Chemical Engineering Science 189 (2018) 191–211

Contents lists available at ScienceDirect

Chemical Engineering Science

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

Process system fault detection and diagnosis using a hybrid technique

Md. Tanjin Amin, Syed Imtiaz, Faisal Khan *

Centre for Risk, Integrity and Safety Engineering (C-RISE), Faculty of Engineering & Applied Science, Memorial University, St. John's, NL A1B 3X5, Canada

HIGHLIGHTS

- A new hybrid methodology for fault detection and diagnosis.
- Methodology is efficient and effective in detecting and locating fault.
- Testing and validation of the methodology using two case studies.
- Recommendation to improve process system performance using the proposed hybrid methodology.

ARTICLE INFO

Article history:
Received 10 July 2017
Received in revised form 7 April 2018
Accepted 23 May 2018
Available online 26 May 2018

Keywords:
Process monitoring
Hybrid methodology
Principal component analysis
Bayesian network
Likelihood evidence

ABSTRACT

This paper presents a hybrid methodology to detect and diagnose the faults in dynamic processes based on principal component analysis (PCA) with T^2 statistics and a Bayesian network (BN). It deals with the uncertainty generated by the multivariate contribution plots and improves the diagnostic capacity by updating the BN with multiple likelihood evidence. It can diagnose the root cause of the process fault precisely as well as identify the fault propagation pathway. This methodology has been applied to the continuous stirred tank heater and the Tennessee Eastman chemical process for twelve fault scenarios. The result shows that it provides better diagnostic performance over conventional principal component analysis with hard evidence-based approaches.

1. Introduction

Monitoring is important in modern process industries due to their complexity, increased safety requirements and product quality demands (Chiang et al., 2000; Dong et al., 2015). Abnormal situations often occur in the process industries, resulting in huge economic loss (Nimo, 1995). An abnormal situation initiates with a fault during operation. A fault can be defined as the deviation of a process variable from an acceptable operational range (Venkatasubramanian et al., 2003a). Fault detection and diagnosis (FDD) is the first step in abnormal situation management (ASM) (Kresta et al., 1991). Data-driven multivariate statistical process monitoring (MSPM) techniques are widely used in process industries due to their effectiveness and ease of development. These techniques can extract features from highly correlated-high dimensional data to detect and diagnose the fault (Bakshi, 1998; Joe Qin, 2003; Kresta et al., 1991).

Principal component analysis (PCA) and partial least square (PLS) are the most widely used data-driven MSPM techniques, and are optimal for monitoring the process variables, following a multivariate Gaussian distribution (Kano et al., 2001; Rhoads and Montgomery, 1996). Independent component analysis (ICA) can capture a non-Gaussian feature by using higher order statistics like kurtosis and negentropy (Kano et al., 2003; Lee et al., 2004b). Modified ICA (MICA) can extract some dominant independent components (ICs), and improves the performance of conventional ICA (Lee et al., 2006; Zhang and Zhang, 2010). Kernel PCA (KPCA) has also been used to handle non-linearity (Cho et al., 2005; Lee et al., 2004a). All these tools need only a few historical data in normal operating condition (NOC) to estimate the control limit (CL), and information of faulty data behavior is not required for a successful performance. The artificial neural network (ANN) and support vector machine (SVM) have also been applied to process monitoring (Chiang et al., 2004; Mahadevan and Shah, 2009; Sorsa and Koivo, 1993; Weerasinghe et al., 1998). These techniques require pre-classified training data of both normal and faulty samples, and are suitable where in-depth fault information is available.

Despite many advancements in the field of MSPM, diagnosis of the root cause of a fault is still a challenge. A BN is an emerging tool in FDD which is becoming popular due to its ability to incorporate process data with expert opinion, and it has many successful

* Corresponding author.
E-mail address: fikhan@mun.ca (F. Khan).

https://doi.org/10.1016/j.ces.2018.05.045

applications in root cause diagnosis. (Liu and Chen, 2009) proposed a Bayesian classification based PCA approach, which can successfully detect and isolate faults. (Weidl et al., 2005) applied an object-oriented BN (OOBN) to digester fiber line to diagnose the root cause. Yu and Rashid, (2013) used a dynamic Bayesian network (DBN) based process monitoring approach for detecting the fault, diagnosing the root cause of the fault, and identifying the fault propagation pathway. They proposed the abnormality likelihood index (ALI) and dynamic Bayesian probability index (DBPI) to detect the fault. (Zhang and Dong, 2014) applied three time-slice DBN with a mixture of Gaussian output (3TDBN-MG) to handle some missing data and non-Gaussian process data.

Two or more techniques have been combined by many researchers to overcome the limitations of an individual method, which are popularly known as hybrid methods (Mylaraswamy and Venkatasubramanian, 1997; Venkatasubramanian et al., 2003b). Fault detection is performed at the first stage using MSPM tools (e.g. PCA, ICA etc.), and diagnosis is performed in the second stage by knowledge-based tools (e.g. BN, causality analysis etc.) utilizing the evidence generated by the first stage detection tool in terms of the multivariate contribution plot. Thus, higher monitoring accuracy is achieved. (Mallick and Imtiaz, 2013) integrated the BN with PCA to improve its diagnosis capacity. (Yu et al., 2014; Yu et al., 2015a, 2015b, 2015c) used MICA and the BN to diagnose the root cause in an unmonitored variable. They determined the causal relationships among process variables and the conditional probability tables (CPTs) from prior knowledge and historical data. (Gharahbagheri et al., 2017) applied KPCA with a BN to capture the non-Gaussian feature of process data, and to diagnose the root cause of a fault. Transfer entropy and Granger causality were used to identify the causal relationships, and prior knowledge was utilized to validate the network. CPTs were estimated using maximum likelihood estimation (MLE). A BN is acyclic in nature. But, chemical processes often have recycling variables. (Gharahbagheri et al., 2017; Yu et al. 2014; Yu et al., 2015a, 2015b, 2015c; Yu and Rashid, 2013) used duplicate dummy node to represent recycling variables in a BN.

In all the developed hybrid methods, first a data-based method (e.g. PCA, KPCA etc.) was used for fault detection and diagnosis. Often, the diagnosis information is incomplete, and points to a group of variables as the probable cause. The current practice is to use heuristic rules to reduce the information to one faulty variable. Usually, the variable with the highest contribution is taken as the faulty variable. Accordingly, a 100% faulty state is assigned (commonly known as hard evidence) to the highest contributing variable from the multivariate contribution plot. In this approach, the contribution of other variables to the fault is ignored. This has several limitations. Firstly, the evidence generated from the multivariate contribution plots is uncertain in nature. The BN is updated considering the observations certain. Secondly, two or more variables can have very close contributions to the fault; selecting a particular variable as faulty is challenging in those cases. Finally, the root cause variable can be an intermediate node, and it can have highest contribution in the multivariate contribution plots. This means that the statistical tool has already accurately done diagnosis. When a statistical tool diagnoses a root node as the root cause, a BN is not required. A BN is used to diagnose the root cause, when a child node has the highest contribution in the multivariate contribution plot. If hard evidence is used in this case, it causes diagnostic error. These limitations can be overcome by updating the BN with multiple uncertain evidence. However, this type of evidence has very limited use in fault diagnosis due to being complex in nature. Uncertain evidence in a BN has not yet been used in process fault diagnosis to authors' best knowledge. There are three main types of uncertain evidence- likelihood or virtual evidence, fixed probabilistic evidence and not-fixed probabilistic evidence (Mrad et al., 2015). In this work, we will focus only on the likelihood or virtual evidence.

The focus of this work is to improve the diagnostic capability of a BN based model using the information received from the multivariate contribution plots. A new hybrid methodology has been proposed integrating, principal component analysis (PCA) with T^2 statistics and the BN model. It is henceforth referred to as PCA-T^2-BN. This methodology considers multiple likelihood evidence to improve the diagnostic capacity of conventional PCA and PCA with BN (using a hard evidence-based approach). The methodology has been applied to two benchmark processes – the continuous stirred tank heater (CSTH) and Tennessee Eastman (TE) chemical process. It is proven that the proposed PCA-T^2-BN methodology performs better than convectional PCA and its derivatives.

This paper is organized as follows: Section 2 presents preliminaries on PCA and BN. Section 3 will discuss the methodology in detail. Section 4 will show the application and suitability of the proposed methodology. Results and discussion will be summarized in Section 5. The contribution, advantages, and future work scope will be discussed in Section 6.

2. Preliminaries

2.1. Principal component analysis (PCA)

PCA is a dimensionality reduction technique, and is also used for process fault detection and diagnosis (Bakshi, 1998). PCA projects data from a high dimensional data space to a lower dimensional subspace, which preserves the maximum variation of the original space in reduced dimensions (Mallick and Imtiaz, 2013). It provides a new set of uncorrelated variables from a set of correlated variables using linear transformation. If a process contains m variable and n samples, the data matrix can be represented as $X \in R^{n\times m}$. The covariance matrix, R, is given by:

$$R = \mathrm{cov}(X) = \frac{X^T X}{n-1} \tag{1}$$

Selection of the number of principal components (PCs) can be done either by SCREE plot or the cumulative percent variance (CPV) approach. SCREE is a graphical method, where eigenvalues are shown on X axis in descending order, and their correspondence variances are shown on Y axis. CPV is a simple approach. Eq. (2) represents it:

$$\mathrm{CPV}(b) = \frac{\sum_{i=1}^{b}\lambda_i}{\mathrm{trace}(R)} \times 100\% \tag{2}$$

b is the selected number of PCs. Usually, b is selected when CPV (b) $\geq$ 90%. The transformation matrix, P, is generated based on b. P contains m number of rows and b number of columns. The columns of P are called loadings. Details on PCA are available in (Garcia-Alvarez, 2009; Jackson, 2005).

Two types of statistics: Hotelling's T^2 and squared prediction error (SPE) statistics are typically used in PCA based process monitoring. T^2 measures the correlated distance between the center of the feature space and projected data samples, while SPE measures the Euclidean distance between the PC feature space and the residual space. More details are available in (Mallick and Imtiaz, 2013; Yu et al., 2015a, 2015b, 2015c; Zadakbar et al., 2012).

For PCA:

$$T_i^2 = t_i t_i^T \tag{3}$$

where $t_i = x_i P \Lambda_b^{-1/2} P^T$ is the contribution of the ith monitored sample. Λ is a diagonal matrix, which contains the eigenvalues in a descending order ($\lambda_1 > \lambda_2 > \ldots\ldots\ldots > \lambda_m$).