Annual Reviews in Control 000 (2016) 1–11

Contents lists available at ScienceDirect

Annual Reviews in Control

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

Review

Perspectives on process monitoring of industrial systems☆,☆☆

Kristen Severson, Paphonwit Chaiwatanodom, Richard D. Braatz*

Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

ARTICLE INFO

Article history:
Received 1 September 2016
Accepted 1 September 2016
Available online xxx

Keywords:
Fault detection
Fault identification
Fault tolerant control
Process monitoring

ABSTRACT

Process monitoring systems are necessary for ensuring the long-term reliability of the operation of industrial systems. This article provides some perspectives on progress in the design of process monitoring systems over the last twenty years. Methods for each step of the process monitoring loop are summarized. The challenges in the field and opportunities for future research are discussed. When looking into the future, it is argued that advances are likely to come from combining different methods to exploit the strengths of various techniques while minimizing their weaknesses.

1. Introduction

Process monitoring is an important component in the long-term reliable operation of any automated controlled system. To distinguish between different types of disruptions on operations, this article adopts the definitions of Isermann and Ballé (1997). A **disturbance** is an unknown and uncontrolled input acting on a system. A **fault** is an unpermitted deviation of at least one characteristic property or parameter of the system from the acceptable/usual/standard operating conditions. A **failure** is a permanent interruption of a system's ability to perform a required function under specified operating conditions. Traditional control systems are designed to return the system to normal operations in the presence of disturbances but not in the presence of faults or failures. *Fault-tolerant control* (FTC) systems refer to control systems that have been designed to explicitly account for some class of specified faults in the closed-loop system. FTC systems must act in the time between a fault and a system failure.

In chemical systems, a fault is an extreme event such as catalyst deactivation, valve blockage or compressor failure. Due to the increasing complexity of facilities, faults are inevitable and occur more often. Monitoring is complicated by recycle streams that cause bidirectional interactions as well as by control systems which can mask the effect of faults. Additionally faults will commonly occur together, known as multiple faults (see Fig. 1). However, even a relatively simple modern facility, in terms of its operations, will have a large sensor network which can be used for process monitoring (see Fig. 2). The key of fault detection and diagnosis (FDD) is how to use these sensors effectively to minimize the impact of faults.

Many process monitoring systems are implemented in the form of a loop that consists of fault detection, fault isolation, fault identification, and process recovery (see Fig. 3). Sometimes the combined steps of fault isolation and identification are referred to as fault diagnosis. The steps are to progressively determine: (1) whether a fault occurred, (2) the location and time of the fault, (3) the magnitude the fault, and (4) how to reverse the effects of the fault (Gertler, 1998).

Process monitoring has been a growing field for nearly a half century. Relevant works on process monitoring in the 1970s include the application by Mehra and Peschon (1971) of systems and statistical decision theory to dynamic systems, the review paper by Willsky (1976) on publications up to the mid 1970s, and the textbook by Himmelblau (1978). Over the years, much of the literature has been focused on particular applications including to aerospace, chemical, nuclear, and automotive systems (Hwang, Kim, Kim, & Seah, 2010). The growing complexity and degree of integration in these systems has increased the possibility that faults occurring locally somewhere in a system can have their effects propagate to other parts of the system, and has made the consequences of designing a poor process monitoring system greater, therefore making the design of process monitoring systems more challenging. As such, many reviews have been published over the last twenty years, e.g. (Alcala & Qin, 2011; Frank & Ding, 1997; Hwang et al., 2010; Isermann, 2005; Isermann & Ballé, 1997; Qin, 2003; Russell, Chiang, & Braatz, 2000a; Venkatasubramanian, Rengaswamy, & Kavuri, 2003a; Venkatasubramanian, Rengaswamy, Kavuri, & Yin, 2003b; Venkatasubramanian, Rengaswamy, Yin, & Kavuri, 2003c; Yin, Ding, Haghani, Hao, & Zhang, 2012).

☆ BP is acknowledged for funding.
☆☆ A preliminary version of this manuscript was presented at the IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes (Severson, Chaiwatanodom, and Braatz, 2015a).

* Corresponding author.
E-mail addresses: kseverso@mit.edu (K. Severson), pchaiwat@mit.edu (P. Chaiwatanodom), braatz@mit.edu (R.D. Braatz).

http://dx.doi.org/10.1016/j.arcontrol.2016.09.001
1367-5788/

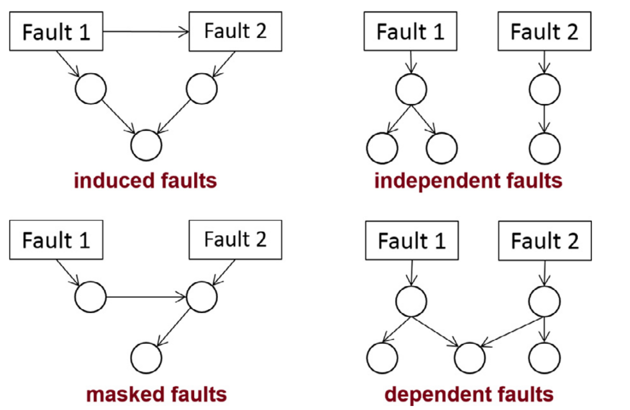

Fig. 1. The four classes of multiple faults (Chiang et al., 2015).

This article does not review the entire process monitoring field which, according to the Web of Science in March 2015, has had over 34,000 publications since the 1970s. This article provides some perspectives on the current state of process monitoring systems as well as current challenges and promising future directions for the field.

2. Process monitoring – background

Modern process monitoring systems are designed based on a model of some form that is developed using process data. The model allows process operators to make informed decisions about whether or not there is a fault. Different fault detection methods provide information of different quality and quantity to the fault diagnosis steps. In this section, each step in the process monitoring loop is presented.

2.1. Fault detection

The design of a fault detection system generally begins with the development of a model that characterizes the normal operating signature of a process. Faults are then typically defined as a deviation from this normal operation above a threshold. As such, the design of a fault detection system can be described as consisting of two steps: building a process model and choosing metrics to test for faults. Active fault detection and identification is an exception to this pattern and is discussed later in the section on process monitoring.

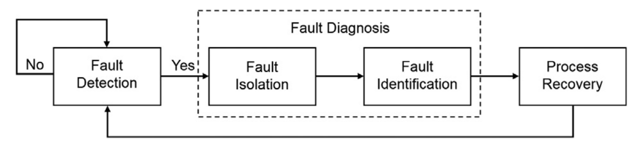

Fig. 3. Process monitoring loop (Isermann & Ballé, 1997; Russell et al., 2000a).

Many types of process models have been employed in fault detection. Principal component analysis (PCA) is one of the most commonly applied fault detection methods for industrial systems. PCA is a linear dimensionality reduction technique that produces lower dimensional representations of the original data that maximize the retained variance (Hotelling, 1933; Jolliffe, 2002). In the absence of noise and disturbances, data from normal operating conditions operate in a much lower dimensional manifold due to physical, chemical, and biological constraints such as Euler's laws of motion, stoichiometry in chemical and/or metabolic reaction networks, and mass, energy, molar species, and fluid momentum balances. In the presence of noise and disturbances, the data from normal operating conditions will approximately lie within a lower dimensional manifold, and data-based dimensionality reduction techniques such as PCA attempt to construct the manifold purely from data.

Variance is a useful metric for fault detection, since it is often reasonable to assume that an outlier as compared to historical operation would indicate a fault. PCA calculates a set of orthogonal vectors, called *loading vectors*, ordered by the amount of variance explained in each loading vector direction using a singular value decomposition. This set of vectors is then truncated, retaining the columns corresponding to the largest singular values. New observations can then be projected into lower dimensional space using the reduced set of loading vectors. The aim of this dimensionality decrease is to keep systematic variations while removing random variations (Wise, Ricker, Veltkamp, & Kowalski, 1990). The technique can be extended to nonlinear systems by us-

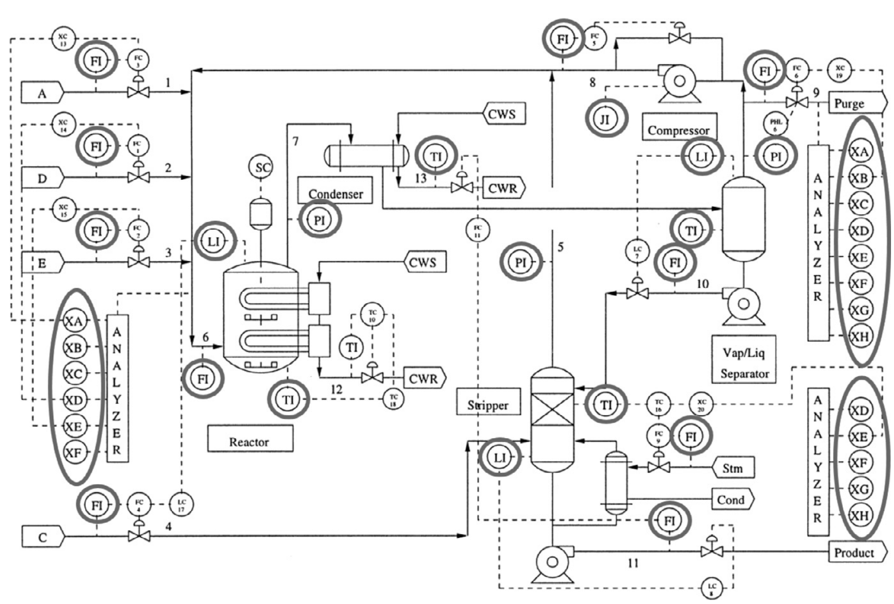

Fig. 2. The process diagram for the Tennessee Eastman (TE) benchmark problem (Downs and Vogel, 1993). The process is a reactor/separator/recycle with two simultaneous gas-liquid exothermic reactions. The process has 12 valves for manipulation and 41 measurements for monitoring and control. The sensors are circled in red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Please cite this article as: K. Severson et al., Perspectives on process monitoring of industrial systems, Annual Reviews in Control (2016), http://dx.doi.org/10.1016/j.arcontrol.2016.09.001