Computers & Industrial Engineering 88 (2015) 63–77

Contents lists available at ScienceDirect

Computers & Industrial Engineering

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

Improved principal component analysis for anomaly detection: Application to an emergency department ☆

Fouzi Harrou [a], Farid Kadri [b,*], Sondes Chaabane [b], Christian Tahon [b], Ying Sun [a]

[a] CEMSE Division, King Abdullah University of Science and Technology, Thuwal 23955-6900 Saudi Arabia
[b] LAMIH, UMR CNRS 8201, University of Valenciennes and Hainaut-Cambrésis, UVHC, Le Mont Houy, 59313 Valenciennes Cedex, France

ARTICLE INFO

Article history:
Received 20 June 2014
Received in revised form 22 June 2015
Accepted 25 June 2015
Available online 3 July 2015

Keywords:
Statistical anomaly detection
Multivariate CUSUM
Emergency department
Abnormal situation

ABSTRACT

Monitoring of production systems, such as those in hospitals, is primordial for ensuring the best management and maintenance desired product quality. Detection of emergent abnormalities allows preemptive actions that can prevent more serious consequences. Principal component analysis (PCA)-based anomaly-detection approach has been used successfully for monitoring systems with highly correlated variables. However, conventional PCA-based detection indices, such as the Hotelling's T^2 and the Q statistics, are ill suited to detect small abnormalities because they use only information from the most recent observations. Other multivariate statistical metrics, such as the multivariate cumulative sum (MCUSUM) control scheme, are more suitable for detection small anomalies. In this paper, a generic anomaly detection scheme based on PCA is proposed to monitor demands to an emergency department. In such a framework, the MCUSUM control chart is applied to the uncorrelated residuals obtained from the PCA model. The proposed PCA-based MCUSUM anomaly detection strategy is successfully applied to the practical data collected from the database of the pediatric emergency department in the Lille Regional Hospital Centre, France. The detection results evidence that the proposed method is more effective than the conventional PCA-based anomaly-detection methods.

1. Introduction

In today's competitive atmosphere, there is growing demand for enhanced process safety to maintain the safe and reliable process operations that are required to meet the higher expectations of process performances and product quality. Process monitoring, such as reliable detection and diagnosis of anomalies, is an important element to process safety and ultimately high quality-products. For example, a survey performed by Nimmo (1995) showed that the petrochemical industry in the USA could increase profits up to 10 billion USD per year if anomalies in their monitored process could be suitably detected and diagnosed. When an anomaly occurs in a monitored process, the monitoring process must immediately detect the anomaly and assist in determining if the process can continue to operate normally (Isermann, 2006).

Management and monitoring in hospital emergency department (ED) systems are among the most growing areas of concern for many countries (Cochran & Broyles, 2010; Aboueljinane, Sahin, & Jemai, 2013). In particular, monitoring patient flow in EDs is a critical issue for many hospital administrations in France and worldwide because often leads to strain situations (Kadri, Chaabane, & Tahon, 2014; Kadri et al., 2013). In France, between 1990 and 1998, the annual number of ED demand increased by 43 (Baubeau et al., 2000), and according to the annual public report of Medical Emergencies (Rapport de la Cour des Comptes, 2006), the 7 million patients that visited EDs in France in 1990 had doubled by 2004. Between 1993 and 2003, the Institute of Medicine of the National Academies (I. of Medicine Committee on the Future of Emergency Care in the US Health System et al., 2006) published a report highlighting a disparity in the US between need and availability of ED facilities: the number of patients who visited EDs increased by approximatively 26%, while the number of EDs decreased approximatively 9% (Kellermann, 2006). Patient influx can generate strain situations that affect building safety and reliability of EDs (Kadri, Harrou, Chaabane, & Tahon, 2014). Therefore, detecting abnormal demands On EDs will contribute to improving the management of patients and medical resources (human and material). The early detection of abnormal demands in EDs promotes reactive control which can help to prevent strain situations, specifically limit the consequences, and allows efficient resource

☆ This manuscript was processed by Area Editor H. Brian Hwarng.

* Corresponding author.
E-mail addresses: fouzi.harrou1@gmail.com (F. Harrou), farid.kadri@univ-valenciennes.fr (F. Kadri), sondes.chaabane@univ-valenciennes.fr (S. Chaabane), christian.tahon@univ-valenciennes.fr (C. Tahon), ying.sun@kaust.edu.sa (Y. Sun).

http://dx.doi.org/10.1016/j.cie.2015.06.020
0360-8352/

allocation. Thus, the goal of this study is to develop an anomaly-detection strategy that detects abnormal ED demands.

An anomaly is defined as an unpermitted deviation of at least one characteristic property of a variable from its acceptable behavior. Therefore, the anomaly is a state that may lead to a malfunction in the system (Isermann, 2005). Two main kinds of anomalies can be distinguished by the way they affect the monitored system: gradual and abrupt anomalies. In an ED, slow or gradual anomalies usually indicate a slow increasing demand or patient flow, while abrupt anomalies, are characterized by sudden increasing demands (patient flow). Here, we address the problem of detecting abrupt and gradual anomalies encountered by various anomaly-detection techniques that have been developed for the safe operation of systems or processes (Harrou, Fillatre, & Nikiforov, 2014; Hwang, Kim, Kim, & Seah, 2010; Qin, 2012; Isermann, 2006; Venkatasubramanian, Rengaswamy, Kavuri, & Yin, 2003). Model-based methods are implemented by measuring the dissimilarity between measured process variables and information obtained from explicit process models. Unfortunately, building a precise model for a monitored process can be challenging. When there is no process model, multivariate latent variable regression (LVR) methods, such as partial least square (PLS) regression and principal component analysis (PCA), have been used successfully in process monitoring because they can effectively deal with highly correlated process variables (Qin, 2012; Harrou, Nounou, Nounou, & Madakyaru, 2013). A number of, the characteristics interest to the operational framework of EDs make it difficult to accurately model their behavior (Kadri et al., 2014; Bhattacharjee & Ray, 2014): (i) they are dynamic and disturbed environments, (ii) some elements that characterize care activity are non-deterministic (e.g. processing time, waiting time, and additional examinations), (iii) each patient requires treatment that is specific to their pathology and involves different routes within the ED, and (iv) no assumptions can be made concerning the types of emergency treatment that patients will require within a given period of time. For these reasons, PCA a well-known multivariate data analysis technique, can be used because it requires no prior knowledge about the process model (MacGregor & Kourti, 1995).

This paper aims to present a statistical anomaly-detection scheme based on a PCA model that can detect abnormal ED demands. Our basis for this approach was conceived by PCA's reputation as a linear dimensionality reduction modeling technique, which is favorable when processing data sets that have a high degree of cross correlation among the variables (Qin, 2012). The basic concept behind PCA is to reduce the dimensionality of highly correlated data, while retaining the maximum possible amount of variability present in the original data set (MacGregor & Kourti, 1995). Detecting an anomaly based on PCA has been widely used in practice because the only information needed is a good historical database describing the normal process operation. In such a framework, PCA and its extensions have successfully been applied for detecting anomalies in various disciplines (Wise & Gallagher, 1996; Simoglou, Martin, & Morris, 1997; Yu, 2011). However, PCA-based monitoring statistics, such as T^2 and Q statistics, are unsuitable for detecting changes resulting from small anomalies (Montgomery, 2005). Unlike PCA-based statistics, multivariate statistical process control charts, such as the multivariate cumulative sum (MCUSUM) (Montgomery, 2005; Bersimis, Psarakis, & Panaretos, 2007; Crosier, 1988), have shown a greater aptitude to detect small anomalies in the process mean. Because the MCUSUM control scheme better detects small faults in the process mean (Montgomery, 2005), the main objective of this paper is to combine the advantages of the MCUSUM and PCA method to enhance their performances and widen their applicability in practice. More specifically, this paper proposes a PCA-based MCUSUM fault detection methodology for identifying signs of abnormal situations caused by abnormal demand for the Pediatric Emergency Department (PED) in the Lille Regional Hospital Centre, France.

The remainder of this paper is organized as follows. Section 2 briefly describes the PCA theory and how it can be used in anomaly detection, and Section 3 explain the MCUSUM control scheme that is commonly used in quality control. Next, the proposed PCA-based MCUSUM anomaly-detection approach that integrates PCA modeling and MCUSUM control scheme is presented in Section 4. Section 5 presents the application of the proposed methodology in the detection of abnormal situations in the PED in the Lille Regional Hospital Centre, France, and describes the practical data set used in the case study. Section 6 presents results of the proposed PCA-based MCUSUM anomaly-detection methodology and compare them with that of conventional PCA-based anomaly-detection. Finally, Section 7 reviews the main points discussed in this work and concludes the study.

2. PCA based statistical monitoring

PCA has a reputation for its usefulness in multivariate statistical techniques for reducing the dimensionality of the process data. Linear PCAs are valued for their ability to manage collinear data with several variables. In its general form, PCAs find the latent variables (not directly observed or measured) from the process data by capturing the largest variability in the data. In this Section we present the PCA theory and how it can be used in anomaly-detection.

2.1. PCA modeling

Let us consider the following raw data matrix $\mathbf{X} = \left[x_1^T, \ldots, x_n^T\right]^T \in \mathbb{R}^{n\times m}$ consisting of n observations and m correlated variables, where $x_n \in \mathbb{R}^n$. Before computing the PCA model, the raw data matrix $\mathbf{X}$ is usually pre-processed by scaling every variable to have zero mean and unit variance. This is because variables are measured with various means and standard deviations in different units. This pre-processing step puts all variables on an equal basis for analysis (Ralston, DePuy, & Graham, 2001). Let $\mathbf{X}_s$ denote the standardized matrix $\mathbf{X}$. By using singular value decomposition (SVD), PCA transforms the data matrix $\mathbf{X}_s$ into a new matrix $\mathbf{T} = [t_1\ t_2\ \cdots\ t_m] \in \mathbb{R}^{n\times m}$ of uncorrelated variable called score or principal components, $t_1 \in \mathbb{R}^n$. Each new variable is a linear combination of the original variables, so that $\mathbf{T}$ is obtained from $\mathbf{X}_s$ by orthogonal transformations (rotations) designed by $\mathbf{P} = [p_1\ p_2\ \cdots\ p_m] \in \mathbb{R}^{m\times m}$, which is given as the following:

$$\mathbf{X}_s = \mathbf{T}\mathbf{P}^T. \tag{1}$$

The column vectors $p_i \in R^m$ of the matrix $\mathbf{P} \in R^{m\times m}$ (also known as the loading vectors) are formed by the eigenvectors associated with the covariance matrix of $\mathbf{X}_s$, i.e., Σ. The covariance matrix, Σ, is defined as follows:

$$\Sigma = \frac{1}{n-1}\mathbf{X}_s^T\mathbf{X}_s = P\Lambda P^T \quad \text{with} \quad PP^T = P^TP = I_n, \tag{2}$$

where $\Lambda = diag(\lambda_1, \ldots, \lambda_m)$ is a diagonal matrix containing the eigenvalues in a decreasing order $(\lambda_1 > \lambda_2 > \cdots > \lambda_m)$, and I_n is the identity matrix (Jackson & Mudholkar, 1979).

For collinear processes, the dimensionality reduction of the m-dimensional space is obtained by retaining only the first (l) largest PCs, which correspond to the largest eigenvalues of the covariance matrix. The first (l) largest PCs normally describe the most of the variance in the data. The smallest PCs are considered noise contributors. An important step in the building of PCA model