model structure

Similar to the traditional LDS model, the supervised LDS model is also a time-series state-space model. The main difference is that the additional quality information has been incorporated into the supervised LDS model, which serves as a supervised item for learning of the LDS model. Therefore, the relationships between quality variables and process variables can be modeled and captured for quality-related fault detection modeling. Generally, the LDS model consists of a latent linear Gaussian dynamic model and a linear Gaussian observation model [15]. The graphical representation of the LDS model is depicted in Fig. 1(a). In contrast, the graphical representation of the supervised LDS model is provided in Fig. 1(b). In these two figures, $\mathbf{h}_t \in R^{H\times 1}$ denotes the latent variable of LDS, which is also known as the state variable in the state space model, $\mathbf{x}_t \in R^{V\times 1}$ is the observed measurement vector of process variables, and $\mathbf{y}_t \in R^{L\times 1}$ is the vector of quality variables, where H, V, and L are numbers of latent variables, process variables and quality variables. It can be seen that additional quality information has been assumed in the supervised LDS model and also connected to the latent variables. Therefore, the supervised LDS model has a latent linear Gaussian dynamic model and two linear Gaussian observation models.

The model structure of SLDS can be described as follows:

$$\mathbf{h}_t = \mathbf{A}\mathbf{h}_{t-1} + \boldsymbol{\eta}_t^h \tag{1}$$

$$\mathbf{x}_t = \mathbf{B}_x\mathbf{h}_t + \boldsymbol{\eta}_t^x \tag{2}$$

$$\mathbf{y}_t = \mathbf{B}_y\mathbf{h}_t + \boldsymbol{\eta}_t^y \tag{3}$$