A Hierarchical Fault Diagnosis Method Using a Decision Support System Applied to a Chemical Plant

Didier Theilliol, Philippe Weber, Marius Ghetie, Hassan Noura

A Hierarchical Fault Diagnosis Method Using a Decision Support System Applied to a Chemical Plant

D. THEILLIOL, P. WEBER, M. GHETIE, H. NOURA
CRAN - CNRS URA 821 - Université Henri Poincaré - Nancy 1, BP 239
54506 Vandœuvre Cedex - France
Phone: (33) 83.91.20.69, Fax: (33) 83.91.20.30, e-mail: theillio@cran.u-nancy.fr

ABSTRACT

Automatic control systems with sophisticated control algorithm can be very large and complex. In order to improve the automatic process control, it is important to develop fault diagnosis strategy. A hierarchical scheme of fault detection and isolation based on Decision Support System (DSS) is presented. For fault diagnosis, a knowledge based procedure is required. In addition to analytic symptoms, heuristic information have to be taken into account. A pattern recognition method, a unified representation of all symptoms, and a fuzzy expert system are integrated in the diagnostic tool. This approach is applied to sensor fault detection in a chemical plant, and compared with different methods.

Keywords: fuzzy decision, on-line fault diagnosis, pattern recognition, fuzzy expert system.

1. INTRODUCTION

The aim of a diagnosis procedure is to solve two main problems: the fault detection which can be done by collecting and analysing the information from the process, and the decision able to localise and give the type and the origin of the fault.

To realise a real-time diagnosis tool, a particular decision method is developed in order to detect sensor failures based on analytical redundancy. This tool is a special element of a diagnostic system applied to a complex process.

To solve these problems, the system can be decomposed into subsystems. The modelisation of these subsystems leads to the residual generation. The residual evaluation provides important information, *analytic symptoms* necessary to solve the diagnostic problem.

Usually, the decision making is realised by an elementary logic. Meanwhile, when multiple faults or false alarms occur, the failures are not isolated [10]. Some specific mathematics algorithms and an additional knowledge about the analytical redundancy can increase the efficiency of the decision making [6][14]. This aim is achieved by the decomposition of the method in four modules:

- the Analytic Symptoms Generation (ASG), provided by classical redundancy residual generation and evaluation,
- the Heuristic Symptoms Generation (HSG), which is generally provided using the human observation,
- the knowledge which contains all information about the process, the fault detection method, the different types of failures and their propagation through the system,
- all the heuristic and analytic symptoms are used to determine the faults affecting the process. The Global Diagnostic (GD) is achieved by the Symptoms Unification (SU) [6], and a procedure of Pattern Recognition (PR)[1]. The outputs of this tool give the type and confidence degree of the faults.

The architecture of such a diagnostic tool is illustrated by the following figure.

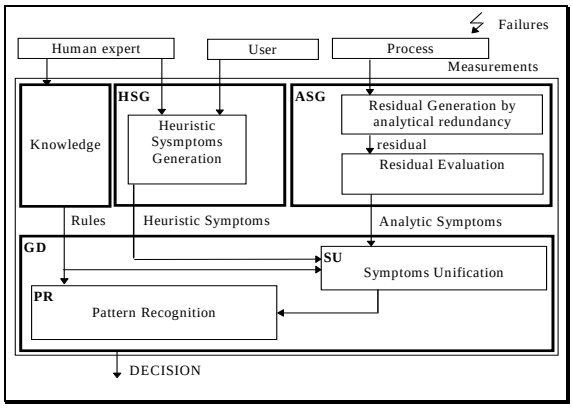

Figure 1: Architecture of a Decision Support System

This paper is organised as follows: in the second and third part, the generation of the analytic and heuristic symptoms are described. The knowledge base is defined in part four. The resolution of the complete diagnosis problem will be presented in part five. Finally, to illustrate our approach, the simulation results of a five tanks process are shown in part six. The concluding remarks are given in the last part.

2. ANALYTIC SYMPTOMS

Classical methods are used to produce analytical information. These methods require several measurable signals. A particular analytical redundancy of multivariable systems and detection algorithm are developed to generate analytic symptoms.

Analytic Symptoms Generation: Redundant Symptoms.
The process models are considered to generate the analytic symptoms. Consider the following linear stochastic system described in a state-space form:

$$\begin{cases} X_{k+1} = AX_k + BU_k + w_k \\ Y_k = CX_k + v_k \end{cases}$$

The process parameters A, B and C are assumed to be known; k represents the sample; X, Y, U, w and v represent respectively the state, the output, the input, the state noise and the output noise vectors where:

$Y = [y_1 ... y_n ... y_N]^T$ and N is the number of sensors.

$$E\{w_k\} = 0,\ E\{v_k\} = 0,\ E\{X_k\} = X_0$$

$$E\left\{\begin{bmatrix} v_k \\ w_k \\ X_0 \end{bmatrix} \cdot \begin{bmatrix} v_l^T & w_l^T & X_0^T \end{bmatrix}\right\} = \begin{bmatrix} R_k\delta_{kl} & 0 & 0 \\ 0 & Q_k\delta_{kl} & 0 \\ 0 & 0 & P_0 \end{bmatrix}$$

with δ_{kl} is the Kronecker symbol.

The Kalman filter is used to estimate the state vector based on the input and output measurements:

$$\hat{X}_{k+1/k+1} = (I - K_{k+1} \cdot C) \cdot A \cdot \hat{X}_{k/k} + (I - K_{k+1} \cdot C) \cdot B \cdot U_k + K_{k+1} \cdot Y_{k+1}$$

$$P_{k+1/k+1} = (I - K_{k+1} \cdot C) \cdot (A \cdot P_{k/k} \cdot A^T + Q_k)$$

$$K_{k+1} = (A \cdot P_{k/k} \cdot A^T + Q_k) \cdot C^T \cdot H_{k+1}^{-1}$$

$$H_{k+1} = R_{k+1} + C \cdot (A \cdot P_{k/k} \cdot A^T + Q_k) \cdot C^T$$

The model-based fault detection using the Dedicated Observer Scheme (DOS) can be achieved according to the following principle: assuming that the system is totally observable, a state observer can be obtained using each measurement variable y_n. The number of observers is equal to the number of measurements. The redundant residuals r_j can be computed [12]. If a bias occurs on sensor $e_{(n)}$, some residuals are equal to zero and others are different from zero. A Generalised Likelihood Ratio test (GLR) is used for the residual evaluation. Therefore the output vector of the GLR method, called coherence vector [1], can be built according to each statistical test applied to each residual:

$\underline{S} = \{ S(s_j) : j = 1 ... J \}$

where J determines the number of residuals and $S(s_j)$ represents the state of the *redundant symptom* as $S(s_j)$ is equal to zero when the statistical characteristic of residuals is a white noise and equal to one in some other case.

Analytic Symptoms Generation: Quality of the GLR test
To improve the decision making, some further information must be considered such as the quality of the residual generation and evaluation, quantified by the *frankness* m_j and the *persistence* p_j. The frankness m_j on the residual r_j represents the jump value of the statistical test with regard to the threshold test:

$$m_j(k) = 1 - \frac{\text{Threshold}}{\text{Jump}_j(k) + \text{Threshold}} \quad , m_j \in [0,1]$$

The symptom p_j represents a persistence degree in time of $S(s_j)$. The persistence is computed by counting the number of occurrence of $S(s_j)=1$ in a sliding window of width d where 1/d is the resolution of p_j.

The frankness and the persistence qualify the magnitude of the faults. The analytic symptoms m_j and p_j increase the number of information on the binary output $S(s_j)$ and must improved the decision making. The following figure presents the generation of the analytic symptoms.

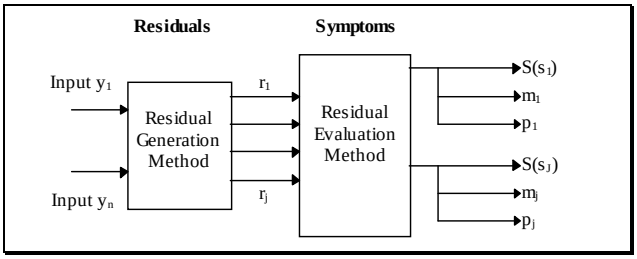

Figure 2: Analytic symptoms generation

3. HEURISTIC SYMPTOMS

Heuristic symptoms are produced on-line using qualitative information in linguistic term defined by human observation, to provide the fault decision.

Model-based Quality
The aim is to obtain a robust decision on uncertainties of the model because the matrix parameters A, B, and C are only estimations of the real system parameters [11]:

$$\begin{cases} X_{k+1} = AX_k + BU_k \\ Y_k = CX_k \end{cases} \qquad \begin{aligned} A &= \overline{A} + \delta A \\ B &= \overline{B} + \delta B \\ C &= \overline{C} + \delta C \end{aligned}$$

with $\overline{A}$, $\overline{B}$, $\overline{C}$ represent the real matrices parameters and $\delta A, \delta B, \delta C$ are the errors associated to each matrix. The human expert can represent a confidence level c_j on the model by a linguistic term [2].

The heuristic symptom v_j is then defined to take into account the age of the system during the time. Some different qualitative rules can be established to represent the quality of the residual generation from c_j and v_j.

4. KNOWLEDGE BASE

Diagnostic Matrix
According to the DOS structure, a bias on sensor $e_{(n)}$ has a signature $\underline{S_{gn}}$. A signature table can be designed for all the N sensors:

	$\underline{S_{g1}}$	$\underline{S_{g2}}$	...	$\underline{S_{gN-1}}$	$\underline{S_{gN}}$
r_1	0	0	...	0	0
r_2	1	1	...	0	0
...	...	...	...	...	...
r_{J-1}	0	0	...	1	0
r_J	0	0	...	0	1

This table is also called ″diagnostic matrix″ [10] or ″structure

matrix″ [4]. In this paper, this table is noted D with different elements D(n,j) where n is the current number of sensors (n=1,...,N) and j is the number of residuals (j=1,...,J).

Membership Functions and Linguistic Terms
The integration of different kinds of diagnosis knowledge with a unified diagnostic strategy is facilitated by the symbolical approach. In order to unify all the symptoms, the frankness and the persistence must be fuzzified. To achieve this task, the definition of the membership functions must be done by the human expert during the system configuration, because he knows all the information necessary to the definition of these functions [7]. These values are approximated and obtained experimentally.

Due to a real-time constraint, triangular or trapezoidal membership function is chosen. An ambiguity degree between the different linguistic classes is defined such that no more than two membership functions are superposed.

Inference Matrix
In order to minimise the false alarm occurrence, we must consider the residual accuracy degree represented by c_j and v_j, and the quality of the residual evaluation represented by m_j and p_j.

It is known that when c_j is small, the residual is probably deviated from zero, which may lead to the increase of false alarm possibility. Moreover, if m_j is small, the fault detected is nothing else but a false alarm. Therefore, in order to increase the decision reliability or the ″robustness″ against false alarms, it seems necessary to take into account these information which may lead to the variation of the vector $\underline{S}$ even in the absence of failures.

Then, the parameters c_j, v_j, m_j and p_j are linked to the symptom quality $\mu(s_j)$ by a fuzzy decision table, also noted inference matrix [13], where each case represents a fuzzy rule as:

IF (c_j is Linguistic A at Fuzzy value W)
AND (v_j is Linguistic B at Fuzzy value X)
AND (m_j is Linguistic C at Fuzzy value Y)
AND (p_j is Linguistic D at Fuzzy value Z)
THEN $\mu(s_j)$ is *MIN* (A, B, C, D) at *min* (W, X, Y, Z)

where *MIN* (Large, Small) = Small and *min* (0.1, 0.8) = 0.1

The decision table makes possible the implementation for fuzzy logic. This table can be considered as a generic table because there is an adjustment of the ranges according to the expert knowledge during the membership functions definition. The rules of this table assume that the confidence of $\mu(s_j)$ is proportional to c_j, v_j, m_j and p_j.

5. GLOBAL DECISION

The fault isolation using the redundant symptoms analysis is based on the following principle [4].

The signature $\underline{S_{gn}}$ representing the state of the redundant symptoms associated to a given failure $e_{(n)}$, makes possible the failure detection and isolation, that is the signature analysis. This kind of methods can be associated to the rules used by the expert system. But, in the case of complex processes, these rules may be insufficient to solve the problem. The reasons are the following [10]:

- Many failures may occur simultaneously. It means that as many rules as failure combinations have to be generated.
- Since the measurement and the identification of the redundant equations parameters are inaccurate, the number of rules must increase in order to achieve an appropriate fault detection method.

While this kind of rules does not allow the failure isolation, a method based on the computation of a distance between the signatures and the redundant symptoms $\underline{S}$ using the pattern recognition is proposed [1].

Unification of Symptoms
Once the membership functions are defined, it is possible to fuzzify the variables m_j and p_j [5] in order to mix them with the heuristic symptoms c_j and v_j. The use of the rules of the fuzzy inference matrix allows the computation of the confidence $\mu(s_j)$ of each $S(s_j)$. These rules are implemented in the rules base of an expert system. The working memory is firstly constituted by the fuzzyfication of m_j and p_j in addition to c_j and v_j. When all the validated rules are executed, this memory contains the linguistic terms describing the confidence $\mu(s_j)$.

The moment or the barycenter method is then used to defuzzify the linguistic terms provided by the expert system [5]. The method is sensitive to the membership functions shape and takes into account all the variable information. The quality on the redundant symptom $S(s_j)$ is defined as follows:

$\mu(s_j) = 0$	$S(s_j)$ is a false alarm.
$0<\mu(s_j) < 1$	The symptom quality on the residual evaluation result is increasing.
$\mu(s_j) = 1$	$S(s_j)$ expresses directly the state of the concerned sensor.

The redundant symptom $S(s_j)$ is then corrected by $\mu(s_j)$ which leads to unified symptom: $u(s_j) = \mu(s_j) \,.\, S(s_j)$

And a set $\underset{\sim}{U}$ is defined as: $\underset{\sim}{U} = \left\{ \underset{\sim}{u_U}(s_j) \;:\; j = 1...J \right\}$

Pattern Recognition: Hamming indicator computation
Consider a set $\underset{\sim}{E}$ of the signatures of the sensor failures which can affect the process:

$$\underset{\sim}{E} = \{ \underline{S_{gn}} : n = 1 \ldots N \} + \underline{S_{g0}}$$

where N is the number of sensors used to construct the observers.

The signatures $\underline{S_{gn}}$ and $\underline{S_{g0}}$ are also defined as the following sets:

$\underline{S_{gn}} = \{ D(j,n) : j = 1 \ldots J \}$ and $\underline{S_{g0}} = \{ 0 ; 0 ; \ldots ; 0 \}$

and D is the diagnostic matrix, J is the number of the redundant symptoms, $\underline{S_{gn}}$ is the failure signature on the sensor $e_{(n)}$ and $\underline{S_{g0}}$ is the signature of the fault free system.

The outputs of the decision method is noted $\underset{\sim}{F_E}(e_{(n)})$. This function represents the state or the confidence of the sensor and takes the following values:

$\underset{\sim}{F_E}(e_{(n)}) = 0$ No failure on the sensor $e_{(n)}$.

$\underset{\sim}{F_E}(e_{(n)}) \in]0,1[$ A value expressing the probability of the failure presence on the sensor $e_{(n)}$.

$\underset{\sim}{F_E}(e_{(n)}) = 1$ The sensor $e_{(n)}$is faulty.

Particularly, $\underset{\sim}{F_E}(e_{(0)})$ is the fault free degree of the whole set of sensors. Therefore, the decision method is achieved comparing the unified symptoms set $\underset{\sim}{U}$ with the signature $\underline{S_{gn}}$.

The Hamming's distance formula [8] is used to make this pattern recognition [1], [14], and [3]:

$$\delta(\underset{\sim}{U}, \underline{S_{gn}}) = \frac{1}{J}\sum_{j=1}^{J} |\underset{\sim}{u_U}(s_j) - D(j,n)| \qquad (1)$$

The resemblance between the redundant symptoms and the signatures of the different failures has to be analysed. Hence, the following function is defined:

$$\underset{\sim}{F_E}(e_{(n)}) = 1 - \delta(\underset{\sim}{U}, \underline{S_{gn}})$$

Notice that, in the case of simultaneous failures on different sensors, it is advantageous to consider only the differences $|\underset{\sim}{u_U}(s_j) - D(j,n)|$ when D(j,n) = 1. Hence, the Hamming distance defined in equation (1) becomes the modified Hamming indicator:

$$\delta_M(\underset{\sim}{U}, \underline{S_{gn}}) = \frac{1}{W_{(n)}}\sum_{j=1}^{J} \{|\underset{\sim}{u_U}(s_j) - D(j,n)|.D(j,n)\} \qquad (2)$$

with W(n) is the number of the elements $D(j,n) \neq 0$

The fuzzy fault indicator F is achieved by:

$$\forall n \neq 0, \underset{\sim}{F_E}(e_{(n)}) = 1 - \frac{1}{W_{(n)}}\sum_{j=1}^{J} \{|\underset{\sim}{u_U}(s_j) - D(j,n)|.D(j,n)\}$$

$$\underset{\sim}{F_E}(e_{(0)}) = 1 - \frac{1}{J}\sum_{j=1}^{J} \underset{\sim}{u_U}(s_j) \qquad (3)$$

In the case of a DOS structure the number of ones per signature is constant such that W=2.(N-1).

6. THE FIVE TANKS PROCESS

This fault detection and isolation method is applied to a simulation process composed of five tanks linked one to another.

In this part, the localisation of the level sensors failures are considered.

The Process Description
This process is described by the following figure:

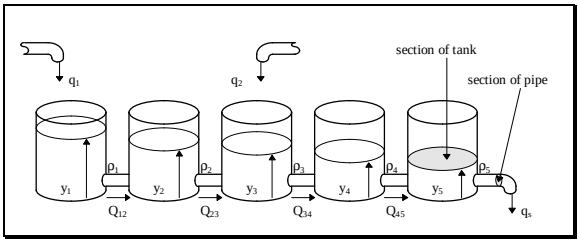

Figure 3: The five tanks process

An operating point is given by the following values:

input flow rates (m^3/s)	$q_1 = 0.6004.10^{-4}$, $q_2 = 0.0929.10^{-4}$
Outflow coefficients	$\rho_1=0.5$, $\rho_2=0.5$, $\rho_3=0.6$, $\rho_4=0.5$, $\rho_5=0.2$
Section (m^2)	$tank = 15.4\ 10^{-4}$ $pipe = 0.5\ 10^{-4}$
levels (m)	$y_{10}=0.5$, $y_{20}=0.47$, $y_{30}=0.42$, $y_{40}=0.37$, $y_{50}=0.32$

A linearized model of this system around an operating point is obtained and described by a continuous state space representation where:

$$A = \begin{bmatrix} -0.0212 & 0.0212 & 0 & 0 & 0 \\ 0.0212 & -0.0383 & 0.0171 & 0 & 0 \\ 0 & 0.0171 & -0.0365 & 0.0194 & 0 \\ 0 & 0 & 0.0194 & -0.0347 & 0.0153 \\ 0 & 0 & 0 & 0.0153 & -0.0179 \end{bmatrix}$$

$$B = \begin{bmatrix} 64.9351 & 0 \\ 0 & 0 \\ 0 & 64.9351 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } C = I_{(5,5)}.$$

Additive gaussian noises are considered for the simulation of the five tanks process. The input noise is equal to 10% around the mean value q_i. The measurement noise is equal to 5% around the mean value y_i.

Before running this method, the knowledge base must be built. The diagnostic matrix is directly obtained by the DOS structure. The membership functions are defined by the human expert according to his knowledge of the system. For this process and for the residual generation and evaluation

methods, the membership functions are chosen as in the following figure:

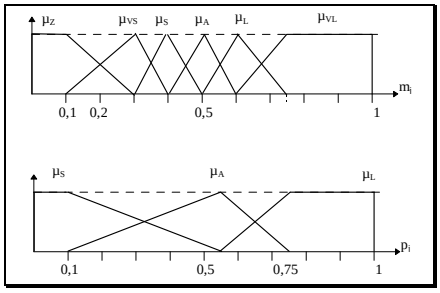

Figure 4: Frankness and persistence

Where the fuzzy sets are represented by Z : Zero, VS: Very Small, S: Small, A: Average, L: Large, VL: Very Large.

In the following study, a comparison between the three fault indicators for the fault decision is achieved. The first one, called ″Hamming Indicator″, is computed by the original Hamming distance formula (1) based directly on the redundant symptoms $S(s_j)$ and not on the unified symptoms $U(s_j)$. The second one, called ″Modified Hamming Indicator″, is computed by (2) always based on the redundant symptoms. The last one is defined in (3) and is called ″Fuzzy Fault Indicator″.

Case 1

Three biases of 4% on sensors y_1, y_3 and y_4 are considered respectively at instants 100, 400 and 600. The comparison between the three fault indicators shows the advantage of our approach with regard to the classical Hamming indicators. As shown in Figure 5, for the first indicator, the decision making is very difficult because the detection is almost impossible.

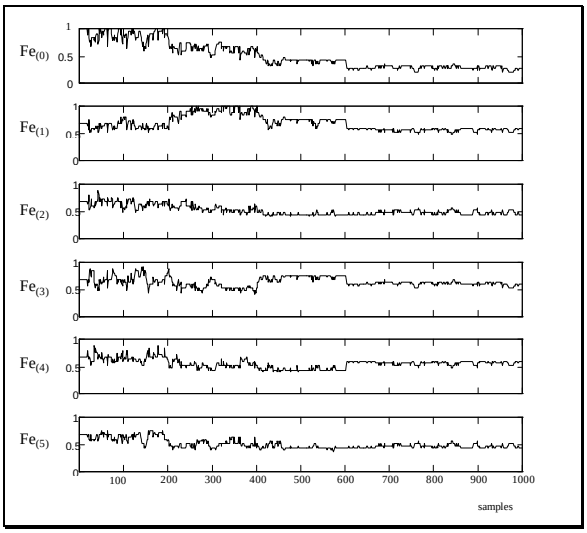

Figure 5: Hamming Indicator

Using ″Modified Hamming Indicator″, the decision making is possible with a large rate of false alarms while the ″Fuzzy Fault Indicator″ makes possible the detection and isolation of different failures with low rate of false alarms. Hence, the decision making is more accurate. At instant 600, 3/5 of sensors are affected. Thus, the decision making is more and more difficult due to the limitation of analytic redundancy method (Figures 6 and 7).

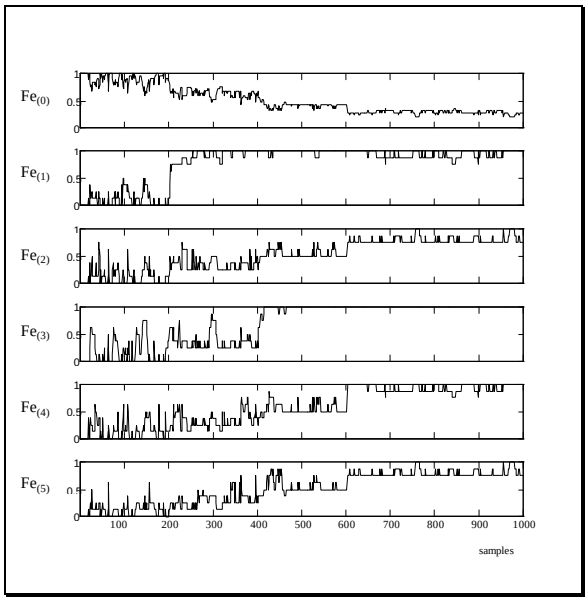

Figure 6: Modified Hamming Indicator

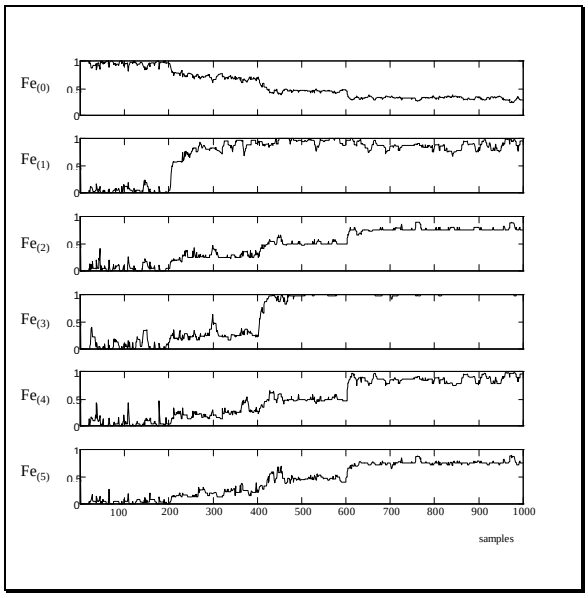

Figure 7: Fuzzy Fault Indicator

Case 2

In this case, only one sensor failure (y_3) is considered. 1000 samples are computed during the study. The aim is to compare the performances of these methods when the quality of the model changes.

In Figure 8, the histograms illustrate the previous remarks about the dispersion of the decision making, while Figure 9 shows again the robustness of the unification approach against model uncertainties. When the quality of the model is bad, with the ″Hamming Indicator″, all the histograms are superposed,

whereas with the ″Modified Hamming Indicator″, the rate of false alarms is very large. But with ″Fuzzy Fault Indicator″, the decision is more ″realistic″.

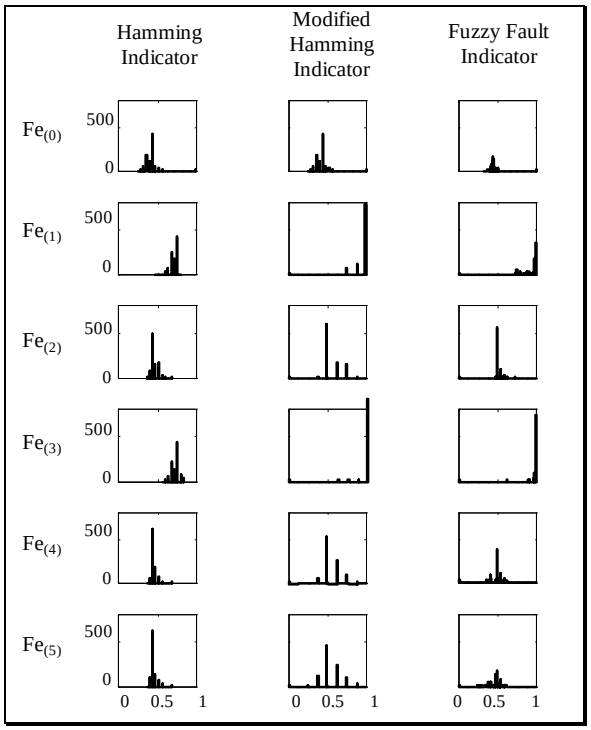

Figure 8: ″Good″ quality of the model

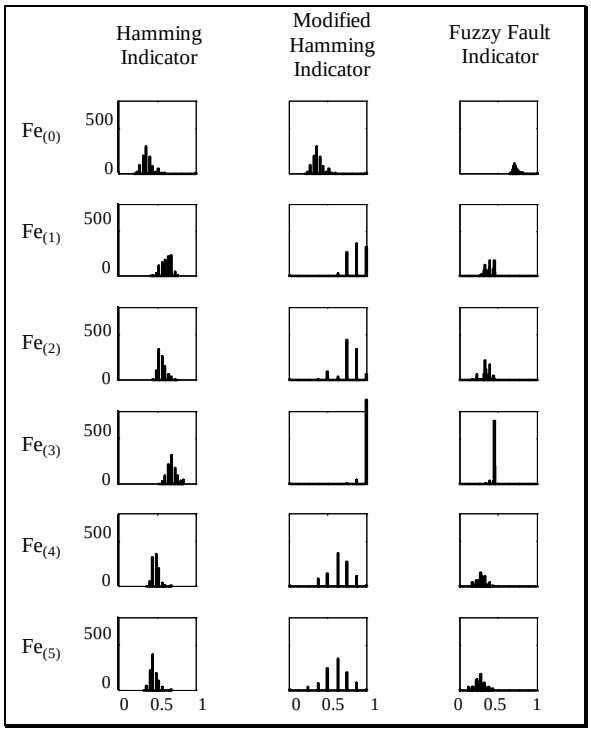

Figure 9: ″Bad″ quality of the model

7. CONCLUSION

In this paper, the Decision Support System (DSS) has been developed. The method used to realise this system is based on the unification of heuristic and analytic symptoms. In fact the unification makes the fault decision more ″**robust**″ against model uncertainties, due to the parameter identification or the ageing of the system. This strategy leads to less false alarms and non-detection. This decision support module is a part of a whole diagnostic system. That is why the output of this module, the fault indicator, must be accurate. The advantage of this approach is emphasised through the comparison with Hamming indicators and applied to the five tanks process. The DSS has been considered as a generic tool. Consequently, different residual generation and evaluation methods can be applied to detect and isolate different types of failures. The results obtained in this paper are encouraging and allow to consider the real-time implementation.

8. REFERENCES

[1] Cassar J.P., Litwak R.G., Cocquempot V., Staroswiecki M., ″Approche structurelle de la conception de systèmes de surveillance pour des procédés industriels complexes″, *Diagnostic et sûreté de fonctionnement*, Vol. 4, No. 2, 1994, pp. 179-202.

[2] Dubois D., Prade H., ″The use of fuzzy number in decision analyses. Fuzzy Information and Decision Processes″, *M.M. Gupta & E. Sanchez (eds.) North-Holland Publishing Company, 1982.*

[3] Gertler J., Anderson K.C., ″An evidential reasoning extension to quantitative model-based failure diagnosis″, *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 22, No. 2, March/April 1992, pp. 275-289.

[4] Gertler J., ″Analytical redundancy methods in fault detection and isolation. Fault Detection, Supervision and Safety for Technical Processes″, *SAFEPROCESS'91 Baden-Baden Germany*, Vol. 1, 1991, pp. 9-22.

[5] Graham I., Jones P.L., ″Expert systems, knowledge, uncertainty and decision″, *London, New York, Chapman and Hall Ltd*, Chapter 5, pp. 117-158.

[6] Isermann R., ″Integration of fault detection and diagnosis methods″, *SAFEPROCESS'94 Espoo Finland*, Vol. 2, 1994; pp. 597-612.

[7] Juuso E.K., ″Fault diagnosis based on linguistic equations framework″ *SAFEPROCESS'94 Espoo Finland*, Vol.1, 1994, pp. 374-379.

[8] Kaufmann A., ″Introduction à la théorie des sous-ensembles flous à l'usage des ingénieurs″, *Tome I et II Masson Paris, New York, Barcelone, Milan 1977.*

[10] Koscielny J.M. ″Method of fault isolation for industrial precesses″, *Diagnostic et sûreté de fonctionnement*, Vol. 3, No. 2, 1993, pp. 205-220.

[11] Monmain J., Gentil S., ″Interprétation qualitative pour le diagnostic en ligne″, *Diagnostic et sûreté de fonctionnement*, Vol. 3, No. 1, 1993, pp. 23-45.

[12] Patton R.J., Frank P.M., Clark R.N., ″Fault diagnosis in dynamic system, theory and applications″, *Prentice Hall International (U.K.) Ltd*, Chapter 2 and 15, 1989.

[13] Pham D.T., Jin G., ″Evolutionary design of an adaptive fuzzy logic controller for processes with time delays″, *IEEE*

Systems, Man & Cybernetics San Antonio, Vol. 1, 1994, pp. 431-436.
[14] Ulieru M., ″A fuzzy logicbased computer assisted fault diagnosis system″, *Diagnostic et sûreté de fonctionnement*, Vol. 4, No. 3, 1994, pp. 405-441.