International Journal of Applied Mathematics in Control Engineering 3 (2020) 151-155

Contents lists available at YXpublications

International Journal of Applied Mathematics in Control Engineering

Journal homepage: http://www.ijamce.com

Contribution Analysis of Sensor Fault Diagnosis Based on Improved Relative Reconfiguration

Hailian Du[a,*], Wenxia Du[*], Ziyi Du, Feng Lv

College of Vocational and Technical, Hebei Normal University, No. 20 Road East, 2nd Ring south, YuHua District, Shijiazhuang, *Heibei, China*

ARTICLE INFO

Article history:
Received 26 June 2020
Accepted 27 August 2020
Available online 28 August 2020

Keywords:
Fault diagnosis
Principal component analysis
Relative contribution value

ABSTRACT

In the use of the principal component analysis method to fault diagnosis of multiple sensors, when a fault happened, the fault data is reconstructed, and the reconstructed contribution rate determined which sensor was failed. But the contribution rate contained variables that were independent of the direction of the fault and the cross terms with other fault directions. There was a small difference in the contribution value of the failure data, which could lead to false report. In order to solve the above problems, after the principal component analysis of the failure data, the improved relative reconstruction contribution rate was calculated, and the correct diagnosis rate was improved by the simulation experiment.

1. Introduction

Nowadays, the controlled objects in industrial production are complex, with many control variables and huge scale. Failure of a component often leads to a chain reaction, which can result in performance degradation or damage to equipment. Therefore, there is a very important social and economic benefit in the key equipment fault detection and diagnosis of the production process. In general, it is impossible for any detection method to detect all kinds of faults of the system with 100% accuracy. Therefore, how to improve the correct detection rate of faults and reduce the failure omission rate and false alarm rate has always been one of the hot spots in the field of fault detection and diagnosis.

Principal Component Analysis (PCA) is a data-driven method for fault detection in industrial production process. A few unrelated variables are used to represent multiple related detection variables of the controlled object, which establishes the principal component model, and calculates the SPE value and T^2 value of the sample[1,2], so as to carry out fault detection. If one or two of the fault detection indicators SPE and T^2 statistics exceed the control threshold, then a fault exists. The contribution graph method is used to judge which variable has failed[3-4]. The criterion of contribution graph method is that the process variable with high contribution rate is the cause of failure, but it is easy to cause false positives. Literature [3] pointed out that when the dimension of the principal component subspace or the residual subspace is 1, multiple fault detection indexes based on reconstruction-based contribution(RBC) are invalid for fault diagnosis, and a single variable fault diagnosis method is constructed by using the moving window. A comprehensive evaluation index based on SPE and T^2 statistics and combined with the reconstruction contribution value diagnoses the fault of the double-capacity water tank in the Literature [5]. In literature [6], the reconstruction contribution ratio method based on the traditional reconstruction contribution diagram is proposed to separate and diagnose the faults of continuous stirred tank reactor.

The traditional fault diagnosis method based on contribution value is to reconstruct the fault sample and analyze the contribution value of the reconstructed data. However, there are two problems:

1) The corresponding contribution values of different sensors include cross terms and independent variables associated with other sensors;

2) The contribution values of each sensor are not compared against the same benchmark in the failure state.

For the above two problems, this paper combined the improved reconstruction contribution with the relative reconstruction contribution to obtain a new reconstruction contribution method.

2. The basic theory

2.1 The method of principal component analysis

* Corresponding author.
E-mail addresses: dwx20040513@163.com (W. Du)

Assuming that the original normal sample is $\boldsymbol{X} \in \boldsymbol{R}_{m \times n}$, where m is the number of samples and n is the number of sensors. The data is standardized, and the covariance matrix $\boldsymbol{S}$ is calculated. The eigenvectors and eigenvalues of $\boldsymbol{S}$ are calculated, and the resulting eigenvector matrix and the diagonal matrix of eigenvalues are arranged from large to small.

Sum the eigenvalues from the largest to the smallest until more than 85%, and the number of added eigenvalues is the number of principal components k. Take the eigenvector $\boldsymbol{P}$ and the eigenvalue $\boldsymbol{T}$ corresponding to the principal components, the eigenvector $\tilde{\boldsymbol{P}}$ and the eigenvalue $\tilde{\boldsymbol{T}}$ corresponding to the residual. The original data $\boldsymbol{A}$ can be decomposed into the principal component and residual parts:

Take the eigenvector and eigenvalue corresponding to the principal element, and the eigenvector and eigenvalue corresponding to the residual. The original data $\boldsymbol{X}$ can be decomposed into the principal component part and the residual part:

$$\boldsymbol{X} = \boldsymbol{T}\boldsymbol{P}^{\mathrm{T}} + \tilde{\boldsymbol{T}}\tilde{\boldsymbol{P}}^{\mathrm{T}} \tag{1}$$

The SPE control limit obtained from the original normal sample is defined as:

$$Q_{\alpha} = \theta_1 \left(\frac{c_{\alpha}\sqrt{2\theta_2 h_0{}^2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1{}^2} \right)^{1/h_0} \tag{2}$$

where, $\theta_i = \sum_{j=k+1}^{m} \lambda_j{}^i (i = 1,2,3)$; $h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2}$; λ_j is the eigenvalue of the covariance matrix $\boldsymbol{S}$; C_{α} is the threshold value of the standard normal distribution at the confidence value of α ; m is the the dimension of sample $\boldsymbol{X}$.

The HotellingT^2 statistic is defined as:

$$T_{\alpha}{}^2 = \frac{k(n\text{-}1)}{n\text{-}k} F_{k,n-k,\alpha} \tag{3}$$

$F_{k,n-k,\alpha}$ is the critical value of F distribution with degrees of freedom $k, n-k$ and confidence α .

When the system fault is determined, it is necessary to separate the variables causing the fault or determine the type of fault. The contribution diagram method is one of the most commonly used fault separation methods in principal component analysis. The statistics commonly used for contribution diagrams are SPE and T2. The contribution diagram based on SPE is defined as follows:

$$\mathrm{Cont}_i^{\mathrm{SPE}} = (\xi_i^{\mathrm{T}} \boldsymbol{M}\boldsymbol{x})^2 \tag{4}$$

$\mathrm{Cont}_i^{\mathrm{SPE}}$ represents the contribution of each variable to SPE, $\boldsymbol{M} = \boldsymbol{I} - \boldsymbol{P}\boldsymbol{P}^{\mathrm{T}}$,and ξ_i represents the i column of the unit matrix $\boldsymbol{I}_m$.

When a fault is detected, the variable with large contribution graph is considered as the probable cause variable of the fault. Although this method is very effective for fault identification, it also has fuzzy effect and can easily cause fault false alarm[7,8].

After the fault is detected, the system will return to normal conditions. Therefore, reconstruction-based contribution(RBC) analysis calculates the contribution value along each sensor direction based on the principal element model, which reduces the false alarm rate [9].

2.2 Failure reconstruction contribution analysis

The fault samples at each moment can be decomposed into:

$$\boldsymbol{x} = \boldsymbol{x}^* + \xi_i f_i \tag{5}$$

Where, x^* is the normal value; $\xi_i = \begin{bmatrix} 0,0 \ldots \underset{i}{1} \ldots 0 \end{bmatrix}^{\mathrm{T}}$, i is the fault direction (i =1,2,3..., n); f_i is the fault amplitude.

Then the normal sample can be expressed as:

$$\boldsymbol{z}_i = \boldsymbol{x} - \xi_i f_i \tag{6}$$

When the fault direction i is correct, the contribution value of $\boldsymbol{z}_i$ can be minimized. So by:

$$\min_{f_i} \left(\boldsymbol{z}_i{}^{\mathrm{T}} \boldsymbol{M} \boldsymbol{z}_i \right) = \min_{f_i} \|\boldsymbol{z}_i\|_{\boldsymbol{M}}^2 = \min_{f_i} \|\boldsymbol{x} - \xi_i f\|_{\boldsymbol{M}}^2 \tag{7}$$

It can be obtained that:

$$\hat{\boldsymbol{f}}_i = (\xi_i{}^{\mathrm{T}} \boldsymbol{M} \xi_i)^{-1} \xi_i{}^{\mathrm{T}} \boldsymbol{M}\boldsymbol{x} \tag{8}$$

The corresponding contribution value of each sensor at each moment is:

$$\mathrm{RBC}_i = \boldsymbol{x}\boldsymbol{M}\xi_i \left(\xi_i{}^{\mathrm{T}} \boldsymbol{M} \xi_i \right)^{-1} \xi_i{}^{\mathrm{T}} \boldsymbol{M}\boldsymbol{x}^{\mathrm{T}} \tag{9}$$

The contribution value is the i sensor's contribution rate to the fault at a certain moment, and the sensor corresponding to the maximum contribution value is the fault direction.

3. Improve the algorithm of relative reconstruction contribution value

3.1 Improved Reconstruction Contribution Value

The element of the ith row and the jth column in the matrix $\boldsymbol{M}$ is denoted by m_{ij} , where $\boldsymbol{B} = \boldsymbol{M}\xi_k \xi_k^{\mathrm{T}} \boldsymbol{M}$, then the contribution value along the direction can be denoted by the following formula[10]:

$$\begin{aligned} \mathrm{RBC}_k &= \frac{\boldsymbol{x}\boldsymbol{B}\boldsymbol{x}^{\mathrm{T}}}{m_{kk}} = \frac{1}{m_{kk}} \left(\sum_{i=1}^{n} \sum_{j=1}^{n} m_{ik} m_{kj} x_i x_j \right) \\ &= \frac{1}{m_{kk}} \left(\sum_{i=1}^{n} m_{ki}^2 x_i^2 + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} m_{ik} m_{kj} x_i x_j \right) \\ &= m_{kk} \boldsymbol{x}_k^2 + \sum_{\substack{i=1 \\ i \neq k}}^{n} \frac{m_{ki}^2 \boldsymbol{x}_i^2}{m_{kk}} + \sum_{\substack{i=1 \\ i \neq k}}^{n} 2 m_{ik} \boldsymbol{x}_i \boldsymbol{x}_k + \sum_{\substack{i=1 \\ i \neq k}}^{n} \sum_{\substack{j=1 \\ j \neq i \\ j \neq k}}^{n} \frac{m_{ik} m_{kj} \boldsymbol{x}_i \boldsymbol{x}_j}{m_{kk}} \end{aligned} \tag{10}$$

By decomposing equation (10), it can be known that the contribution value consists of four parts:

1) Square term containing $\boldsymbol{x}_k$: $m_{kk} \boldsymbol{x}_k^2$;

2) Cross terms containing $\boldsymbol{x}_k$: $\sum_{\substack{i=1 \\ i \neq k}}^{n} 2 m_{ik} \boldsymbol{x}_i \boldsymbol{x}_k$;

3) Square term excluding $\boldsymbol{x}_k$: $\sum_{\substack{i=1 \\ i \neq k}}^{n} \frac{m_{ki}^2 \boldsymbol{x}_i^2}{m_{kk}}$;

4) Cross-terms that do not contain $\boldsymbol{x}_k$: $\sum_{\substack{i=1 \\ i \neq k}}^{n} \sum_{\substack{j=1 \\ j \neq i \\ j \neq k}}^{n} \frac{m_{ik} m_{kj} \boldsymbol{x}_i \boldsymbol{x}_j}{m_{kk}}$;

As for the contribution rate in the ξ_k direction, weight factors $\alpha_{k,l}$ are added to the part of the second $\boldsymbol{x}_k$ cross term, and the third and the fourth irrelevant terms are eliminated. Finally, the improved reconstruction contribution index is as follows [11]:

$$\mathrm{RBCI}_{\,k} = m_{kk}\boldsymbol{x}_k^2 + \sum_{\substack{i=1 \\ i \neq k}}^{n} 2\alpha_{k,i} m_{ik} m_{kj} \boldsymbol{x}_i \boldsymbol{x}_j \tag{11}$$

Where

$$\alpha_{k,l} = \frac{|\boldsymbol{x}_k|}{|\boldsymbol{x}_k| + |\boldsymbol{x}_l|} \tag{12}$$

3.2 Relative reconstruction contribution

When there is no fault in the system, the minimum requirement for fault diagnosis is that the contribution rate of all variables should be statistically equal, so as to provide a benchmark for the comparison of the contribution rate of all variables in the case of fault. In order to make the contribution rate of each variable statistical equal, the contribution rate of each variable is divided by the corresponding mean value to obtain the relative contribution rate [11,12].

$$\mathrm{Cont}_{r_i}^{\mathrm{SPE}} = \frac{\mathrm{Cont}_i^{\mathrm{SPE}}}{\mathrm{E}(\mathrm{Cont}_i^{\mathrm{SPE}})} \tag{12}$$

The mathematical expectation (mean) of relative contribution rate is

$$\mathrm{E}(\mathrm{Cont}_{r_i}^{\mathrm{SPE}}) = \mathrm{E}\left(\frac{\mathrm{Cont}_i^{\mathrm{SPE}}}{\mathrm{E}(\mathrm{Cont}_i^{\mathrm{SPE}})}\right) = \frac{\mathrm{E}(\mathrm{Cont}_i^{\mathrm{SPE}})}{\mathrm{E}(\mathrm{Cont}_i^{\mathrm{SPE}})} = 1 \tag{13}$$

The relative reconstruction contribution rate is

$$\mathrm{RBC}_{r_k} = \frac{\mathrm{RBC}_k}{E(\mathrm{RBC}_k)} \tag{14}$$

Then the mathematical expectation (mean) of the relative reconstruction contribution rate is

$$\mathrm{E}(\mathrm{RBC}_{r_k}) = \mathrm{E}\left(\frac{\mathrm{RBC}_k}{E(\mathrm{RBC}_k)}\right) = \frac{E(\mathrm{RBC}_k)}{E(\mathrm{RBC}_k)} = 1 \tag{15}$$

It can be seen from the above analysis that, when there is no fault, the mathematical expectation of the relative contribution rate or the relative reconstruction contribution rate of each variable is the same, which can provide a comparison benchmark for the reconstruction contribution rate of each variable in the fault state.

3.3 Algorithm based on improved relative reconstruction contribution value

Through principal component analysis and improved reconstruction contribution method, the reconstruction contribution value RBCI_k after the elimination of irrelevant items in the fault sample matrix $\boldsymbol{M}$ is calculated, as shown in Equation (11). Then, the contribution rate of each variable is divided by the average contribution rate of each sensor direction corresponding to the same sample, and the relative contribution rate of each variable is obtained as:

$$\mathrm{RBCI}_{r_k} = \frac{\mathrm{RBCI}_k}{E(\mathrm{RBCI}_k)}$$

The contribution rate of each sensor direction obtained by this algorithm is not only more accurate than the traditional reconstruction method, but also combines the advantages of relative reconstruction contribution rate, so that the contribution rate of each variable has a comparable benchmark.

4. Simulation experiment

It is assumed that the number of sensors in the multivariable control system is 6, that is, $\boldsymbol{x}(k) = [\boldsymbol{x}_1(k), \boldsymbol{x}_2(k), \boldsymbol{x}_3(k), \boldsymbol{x}_4(k), \boldsymbol{x}_5(k), \boldsymbol{x}_6(k)]^{\mathrm{T}}$. Any sensor fails at each sampling moment, that is, the direction ξ_i may be any of them ξ_1、ξ_2、ξ_3、ξ_4、ξ_5、ξ_6 . Assume the normal data of the system is as follows:

$$\boldsymbol{x}_1(k) = \mathrm{randn}(1) \quad k = 1,2\cdots500$$

$$\boldsymbol{x}_2(k) = 0.8\,\mathrm{randn}(1) \quad k = 1,2\cdots500$$

$$\boldsymbol{x}_3(k) = 0.5\boldsymbol{x}_1(k) + (\boldsymbol{x}_1(k) + \boldsymbol{x}_2(k))/2 \quad k = 1,2\cdots500$$

$$\boldsymbol{x}_4(k) = 0.5\boldsymbol{x}_1(k) + (\boldsymbol{x}_1(k) - \boldsymbol{x}_2(k))/2 \quad k = 1,2\cdots500$$

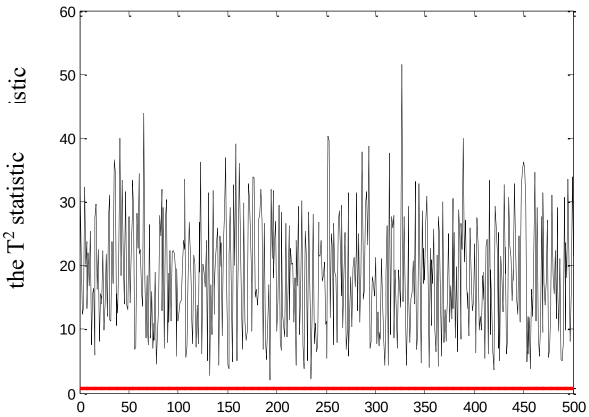
$$\boldsymbol{x}_5(k) = 0.6\,rand(1) \quad k = 1,2\cdots500$$

$$\boldsymbol{x}_6(k) = 0.4\,rand(1) + 0.9\,\boldsymbol{x}_5(k) \quad k = 1,2\cdots500$$

Under this data, the fault model was simulated by formula (5), and the magnitude of the fault f_i was subject to the uniform distribution of [2,5]. There was 500 data in the fault model, and then RBC, RBCI_k , RBCI_{r_k} was used to diagnosis respectively. The diagnosis effect was evaluated by the correct diagnosis rate.

Fig.1 is the monitoring chart of SPE statistics of the data after the introduction of fault. Fig.2 is the monitoring chart of T^2 statistics of the data after the introduction of fault. It can be seen from the figure that both SPE and T^2 statistics exceeded the corresponding control limits, indicating that there was a fault direction detectable in this group of data.

Fig. 1 the SPE monitoring results



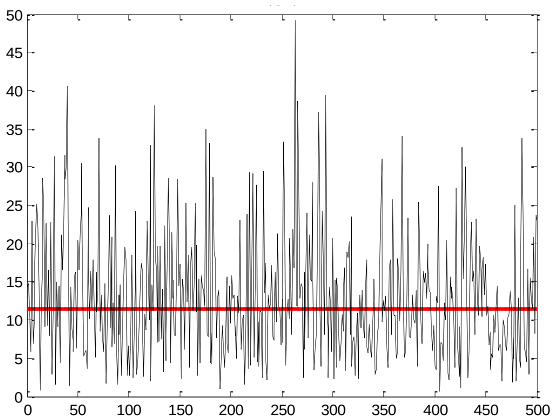


Fig. 2 T^2 monitoring results

After the fault is judged to exist, three kinds of reconstruction are carried out on the fault data respectively to obtain the accuracy rate calculated based on SPE as shown in Table 1. As can be seen from Table 1, the accuracy of improved reconstruction and improved relative reconstruction is the same, but due to differences in original data, the accuracy of improved reconstruction and improved relative reconstruction is not always higher than that of basic reconstruction.

Tab.1 Correct diagnostic rates of three indicators based on SPE

Indicators	RBC	$RBCI_k$	$RBCI_{rk}$
Correct / %	0.60	0.86	0.86

Representative sets of data are selected from 500 sets of data, and their contribution values in the direction of each sensor are shown in Tab.2 and Tab.3. The direction of the sensor with the largest contribution value in the same group of data is the determined fault direction.

Tab.2 Diagnostic results of group 79 data

The fault direction	RBC	$RBCI_k$	$RBCI_{r_k}$
L1	0.002658088	0.245919061	0.417788753
L2	0.020666514	0.095365593	0.162015387
L3	0.007257417	0.226406799	0.38463962
L4	0.000303904	0.153877302	0.261420183
L5	9.312616744	2.663513057	4.525008317
L6	9.312616744	0.146641717	0.249127741
Fault direction	L5	L5	L5
Actual fault direction	L5	L5	L5

Tab.3 Diagnostic results of group 88 data

The fault direction	RBC	$RBCI_k$	$RBCI_{r_k}$
L1	0.008681967	1.648047371	0.550471192
L2	0.067501893	0.566940731	0.189366244
L3	0.023704501	1.477350914	0.493456155
L4	0.000992626	1.530096881	0.511074056
L5	30.41728610	9.312544562	3.110521942
L6	30.41728610	3.428328729	1.145110411
Fault direction	L5 and L6	L5	L5
Actual fault direction	L5	L5	L5

Tab.4 Diagnostic results of group 468 data

The fault direction	RBC	$RBCI_k$	$RBCI_{r_k}$
L1	0.007921	0.658696	0.417392
L2	0.061584	0.01416	0.008973
L3	0.021626	0.308077	0.195217
L4	0.000906	0.696736	0.441497
L5	27.750678	6.224378	3.944166
L6	27.750678	1.566689	0.992755
Fault direction	L6	L5	L5
Actual fault direction	L5	L5	L5

It can be seen from Table 2, Table 3 and Table 4 that when RBC's result data are very similar, it is possible that RBC cannot correctly determine the fault direction, while $RBCI_k$ and $RBCI_{r_k}$ can correctly diagnose the fault direction.

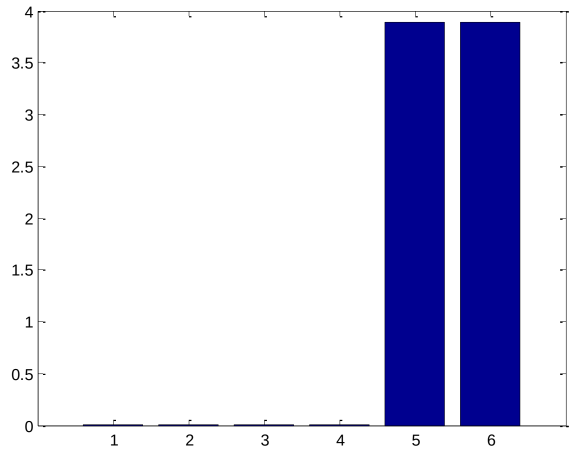


Fig. 3 Contribution values in all directions of the 227th group of data based on reconstruction

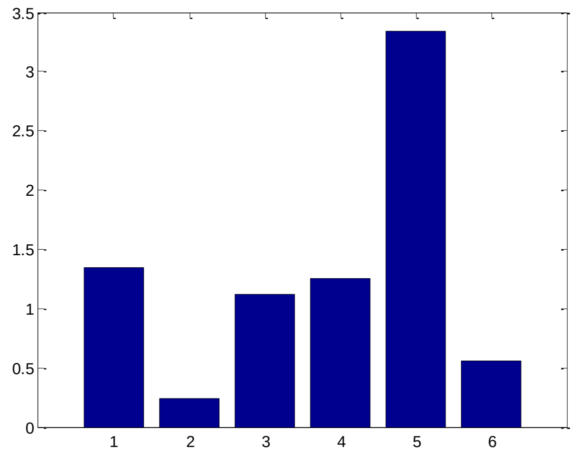


Fig. 4 Contribution values in all directions of the 227th group of data based on improved reconstruction

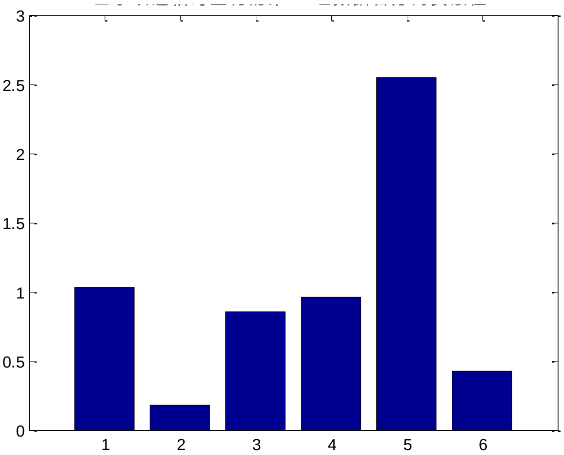


Fig. 5 Contribution values in all directions of the 227th group of data based on improved relative reconstruction

It can be seen from Fig. 3 that there was a small difference in the contribution value obtained by basic reconstruction, which was not easy to distinguish. As can be seen from Fig. 4 and Fig. 5 the results obtained through improved reconstruction or improved relative reconstruction increased the difference between the directions of each sensor and highlight the contribution value in the fault direction. In addition, the contribution value in Fig. 5 had the same reference value and could be compared with each other to obtain the influence of each sensor on the data.

5. Conclusion

To solve the problem of traditional reconstruction method, an improved relative reconstruction algorithm was proposed. In the calculation of the contribution value of the algorithm, the unrelated terms were first eliminated, and then the contribution value of each variable was divided by the mean of the directional contribution rate of each sensor corresponding to the same sample. The simulation results showed that the algorithm not only improved the accuracy, but also provided a benchmark for the comparison of contribution values, so that the contribution values of the same sample could be compared with each other.

Acknowledgements

This work has been supported by fund projects: the national natural science Foundation of China (code: 60974063, 61175059), the nature science foundation of Hebei province(code: F2014205115), Hebei province department of education (code: QN2018087).

References

[1] Bo Zhou, Hao Ye, Haifeng Zhang, Mingliang Li. Process monitoring of iron-making process in a blast furnace with PCA-based methods Original Research Article[J]. *Control Engineering Practice*, 2016, 47:1-14.

[2] Yijun Pan, Chunjie Yang, , Ruqiao An, Youxian Sun.Fault detection with improved principal component pursuit method[J] *Chemometrics and Intelligent Laboratory Systems,*Volume 157, 15 October 2016, Pages 111–119.

[3] Hongquan Jia, Xiao Hea, Donghua Zhou.On the use of reconstruction-based contribution for fault diagnosis[J].Journal of Process Control, 2016(40): 24-34.

[4] YANG Jie-ming,WANG Jing,QU Zhao-yang. Feature Selection Method Based on the Relative Contribution[J]. *Journal of Northeast Dianli University*，2014, 34(4): 62-68.

[5] ZHANG Yewei，FAN Jialu. Fault diagnosis of dual-tank system based on reconstruction-based contribution[J]. Journal of PLA University of Science and Technology(Natural Science Edition).2014（15）2:103-108.

[6] Baligh Mnassri, El Mostafa El Adel, Mustapha Ouladsine. Reconstruction-based contribution approaches for improved faultdiagnosis using principal component analysis[J].Journal of Process Control , 2015(33): 60-76.

[7] Johan A Westerhuis, Stephen P Gurden, Age K Smilde. Generalized contribution plots in multivariate statistical process monitoring[J]. *Chemometrics and Intelligent Laboratory Systems*, 2000,51(1):95-114.

[8] Carlos F Alcala, S Joe Qin. Analysis and generalization of fault diagnosis methods for process monitoring[J]. *Journal of Process Control* ,2011,21(3):322-330.

[9] Alcala C F, Qin S J.Reconstruction-based contribution for process monitoring[J]. *Automatica*, 2009, 45: 1593- 1600.

[10]YE Hao, XU Haipeng, Reconstruction-Based contribution analysis for sensor fault diagnostics[J], *Tsinghua Univ(Sci & Tech)*,2012 ，52（1）：36-39

[11]Alcala C F, Qin S J.Reconstruction-based contribution for process monitoring[J]. *Automatica*, 2009, 45: 1593- 1600.

[12]TANG Yongbo, GUI Weihua, OUYANG Wei. Transformer Fault Diagnosis Method Based on Relative Reconstruction-Based Contribution[J]. *High Voltage Engineering*. 2012, 38(11): 2964-2971.

Du Hailian obtained her master degree control science and control engineering major from the Department of Automation of North China Electric Power University in 2006. She is currently working in Hebei Normal University as associate professor. Her main research interests are in area of neural network control, information fusion, and the fault diagnosis.

Du Wenxia obtained her master degree and doctor degree of control theory and control engineering major from the institute of electrical engineering of Yan Shan University in 2001 and 2012 respectively. She is currently working in Hebei Normal University. Her main research interests are in area of neural network control and predictive control of nonlinear system, the fault diagnosis based on data driven is her new research direction in recent years.

Du Ziyi obtained her bachelor degree electrical engineering and automation major from the College of Vocational and Technical of Hebei Normal University in 2018. She is currently working in State Grid Shahe City Power Supply Company. Her main research interests are the fault diagnosis.

Lv Feng obtained her bachelor's degree from electrical automation major of Hebei Electromechanical College in 1982. She obtained her Master's degree from test measurement technology and instrument major of Yan Shan University in 2006. She is currently working in Hebei Normal University as professor and master supervisor. Her main research interests are in area of industrial intelligent control, information processing, fault detection and diagnosis technology and so on.