Distance Rejection in a Bayesian Network for Fault Diagnosis of Industrial Systems

Sylvain Verron, Teodor Tiplica, Abdessamad Kobi

Distance Rejection in a Bayesian Network for Fault Diagnosis of Industrial Systems

Sylvain VERRON, Teodor TIPLICA, Abdessamad KOBI

Abstract— The purpose of this article is to present a method for industrial process diagnosis with bayesian network. The interest of the proposed method is to combine a discriminant analysis and a distance rejection in a bayesian network in order to detect new types of fault. The performances of this method are evaluated on the data of a benchmark example: the Tennessee Eastman Process. Three kinds of fault are taken into account on this complex process. The challenging objective is to obtain the minimal recognition error rate for these three faults and to obtain sufficient results in rejection of new types of fault.

I. INTRODUCTION

Nowadays, industrial processes are more and more complex. So, they include a lot of sensors. Consequently, an important amount of data can be obtained from a process. A process dealing with many variables can be named multivariate process. But, the monitoring of a multivariate process cannot be reduced to the monitoring of each process variable because the correlations between the variables have to be taken into account. Process monitoring is an essential task. The final goal of the process monitoring is to reduce variability, and so, to improve the quality of the product [1]. The process monitoring comprises four procedures: fault detection (decide if the process is under normal condition or out-of-control); fault identification (identify the variables implicated in an observed out-of-control status); fault diagnosis (find the root cause of the disturbance); process recovery (return the process to a normal status).

Three major kinds of approaches exists for process monitoring [2]: data-driven, analytical and knowledge-based. Theoretically, the best method is the analytical one because this method constructs mathematic models of the process. But, for large systems (lots of inputs, outputs and states), obtaining detailed models is almost impossible. In the knowledge-based category are placed methods that are based on qualitative models like fault tree, FMECA, expert systems [3]. Finally, data-driven methods are techniques based on rigorous statistical development of process data. Our interest is to monitor large systems, and so, we are concerned with data-driven methods.

In literature, we can find many different data-driven techniques for process control. For the fault detection of industrial processes many methods have been submitted: univariate statistical process control (Shewhart charts) [4], [1], multivariate statistical process control (T^2 and Q charts) [5], [6], and some PCA (Principal Component Analysis)

Three authors are with LASQUO/ISTIA, University of Angers, 49000 Angers, France `sylvain.verron@univ-angers.fr`

based techniques [7] like Multiway PCA or Moving PCA [8]. Kano et al. [9] make comparisons between these different techniques. For the fault identification procedure, one of the better statistical techniques is the MYT decomposition of the T^2 statistic [10]. Finally, for the fault diagnosis techniques we can cite the book of Chiang, Russell and Braatz [2] which presents a lot of them (PCA based techniques, Fisher Discriminant Analysis, PLS based techniques, etc).

The purpose of this article is to present a new method for the diagnosis of faults in large industrial systems. This method is based on bayesian networks and particularly bayesian network classifiers. The major interest of this method is the combination of a discriminant analysis and distance rejections in a bayesian network in order to detect new types of fault of the system.

The article is structured in the following manner. In section II, we introduce the classical method to diagnose faults with bayesian network classifiers. The section III gives a method to apply a distance rejection on a fault of a system with a bayesian network, and presents the combination of this distance rejection with the classical diagnosis with bayesian network. The section IV presents an application of the proposed method for diagnosis of three types of fault on the benchmark Tennessee Eastman Problem. Finally, we conclude on interests and limitations of this method, and we present some perspectives of the fault diagnosis with bayesian network classifiers.

II. BAYESIAN NETWORK FOR FAULT DIAGNOSIS

A. Classification task

Once a problem (fault) has been detected in the evolution of the process by the mean of a detection method, we need to identify (diagnosis) the belonging class of this fault. Thereby, the diagnosis problem can be viewed as the task to correctly classify this fault in one of the predefined fault classes. The classification task needs the construction of a classifier (a function allocating a class to the observations described by the variables of the system). Two types of classification exist: unsupervised classification which objective is to identify the number and the composition of each class present in the data structure; supervised classification where the number of classes and the belonging class of each observation are known in a learning sample and whose objective is to class new observations to one of the existing classes. For example, given a learning sample of a bivariate system with three different known fault as illustrated in the figure 1, we can easily use supervised classification to classify a new faulty

observation. In this study, we use the bayesian network as a supervised classification tool.
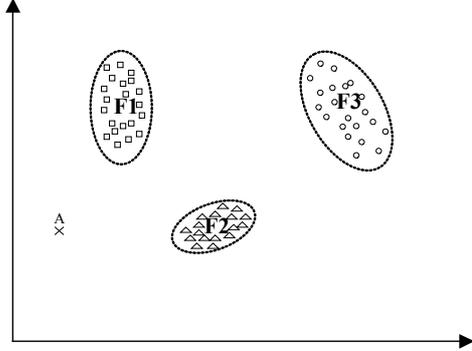

Fig. 1. Bivariate system with three different known fault

B. Bayesian network classifiers

A Bayesian Network (BN) [11] is an acyclic graph where each variable is a node (that can be continuous or discrete). Edges of the graph represent dependences between linked nodes. A formal definition is given here:

A bayesian network is a triplet $\{\mathbf{G}, \mathbf{E}, \mathbf{D}\}$ where:

$\{\mathbf{G}\}$ is a directed acyclic graph, $\mathbf{G} = (V, A)$, with V the ensemble of nodes of $\mathbf{G}$, and A the ensemble of edges of $\mathbf{G}$,

$\{\mathbf{E}\}$ is a finite probabilistic space (Ω, Z, p), with Ω a non-empty space, Z a collection of subspace of Ω, and p a probability measure on Z with $p(\Omega) = 1$,

$\{\mathbf{D}\}$ is an ensemble of random variables associated to the nodes of $\mathbf{G}$ and defined on $\mathbf{E}$ such as:

$$p(V_1, V_2, \ldots, V_n) = \prod_{i=1}^{n} p(V_i | C(V_i)) \qquad (1)$$

with $C(V_i)$ the ensemble of causes (parents) of V_i in the graph $\mathbf{G}$.

Bayesian network classifiers are particular BN [12]. They always have a discrete node C coding the k different classes of the system. Thus, other variables $X_1, \ldots, X_p$ represent the p descriptors (variables) of the system.

A famous bayesian classifier is the Naïve Bayesian Network (NBN), also named Bayes classifier [13]. This bayesian classifier makes the strong assumption that the descriptors of the system are class conditionally independent. Assuming the hypothesis of normality of each descriptor, the NBN is equivalent to the classification rule of the diagonal quadratic discriminant analysis. But, in practice, this assumption of independent and non correlated variables is not realistic. In order to deal with correlated variables, several approaches have been developed. We can cite the Tree Augmented Naïve bayesian networks (TAN) [12]. These BNs are based on a NBN but a tree is added between the descriptors. An other interesting approach is the Kononenko one [14], which represent some variables in one node. As in [15] the assumption we will make is that this variable follows a normal

multivariate distribution (conditionally to the class) and we will refer to this kind of BN as Condensed Semi Naïve Bayesian Network (CSNBN). The CSNBN is equivalent to the discriminant rule of the quadratic discriminant analysis.
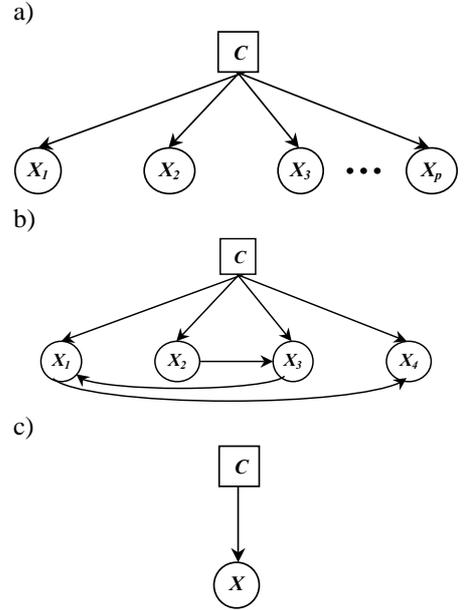

Fig. 2. Different bayesian network classifiers: NBN (a), TAN (b) and CSNBN (c).

C. Bayesian network for fault diagnosis

In the context of the diagnosis of industrial systems, bayesian networks have been already used and give convenient results compared to other classification tools like support vector machines, neural networks or k-nearest neighborhoods [16], [17], [18], [19], [20]. As the performances of the CSNBN have been previously demonstrated [17], [18], we choose this classifier in this article. It is equivalent to a Discriminant Analysis (DA). So, we name the class node DA (coding the different known faults of the system), and the observation node $\boldsymbol{X}$ (a normal multivariate node). The figure 3 presents the bayesian network equivalent to a discriminant analysis, with the probability tables associated to each node. To simplify, the a priori probability of each class F_i is fixed to $p(F_i) = \frac{1}{k}$. The node $\boldsymbol{X}$ follows the different normal probability densities ($\mathcal{N}$) conditionally to the class of DA, where $\boldsymbol{\mu}_i$ is the mean vector of the fault F_i, $\boldsymbol{\Sigma}_i$ is the covariance matrix of the fault F_i. $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}_i$ are estimated on the fault database by Maximum Likelihood Estimation (MLE). On the simple example of the figure 1, the bayesian network gives the different areas of classification of the figure 4.

As we mentioned previously, an objective of a fault diagnosis method is to classify new observations to one of the existing classes. But, in certain cases, the observation may be a new type of fault (unknown or unseen before). This is the case when the observation is distant of any known classes of fault (example of the point A in figure 1). In order to

DA		
F_1	$\ldots$	F_k
$\frac{1}{k}$	$\ldots$	$\frac{1}{k}$

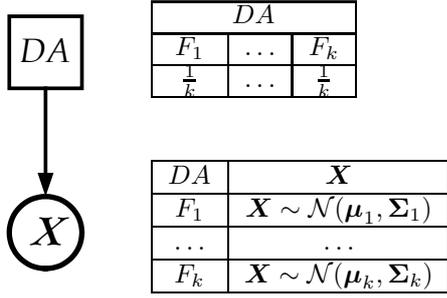| DA	X
F_1	$X \sim \mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$
$\ldots$	$\ldots$
F_k	$X \sim \mathcal{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$

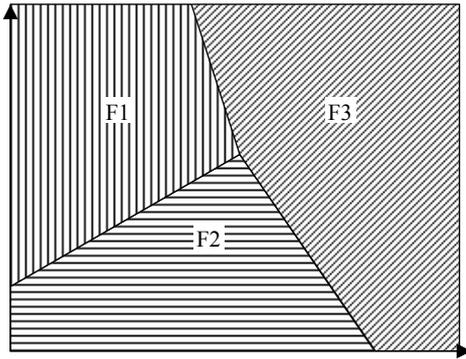Fig. 3.   Bayesian network equivalent to a discriminant analysis

Fig. 4.   Classification areas of the bivariate system

detect these new types of fault, we have to use a criterion called distance rejection (see [21], [22]). Now, we will see how to take into account this criterion in a bayesian network, and how to combined it with the diagnosis of the previous bayesian network.

III. INTEGRATION OF DISTANCE REJECTION

A. Distance rejection in a bayesian network

Assuming one class of fault (gaussian), the rejection of an observation is equivalent to a T^2 control chart [5] on the data of this fault. In previous works [23], [24], we have demonstrated that a T^2 control chart could be modelized with a bayesian network, and more particularly with a CSNBN. Assuming $\boldsymbol{\mu}_i$ the mean vector and $\boldsymbol{\Sigma}_i$ the covariance matrix of the fault F_i, we can monitor the fault F_i with the following rule : if $p(F_i|\boldsymbol{x}) < p(F_i)$ then the observation $\boldsymbol{x}$ can not be attribute to the fault F_i (distance rejection for the fault F_i). For that, it can be used the CSNBN of the figure 5.

F_i	
True	False
$1 - \alpha$	α

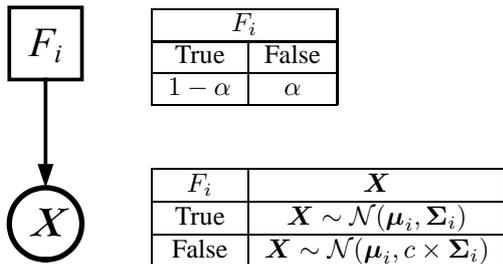| F_i	X
True	$X \sim \mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$
False	$X \sim \mathcal{N}(\boldsymbol{\mu}_i, c \times \boldsymbol{\Sigma}_i)$

Fig. 5.   Distance rejection of fault F_i

In the figure 5, we can see that a coefficient c is implicated in the distance rejection. This coefficient is the root (non equal to 1) of the following equation:

$$1 - c + \frac{pc}{CL} \ln(c) = 0 \qquad (2)$$

where p is the dimension of the system, and CL is the control limit of the T^2 control chart. In numerous cases, CL is equal to $\chi^2_{\alpha,p}$, the quantile to the α value of the χ^2 distribution with p degree of freedom [1]. So, α allows tuning the distance rejection: the higher is α, the stronger is the distance rejection.

B. Fault diagnosis with distance rejection in a bayesian network

It is possible to combine a discriminant analysis and the distance rejection notion in a bayesian network. Indeed, we have the probabilities associated with the different known fault (discriminant analysis), and we can know if a suspected observation can be attribute to none, one or several types of fault (distance rejection on each type of fault). But, for this second point, a decision step is necessary to do conclusions (test if $p(F_i|\boldsymbol{x}) < p(F_i)$). Unfortunately, classical inference in a bayesian network is not able to realize this step. So, we will make two successive inferences. Between these two inferences, soft probabilities of the fault F_i will be transformed in hard evidence (conclusion of the test for each fault) for the second inference. The figure 6 presents the bayesian network proposed for the fault diagnosis with integration of distance rejection.
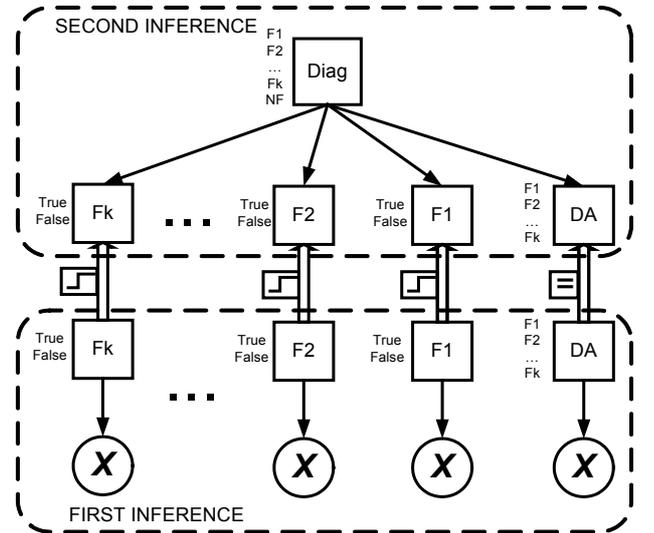
Fig. 6.   Fault diagnosis with distance rejection in a bayesian networks

In the previous section, we have already detailed the different conditional probabilities tables implicated in the first inference. We will study the conditional probabilities tables implicated in the second inference.

Firstly, we can see that in the second inference a node has been added: the node "$Diag$". This node represents the final decision concerning the suspected observation. This node has

$k + 1$ modalities (k is the number of types of fault), one for each fault (the k F_i) and one for a New type of Fault NF. The a priori probabilities table of this node is fixed in order to not advantage any modality. As we can see on the table I, each modality has the same a priori probability : $\frac{1}{k+1}$.

Diag			
F_1	$\ldots$	F_k	NF
$\frac{1}{k+1}$	$\ldots$	$\frac{1}{k+1}$	$\frac{1}{k+1}$

TABLE I

A PRIORI PROBABILITIES TABLE OF THE NODE $Diag$

Each node F_i has the conditional probabilities table of the table II. We have set the different probabilities in order to respect the following rules :

- if $Diag = F_i$, then it is certain that the observation is from fault F_i,
- if $Diag = NF$, then it is certain that the observation is not from fault F_i,
- if $Diag = F_j$ (with $F_j \neq F_i$), then no knowledge is learned on the attribution of the observation to the fault F_i. So, we fix $p(F_i = True|Diag = F_j) = p(F_i = False|Diag = F_j) = 0.5$ (this value can be tuned to advantage some faults of the system).

	F_i	
$Diag$	True	False
F_1	0.5	0.5
$\ldots$	$\ldots$	$\ldots$
F_i	1	0
$\ldots$	$\ldots$	$\ldots$
F_k	0.5	0.5
NF	0	1

TABLE II

CONDITIONAL PROBABILITIES TABLE OF NODES F_i

The table III presents the conditional probabilities table of the node AD during the second inference. We can see that the knowledge of a fault F_i at the node $Diag$ allows to set the knowledge of the node DA, expressed by $P(DA = F_i|Diag = F_i) = 1$. But, the knowledge on $Diag$ of a new type of fault NF do not give any information for the discrimination between the different faults F_i of the node DA.

	DA		
$Diag$	F_1	$\ldots$	F_k
F_1	1	$\ldots$	0
$\ldots$	$\ldots$	$\ldots$	$\ldots$
F_k	0	$\ldots$	1
NF	$\frac{1}{k}$	$\ldots$	$\frac{1}{k}$

TABLE III

CONDITIONAL PROBABILITIES TABLE OF THE NODE AD (SECOND INFERENCE)

The interest of the second inference is to add the different results of distance rejection to the result of the discriminant analysis of the first inference. The application of the network of the figure 6 to the bivariate example gives the figure 7.
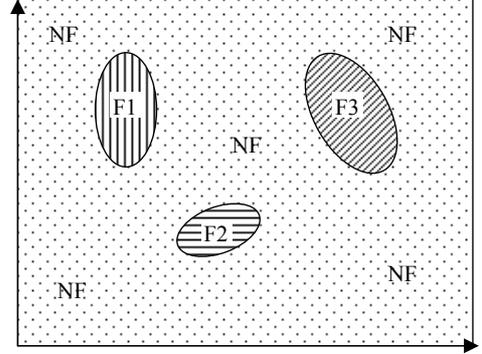

Fig. 7. Fault diagnosis with distance rejection for the bivariate example

Now, we will see an application of this approach on a benchmark problem: the Tennessee Eastman Process (figure 8).

IV. APPLICATION TO THE TEP

A. Presentation of the TEP

The Tennessee Eastman Process (TEP) is a chemical process. It is not a real process but a simulation of a process that was created by the Eastman Chemical Company to provide a realistic industrial process in order to evaluate process control and monitoring methods. Article of Downs and Vogel [25] entirely describes this process. Authors also give the Fortran code of the simulation of the process. Ricker [26] has implemented the simulation on Matlab. The TEP is composed of five major operation units: a reactor, a condenser, a compressor, a stripper and a separator. Four gaseous reactant A, C, D, E and an inert B are fed to the reactor where the liquid products F, G and H are formed. This process has 12 input variables and 41 output variables. The TEP has 20 types of identified faults. So, this process is ideal to test monitoring methods. But, it is also a benchmark problem for control techniques because it is open-loop unstable. Many articles (authors) present the TEP and test their approaches on it. For example, in fault detection, we can cite [9], [27]. Some fault diagnosis techniques have also been tested on the TEP [2], [28], [29] with the plant-wide control structure recommended in Lyman and Georgakis [30]. In [28], [29], authors focus on only 3 types of fault and give the datasets they used. For this reason, we will take the same data that in these articles and compare our approach to those of the others.

As we said, we have taken into account 3 types of faults: fault 4, 9 and 11 (see table IV). These three types of fault are good representations of overlapping data and so, are not easy to classify. As indicated on the table IV, each type of fault have 2 datasets: a training sample and a testing sample, containing respectively 480 and 800 observations. We have tested our approach on the Tennessee Eastman Process. We precise that all computations have been made on Matlab with the BNT (BayesNet Toolbox) developed by Murphy [31].
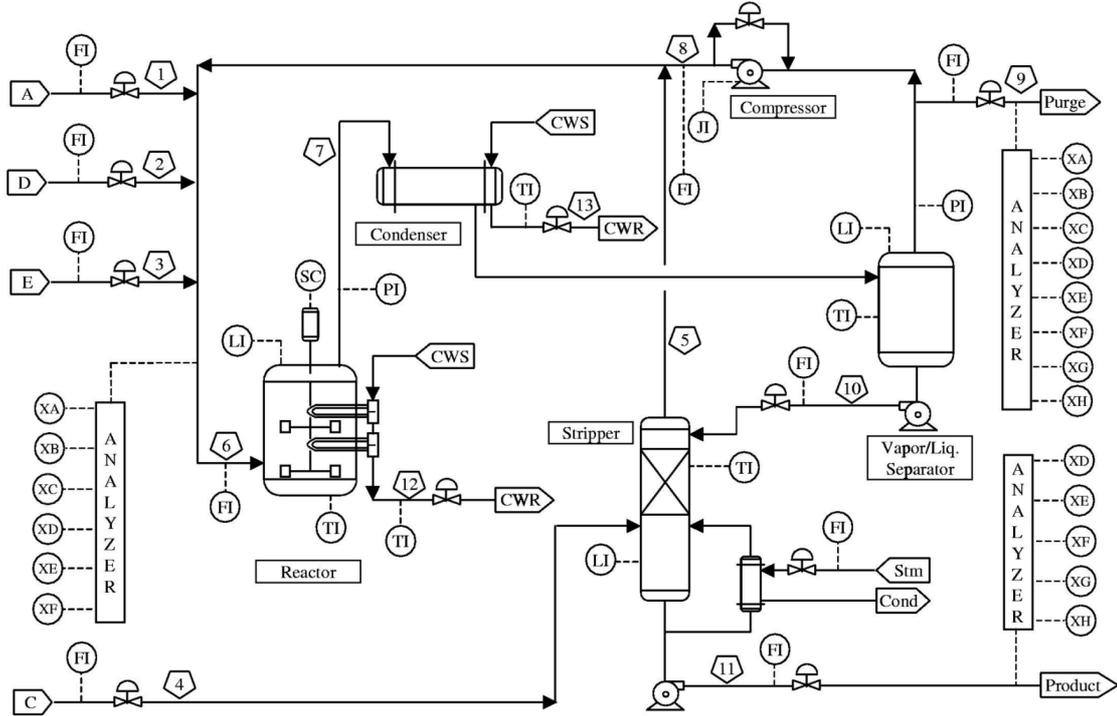
Fig. 8. Process flowsheet of the TEP

Class	Fault type	Train data	Test data
1	Fault 4: step change in the reactor cooling water inlet temperature	480	800
2	Fault 9: random variation in D feed temperature	480	800
3	Fault 11: random variation in the reactor cooling water inlet temperature	480	800

TABLE IV

DESCRIPTION OF FAULT DATASETS

Method	FDA	SVM	PSVM	ISVM	**BN**
Misclassification rate	38%	44%	35%	29.86%	**18.83%**

TABLE V

MISCLASSIFICATION RATE OF THE DIFFERENT METHODS

B. Proposed approach without distance rejection

In the first part of this application, we have applied the proposed approach without the integration of the distance rejection. In an objective evaluation purpose of our procedure and to compare it with the results of other published methods (like Support Vector Machines), we classified 2400 new observations (800 of each type of fault) of the TEP. The results are given in the table V. For the bayesian network (BN) approach, we compute the misclassification rate (percentage of observations which are not well classified). We are also giving the results of other methods on the same data. The results for the FDA (Fisher Discriminant Analysis), SVM (Support Vector Machines), PSVM (Proximal Support Vector Machines) and ISVM (Independent Support Vector Machines) methods are extracted from [28] and [29].

On the table V, we can observe that the BN approach outperforms all the others methods. The confusion matrix

for the bayesian network is given on table VI and gives us the possibility to see how the discrimination of the different faults is done. Each column of the matrix represents the instances in a predicted class, while each row represents the instances in an actual class. For example, for 800 tested observations of fault 4, the diagnosis procedure gives 141 observations as the fault 11, and 659 observations as the fault 4, so 17.62% (141/800) of misclassified observations for the fault 4.

Class	Fault 4	Fault 9	Fault 11	Total
Fault 4	659	0	28	687
Fault 9	0	582	66	648
Fault 11	141	218	706	1065
Total	800	800	800	2400

TABLE VI

CONFUSION MATRIX OF BN

C. Integration of distance rejection

In this second part, the distance rejection is taken into account and we fix α to 0.001 for each known fault. The network has to classify 800 new observations of each of the 17 other faults of the TEP (so, a total of 13600 new

observations). The table VII represents the confusion matrix for this case where the label NF means New type of Fault.

Class	Fault 4	Fault 9	Fault 11	NF	Total
Fault 4	653	0	28	0	681
Fault 9	0	573	64	1596	2233
Fault 11	145	221	688	877	1931
NF	2	6	20	11127	11155
Total	800	800	800	13600	16000

TABLE VII

CONFUSION MATRIX OF BN WITH DISTANCE REJECTION

The table VII shows the results of the integration of the distance rejection in the bayesian network. We can see that the discrimination of the three known faults is not really affected by this integration. Indeed, the misclassification rate for this 3 faults is 20.25%, instead of 18.87% previously. However, this table shows correctly the advantage of this approach since 11155 observations on 13600 have been correctly classified as NF (rate of 81.8%). In this case, an unsupervised classification tool as the K-mean algorithm would be able to identify the different classes of these observations. Finally, we can say that the proposed bayesian network allows to detect new types of fault, without losing performances of discrimination between known faults.

V. CONCLUSION AND OUTLOOKS

The main interest of this article is the presentation of a new method for the fault diagnosis of industrial processes. This method uses a faults database to construct a bayesian network. This bayesian network is able to discriminate between the different known faults of the system, but it is also able to recognize some new types of fault on the system. The performances of this approach have been tested on a concrete example: the Tennessee Eastman Process. The results of the method are good and outperform some previous results of other published methods.

First outlook of the proposed method will be application of some notions of the decision theory, in order to avoid the separation of the network in two inferences. So, this application will make the network faster and more simple. A second outlook of the proposed approach is the extension to the gaussian models mixture in order to take into account some non normal classes.

REFERENCES

[1] D. C. Montgomery, *Introduction to Statistical Quality Control, Third Edition*. John Wiley and Sons, 1997.
[2] L. H. Chiang, E. L. Russell, and R. D. Braatz, *Fault detection and diagnosis in industrial systems*. New York: Springer-Verlag, 2001.
[3] V. Venkatasubramanian, R. Rengaswamy, and S. Kavuri, "A review of process fault detection and diagnosis part ii: Qualitative models and search strategies," *Computers and Chemical Engineering*, vol. 27, no. 3, pp. 313–326, 2003.
[4] W. A. Shewhart, *Economic control of quality of manufactured product*. New York : D. Van Nostrand Co., 1931.
[5] H. Hotelling, "Multivariate quality control," *Techniques of Statistical Analysis*, vol. , pp. 111–184, 1947.
[6] J. Westerhuis, S. Gurden, and A. Smilde, "Standardized q-statistic for improved sensitivity in the monitoring of residuals in mspc," *Journal of Chemometrics*, vol. 14, no. 4, pp. 335–349, 2000.
[7] E. J. Jackson, "Multivariate quality control," *Communication Statistics - Theory and Methods*, vol. 14, pp. 2657 – 2688, 1985.
[8] B. R. Bakshi, "Multiscale PCA with application to multivariate statistical process monitoring," *AIChE Journal*, vol. 44, pp. 1596–1610, 1998.
[9] M. Kano, K. Nagao, S. Hasebe, I. Hashimoto, H. Ohno, R. Strauss, and B. Bakshi, "Comparison of multivariate statistical process monitoring methods with applications to the eastman challenge problem," *Computers and Chemical Engineering*, vol. 26, no. 2, pp. 161–174, 2002.
[10] R. Mason, N. Tracy, and J. Young, "A practical approach for interpreting multivariate t^2 control chart signals," *Journal of Quality Technology*, vol. 29, no. 4, pp. 396–406, 1997.
[11] F. V. Jensen, *An introduction to Bayesian Networks*. Taylor and Francis, London, United Kingdom, 1996.
[12] N. Friedman, D. Geiger, and M. Goldszmidt, "Bayesian network classifiers," *Machine Learning*, vol. 29, no. 2-3, pp. 131–163, 1997.
[13] P. Langley, W. Iba, and K. Thompson, "An analysis of bayesian classifiers," in *National Conference on Artificial Intelligence*, 1992.
[14] I. Kononenko, "Semi-naive bayesian classifier," in *EWSL-91: Proceedings of the European working session on learning on Machine learning*, pp. 206–219, 1991.
[15] A. Perez, P. Larranaga, and I. Inza, "Supervised classification with conditional gaussian networks: Increasing the structure complexity from naive bayes," *International Journal of Approximate Reasoning*, vol. 43, pp. 1–25, 2006.
[16] T. Tiplica, S. Verron, A. Kobi, and I. Nastac, "Fdi in multivariate process with naïve bayesian network in the space of discriminant factors," in *International Conference on Computational Intelligence for Modelling, Control and Automation*, (Sydney, Australia), 2006.
[17] S. Verron, T. Tiplica, and A. Kobi, "Fault diagnosis of industrial systems with bayesian networks and mutual information," in *European Control Conference*, 2007.
[18] S. Verron, T. Tiplica, and A. Kobi, "Procedure based on mutual information and bayesian networks for fault diagnosis of industrial systems," in *American Control Conference*, 2007.
[19] F. Pernkopf, "Bayesian network classifiers versus selective k-nn classifier," *Pattern Recognition*, vol. 38, no. 1, pp. 1–10, 2005.
[20] M. Perzyk, R. Biernacki, and A. Kochanski, "Modeling of manufacturing processes by learning systems: The naive bayesian classifier versus artificial neural networks," *Journal of Materials Processing Technology*, vol. 164-165, pp. 1430–1435, 2005.
[21] T. Denoeux, M. Masson, and B. Dubuisson, "Advanced pattern recognition techniques for system monitoring and diagnosis : A survey," *Journal Europeen des Systemes Automatises*, vol. 31, no. 9-10, pp. 1509–1539, 1997.
[22] B. Dubuisson, *Diagnostic, intelligence artificielle et reconnaissance des formes*. Trait IC2 information. Srie productique, Herms sciences publications, 2001.
[23] S. Verron, T. Tiplica, and A. Kobi, "Multivariate control charts with a bayesian network," in *4th International Conference on Informatics in Control, Automation and Robotics (ICINCO)*, 2007.
[24] S. Verron, *Diagnostic et surveillance des processus complexes par rseaux baysiens*. PhD thesis, University of Angers, 2007.
[25] J. Downs and E. Vogel, "Plant-wide industrial process control problem," *Computers and Chemical Engineering*, vol. 17, no. 3, pp. 245–255, 1993.
[26] N. Ricker, "Decentralized control of the tennessee eastman challenge process," *Journal of Process Control*, vol. 6, no. 4, pp. 205–221, 1996.
[27] U. Kruger, Y. Zhou, and G. Irwin, "Improved principal component monitoring of large-scale processes," *Journal of Process Control*, vol. 14, no. 8, pp. 879–888, 2004.
[28] L. Chiang, M. Kotanchek, and A. Kordon, "Fault diagnosis based on fisher discriminant analysis and support vector machines," *Computers and Chemical Engineering*, vol. 28, no. 8, pp. 1389–1401, 2004.
[29] A. Kulkarni, V. Jayaraman, and B. Kulkarni, "Knowledge incorporated support vector machines to detect faults in tennessee eastman process," *Computers and Chemical Engineering*, vol. 29, no. 10, pp. 2128–2133, 2005.
[30] P. Lyman and C. Georgakis, "Plant-wide control of the tennessee eastman problem," *Computers and Chemical Engineering*, vol. 19, no. 3, pp. 321–331, 1995.
[31] K. P. Murphy, "The bayes net toolbox for matlab," in *In Computing Science and Statistics : Proceedings of Interface*, 2001.