Chemometrics and Intelligent Laboratory Systems 148 (2015) 51–59

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

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

Hidden Markov model-based approach for multimode process monitoring

Fan Wang, Shuai Tan, Hongbo Shi *

Key Laboratory of Advanced Control and Optimization for Chemical Processes of Ministry of Education, East China University of Science and Technology, Shanghai 200237, PR China

ARTICLE INFO

Article history:
Received 20 May 2015
Received in revised form 6 August 2015
Accepted 29 August 2015
Available online 6 September 2015

Editor: P.K. Hopke

Keywords:
Hidden Markov model
Fault detection
Multimode
Mode identification

ABSTRACT

Many industrial processes have multiple operating modes due to various factors, such as alterations of feedstock and compositions, different manufacturing strategies, fluctuations in the external environment, and various product specifications. There are just a few literatures concerning hidden Markov model in multimode process monitoring. And HMM has not been explored to deal with transitional modes. Besides, those monitoring methods fail to take advantage of internal elements of HMM. In this article, a novel monitoring scheme based on hidden Markov mode (HMM) is proposed for multimode process with transitions. To begin with, a hidden Markov model is trained on the basis of the measurement data. Then a probability ratio strategy based on HMM is developed to identify stable modes and transitional modes. Further, the Viterbi algorithm classifies samples into various modes and a new monitoring indication is built based on the elements of HMM in each mode for fault detection. At last, the effectiveness of the proposed method is demonstrated through a numerical simulation and the Tennessee Eastman process.

1. Introduction

Fault detection is one of the most important tasks for the successful operation of any chemical process as it is essential to ensure the stable operation of chemical plants, maintain product quality at desired grades, optimize production profit, and improve plant safety and environmental sustainability. Multivariate statistical process monitoring (MSPM) is growing popular due to the availability of big data generated by large-scale chemical processes. MSPM techniques, such as principal component analysis (PCA) and partial least-squares (PLS), have been intensively investigated and widely applied to chemical process monitoring with success [1–3]. To overcome the limitations, including nonlinearity and dynamics of processes, improvements on traditional methods and other complementary MSPM have been proposed, for example, ICA, KPCA, DPCA, SVDD, and manifold-learning methods [4–9]. However, it should be noted that these approaches often assume that the process operates within a single mode. In fact, complex chemical manufacturing processes frequently work at multiple operating modes due to various factors, such as alterations of feedstock and compositions, different manufacturing strategies, fluctuations in the external environment, and various product specifications. Consequently, mode shifts greatly confine the scope of applications of the conventional methods.

In order to monitor multiple operating modes processes, some research efforts have been reported. Chen and Liu [10] proposed a mixture principal component analysis (MixPCA) model detector, which is a group of PCA models. Zhao [11,12] handled processes with multiple operating modes through multiple PCA models and multiple PLS models. These approaches constructed multiple models for different operating modes and the monitoring task is carried out then. Based on this idea, Liu [13], Yoo et al. [14], Ng et al. [15], Khediri et al. [16], and Zhu et al. [17] came up with other techniques. Recently, the Bayesian framework and the combination strategy were used. Yu and Qin [18,19] developed a finite Gaussian mixture model (FGMM) and used the Bayesian inference-based probability (BIP) index for process monitoring. Ge and Song [20] presented a mixture Bayesian regularization method of PPCA. These methods integrated monitoring results of different operating modes in a probabilistic manner. However, transitional mode between two stable modes is ignored by the above techniques. Zhao et al. [21] put forward a soft-transition multiple PCA (STMPCA) modeling method to overcome the hard-partition and misclassification problems of the stage-based sub-PCA modeling method. This is based on the idea that process transition can be detected by analyzing changes in the loading matrices. After the initial stage division, a transition model is expressed as a weighted sum of two phase models. Then different monitoring models are established in transition regions. Yao et al. [22] improved STMPCA by using angle-based transition identification and modeling strategy. Tan et al. [23] employed the similarity of data characteristics of the sampling window to realize mode identification. Then a transition can be modeled by a series of sub-models for monitoring.

In contrast to traditional approaches, hidden Markov models (HMM)-based methods have not been explored widely and deeply in multimode process monitoring. Yu [24] applied HMM to model complicated data distribution with nonlinear and multimodal features and developed two HMM-based process monitoring quantification indications. M. Rashid and Yu [25] presented an HMM-based ICA approach

* Corresponding author. Tel: +86 021 64252189; fax: +86 021 64252349.
E-mail address: hbshi@ecust.edu.cn (H. Shi).

http://dx.doi.org/10.1016/j.chemolab.2015.08.025
0169-7439/

for chemical processes with multiple operating conditions and inherent system uncertainty. The hidden Markov model is built from measurement data to estimate dynamic mode sequence. Further, the localized ICA models are developed to characterize different operating modes for online process monitoring purpose. The monitored samples can be classified into the corresponding modes by HMM-based hidden state estimation. Ning et al. [26] incorporated HMM and the window-based SPA monitoring method in a single framework where HMM is also adopted to identify the operating mode.

As far as we know, there exists no work applying HMM to deal with transitional mode identification problem. In this paper, a probability ratio strategy based on HMM is proposed to identify stable modes and transitional modes. Then the Viterbi algorithm is employed to carry out mode identification of stable modes. After mode identification, a novel monitoring indictor based on HMM is used to commit the monitoring task for stable modes. It combines the Mahalanobis distance and the observation probability matrix. Some other traditional methods are responsible for modeling and monitoring transitional modes.

The remainder of this article is organized as follows. First, an introduction to hidden Markov models is presented. The elements of this mathematical model are described. Subsequently, the proposed probability ratio strategy is developed using factors of HMM for mode identification. Then, the new monitoring indication is mathematically formulated with the related analyses. So far, the proposed monitoring framework is elaborated. After that, two examples, including a numerical simulation and the Tennessee Eastman (TE) process, are employed to show the effectiveness of the proposed method. Finally, conclusions of this work are given in the last section.

2. Hidden Markov model

HMM is a kind of probabilistic model extended from Markov chains to generate the statistically inferential information on a series of state sequences [27]. In general, it contains finite numbers of hidden states, where each state outputs an observation at a certain time point. Each hidden state is characterized by two sets of probabilities: a transition probability between two states and an observation probability distribution. Not like Markov chains, HMMs are doubly stochastic processes: the stochastic transition between one state to another state and stochastic output observations generated at each state. Although the actual sequence of states cannot be directly observable and is hidden from the observer, it can be inferred from the observations. An HMM has the following key ingredients [28].

(1) The hidden states:

$$\boldsymbol{S} = \{\boldsymbol{S}_1, \boldsymbol{S}_2, \ldots, \boldsymbol{S}_N\} \tag{1}$$

where $\boldsymbol{N}$ denotes the number of hidden states.

(2) State transition probability distribution:

$$\boldsymbol{A} = \{\boldsymbol{a}_{ij}\}, 1 \le \boldsymbol{i}, \boldsymbol{j} \le \boldsymbol{N} \tag{2}$$

where $\boldsymbol{a}_{ij} = \Pr(\boldsymbol{q}_{t+1} = \boldsymbol{S}_j | \boldsymbol{q}_t = \boldsymbol{S}_i)$, $1 \le \boldsymbol{i}, \boldsymbol{j} \le \boldsymbol{N}$, and $\boldsymbol{q}_t$ is the hidden state at time $\boldsymbol{t}$.

(3) The observations:

$$\boldsymbol{O} = \{\boldsymbol{O}_1, \boldsymbol{O}_2, \ldots, \boldsymbol{O}_M\} \tag{3}$$

where M represents the number of distinctive observation per state. M is defined to be infinite when the observation space is continuous.

(4) The observation probability distribution:

$$\boldsymbol{B} = \{\boldsymbol{b}_i(\boldsymbol{k})\} \tag{4}$$

where $\boldsymbol{b}_i(\boldsymbol{k}) = \Pr(\boldsymbol{O}_k | \boldsymbol{q}_t = \boldsymbol{S}_i)$, $1 \le \boldsymbol{i} \le \boldsymbol{N}$, $1 \le \boldsymbol{k} \le \boldsymbol{M}$.

(5) Initial hidden state probability distribution:

$$\boldsymbol{\pi} = \{\boldsymbol{\pi}_i\} \tag{5}$$

where $\boldsymbol{\pi}_i = \Pr(\boldsymbol{q}_{t=1} = \boldsymbol{S}_i)$, $1 \le \boldsymbol{i} \le \boldsymbol{N}$.

As listed above, the three main components of an HMM are the state transition probability matrix $\boldsymbol{A}$, the measurement probability distribution matrix $\boldsymbol{B}$, and the initial state probability distribution $\boldsymbol{\pi}$. Then, for convenience, the compact notation is used to indicate the complete parameter set of the model:

$$\boldsymbol{\lambda} = (\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\pi}) \tag{6}$$

In an HMM, the following three basic problems need to be solved: training parameters of HMMs, calculating the probability of one observation sequence, and finding a state sequence that matches the observation sequence perfectly. Therefore, there are three basic algorithms for the above problems, respectively, namely, the Baum–Welch algorithm, the forward–backward procedure, and the Viterbi algorithm [28].

3. Hidden Markov model-based approach for multimode process monitoring

3.1. Mode identification

For multimode processes, mode identification refers to the assignment of mode labels to samples. It contains two parts: mode identification for offline modeling data, which means classifying modeling data into different modes, and mode identification for online data, which is selecting the proper model for online sampling [23]. It is important to divide data with different characteristics into different modes. Here, a probability ratio strategy based on HMM is proposed to identify stable modes and transitional modes. Then, the Viterbi algorithm classifies stable modes data into distinct operating modes.

In this work, an HMM model is employed to characterize multimode processes with transitional modes. Each hidden state is representative of a mode of operation as follows (the mode number N is determined according to process knowledge):

$$\begin{aligned}
&\boldsymbol{S}_1 : \text{the 1st oprating mode} \\
&\boldsymbol{S}_2 : \text{the 2nd oprating mode} \\
&\vdots \\
&\boldsymbol{S}_N : \text{the } \boldsymbol{N}\text{-th oprating mode}
\end{aligned} \tag{7}$$

The HMM model is trained by the Baum–Welch algorithm from the observations which are measurement data. The training data consist of data from stable modes and data from transitional modes.

$$\boldsymbol{Y}(\boldsymbol{T} \times \boldsymbol{J}) = [\boldsymbol{Y}_1, \boldsymbol{Y}_2, \ldots, \boldsymbol{Y}_T]^{\mathrm{T}} \tag{8}$$

where $\boldsymbol{T}$ means the number of observations and $\boldsymbol{J}$ means the number of variables.

In order to do the calculation, define a forward variable $\boldsymbol{\alpha}_t(\boldsymbol{i})$ as:

$$\boldsymbol{\alpha}_t(\boldsymbol{i}) = \Pr(\{\boldsymbol{Y}_1, \boldsymbol{Y}_2, \ldots, \boldsymbol{Y}_t\}, \boldsymbol{q}_t = \boldsymbol{S}_i | \boldsymbol{\lambda}) \tag{9}$$

which can be solved by the inductive steps as follows.

(1) Initialization

$$\boldsymbol{\alpha}_1(\boldsymbol{i}) = \boldsymbol{\pi}_i \boldsymbol{b}_i(\boldsymbol{Y}_1) \tag{10}$$

with $1 \le \boldsymbol{i} \le \boldsymbol{N}$.