

Fault Detection and Diagnosis for Industry Process Based on Support Vector Data Description

Shuning Zhang*, Hongyong Yang and Guanlong Deng

School of Information and Electrical Engineering, Ludong University, Yantai, 264025, Shandong Province, P R China

*Corresponding author

Abstract—A new approach for fault detection and diagnosis based on Support Vector Data Description (SVDD) has been proposed in this paper. Similar to the T^2 and SPE statistic in principal components analysis (PCA), an appropriate nonlinear distance metric measured in feature space and threshold have been developed for fault detection. Once the fault is detected, fault diagnosis is then carried out using SVM based method. The fault diagnosis procedure is based on SVM and mutual information. The idea and effectiveness of the proposed algorithm are illustrated with respect to the simulation data collection from an illustrative example and the well-known Tennessee Eastman benchmark chemical process. Both the results show that the proposed approach works well to capture the underlying nonlinear process correlation thus providing a feasible and promising solution for nonlinear process monitoring.

Keywords—fault detection; fault diagnosis; support vector data description; mutual information

I. INTRODUCTION

The increasing requirements on product quality and process safety has encouraged the development of the studies on fault detection and diagnosis for industrial processes [1]–[3]. Existing process monitoring methods are roughly classified into two types: model-based methods [4], [5] and data-driven methods [6]–[8]. Owing to the complexity of model-based approach, data-driven methods have been widely utilized in process monitoring. Most existing data-driven approaches are based on principal component analysis (PCA) [9], partial least square (PLS) [10], and independent component analysis (ICA) [11].

Though the methods perform well for process monitoring and fault detection, another important unit in process monitoring is fault diagnosis which means how to identify faulty process variables and diagnose the root cause of faults. Supervised classification can be used for fault diagnosis. Many methods have developed for supervised classification such as fisher discriminant analysis (FDA) [12], artificial neural networks (ANN) [13], support vector machine (SVM) [14], et al. In recent years, the data domain description problem (also called one-class classification), an extension of the classification problem, is proposed by Tax and Duin in [15]. The method is inspired by the SVM by Vapnik, called the support vector data description (SVDD). This data description can be used for novelty or outlier detection.

These pioneer works have provided abundant theoretical bases for our following work. In this paper, a new data-based method to diagnose the faults of a system in steady state

conditions is presented. The approach is based on a non-linear distance metric measured in a feature space. Then, SVDD is used to separate all the possible faulty samples (outliers) from the normal operating samples. This paper is organized as follows. In section 2, the basis of SVDD is outlined and the SVDD is depicted. Consequently, details of the new fault diagnosis based on distance metric using the SVDD is established in section 3. The efficacy of this method is demonstrated by applying it on a simulation example and a benchmark problem - Tennessee Eastman benchmark chemical process in section 4. Finally, conclusions are drawn in section 5.

II. SUPPORT VECTOR DATA DESCRIPTION

The main idea of support vector data description (SVDD), which is a variant of the standard SVM, is to map data points by means of a nonlinear transformation to a high dimensional feature space and to find the smallest sphere that contains most of the mapped data points in the feature space [15]. Of a data set containing n data objects, $x_i, i = 1, 2, \dots, n$, a description is required. We try to find a sphere with minimum volume, containing all (or most of) the data objects. This is very sensitive to the most outlying object in the target data set. When one or a few very remote objects are in the training set, a very large sphere is obtained which will not represent the data very well. Therefore, we allow for some data points outside the sphere and introduce slack variables ξ_i (analogous to [16]).

The minimum enclosing sphere, described by center a and radius R , can be found by solving the following constrained quadratic optimization problem:

$$\begin{aligned} \min F(R, a, \xi_i) &= R^2 + C \sum_{i=1}^n \xi_i \\ \text{subject to } (x_i - a)^T (x_i - a) &\leq R^2 + \xi_i \\ \xi_i &\geq 0, i = 1, 2, \dots, n \end{aligned} \quad (1)$$

where the constant C gives the tradeoff between simplicity (or volume of the sphere) and the number of errors (number of target objects rejected).

Using the Lagrange multiplier method, the above optimization problem can be solved. To test the point x , the distance to the center of the sphere has to be calculated. A test point x is accepted when this distance is smaller or equal than the radius:

$$\begin{aligned} \|\varphi(x) - a\|^2 &= K(x, x) - 2 \sum_{i=1}^n \alpha_i K(x_i, x) \\ &+ \sum_{i,j=1}^n \alpha_i \alpha_j K(x_i, x_j) \leq R^2 \end{aligned} \quad (2)$$

where R^2 is the distance from the sphere center a to any support vector x_i on the boundary.

III. FAULT DETECTION AND DIAGNOSIS BASED ON SVDD

The proposed approach composes two main steps: 1) SVDD model was firstly used to detect process faults caused by improper operation, equipment malfunction, and so on. 2) The SVM and mutual information methods were used to diagnose the root causes of faults. The remainder of the section will focus on the two steps respectively.

A. Fault Detection

The SVDD technique proposed by Tax and Duin [19], is a variant of the standard SVM algorithm. Let $X = [x_1, \dots, x_N] \in R^{N \times m}$ represent the normalized training data set consisting of N samples and m process variables. A PCA model which is used to extract the latent variables is built as

$$T = XP \quad (3)$$

Where T and P represent a score matrix and a loading matrix respectively.

Based on the matrix T , a SVDD model is established to determine a hypersphere with minimum volume that holds all training data. The radius R^2 of hypersphere according Eq.(2) can be calculated. When the new data x_{new} arrives, the score vector t_{new} is computed using Eq.(3). Then the distance d_{new} can be obtained, i.e.

$$d_{new} = \sqrt{1 - 2 \sum_{i=1}^N \alpha_i K(t_i, t_{new}) + \sum_{i,j=1}^N \alpha_i \alpha_j K(t_i, t_j)} \quad (4)$$

If d_{new} is beyond a control limit d_{lim} , the sample x_{new} is outside the hypersphere, which means that x_{new} is a fault.

B. Fault Diagnosis Using SVM

Fault diagnosis aims to determine which specific variables responsible for the fault and infer the root cause of the fault. Therefore a selection of the significant variables for the fault before the classification is very important. The SVM Recursive Feature Elimination (SVM-RFE) algorithm is used to select the features which are the optimum subset of the variables in the data set that gives the best separation or classification accuracy between the normal samples and the faulty samples[17].The weight magnitude as an ranking criterion is used in SVM-RFE

method [17], [18]. However, the relationships among variables in real industrial plant is very complicated which make the SVM-RFE is time-consuming. Therefore the mutual information (MI) [18] is used to solve the variable correlations problem.

Let $X = [x_1, \dots, x_N] \in R^{N \times m}$ denotes the scaled matrix, the relationship between any two variables x_1 and x_2 with the joint density $p(x_1, x_2)$ can be given by MI:

$$I(x_1, x_2) = \int \int p(x_1, x_2) \log \frac{p(x_1, x_2)}{p(x_1)p(x_2)} dx_1 dx_2 \quad (5)$$

Where $p(x_1)$ and $p(x_2)$ represent the marginal densities of variables x_1 and x_2 respectively. And the two marginal densities can be defined as $p(x_1) = \int_{x_2} p(x_1, x_2) dx_2$ and

$$p(x_2) = \int_{x_1} p(x_1, x_2) dx_1 \text{ respectively.}$$

Since the MI method is used to supervised classification, the classes can be considered as a multinomial variable with K possible values (where K is the classes of the system, in our paper $K = 2$ denotes the normal and faulty respectively). The MI between a multivariate Gaussian variable s and a multinomial variable k can be calculated as [18]:

$$I(s, k) = \frac{1}{2} (\log(|\sum|)) - \sum_{k=1}^K P(k) \log(|\sum_k|) \quad (6)$$

With

$$p(s) = \frac{1}{(2\pi)^{\frac{m}{2}} |\sum|^{\frac{1}{2}}} \exp(-\frac{1}{2} (s - \mu)^T |\sum|^{-1} (s - \mu)) \quad (7)$$

μ and \sum represent the mean vector and the covariance matrix respectively, $|\sum|$ represents the determinant of the matrix, $p(k)$ represents the probability distribution for class k , and the normal variable s conditionally to a class k can be written as in Eq.(8), where μ_k and \sum_k represent the mean vector of the class k and the covariance matrix of the class k respectively.

$$p(s|k) = \frac{1}{(2\pi)^{\frac{m}{2}} |\sum_k|^{\frac{1}{2}}} \exp(-\frac{1}{2} (s - \mu_k)^T |\sum_k|^{-1} (s - \mu_k)) \quad (8)$$

The MI computed by Eq.(6) for all the variables can be obtained. And the variable with the largest value I is selected as the important group variables for classification. However, all the selected variables may correlate with each other. To solve this problem, the sorting algorithm for variables proposed in [18] is used in our paper. The selection procedure for sorting variables is described in Figure 1. The C_i represents the variable selected at step i and the parameter m represents the process variables number. For more detailed description of the sorting algorithm of variables, the interested is referred to the article by Verron [18].

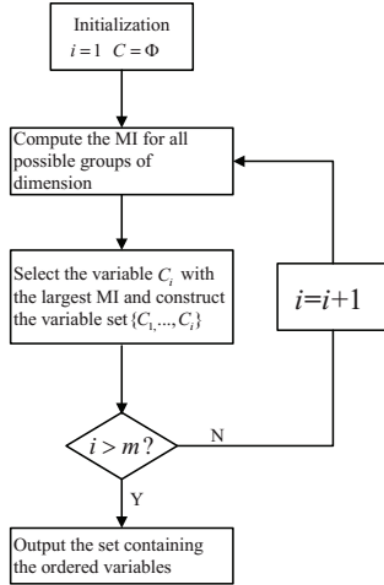


FIGURE 1. THE PROCEDURE FOR VARIABLE SELECTION

The ordered variables C_1, \dots, C_m can be obtained, where C_1 is the most important variable for the process, C_2 is the second important variable for the process, and so on. Then the group G_1, \dots, G_m are obtained, where $G_1 = \{C_1\}$, $G_2 = \{C_1, C_2\}, \dots, G_m = \{C_1, \dots, C_m\}$. For each group G_g , SVM with the k -fold cross validation [19] is used for fault diagnosis. Let $D = (x_i, y_i) (i = 1, \dots, N)$ denotes the dataset obtained from the process, and D is partitioned into k subsets F_1, \dots, F_k such that $F_i \cap F_j = \Phi$ for any $i \neq j$. One of the subset is selected as the testing set while the $k-1$ other subsets are used as the training set firstly. Suppose the i th subset F_i was selected as the testing set. Then the SVM model is built using the subsets $F_1, \dots, F_{i-1}, F_{i+1}, \dots, F_k$. And the misclassification rate e_g was computed for the testing set. At last the average error for group g can be obtained

$$e_g = \frac{1}{k_g} \sum_{i=1}^{k_g} e_{g,i} \quad (9)$$

where $k_g = \frac{N}{k}$ is the number of samples for the testing set.

Therefore for all the m groups, the error vector $e = [e_1, \dots, e_m]$ was obtained. The group with the minimum error was selected and the number of variables N_{\min} in the group is obtained.

Now, we will see an illustrative example and an application of this approach on a benchmark problem: the Tennessee Eastman Process.

IV. ILLUSTRATION AND DISCUSSION

The proposed fault detection and diagnosis approach is tested with an illustrative example and a benchmark problem: the Tennessee Eastman Process. The previous focuses on analyzing the advantage for fault detection of the approach whereas the latter emphasizes particularly on verifying the superiority of the method over PCA method when online monitoring and fault diagnosing for relatively complex process.

A. Illustrative Example

Consider the following model

$$z_2 = -2z_1^2 + \varepsilon \quad (10)$$

where $\varepsilon \sim N(0, 0.1)$ is random gaussian noise. For $z_1 \in [-1.5, 1.5]$, 495 normal samples $Z = [z_1, z_2] \in R^{495 \times 2}$ are obtained by the model. Then five faulty samples which is not obey the model are added. The plot of 500 sample points is shown in Figure 2.

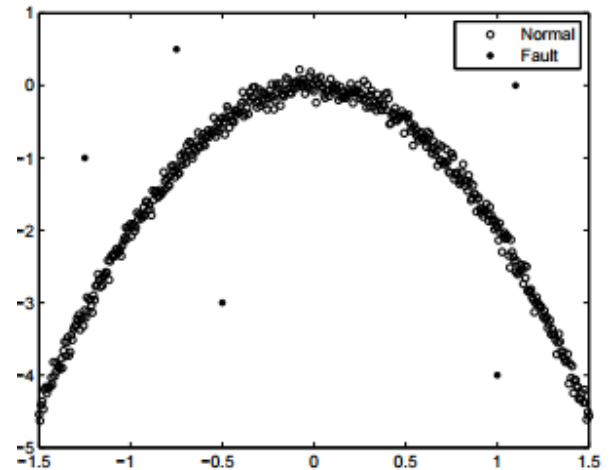


FIGURE II. ILLUSTRATIVE EXAMPLE-VISUALIZATION OF SAMPLES.

Of the 495 normal samples, the first 300 samples are selected as the training set while the 195 other samples are selected as the validation set. The proposed method and PCA is applied to the model. The confidence levels of the two

approaches are set as 0.99. The parameters $\nu = 0.1$ and $\sigma^2 = 10$ were chosen for SVDD. The T^2 of the PCA and the distance metric of SVDD are shown in Figure 3 and Figure 4 respectively.

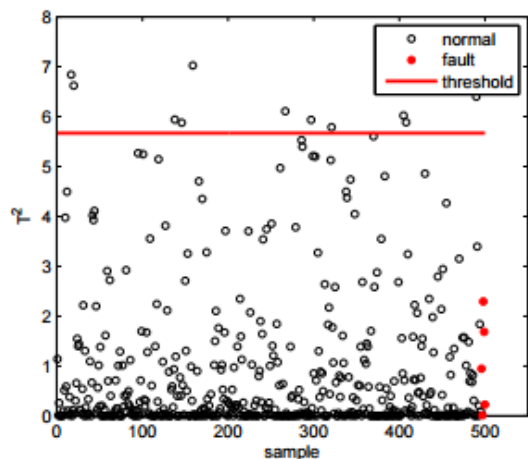


FIGURE III. ILLUSTRATIVE EXAMPLE-PCA T^2 PLOT

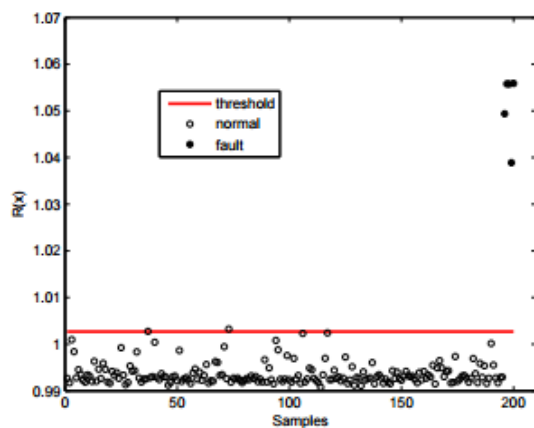


FIGURE IV. ILLUSTRATIVE EXAMPLE-SVDD PLOT.

It can be seen from Figure 3 and Figure 4 that PCA cannot detect the five faulty points while the proposed approach can detect all the faulty points. The proposed method shows good performance.

B. Tennessee Eastman Process

In this section, the proposed method, as well as PCA, is applied to TE benchmark chemical process which has been widely used for examining the efficiency of various process monitoring methods. The TE process was first introduced by Downs and Vogel [20]. The process comprises five major operation units, i.e. reactor, product condenser, vapor-liquid separator, recycle compressor and product stripper. The details of this industrial process model have been described by Downs and Vogel [20] and the open-loop code is provided at <http://brahms.scs.uiuc.edu>. Four hundred and eighty normal data samples are generated to train the monitoring model and nine hundred sixty normal data samples are generated to valid the monitoring model. Moreover, 21 different kinds of faulty

classes (there are 960 samples for each faulty class) are introduced in Table 1. The proposed method and the PCA method is applied to the TE process. The confidence level for the two methods is set as 0.95. And based on the validation data, the number of PCs retained in PCA is set as 11 while the parameter σ^2 of the proposed method empirically chosen as 150.

TABLE I. SUMMARY OF FAULT TYPES INTRODUCED IN TE PROCESS.

Fault no.	Process variable	Fault type
1	A/C feed ratio, B composition constant (stream 4)	Step
2	B composition, A/C ration constant (stream 4)	Step
3	D feed temperature (stream 2)	Step
4	Reactor cooling water inlet temperature	Step
5	Condenser cooling water inlet temperature	Step
6	A feed loss (stream 1)	Step
7	C header pressure loss-reduced availability (stream 4)	Step
8	A, B, C, feed composition (stream 4)	Random
9	D feed temperature (stream 2)	Random
10	C feed temperature (stream 4)	Random
11	Reactor cooling water inlet temperature	Random
12	Condenser cooling water inlet temperature	Random
13	Reaction kinetics	Slow drift
14	Reactor cooling water valve	Sticking
15	Condenser cooling water valve	Sticking
16	Unknown	Unknown
17	Unknown	Unknown
18	Unknown	Unknown
19	Unknown	Unknown
20	Unknown	Unknown
21	The valve for stream 4 was fixed at the steady-state position	Constant position

The fault detection results for two fault cases are illustratively shown in Figures 5-8 using the proposed methods in comparison with PCA algorithm. For fault 8 and fault 13, PCA cannot detect the faulty data whereas the proposed method can detect the the two faults easily. Moreover, compared to PCA, the proposed method was more sensitive to the two faults and detected them earlier. Therefore using SVDD method, one can easily track the process operation progress and detect the occurrence of process faults via a simple monitoring chart. Moreover, the comparison results of fault detection performance between SVDD and PCA methods have been tabulated in Table 2. The fault detection rate, which is the ratio of the fault alarm number to overall fault duration, is used to evaluate the fault detection performance.

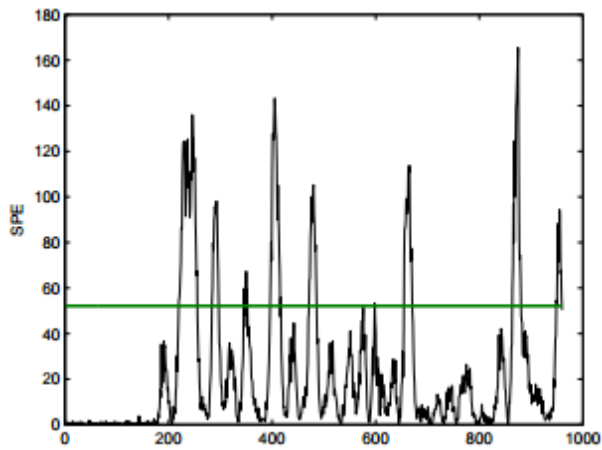


FIGURE V. TE: PCA plot for fault 8.

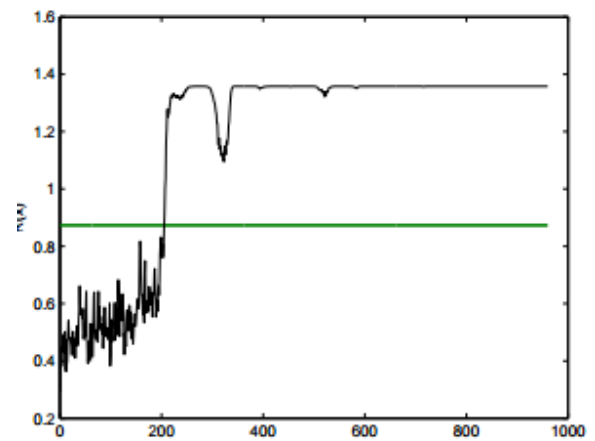


FIGURE VIII. TE: SVDD PLOT FOR FAULT 13.

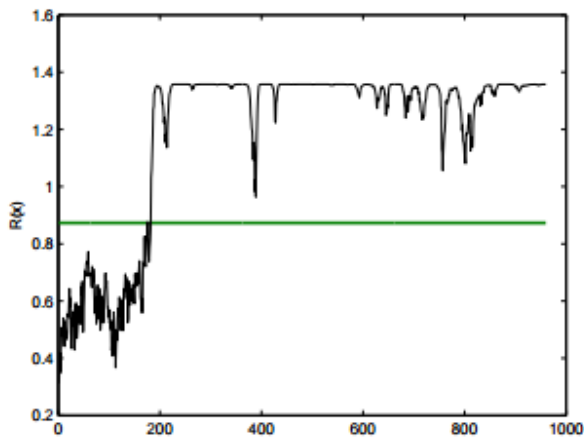


FIGURE VI. TE: SVDD PLOT FOR FAULT 8.

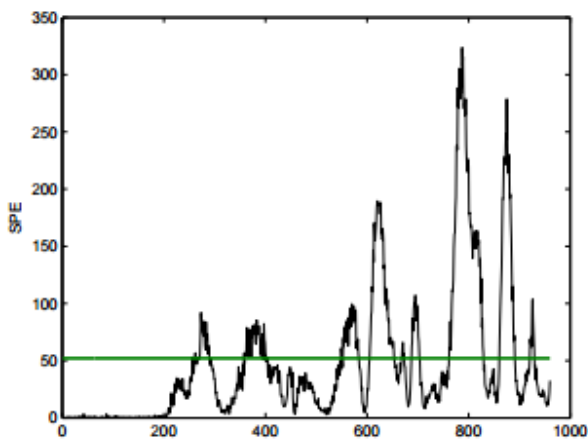


FIGURE VII. TE: PCA PLOT FOR FAULT 13.

From the table, generally SVDD shows overwhelming superiority over PCA. The proposed method can detect the faults much earlier and much higher detection rate than the PCA model. In conclusion, the above illustrations and discussions have demonstrated those theoretical analysis presented in section 3.1, concerning the effectiveness of SVDD.

TABLE II. SUMMARY OF FAULT DETECTION RATE FOR TEST DATA IN TE PROCESS.

Fault no.	PCA-SPE	SVDD
1	99.8	99.9
2	98.6	99.7
4	96.2	98.6
5	25.4	71.4
6	100	100
7	100	100
8	97.6	99.6
10	34.1	44.4
11	64.4	53.8
12	97.5	100
13	95.5	99.0
14	100	100
16	24.5	24.9
17	89.2	93.0
18	89.9	98.0
19	12.7	85.6
20	45.0	67.2
21	43.0	55.4

After a process fault is detected, the PCA contribution plot and the SVM based method is used for fault diagnosis. If 6 continuous SPE of PCA are beyond the confidence level, a fault can be seen as the real fault. Then the SPE is decomposed to identify the individual contribution of each process variables to the fault. Later, the contribution plots are obtained through summing the contribution of each process variable to SPE. In the case of SVM based method, the set of variables which gives the lower error but with the lower number of variables is

reported. The results of PCA-based contribution plots along with SVM-based method are tabulated as in Table 3.

TABLE III. SUMMARY OF FAULT DIAGNOSIS IN TE PROCESS.

Fault no.	PCA-SPE	SVDD
1	16,20,46	1,16
2	30,24,16,40	10,34
4	51,21,9	51
5	11,9,35,22,18	7,50,18
6	44,1	1
7	4,16,36,38,31,25	45
8	3,46,29,16,20	16,20
10	19,34,35	19,18
11	51,9,21	51
12	11,37,22,4	7
13	37,32,26,42	16,50,38
14	21,51,9,2,42	51
16	50,3,27	50,19
17	21,2,42	21,9
18	22,11,40	43,51
19	5,27,37,46,38	50,19,18
20	46,39,13,23	13,20,11,18
21	39,26,5	16,8,20,11

In order to describe the procedure of fault diagnosis clearly, we will take two faults (fault 7 and fault 10) as examples. For fault 7, after sorting the variables of TE process with mutual information, SVM with 10-fold cross validation was applied to identify the variables related to the fault. The result is given in Figure 9.

In Figure 9, we can clearly see that the average error is not increased with the increased number of variables. Therefore the best group with one variable (the C header pressure loss in stream 4) is chosen. As shown in Table 1, there is a step change for root cause variable (the C header pressure loss in stream 4) for fault 7. Hence the proposed approach can identify the right variable. In the process, to correct the process back to the normal condition, the flow rate valve should be opened. Therefore the total feed flow valve in stream 4 is identified as the faulty variable for fault 7 by PCA. However, the variable 4 is not the root cause variable. In order to illustrate the result clearly, the variables 45 and 4 were extracted from the dataset and plotted in Figure 10.

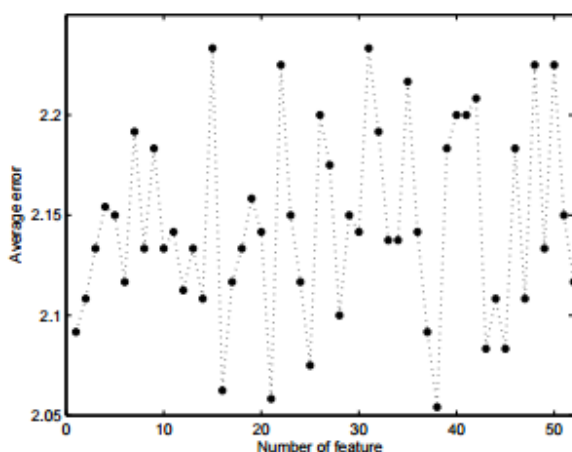


FIGURE IX. RESULT OF THE 10-FOLD CROSS VALIDATIONS FOR FAULT 7.

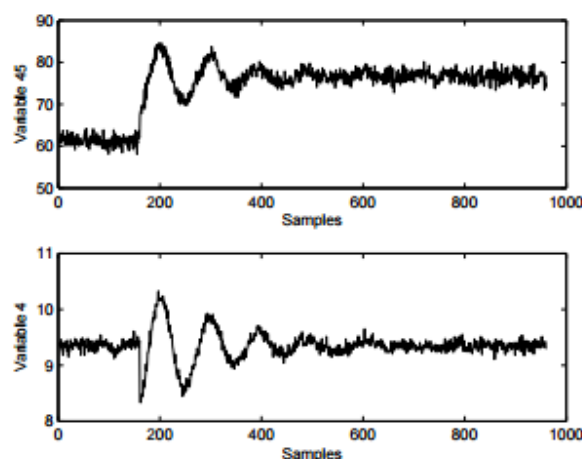


FIGURE X. VARIABLES 45 AND 4 PLOTS IN FAULT 7.

It can be seen from Figure 9 that the curve of variable 4 changed with an oscillatory behavior firstly. And then the curve eventually dies down after 550 samples which looks like the normal operating data. Hence the variable 4 is not a root cause variable for fault 7. However, there is a step change for variable 45, which is the nature for fault 7. Moreover, PCA contribution also highlights many other variables such as variables 16, 36, 38, 31 and so on. The four plots of the trend of these variables in fault 7 are shown in Figure 11. As shown in Figure 11, there are oscillatory behaviors for these four variables. The comparison between the four variables and the variable 45 turns out that they share different changing trends, which indicate that these variables are not faulty variables. Hence, it can demonstrate that the proposed approach gives a better performance than the PCA method.

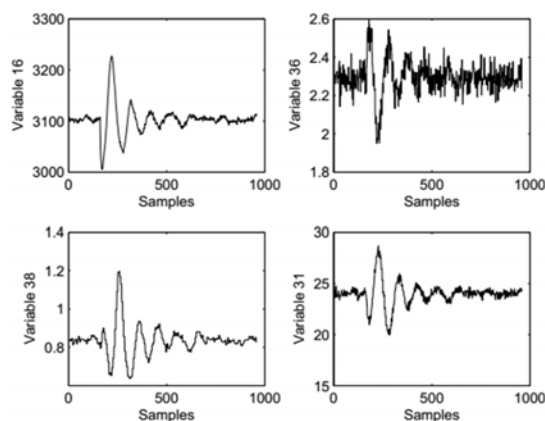


FIGURE XI. VARIABLES OF PCA CONTRIBUTION PLOTS IN FAULT 7.

Next, we consider fault 10, after obtaining the ordered variables using mutual information, SVM with 10-fold cross validation was applied to identify the variables related to the fault. The result is given in Figure 12. As shown in the figure, two variables (variable 19 and 18) are selected according to the proposed method.

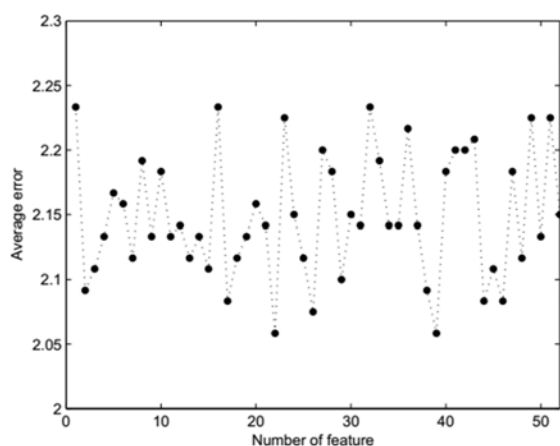


FIGURE XII. RESULT OF THE 10-FOLD CROSS VALIDATIONS FOR FAULT 10.

Fault 10 in Table 1 is actually caused by random variation of the C feed temperature in stream 4. The variable 19 and 18 were identified by SVM based method whereas the variables 19, 34 and 35 were selected through the by PCA-based contribution plots. For the two approaches, the variable 19 (stripper steam flow) is chosen as the main faulty variable responsible for fault 10. To elaborate the benefits of the proposed method (SVM based method), a scores plot with the feature extracted from the PCA-based contribution and a scatter plot with the feature extracted from the SVM based method are shown in Figure 13 and Figure 14 respectively.

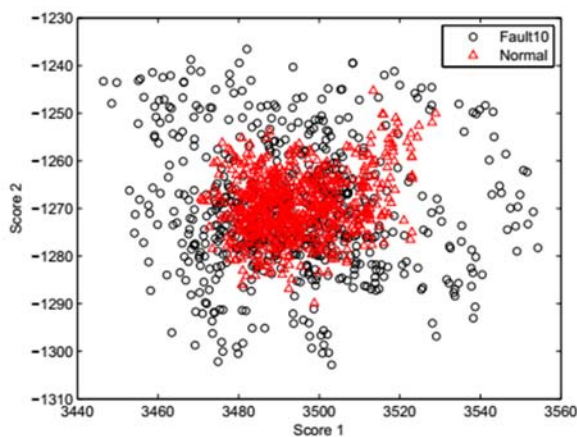


FIGURE XIII. PCA SCORES PLOT FOR FAULT 10.

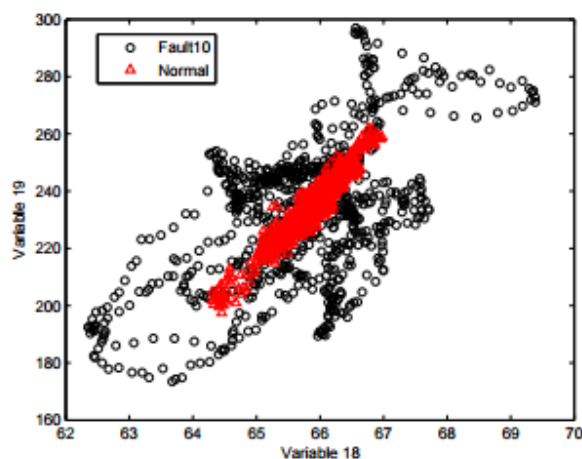


FIGURE XIV. SVDD SCATTER PLOT FOR FAULT 10.

As shown in the two figures, the SVM based method gives a much better performs for classifying the normal and faulty data. Variable 18 (stripper temperature) selected by the proposed method is closed related to the real faulty variable (C feed temperature) for fault10. However, according to PCA-based method contribution, variables 34 and 35 (F,G compositions in stream purge steam) are not related to the real faulty variable. Therefore diagnosis result for SVM based method is close to the real root cause of fault 10, however, the PCA-based method is not helpful for diagnosing the root cause of fault 10.

In conclusion, the above illustration and discussions have demonstrated the necessary and effectiveness of SVDD method proposed in the paper. Although the short communication has not explored how well the method can function in other nonlinear industrial processes, however, the present case studies have successfully promised the potential of SVDD method.

V. CONCLUSIONS

A new method based on SVDD and SVM classifier for fault detection and diagnosis has been proposed in the paper. The approach is based on a non-linear distance metric measured in a feature space. Just as in PCA, appropriate distance metrics and thresholds have been developed for fault detection. Then fault diagnosis is done by SVM based method. This method, after a feature selection based on mutual information, makes classification of faulty samples with SVM classifier. A simulated problem and the simulated TE benchmark process have been used to demonstrate the efficacy of the proposed method. The simulation results show that the proposed approach is effective and reasonable, indicating that it is competent to detect process faults with greater reliability.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China under Grant 61603172, the Natural Science Foundation of Shandong Province under Grant BS2015DX018, the Scientific Research Foundation of Ludong University under Grant LY2015024, and the Funds for Creative Research Groups of China (No.61621004).

REFERENCES

- [1] T. Kourti, J.F. MacGregor, "Process analysis, monitoring and diagnosis, using multivariate projection methods," *Chemometr. Intell. Lab.*, vol. 28, no. 1, pp. 3–21, Apr. 1995.
- [2] S.J. Qin, "Statistical process monitoring: basics and beyond," *J. Chemometrics*, vol. 17, no. 8-9, pp. 480–502, Aug. - Sep. 2003.
- [3] D. Liefucht, U. Kruger, G.W. Irwin, "Improved reliability in diagnosing faults using multivariate statistics," *Comput. Chem. Eng.*, vol. 30, no. 5, 901–912, May. 2006.
- [4] M.S. Choudhury, S. Shah, N. Thornhill, D.S. Shook, "Automatic detection and quantification of stiction in control valves," *Control Eng. Pract.*, vol. 14, no. 12, pp. 1395–1412, Dec. 2006.
- [5] N.F. Thornhill, A. Horch, "Advances and new directions in plant-wide disturbance detection and diagnosis," *Control Eng. Pract.*, vol. 15, no. 10, pp. 1196–1206, Oct. 2007.
- [6] S. Gajjar, M. Kulahci, A. Palazoglu, "Real-time fault detection and diagnosis using sparse principal component analysis," *J. Process Contr.*, <http://dx.doi.org/10.1016/j.jprocont.2017.03.005>, Article in press, 2017.
- [7] L. Luo, S. Bao, J. Mao, D. Tang, "Fault detection and diagnosis based on sparse PCA and two-level contribution plots," *Ind. Eng. Chem. Res.*, vol. 56, no. 1, pp. 225–240, Jan. 2017.
- [8] J. Huang, X. Yan, "Relevant and independent multi-block approach for plant-wide process and quality-related monitoring based on KPCA and SVDD," *ISA Tran.*, <https://doi.org/10.1016/j.isatra.2018.01.003>, Article in press, 2017.
- [9] R. Raveendran, B. Huang, "Two layered mixture Bayesian probabilistic PCA for dynamic process monitoring," *J. Process Contr.*, vol. 57, no. 9, 148–163, Sep. 2017.
- [10] L. Luo, S. Bao, J. Mao, D. Tang, "Quality prediction and quality-relevant monitoring with multilinear PLS for batch processes," *Chemometr. Intell. Lab.*, vol. 150, no. 1, pp. 9–22, Jan. 2016.
- [11] X. Tian, L. Cai, S. Chen, "Noise-resistant joint diagonalization independent component analysis based process fault detection," *Neurocom.*, vol. 149, no. 2, pp. 652–666, Feb. 2015.
- [12] R.O. Duda, P.E. Hart, D.G. Stork, "Pattern Classification," second ed., Wiley, 2001.
- [13] M.J. Fuente, P. Vefa, "Neural networks applied to fault detection of a biotechnological process," *Eng. Appl. Artif. Intel.*, vol. 12, no. 5, pp. 569–584, May. 1999.
- [14] V.N. Vapnik, "The Nature of Statistical Learning Theory," Springer, 1995.
- [15] D.M.J. Tax, R.P.W. Duin, "Support vector data description," *Mach. Learn.*, vol. 54, no. 1, pp. 45–66, Jan. 2004.
- [16] B.E. Boser, I.M. Guyon, V.N. Vapnik, "A training algorithm for optimal margin classifiers," in: *Proceedings of the Annual Workshop on Computational Learning Theory*, Pittsburgh, 1992, 144–152.
- [17] S. Mahadevan, S.L. Shah, "Fault detection and diagnosis in process data using one-class support vector machines," *J. Process Contr.*, vol. 19, no. 10, pp. 1627–1639, Oct. 2009.
- [18] S. Verron, T. Tiplica, A. Kobi, "Fault detection and identification with a new feature selection based on mutual information," *J. Process Contr.*, vol. 18, no. 5, pp. 479–490, May 2008.
- [19] J.D. Rodríguez, A. Pérez, J.A. Lozano, "Sensitivity Analysis of k-Fold Cross Validation in Prediction Error Estimation," *IEEE Trans Pattern Anal. Mach. Intell.*, vol. 32, no. 3, pp. 569–575, Mar. 2010.
- [20] J.J. Downs, E.F. Vogel, "A plant-wide industrial process control problem," *Comput. Chem. Eng.*, vol. 17, no. 3, pp. 245–255, Mar. 1993.