Chemometrics and Intelligent Laboratory Systems 138 (2014) 30–40

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

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

Fault detection and diagnosis for missing data systems with a three time-slice dynamic Bayesian network approach

Zhengdao Zhang *, Feilong Dong

Key Laboratory of Advanced Process Control for Light Industry, Ministry of Education, Jiangnan University, Wuxi 214122, China

ARTICLE INFO

Article history:
Received 9 April 2014
Received in revised form 1 July 2014
Accepted 9 July 2014
Available online 17 July 2014

Keywords:
Dynamic Bayesian network
Mixture Gaussian output
Fault detection and identification
Missing data
EM algorithm
Non-imputation

ABSTRACT

A multi-time-slice dynamic Bayesian network with a mixture of the Gaussian output (MT-DBNMG) based data-driven fault identification method is proposed to handle the missing data samples and the non-Gaussian process data. First, via introducing more time slices, a new dynamic Bayesian network structure with multi-time-slice is constructed which can describe the dependence between the current state and historic states. Second, a parameter learning strategy based on expectation maximization algorithm is deduced, from the complete historical data with the non-Gaussianity, to train the parameters of MT-DBNMG. Subsequently, for the missing measurements, an online non-imputation inference method for MT-DBNMG is proposed to conduct fault detection and identification. The effectiveness of the proposed approach is demonstrated by the continuous stirred tank reactor system and the Tennessee Eastman chemical process. The results show that the presented approach can accurately detect abnormal events, identify the fault, and is also robust to unknown noise.

1. Introduction

With the development of the industrial manufacturing alone with the advanced automation and control system, the complexity of systems is increased. Thus, process monitoring and fault detection are very important in modern industry. Traditional fault detection and diagnosis methods proposed in the literature can be classified as quantitative model-based approaches [1,2], qualitative knowledge-based approaches [3] and process data-driven approaches. Comparing with other methods, data-driven methods, especially multivariate statistical process monitoring methods, are developed and have attracted growing attention in the field [4,5]. Although the data-driven methods will have difficulty in diagnosing on-line faulty data with a much different magnitude or new faulty data, it is worth to implement them due to their well-known excellent properties which include no requirement of the in-depth process knowledge or the first principle of controlled systems, easy to collect mass data, and easy to apply to real processes of a rather large scale compared to other methods based on systems theory or rigorous process models.

Among the data-driven methods, principal component analysis (PCA) and partial least squares (PLSs) are the two most well-known techniques, and many extensions are further developed based on them (see [4,6,7] and references therein). However, the PCA/PLS methods depend on the assumption that the process data follow an approximate multivariate Gaussian distribution, which may not be unsatisfied in real industry such that the traditional PCA/PLS monitoring approaches become inappropriate [8]. The Gaussian mixture model (GMM) based monitoring approach can nicely handle the multi-Gaussianity which is approximated by multi-Gaussian distributions [8,9]. Compared with neural networks and other methods to handle the non-Gaussianity, the GMM only uses the historical data of process and avoids the performance degradation caused by the initial parameter selection. Moreover, the GMM can deal with the partly missing measurements. Therefore, the GMM algorithm has been introduced to automatically detect, isolate, and even forecast the faults [10,11].

From the perspective of statistical inference, fault detection and identification can be treated as an uncertain evidence inference problem, and Bayesian methods are the best tools to infer and formulate uncertainty of evidence [12,13]. Among Bayesian methods, static Bayesian network (BN) is suitable for dealing with conditional-dependent uncertain modeling and inference [14], and has been applied to different areas including fault detection and diagnosis (FDD) [6,15–19]. Moreover, BN can be combined with GMM to deal with the non-Gaussianity problem [20–22]. Nevertheless, since static BN does not

* Corresponding author. Tel.: +86 136 1618 7667.
E-mail address: wxzzd.dr@gmail.com (Z. Zhang).

http://dx.doi.org/10.1016/j.chemolab.2014.07.009
0169-7439/

Nomenclature

C_t	state node
M_t	mixture Gaussian node
$P(C_0)$	initial state probability distribution
$P(Y_t\|C_t)$	observation variable probability distribution
$P(C_t\|C_{t-1})$	state transition probability distribution
$y_{t,u}$	missing part of y_t
C_{Af}	feed concentration
T_c	coolant inlet temperature
q	reactor feed flow rate
ρ	density
$-\Delta H$	heat of reaction
$\phi_c(t)$	deactivation coefficient
Y_t	observed node
t	time instant
$y_{a,t}$	ath element of y_t
y_t	measurement vector at time t
$y_{t,o}$	observable part of y_t
C_A	effluent concentration
T	reactor temperature
T_f	feed temperature
V	reactor volume
k_0	pre-exponential factor
C_p,C_{pc}	heat capacity
$\phi_h(t)$	fouling coefficient

consider the temporal relationship among states of dynamic system [23], it is not suitable for expressing and dealing with the sequence relationship between samples in the output measurement time series of industrial process. As an extension of static BN, dynamic Bayesian network (DBN) combines static network with temporal information, and forms a probability model which can deal with timing sequence data [24]. Even if the study of the dynamic Bayesian network is not very mature, it has been applied in FDD [25–30]. Yu et al. developed a novel DBN based networked process monitoring approach which can accurately detect abnormal events, identify the fault propagation pathways, and diagnose the root cause variables [31]. At present, the major limitations of DBN based FDD methods lie in the fact that the network structure is designed depending on the prior process knowledge and process flow diagram.

Although data-driven approaches have been widely applied for FDD, the missing data output problem is the major challenge of these methods. Caused by a sudden mechanical breakdown, hardware sensor failure or data acquisition system malfunction, etc., missing data or irregularly sampled data is a common phenomenon in industrial practice [32,33]. The data losses and packet dropouts in communication networks are the increasing common sources for this missing data problem. However, the process monitoring and fault detection techniques in the presence of missing observation have not been well studied. Most of the existing data-driven methods, including neural networks, k-nearest neighbors and decision trees, are designed for well-conditioned data sets and cannot treat the incomplete data. Therefore, these methods will result in detection delays or failures in FDD with incomplete data.

In order to meet the requirements of the real-time fault detection and diagnosis, the problem of missing data should be considered, in other words, we must use the partly observed data to detect and diagnose fault if the data are missing at a certain moment. To achieve this end, some common data imputation approaches are used [13], such as mean substitution, regression imputation, multiple imputation, nearest neighborhood shift, support vector machine (SVM) and expectation maximization, to make the missing samples complete [34,35], then the complete estimated data is used to detect and diagnose fault. However, the variances of the data may be considerably changed with imputation, which was pointed out by Khatibisepehr et al. [13]. Moreover, there are three kinds of missing data mechanisms, i.e. missing at random, missing completely at random and not missing at random [35]. Unfortunately, neither a single imputation approach is suitable for all of the missing data mechanism assumptions. On the other hand, the imputed value is an approximation of the real value and the imputation error increases with the increase of missing rate. Therefore, the imputed value cannot take the place of the original one for fault detection and identification, because a bias repaired value may lead to a false alarm or missing alarm. Also, there were only limited literatures reported on the use of Bayesian networks for process fault detection and diagnosis with missing data. Unlike those heuristic schemes which deal with special problems, the concept of correntropy has been applied to develop more general methods based on existing models without resort to unnecessary efforts for outlier detection [37]. A similar idea also can be found in [38]. Therefore, we also focus on direct fault detection and diagnosis without imputation of missing data.

This paper proposes a multi-time-slice dynamic Bayesian network with mixture of Gaussian output (MT-DBNMG), and then, achieves fault detection and identification with the partially observed data for those systems that have the missing data output problem and the non-Gaussianity. The research of this paper is the expansion of our prophase research work [7,36]. In these literature, the two-time-slice dynamic Bayesian network with a mixture of Gaussian output (2T-DBNMG) is proposed to solve the problem of incomplete data and non-Gaussianity in processes. But it is not effective to detect incipient fault and has a large delay alarm rate for this incipient fault detection. Inspired by the high-order Markov model, we introduce more time slices into DBN, which can relate the current state with more historic data. The proposed algorithm can be divided into two steps. First, using the complete historical data, the parameter learning algorithm of MT-DBNMG is deduced based on expectation maximization method. Second, based on the trained MT-DBNMG, the inference algorithm is developed with partly missing data to accomplish the fault detection and identification. At last, the proposed approach is applied to monitor the continuous stirred-tank reactor (CSTR) and the Tennessee Eastman (TE) chemical process in this study and the presented method is demonstrated to be effective in monitoring and diagnosing for these two benchmark processes.

The remainder of this paper is organized as follows. In Section 1, after a brief introduction of the DBNMG model and missing data processing, the parameter learning algorithm of MT-DBNMG and the inference algorithm are deduced in detail. Then, Section 2 introduces the process of fault detection and identification based on the proposed MT-DBNMG. The presented method is applied to the CSTR and the TE process in Section 3. Finally, the conclusions are summarized in Section 4.

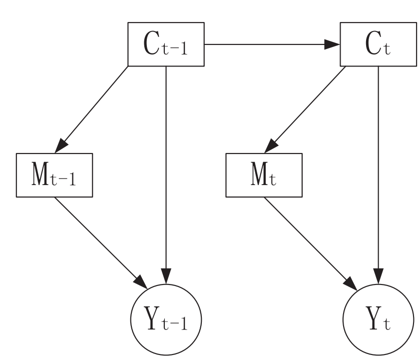

Fig. 1. The structure of the 2T-DBNMG models.