The 33rd Youth Academic Annual Conference of Chinese Association of Automation (YAC)
May 18-20, 2018, Nanjing, China

Distributed Bayesian Network with Slow Feature Analysis for Fault Diagnosis

Jie Gao
College of Control Science and Engineering, Zhejiang University
State Key Laboratory of Industrial Control Technology
Hangzhou, China
jiegao@zju.edu.cn

Chunhui Zhao
College of Control Science and Engineering, Zhejiang University
State Key Laboratory of Industrial Control Technology
Hangzhou, China
(corresponding author: chhzhao@zju.edu.cn)

Abstract—**Although the conventional Bayesian network classifier (BNC) has been widely reported in the field of fault diagnosis, sometimes, it may not always work well to separate all faults with a single network. Besides, BNC only considers static distribution information but ignores the dynamic information. The dynamic nature in fault data, which is reflected in temporal behaviors of process data, can describe some meaningful underlying fault characteristics. Here, a novel distributed Bayesian network based on slow feature analysis (SFA-DBN) is proposed for fault diagnosis of complex process. The purpose of this work is to improve the performance of BNC for inseparable faults. The first step of this model is to build a global Bayesian network (GBN) to distinguish the inseparable faults from the separable faults. Second, the inseparable faults are divided into several fault subsets on basis of slow feature analysis (SFA) in which, dynamic variations are similar within the same subset while they are significantly different for different subsets. Third, a set of parallel Bayesian networks (PBNs) are designed to construct the distributed BN for fault diagnosis with different BNs for different fault subsets. Two credible criteria are also presented for offline reasonable partition and online correct identification. Finally, Tennessee Eastman (TE) Process verifies the validity of this proposed model.**

Keywords—Fault diagnosis, Bayesian networks classifiers, SFA, Feature extraction, Industrial process

I. INTRODUCTION

With the increase of complexity in industrial processes, it is necessary to ensure process safety and efficiency so that fault detection and diagnosis (FDD) plays an important role in process system. However, the characteristics of process data (high dimension and significant correlations) bring a lot of difficulties in FDD, because fault information is buried in the huge historical data. Hence, data-driven methods [1] [2] [3] are developed rapidly in this field.

Generally, data-driven methods [4] such as principal component analysis (PCA) [5], partial least squares (PLS) [6] and fisher discriminant analysis (FDA) [7] can transform the highly correlated data into a low-dimensional subspace. With the evolution of data-driven methods, more and more machine learning algorithms have been introduced in FDD research area. Indeed, those algorithms exhibit improved performance in dealing with a large scale industrial process data. Among machine learning algorithms, Bayesian network classifier (BNC) [8] is famous for its capacity of causal analysis under system uncertainty. The Bayesian network (BN) [9] is a type of graphical model and able to model probabilistic influence, which based on a strong mathematical theory, named Bayesian Theorem. The application of BNC for FDD achieves some successes. Parteepasen [10] demonstrates the ability to combine signals from acoustic emission and vibration sensors for tool wear monitoring. Sylvain Verron [11] applied the Bayesian network in multivariate process for fault diagnosis.

However, classical BNC only takes static distribution of fault data into consideration to complete the classification task in fault diagnosis, while ignores the importance of dynamic nature in fault data. The dynamic nature in fault data is reflected in temporal behaviors of process data. Therefore, the information extracted from this aspect is called dynamic information, which explicitly depicts some meaningful underlying fault characteristics [12]. More deeply, dynamic information and static information carry the different fault information, analogous to the concepts of position and velocity in physics, respectively. In terms of complex industrial process, there is usually a phenomenon where some faults are unable to be well classified by classifiers when they are put together. These faults are termed as inseparable faults. The reason of this phenomenon is various, such as co-linearity and the similar position distribution. Fortunately, extracting the dynamic information, as an alternative method, can tackle this problem. Because each fault has its own dynamic nature, the dynamic information can be used to divide the inseparable faults into different classifiers in order to achieve a better classification accuracy.

One of the methods to extract dynamic information is slow feature analysis (SFA). The SFA [13] is a promising dimension reduction methodology that extracts slow varying features from temporal data. From 2002, SFA has received increasing attention in image recognition [14] and human action recognition [15]. Until 2015, it was found to be used in fault detection by Shang [16]. He has defined two new indices to detect anomalies in process dynamics through the SFA. It is notable that, in his work, slow features are be partitioned into two groups, i.e., $\boldsymbol{s}_d = [\boldsymbol{s}_1 ... \boldsymbol{s}_M]^{\mathrm{T}} \in \boldsymbol{R}^M$ and $\boldsymbol{s}_e = [\boldsymbol{s}_{M+1} ... \boldsymbol{s}_m]^{\mathrm{T}} \in \boldsymbol{R}^{M_e}$, where the M slowest features in $\boldsymbol{s}_d$ represent dominant slowlyvarying trends that underlie processes, and the M_e fastest ones in $\boldsymbol{s}_e$ contain short-term fluctuations. Thus, it can be seen that SFA is a good tool to divide faults into different groups by extracting dynamic varying information.

In a complex process system, there are two types of faults: separable faults and inseparable faults. In order to improve the performance of BNC for inseparable faults, this article proposes a novel distributed Bayesian network based on SFA (SFA-DBN) for fault diagnosis. In this algorithm, the first step is to construct a global Bayesian network (GBN) to distinguish the separable faults, which are denoted as F_S and the inseparable faults, which are denoted as F_I. Secondly, for the fault-set F_I, they are divided into several fault subsets based on the difference of fault varying rate. In this step, we design a partition criterion based on SFA (SFA-PC) to divide faults reasonably. Then, after the partition, parallel Bayesian networks (PBN) for different fault subsets will be established. As for now, the entire model including a GBN and a set of PBNs is ready for online diagnosis. Whenever an observation is available, the first step is to put it into the GBN. Next, if it is classified to F_I set, an online discriminant criterion based on SFA (SFA-DC) presented in this paper starts to work, which provides a fast and effective method to precisely discriminate the right fault subset for the observation. Finally, the PBN developed for the chosen subset is used to classify this observation and output the classification result. In this way, both separable faults and inseparable faults can be well classified. Meanwhile, besides the improved performance of BNC, the situation of fault dynamic varying in F_I is also acquired for further analysis of the fault characteristics.

The major contribution is summarized as below:

- Besides the global Bayesian network, a set of parallel Bayesian networks (PBNs) are designed by further separating inseparable faults into different subsets so that the local fault information can be elaborated for fault diagnosis.
- Dynamic features in fault data are effectively extracted with SFA, which reveal some significant underlying fault characteristics and are thus beneficial to distinguish different faults.

The reminder of this paper is organized as follows. Section 2 presents the proposed algorithm. Then, the proposed algorithm is verified on Tennessee Eastman process in Section 3 and a conclusion is drawn in Section 4.

II. Methodology

Aiming to promoting the performance of Bayesian network classifier for inseparable faults, a distributed Bayesian network model based on slow feature analysis (SFA-DBN) is presented. The dynamic nature in fault data plays a significant role on describing the underlying fault characteristics, which is reflected in temporal behaviors of process data. Because each fault has its own dynamic nature, extracting the dynamic information by SFA is a reasonable fashion to divide the inseparable faults into different classifiers for achieving a better result of classification. Hence, besides the proposed model containing a GBN and a set of PBNs, two useful criteria are provided: one is the partition criterion based on SFA (SFA-PC) for dividing faults reasonably; another is the online discriminant criterion based on SFA (SFA-DC) for making sure that online data is identified into the right fault subset efficiently. one is training the conditional probability tables (CPTs) as parameters for a GBN.

A. The Construction of Global Bayesian Network

Considering the performance and simple applicability of the classifier, we choose the Naive Bayesian classifier as the classification network. In fact, the global Bayesian network (GBN) is a global Naïve Bayesian network. To construct a GBN, it has two steps. The first one is building the architecture, which relies on the number of nodes. The second one is training the conditional probability tables (CPTs) as parameters for a GBN.

1) Variables selection for architecture building

Before putting variables into the network, an input sequence of variables is needed because the input way is one by one. Here, a variable sorting algorithm [17] based on mutual information is used to deal with the input variables. Then, a sequence in order from the most informative variables to the noisiest variables is attained. Thirdly, start to pick up the variables into the network by sequence order and test classification performance to select optimized variables, yielding the optimal NBN model followed.

The class node of GBN contains all of types of faults to be classified and the number of child nodes is determined by the result of optimal model selection in Fig.1.

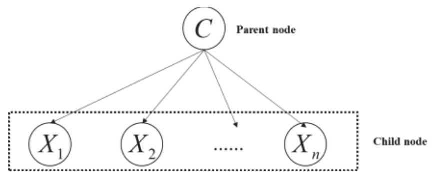


Fig. 1. The Construction of GBN.

2) Training of the conditional probability tables

After building the architecture, training data, which for obtaining the CPTs, should be pre-processed by Independent Component Analysis (ICA) [18] and Entropy Minimum Discretization (EMD) [19]. The former step is to satisfy the special assumption of Naive Bayesian classifier that all the variables are conditionally independent. And the second step can well catch the distribution of each feature.

In the second step, the CPTs are acquired via computing the likelihood function with pre-processed training data. let P be a joint probability distribution over the variables in our network and let $C\{c_1, c_2, \ldots, c_k\}$ be the class node coding r different faults (classes) of the system. The child nodes are indicated by $X_1, X_2, \ldots, X_n$, representing there are n variables or n features in the system, and all the possible child nodes constitute the evidence X . We view the fault with the maximum posterior probability as the right fault type for this new observation X .

The objective function is indicated as follows:

$$\arg\max_{c_i} P(C=c_i \mid X) = \frac{P(X \mid C=c_i)\cdot P(C=c_i)}{P(X)} \tag{1}$$

where $P(C=c_i \mid X)$ is the posterior probability, $P(X \mid C=c_i)$ is likelihood function, extracted from history data, and $P(C=c_i)$ is the prior probability. To simplify, the prior probability of each fault is fixed to $\frac{1}{r}$ (having r different fault types).

3) Setting the threthholds

When the training of GBN is finished, the training data of all faults is divided into F_S portion and F_I portion by a threshold θ . The outputs of this construction are the GBN model and two fault-sets (F_S and F_I).

Furthermore, the threshold θ is defined on the basis of accurate classification rate (ACR) of training data, which depends on individual demand. The detailed discriminant standard is as below:

Firstly, the definition of ACR is,

$$\text{ACR} = \frac{p}{H} \tag{2}$$

where p is the count of this state samples that are classified to this state and H is the total count of this state samples.

Secondly, the discriminant standard is,

$$\begin{cases} \text{if ACR} \ge \theta, & F_i \in F_S \\ \text{if ACR} < \theta, & F_i \in F_I \end{cases} \tag{3}$$

where F_i represents the type of fault.

Specially, there is a misclassification index λ , derived from the average of maximum posterior probability of the right samples to be classified in each F_S type. For online data, if an online sample is determined to a separable fault but its value of maximum posterior probability is lower than λ , this sample will be turned back to the inseparable fault types; otherwise, diagnosis result of this observation in GBN will be output.

B. The Partition Criterion Based on SFA

For F_I set, this article proposes a partition criterion based on SFA (SFA-PC) in order to divide the faults into different fault subsets. In this part, the input is F_I set and the outputs are several fault subsets $F_{V_i}\{F_1, F_2, ..., F_m\}$.

The approach of SFA-PC is as follows.

1) The computation of the average-slow-varying index

Let $\mathbf{X}(t) = \{\boldsymbol{x}_1(t), ..., \boldsymbol{x}_n(t)\}^{\mathrm{T}}$ is the input data. Utilize SFA to process all inseparable fault data at once and an average slow varying matrix $\mathbf{w}^n{}_{sf\_ave}$ will be gained.

Then choose the slowest feature, meaning that pick up the slowest vector $v^n{}_{sf\_ave}$ as the projection vector to extract the average-slowest-varying curve. Thirdly, to calculate the variance $\Delta^n{}_{sf\_ave}$ under the slowest projective vector as the average-slow-varying index, which represents the varying trend of all fault class.

$$\Delta^n{}_{sf\_ave} = \alpha \cdot \sum_{n=1}^{C} \sum_{\substack{k=1,l=2 \\ k<l}}^{M} \left(v^n{}_{sf\_ave} \cdot x_k(t) - v^n{}_{sf\_ave} \cdot x_l(t) \right)^2 \tag{4}$$

$$\alpha = \frac{1}{\sum_{n=1}^{C} (R-1)} \tag{5}$$

where $x_k(t), x_l(t)$ are input data from time series $\mathbf{x}(t)$ and have strict sequential order ($k<l$), α is a normalization constant. R depends on the quantity of samples in the fault data, so if the number of fault types change, R will change too.

2) The computation of single-fault-slow-varying index

Choose the same slowest feature $v^n{}_{sf\_ave}$ as the projection vector for each type of fault. Then, calculate the variance $\Delta^n{}_{sf\_f_i}$ under the slowest projective vector as single-fault-slow-varying index, which represents the varying trend of each fault class.

$$\Delta^n{}_{sf\_f_i} = \beta \cdot \sum_{\substack{k=1,l=2 \\ k<l}}^{M} \left(v^n{}_{sf\_ave} \cdot x_k(t) - v^n{}_{sf\_ave} \cdot x_l(t) \right)^2 \tag{6}$$

$$\beta = \frac{1}{R-1} \tag{7}$$

where $x_k(t), x_l(t)$ are input data from time series $\mathbf{x}(t)$ and have strict sequential order ($k<l$), β is a normalization constant, R depends on the quantity of samples in the fault data.

It is notable that detection system generally begins to produce alarming signal from the sixth point, so fault-slow-varying variance can also be computed in an early time for online stage.

3) Define a discriminant index

Define a discriminant index $\Delta^n{}_{DP\_fi}$ to distinguish the relative fast varying faults and the relative slow varying faults.

$$\Delta^n{}_{DP\_fi} = \Delta^n{}_{sf\_f_i} - \Delta^n{}_{sf\_ave} \tag{8}$$

Where i denotes fault class, n denotes the number of iterations, $\Delta^n{}_{sf\_ave}$ is the average-slow-varying index of all faults to be divided, $\Delta^n{}_{sf\_f_i}$ is the single-fault-slow-varying index. If $\Delta^n{}_{DP\_fi} > 0$, the fault F_i is a relative fast varying fault; otherwise, if $\Delta^n{}_{DP-fi} \le 0$, the fault F_i is a relative slow varying fault.

4) Divide F_I set into several FVR subsets

Considering the accuracy of classification, a hypothesis is made that the proper quantity of fault class equals z in one FVR subset. When it comes to real-time application, parameter z depends on the performance of classifier under the training data.

Each discriminant index $\Delta^{n}{}_{DP\_fi}$ cut the faults set into two portions, the relative fast varying portion $P_F\{F_{M+1},...,F_m\}$ and the relative slow varying portion $P_S\{F_1,...,F_M\}$. If $m-M-1>z$ or $M>z$, repeat to calculate $\Delta^{n}{}_{DP\_fi}$ and $\Delta^{n}{}_{sf\_ave}$ in P_F or P_S until fulfilling the cut-off condition:

$$(M \le z) \text{ or } (m-M-1 \le z) \tag{9}$$

Save each $\Delta^{n}{}_{sf\_ave}$ in every iteration as discriminant index for online application.

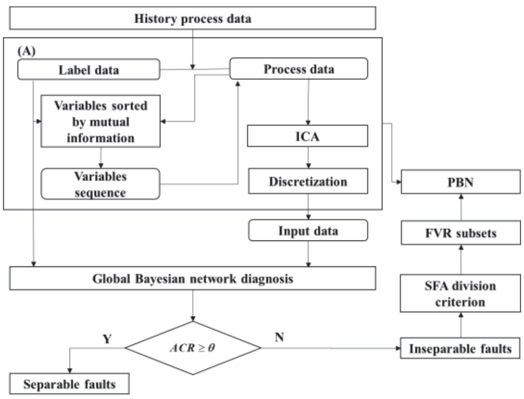


Fig. 2. Framework of fault diagnosis model offline building. Part (A) is the pre-processing of input data for NBNs.

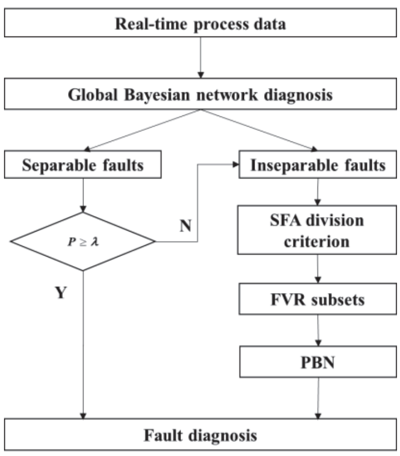


Fig. 3. Framework of online fault diagnosis model.

C. The Construction of Parallel Bayesian Networks

Parallel Bayesian networks (PBNs) are constructed corresponding to the fault subsets. In other words, different BNs are developed for different fault subsets. It is noted that the procedures of the construction of PBNs is the same with the construction of GBN. The outputs of this part are several PBNs. The whole offline modelling process is show in Fig. 2.

In summary, the step-by-step procedure of the proposed approach is given below:

1) Determine the GBN structure based on the optimal model selection result.
2) Set the thresholds { θ , λ } of GBN for online application.
3) Based on θ, distinguish F_I set and F_S set.
4) Then, divide F_I set into different fault subsets $F_{V_i}\{F_1,F_2,...,F_m\}$ by SFA-PC.
5) Construct PBNs by developing a BN for a subset.

D. Online Fault Diagnosis

Firstly, whenever a new observation $\mathbf{x}(\mathbf{t})=\{x_1(t),...,x_n(t)\}^{\mathrm{T}}$ is available, GBN is firstly implemented to detect whether the data belongs to inseparable faults or not. If the data can be well classified in the GBN and the value of posterior probability exceeds the threshold λ (indicated in Section A), the classification result is output as diagnosis result of this observation. However, if this observation is classified in an inseparable fault, it will be put into the next step.

The second step is to determine which fault subset $F_{V_i}\{F_1,F_2,...,F_m\}$ is the right one for the observation by SFA-DC. The inference of SFA-DC is as follow:

1) Calculate the index $\Delta^{n'}{}_{sf\_f_i}$ for online data.
2) Compare $\Delta^{n'}{}_{sf\_f_i}$ with each $\Delta^{n}{}_{sf\_ave}$ saved from offline modelling to get $\Delta'^{n}{}_{DP\_fi}$, then laid this observation into its belonging F_{V_i} according to the $\Delta'^{n}{}_{DP\_fi}$.

Finally, this observation is classified by the PBN related to the chosen F_{V_i} and output the result of this classification of the observation. The framework of online fault diagnosis is shown in Fig. 3.

III. Illustration

In this section the proposed SFA-DBN model is applied to the well-known Tennessee Eastman (TE) process to test the practicability in a chemical industrial system.

A. Process Description

The TE process is further used to evaluate the effectiveness of methods of FDD. The article of Downs [20] entirely describes this process. The process consists of five major units: an exothermic two-phase reactor, a flash separator, a recycle compressor, a re-boiled stripper, and a product condenser. There are total 41 measurement variables and 12 manipulated variables in TE process. Moreover, dataset is sampled with an interval of 3 minutes.

On the purpose to objectively evaluating our SFA-DBN model, the input data of each type of fault includes 600 training samples, 480 validate samples, and 120 test samples. In this article, we use 10 types of faults F {#1, #2, #3, #4, #6, #7, #10, #11, #12, #18} and 52 variables for simulation.

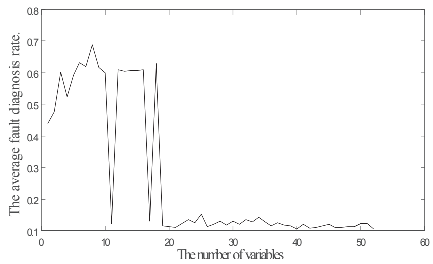


Fig. 4. The optimal model selection of the global Bayesian network.

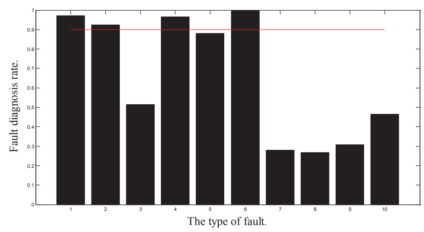
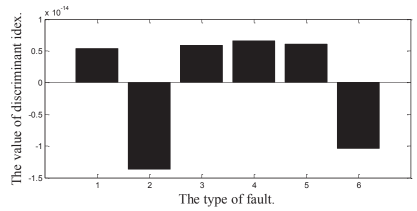
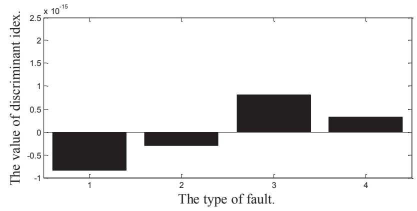


Fig. 5. The diagnosis result of the global Bayesian network.



(a)



(b)

Fig. 6. The values of discriminant index.

B. SFA-DBN for Fault Diagnosis

1) The construction of SFA-DBN

The first step is to design the structure of the global Bayesian network (GBN) based on result of optimal model selection shown in Fig. 4. The designed GBN consist of a parent node with ten fault states and eight child nodes referring to the selected informative variables. The CPT of GBN is estimated from the pre-processed training data. Furthermore, other parameters such as θ and λ are set to 0.9 and 0.85, respectively. The diagnosis result of GBN is depicted in Fig. 5. Based on θ, faults in F_S are selected, that are F {#1, #2, #4, #7}. And the rest of faults belong to F_I set.

TABLE I. THE DISCRIMINANT RESULT OF SFA-DC

	Fault	PBN1	PBN2	PBN3	ACR
	#6	120	0	0	1.00
PBN1	#18	101	13	6	0.84
	#03	0	100	12	0.83
PBN2	#10	8	112	0	0.93
	#11	0	0	120	1.00
PBN3	#12	0	0	120	1.00
	Average ACR				0.93

TABLE II. THE DIAGNOSIS RESULT FOR TEP EVALUATED BY FDR USING (A) THE OPTIMIZED VARIABLE SELECTION METHOD BASED PCA (GHOSH ET AL., 2014); (B) THE BAYESIAN METHOD (JIANG ET AL., 2016, AND (C) THE PROPOSED METHOD.

Fault No.	(a)	(b)	(c)
#1	99.87	100.00	99
#2	97.87	99.00	98
#4	100.00	100.00	90
#7	100.00	100.00	100
#3	2.37	6.00	80
#6	99.50	100.00	100
#10	82.25	84.00	93
#11	64.75	82.00	95
#12	99.00	100.00	100
#18	89.50	90.00	84

Within the identified F_I periods, the partition based on SFA is further carried out. Firstly, the scale of each fault subset is set to holding 3 fault states for a more correct classification, meaning z is 3. Then, use training data in F_I to extract the average slowest feature vector v_{sf_ave}. According to the v_{sf_ave}, the values of discriminant index $\Delta_{DP\_f_i}$ for each type of fault are computed. In Fig. 6(a), the x-axis is Fault #3, Fault #6, Fault #10, Fault #11, Fault #12, Fault #18 from the left to right. It is obvious that $\Delta_{DP\_f_6}$ and $\Delta_{DP\_f_{18}}$ are significantly lower than zero, which means Fault #6 and Fault #18 have a relative slower varying rate among these six faults.

Subsequently, $P_F\{\#3,\#10,\#11,\#12\}$ is the relative fast varying portion, partitioned in the same way. As shown in Fig. 6(b), $\Delta^2_{DP\_f_3}$ and $\Delta^2_{DP\_f_{10}}$ are lower than zero, so the relative slow varying portion is $P_S^{'}$ {#3, #10} and the relative fast varying portion is $P_F^{'}$ {#11, #12}. Now we have three fault subsets:

F_{V_1} {#6, #18}, F_{V_2} {#3, #10}, and F_{V_3} {#11, #12} for the next step.

Three fault subsets correspond to three PBNs, respectively. The construction of PBNs is in the same way with GBN.

2) SFA-DBN for online fault diagnosis

When the online input observation is a fault data in F_I, the SFA-DC is used to discriminate which fault subset should be applied to. As shown in Table Ⅰ, the accuracy of SFA-DC is reliable because its average correct classification rate is 0.93. Finally, use the corresponding PBN to output the classification result.

Compared with other diagnosis models, the performance of our model has much improvement, especially for Fault #3, Fault #10 and Fault #11. The improved results are shown in Table Ⅱ.

The accurate classification rate (ACR, mentioned in Section A) of Fault #3 reaches 83%, the ACR of Fault #10 reaches 93%, and the ACR of Fault #11 reaches 100%. Moreover, the order of varying rate in F_I can also be acquired, which is {#6, #18, #3, #10, #12, #11}, revealing the lowest varying fault to fastest varying fault.

IV. Conclusion

To improve the accuracy of Bayesian network classifier for complex processes, an effective SFA-DBN algorithm is proposed in this paper. By dividing inseparable faults into different subsets and analyzing the dynamic features, the inseparable faults can be well classified by development of PBNs. Furthermore, future work may focus on how to extend the proposed method to trace the fault root variables.

Acknowledgment

This work is supported by NSFC-Zhejiang Joint Fund for the Integration of Industrialization and Informatization (No. U1709211), the Research Project of the State Key Laboratory of Industrial Control Technology, Zhejiang University, China (ICT1802) and the Open Research Project of the State Key Laboratory of Industrial Control Technology, Zhejiang University, China (No. ICT1800397).

References

[1] C. F. Alcala, S. J. Qin, "Reconstruction-based contribution for process monitoring with kernel principal component analysis," Industrial & Engineering Chemistry Research, 2010, 49(17), pp.7849-7857.

[2] C. C. Hsu, C. T. Su, "An adaptive forecast-based chart for non-Gaussian processes monitoring: with application to equipment malfunctions detection in a thermal power plant," IEEE Transactions on Control Systems Technology, 2011, 19(5), pp.1245-1250.

[3] C.H. Zhao, F.R. Gao, "Fault-relevant principal component analysis (FPCA) method for multivariate statistical modeling and process monitoring," Chemometrics and Intelligent Laboratory Systems, 2014, 133, pp.1-16.

[4] C. H. Zhao, F. R. Gao, "Critical-to-fault-degradation variable analysis and direction extraction for online fault prognostic," IEEE Transactions on Control Systems Technology, 2017, 25.3, pp.842-854.

[5] J. V. Kresta, J. F. MacGregor, T. E. Marlin, "Multivariate statistical monitoring of process operating performance," The Canadian journal of chemical engineering, 1991, 69(1), pp.35-47.

[6] M. J. Piovoso, K. A. Kosanovich, "Applications of multivariate statistical methods to process monitoring and controller design," International Journal of Control, 1994, 59(3), pp.743-765.

[7] L. H. Chiang, E. L. Russell, R. D. Braatz, "Fault diagnosis in chemical processes using Fisher discriminant analysis, discriminant partial least squares, and principal component analysis," Chemometrics and intelligent laboratory systems, 2000, 50(2), pp.243-252.

[8] N. Friedman, D. Geiger, M. Goldszmidt, "Bayesian network classifiers," Machine learning, 1997, 29(2-3), pp.131-163.

[9] J. Pearl, "Fusion, propagation, and structuring in belief networks," Artificial intelligence, 1986, 29(3), pp.241-288.

[10] A. Prateepasen, Y. H. J. Au, B. E. Jones, "Acoustic emission and vibration for tool wear monitoring in single-point machining using belief network," Instrumentation and Measurement Technology Conference, 2001, Proceedings of the 18th IEEE, 2001, Vol. 3. IEEE.

[11] S. Verron, J. Li, T. Tiplica, "Fault detection and isolation of faults in a multivariate process with Bayesian network," Journal of Process Control, 2010, 20(8), pp.902-911.

[12] C. H. Zhao, H. Biao, "A full-condition monitoring method for nonstationary dynamic chemical processes with cointegration and slow feature analysis," AIChE Journal, 2018, 64.5, pp.1662-1681.

[13] L. Wiskott, T. J. Sejnowski, "Slow feature analysis: Unsupervised learning of invariances," Neural computation, 2002, 14(4), pp.715-770.

[14] L. Zafeiriou, M. A. Nicolaou, S. Zafeiriou, S. Nikitidis, M. Pantic, "Learning slow features for behaviour analysis," Computer Vision (ICCV), 2013 IEEE International Conference on. IEEE, 2013, pp.2840-2847.

[15] Z. Zhang, D. Tao, "Slow feature analysis for human action recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(3), pp.436-450.

[16] C. Shang, B. Huang, F. Yang, D. Huang, "Slow feature analysis for monitoring and diagnosis of control performance," Journal of Process Control, 2016, 39, pp.21-34.

[17] S. Verron, T. Tiplica, A. Kobi, "Fault detection and identification with a new feature selection based on mutual information," Journal of Process Control, 2008, 18(5), pp.479.-490.

[18] C. Jutten, J. Herault, "Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture," Signal processing, 1991, 24(1), pp.1-10.

[19] U. M. Fayyad, K. B. Irani, "On the handling of continuous-valued attributes in decision tree generation," Machine learning, 1992, 8(1), pp.87-102.

[20] J. J. Downs, E. F. Vogel, "A plant-wide industrial process control problem," Computers & chemical engineering, 1993, 17(3), pp.245-255.