



An improved SVM integrated GS-PCA fault diagnosis approach of Tennessee Eastman process



Xin Gao, Jian Hou*

College of Engineering, Bohai University, Jinzhou 121013, China

ARTICLE INFO

Article history:

Received 1 August 2015

Received in revised form

22 September 2015

Accepted 5 October 2015

Communicated by Yang Tang

Available online 23 October 2015

Keywords:

Support vector machine (SVM)

Fault diagnosis

Tennessee Eastman (TE) process

Principal component analysis (PCA)

Gridsearch (GS)

Genetic algorithm (GA)

ABSTRACT

In modern industry, fault diagnosis and process supervision are very important in detecting machinery failures and keeping the stability of production systems. In this paper, a multi-class support vector machine (SVM) based process supervision and fault diagnosis scheme is proposed to predict the status of the Tennessee Eastman (TE) Process. After preprocessing the collected data, principal component analysis (PCA) is firstly used to reduce the feature dimension. Then, to increase prediction accuracy and reduce computation load, the optimization of SVM parameters is accomplished with the grid search (GS) method, which generates comparable classification accuracy to genetic algorithm (GA) and particle swarm optimization (PSO) while being more efficient than the latter two algorithms. Finally, to demonstrate the effectiveness of the proposed SVM integrated GS-PCA fault diagnosis approach, a comparison is made with other related fault diagnosis methods.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Nowadays, the modern industrial process is usually complicated, and various uncertainty and nonlinearity are involved. In this case, a seemingly small fault may cause unimaginable consequences. Therefore, the importance of fault detection and diagnosis in the industry can not be ignored any more. As fault diagnosis is a potential approach to improve productivity, increase production process utilization and reduce the maintenance costs, it has attracted much attention in recent years. Some of the newly proposed fault detection and diagnosis methods can be found in [1–6]. Most of the existing fault diagnosis schemes rely on the availability of process analytical models and only a few papers, e.g., [1,7,8], are devoted to data-driven fault detection and isolation approaches. Data driven techniques can be quite effective in the case that the analytical process models are not available. Especially, for the large-scale plants, whose physical models are generally difficult to be established, data driven approaches offer an effective alternative solution for process monitoring. Therefore, the production process monitoring and fault diagnosis schemes with data based fault identifying techniques have prospered and begun to be applied in the industry to protect the production system from disastrous accidents and keep stability [9,34].

With the availability of statistical learning theory [10–13] and convenient process signals collecting techniques, data based techniques have been developed quickly, and some works in this area include [8,14,15,7]. Furthermore, machine learning techniques have been applied widely in many domains ranging from image processing, face and object categorization, information and image retrieval, etc. Several machine learning algorithms, e.g., artificial neural networks (ANN), principal component analysis (PCA), fuzzy expert system, lazy learning, random forest, have been applied in industrial production systems to monitor machine condition and detect process faults [16,17]. For example, PCA was used to analyze product quality for a pilot plant [18] and in [19] a key performance indicator prediction scheme was designed to be used in a hot strip mill. However, Support vector machine (SVM) is rarely used in machine condition monitoring and fault diagnosis. Sometimes, it is used in combination with some other technologies, e.g., wavelet package transform, genetic algorithm (GA) and fractal dimension [20–22]. And it is usually expertise-oriented and problem-oriented [23].

In machine learning, SVM is an important classification approach. SVM is designed based on the Vapnik–Chervonenkis theory [24,25] and exhibits superior generalization ability in practice. It is able to handle the classification problem with finite (or small) samples as well as large feature space [23,26]. Since in the manufacture field the number of faulty samples is relatively small and signal data can be easily collected, SVM has been introduced to conduct the machine condition monitoring and fault diagnosis.

* Corresponding author.

E-mail addresses: xin.gao1991@qq.com (X. Gao), dr.houjian@gmail.com (J. Hou).

In fault diagnosis, data analysis technique is one of the most preferable and reliable methods. Generally, a diagnosis scheme is composed of three stages. The first stage is signal data acquisition with appropriate sensors. Then data processing including normalization, dimension reduction and statistical feature extraction are executed in the next step. Some feature selection techniques including Genetic algorithms (GAs), PCA and Particle swarm optimization technique (PSO), are commonly used to reduce the dimension of the recorded data set as well as save computation load [9]. In the final step, the operation status of the monitored process is predicted by maintenance personnel manually or by some automated computational intelligence methods such as ANN and SVM. SVM is preferred here as it is a representative nonlinear classifier and possesses superior generalization ability [27]. These nice properties make it extensively used in the field of classification analysis, such as process monitoring and fault diagnosis [28]. In this paper, after methods comparison, an improved SVM based industrial process faults diagnosis scheme is proposed. Fault diagnosis experiments are conducted with the TE process model, and the results show that the proposed SVM integrated GS-PCA fault diagnosis scheme performs better in both classification accuracy and computation efficiency.

The remainder of this paper is organized as follows. The SVM classification algorithm, the PCA algorithm, and three optimization algorithms, i.e., GA, PSO and GS, are briefly introduced in the next section. In Section 3, simulations are conducted to predict the status of the Tennessee Eastman process and the results validate the effectiveness of the proposed approach. Finally, Section 4 summarizes the conclusions.

2. Methods

2.1. Support vector machine (SVM)

SVM is a popular classification technique widely used in various fields. In the following we briefly introduce the theoretical foundation of SVM. Generally speaking, a classification task involves training data and testing data. The training data is used to train a prediction model, which is then used to predict the labels of testing data. In industrial applications, the collected historical data usually serve as the training data. Every sample in the training set consists of two parts, i.e., the observed variables (called attributes) and the corresponding class labels. Since this paper focuses on the industrial fault diagnosis, the class labels here refer to the statuses of the industry process, i.e., normal status or the faulty statuses (needed to be specific to the serial number of the faults). Then, the training data are used to build a prediction model. The testing data generally contain only the attributes. The missing class labels will be predicted by feeding the testing data to the prediction model. In this way, statuses of the process can be identified based on the collected data.

The key of SVM is to find out linear hyperplanes to separate the data in different classes with maximum margins. Taking the binary classification problem for example, the distance between two classes of training data should be as large as possible. The separating hyperplane should have maximal distances with the closest data points belonging to the both sides and it is located in the middle of the margin. The nearest points to the hyperplane are called support vectors (SVs), and they are representative data points because they contain almost all the information needed to determine the classifier.

In training set, an instance is denoted as $(x_i, y_i), i = 1, \dots, m$ where $x_i \in R^n$ and $y \in \{1, -1\}$. m is the number of the observed samples while n represents the number of attributes. The hyperplane is denoted by $w^T x_i + b = 0$, and the SVM classification

problem is transformed to finding out optimal parameters w^T and b to maximize the distance between hyperplanes $w^T x_i + b = 1$ and $w^T x_i + b = -1$. According to [25], the problem translates to the following optimization problem:

$$\min_{w,b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \xi_i \quad (1)$$

$$\text{s.t. } y_i(w^T x_i + b) \geq 1 - \xi_i, \quad (2)$$

$$\xi_i \geq 0, \quad (3)$$

In Eq. (1), the parameter C is a user-specified penalty parameter of error term. The error term is represented by the parameter ξ_i .

In some cases, the original input data can not be linearly separated. Then the mapping function ϕ is introduced to map the input data into a new feature space with higher dimension, where the input data become linearly separable. Using the corresponding mapping function, the kernel function is denoted as

$$K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle \quad (4)$$

Usually the kernel function is used more extensively than the mapping function, as it is less expensive to calculate. It enables the optimization problem Eq. (1) to be solved in the input space itself instead of the higher dimensional mapping space [25]. Table 1 presents four commonly used kernels.

In this paper, we adopt the Gaussian radial basis function (RBF) in the SVM classifier. Generally speaking, the RBF kernel function is usually the first option, especially when the data set has large numbers of instances. The main reason is as follows. Firstly, by adjusting the parameters, the RBF kernel can perform similarly with sigmoid kernel and linear kernel. Secondly, the RBF kernel involves only penalty error parameter C and kernel parameter γ , and has relatively small numerical difficulty [29].

2.2. Dimension reduction with PCA

It is very important to scale the attribute values in the training data and testing data into a small range before they are put into the SVM prediction model. On one hand, it can avoid the large values dominating the small ones. On the other hand, scaling attribute values into a small range will avoid dimension disaster and decreasing calculation load [35]. In this paper, each attribute value will be linearly scaled to the range between 0 and 1 to avoid the numerical difficulty caused by large attribute values. What is more, the training set and testing set are scaled at the same time.

As a commonly used dimension reduction algorithm, PCA extracts a set of uncorrelated variables and store them in features of smaller dimensions. Due to its efficiency and simplicity in dealing with the data sets with high dimensions, PCA has been widely applied in various fields in practice, especially for data compression. The procedure of standard PCA approach to reduce the dimension of data is briefly presented as follows.

The original data set with N samples denoted as $Z^T = [z_1, \dots, z_N] \in R^{m \times N}$, where m is the number of measurement signals in each sample. Then, singular value decomposition (SVD) is applied to the covariance estimation matrix of the observed data

Table 1
Formulation of kernel functions.

linear	$rx_i^T \cdot x_j$
polynomial	$(\gamma x_i^T \cdot x_j + r)^d$
radial basis function (RBF)	$e^{-\gamma \ x_i - x_j\ ^2}$
sigmoid	$\tanh(\gamma x_i^T \cdot x_j + r)$

γ , r and d are parameters.

Download English Version:

<https://daneshyari.com/en/article/411624>

Download Persian Version:

<https://daneshyari.com/article/411624>

[Daneshyari.com](https://daneshyari.com)