Computers and Chemical Engineering 64 (2014) 167–177

Contents lists available at ScienceDirect

Computers and Chemical Engineering

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

Root cause analysis in multivariate statistical process monitoring: Integrating reconstruction-based multivariate contribution analysis with fuzzy-signed directed graphs

Bo He [a,c], Tao Chen [b,*], Xianhui Yang [a]

[a] Department of Automation, Tsinghua University, Beijing 100084, China
[b] Department of Chemical and Process Engineering, University of Surrey, Guildford GU2 7XH, UK
[c] State Grid Energy Research Institute, Beijing 102209, China

ARTICLE INFO

Article history:
Received 16 May 2013
Received in revised form 6 January 2014
Accepted 16 February 2014
Available online 1 March 2014

Keywords:
Fault diagnosis
Fuzzy logic
Multivariate statistical process monitoring
Reconstruction-based multivariate contributions
Signed directed graph

ABSTRACT

Root cause analysis is an important method for fault diagnosis when used with multivariate statistical process monitoring (MSPM). Conventional contribution analysis in MSPM can only isolate the effects of the fault by pinpointing inconsistent variables, but not the underlying cause. By integrating reconstruction-based multivariate contribution analysis (RBMCA) with fuzzy-signed directed graph (SDG), this paper developed a hybrid fault diagnosis method to identify the root cause of the detected fault. First, a RBMCA-based fuzzy logic was proposed to represent the signs of the process variables. Then, the fuzzy logic was extended to examine the potential relationship from causes to effects in the form of the degree of truth (DoT). An efficient branch and bound algorithm was developed to search for the maximal DoT that explains the effect, and the corresponding causes can be identified. Except for the need to construct an SDG for the process, this method does not require historical data of known faults. The usefulness of the proposed method was demonstrated through a case study on the Tennessee Eastman benchmark problem.

1. Introduction

As competition becomes more and more intense in the process industry, coupled with the manufacturing plants getting more complex, the fault detection and diagnosis (FDD) technology has been recognised as an important tool to enable safe, efficient and environmentally benign operation of industrial processes (Frank, 1990; Venkatasubramanian, Rengaswamy, Yin, & Kavuri, 2003a). Meanwhile, with the rapid progress in sensor technology, distributed process control and data acquisition systems, more and more process variables can be measured on a routine basis. As a result, data-driven FDD methods have attracted substantial attention in both academia and industry (MacGregor & Cinar, 2012; Venkatasubramanian, Rengaswamy, Kavuri, & Yin, 2003b). Multivariate statistical process monitoring (MSPM) is a well-known data-driven FDD method (Qin, 2012; Yao & Gao, 2009). Traditional MSPM uses latent variable models, such as principal component analysis (PCA) and partial least squares (PLS), to detect faults and disturbances (Martin, Morris, & Zhang, 1996; Qin, 2003). The extensions of PCA/PLS have enabled the monitoring of processes with more complex behaviours (e.g. non-Gaussian, non-linear and non-stationary); a sample of these methods may include kernel PCA (Choi, Lee, Lee, Park, & Lee, 2005), independent component analysis (Kano, Tanaka, Hasebe, & Hashimoto, 2003), PCA with one-class support vector machine (Ge & Song, 2011), dynamic PCA (Ku, Storer, & Georgakis, 1995; Yao & Gao, 2007), among others. In practice, MSPM has been successfully implemented in many industrial applications (AlGhazzawi & Lennox, 2008; Kano & Nakagawa, 2008).

Following fault detection in MSPM are fault isolation and root cause analysis, which are important steps to help find the sources of the detected anomalies. To this end, if historical data are available for known types of fault, a wide range of data-drive methods can be used; for example the fault signatures (Yoon & MacGregor, 2001), fuzzy IF-THEN rules (Musulin, Yelamos, & Puigjaner, 2006), statistical distances and angles (Raich & Cinar, 1997), fault subspaces (Dunia & Qin, 1998), fisher discriminant analysis and support vector machine (Chiang, Kotanchek, & Kordon, 2004). He, Qin, and Wang (2005) introduced a new method, in which discriminant analysis was used to obtain fault directions, and the directions are combined with contribution analysis for fault diagnosis. However, the

* Corresponding author. Tel.: +44 1483 686593.
E-mail addresses: t.chen@surrey.ac.uk, tc.tao.chen@gmail.com (T. Chen).

http://dx.doi.org/10.1016/j.compchemeng.2014.02.014
0098-1354/

faulty data are usually difficult to acquire during process operations; after all, abnormal process operation is a rare event by definition. Furthermore, these data-driven fault diagnosis methods may face difficulty in handling unknown faults, which have not occurred in the limited historical data.

An alternative approach is contribution analysis, which aims to isolate the variables that are the most responsible to the detected fault. Contribution analysis does not need prior knowledge of the faults, and thus is suitable for analysing unknown fault types (Miller, Swanson, & Heckler, 1998). In addition, confidence limits can be established for contribution plots to further help discriminate the faulty variables from non-faulty ones (Westerhuis, Gurden, & Smilde, 2000). However, the conventional contribution analysis is known to suffer from the fault "smearing" effect (Alcala & Qin, 2009). To address this issue, the reconstruction-based (Alcala & Qin, 2009) and missing variable (Chen & Sun, 2009) contribution methods have been proposed. In addition, the traditional contribution analysis focuses on the effect of individual variables, ignoring the interaction between variables. This limitation has recently been recognised, and the joint analysis of multiple variables under PCA models has been reported by using variable reconstruction (Liu, 2012), and missing variable analysis with a branch and bound (BnB) optimisation algorithm (Kariwala, Odiowei, Cao, & Chen, 2010). We have followed this line of research to develop a generic, reconstruction-based multivariate contribution analysis (RBMCA) framework, which can be used with any process monitoring model (He, Yang, Chen, & Zhang, 2012). Later, this generic RBMCA approach has been further improved by introducing an L_1-penalty term for variable reconstruction (PRBMCA), prior to using BnB for optimisation (He, Zhang, Chen, & Yang, 2013).

Despite being effective in identifying the most responsible variables to the detected fault, contribution analysis still cannot directly reveal what fault has occurred and/or the root cause, which would be important information to help decide appropriate course of actions. Therefore, an approach to automatically interpret the variable contribution and to identify the root cause is needed to fill this gap. It appears that such methods would require the knowledge about the process being monitored, and the knowledge can be effectively represented in qualitative models.

Combing data-driven methods with qualitative modelling is not a new concept. Typical qualitative models include signed directed graph (SDG) (Lee, Han, & Yoon, 2004; Vedam & Venkatasubramanian, 1999), expert system (Norvilas, Negiz, DeCicco, & Cinar, 2000), cause–effect models (Chiang & Braatz, 2003; Leung & Romagnoli, 2002), plant connectivity modelling using extensible markup language (XML) (Thambirajah, Benabbas, Bauer, & Thornhill, 2009), among others. To handle the multivariate nature of process variables, the PCA-SDG hybrid (Vedam & Venkatasubramanian, 1999) is a well-known and effective method for fault detection and diagnosis. PCA-SDG relies on PCA to detect the fault, and then performs contribution analysis to determine the signs of the nodes (process variables) to be used in SDG. Then, a backward–forward propagation algorithm is used to search for the root cause. However, as discussed previously, the conventional contribution plot suffers from "fault smearing" (Alcala & Qin, 2009) and faces difficulty in handling multivariate faults (He et al., 2012). In addition, the contribution threshold, which is used for determining the signs of nodes in SDG, is a tuning parameter and has strong impact on results. Furthermore, the crisp logic (as opposed to fuzzy logic) used in PCA-SDG, i.e. labelling the nodes as "positive", "normal" or "negative", does not reflect the quantitative status of a variable. For example, one node labelled "positive" may be closer to the "normal" situation than another "positive" node; however this cannot be modelled in crisp logic, resulting in poor resolution of diagnosis (Han, Shih, & Lee, 1994).

In this paper, a root cause analysis method integrating PRBMCA with a fuzzy-SDG based reasoning scheme is proposed. PRBMCA determines the effect of the fault, in terms of the combinations of isolated faulty variables with corresponding signs; it inherits the advantage of MCA to avoid "fault smearing" and tuning the threshold for variable contribution. A reconstruction-based fuzzy logic is adopted to represent the signs of nodes in the form of degree of truth (DoT). The DoT function is further expanded to evaluate the "consistency" of "paths" between nodes, serving as the measure of likelihood of going from the cause node to the effect in SDG. A depth-first search algorithm with BnB is developed to search for the paths from candidate cause nodes to effect nodes with respect to the DoT, and the candidate cause node with the maximal DoT of consistency is considered as the root cause. The use of fuzzy logic provides the quantitative information, in terms of DoT, of each cause. Moreover, in comparison with pure data-driven diagnosis methods, the proposed approach does not require historical data of known faults.

The hybrid fault diagnosis method also inherits another important feature of PRBMCA, i.e. it is a generic framework that is applicable to any process monitoring model, provided that a variable reconstruction algorithm can be developed for that particular model. Since the primary focus of this study is not fault detection but diagnosis, a simple probabilistic PCA (PPCA) is used as the process monitoring model (Chen & Sun, 2009; Kim & Lee, 2003) to demonstrate how the proposed method can help identify the root cause. Clearly, diagnosis is only possible for the faults that can be detected by the monitoring model. We will briefly discuss how other MSPM models can be incorporated in the proposed framework.

The rest of this paper is organised as follows. A brief review of RBMCA and PRBMCA is presented in Section 2. Section 3 presents the detailed the root cause analysis method, which combines PRBMCA and fuzzy-SDG. Section 4 discusses the results from the case study on the simulated Tennessee Eastman benchmark problem, and Section 5 concludes this paper.

2. Fault isolation using reconstruction-based multivariate contribution analysis

RBMCA is a general framework for isolating faulty variables and is applicable to any process monitoring model, provided that a variable reconstruction algorithm can be developed for that model (He et al., 2012). The basic idea of RBMCA is to search for the combination of process variables that contribute the most to the detected fault. A BnB algorithm was proposed to efficiently solve the combinatorial optimisation problem. Later, an L_1 penalty was introduced to the variable reconstruction step of RBMCA (thus named PRBMCA), before using the BnB algorithm (He et al., 2013). PRBMCA tends to reduce the number of isolated faulty variables thus to give clearer interpretation of the results. The computation demand of PRBMCA is also significantly less than that of RBMCA, because the L_1 penalty removes the improbable combinations of variables that would have to be evaluated in RBMCA. In this section, both versions are briefly reviewed.

2.1. Reconstruction-based multivariate contribution analysis

Suppose that a statistical model ($\mathbf{M}$) has been developed to represent the normal process operations, and a monitoring statistic (D) with appropriate control limit CL has been established to distinguish normal samples from faulty ones. We use $D(\mathbf{x}|\mathbf{M})$ to denote the monitoring statistic of an n-dimensional sample $\mathbf{x}$ under model $\mathbf{M}$. Thus, the process is considered out-of-control if $D(\mathbf{x}|\mathbf{M}) > \mathrm{CL}$, and in-control if $D(\mathbf{x}|\mathbf{M}) \le \mathrm{CL}$. To enable variable reconstruction, a faulty