Fuzzy reasoning based interactive diagnosis of a grid network of rain gauge sensors

Quang-Huy Giap * **Olivier Adrot** * **Stéphane Ploix** *
Christian Depraetere † **Jean-Marie Flaus** *

* *Laboratoire G-SCOP, CNRS UMR5272/Institut Polytechnique de Grenoble, 46 avenue Flix Viallet, 38 031 Grenoble Cedex 1, France*
† *Laboratoire d'Etudes des Transferts en Hydrologie et Environnement (LTHE), BP 53 8041, Grenoble cedex 9, France*

Abstract: This paper deals with the design of a diagnosis tool for a network of rain gauge sensors in the context of human-machine cooperation. The model of the whole system is difficult to be completely established before the diagnosis analysis. A part of the expert's knowledge is tacit and it'll be exploited during the diagnostic process e.g. expert analyzes the hyetograph of rainfall of a cluster of rain gauges to collect symptoms. Diagnosis becomes therefore an interactive process. At each step, the role of diagnosis tool is to accompany the expert to establish a diagnosis. The way of handling such a process is presented in this paper.

Keywords: Fault diagnosis, diagnosis analysis, detection test design, symptom generation.

1. INTRODUCTION

Performing diagnoses relies on a complex process, which can be decomposed into a design process followed by a running process. The following design tasks may be distinguished: *system modeling* Reiter (1987); de Kleer and Williams (1992), *detection test* design Blanke et al. (2006); Ploix et al. (2005), *isolation algorithm design* Nyberg and Krysander (2003); Ploix et al. (2003) and *sensor placement* Yassine et al. (2008); Frisk and Krysander (2007).

The running process is closely related to the design task: *symptom generation, diagnostic analysis and possibly backward analysis*. In scientific literature, most of the contributions aim at automating the running process assuming that models can be completely established before the diagnosis analysis. But, Ploix and Chazot (2006) points out that, in many practical contexts, this prerequisite cannot be satisfied. Consequently, new problems arise because:

- models may be considered at different levels of detail. So, it is difficult to model a whole complex system before the diagnosis analysis. Usually, modeling and diagnosis are iterative processes where systems are partially depicted and some parts are refined step by step.
- in addition to the modeled part of a system, there are other types of non-formalized knowledge, often qualified as implicit, which may not be formalized by expert because either
  - · the system is too complex to achieve detailed modeling of the whole system,
  - · the expert does not have a detailed model of the system at the beginning of the diagnosis process,
  - · or some knowledge cannot be easily formalized.

In order to tackle theses difficulties, the interactions between experts and computer-aided diagnosis systems are obviously needed during the running process. Then, the practical diagnosis processes may be classified into three categories according to three tasks mentioned in the design process.

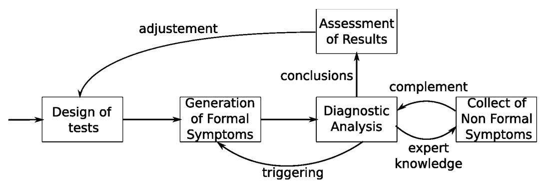

Fig. 1. The interactive diagnosis process

This paper focuses on the diagnosis with interactions during the diagnostic analysis (figure 1). This problem is proposed in the context of the Hydrodiag project, which deals with the diagnosis of a grid network of rain gauge sensors. For this purpose, a *Hydrodiag computer-aided diagnosis system* has been designed to guide experts during the diagnosis process. This problem will be detailed in the next sections.

2. PROBLEM FORMULATION

The problem is to set up a computer-aided diagnosis process to determine the faults in a network of rain gauge sensors set up in the Upper Oueme Valley in Benin, with an area of 47536 km^2. 46 bucket rain gauges are distributed over the basin.

The first task of a diagnostic process is generally the design of detection tests. The main idea for testing the rain gauges is to compare data from nearby sensors. Each couple of sensors can be tested using the average correlation level within each month, providing that the distance between two sensors is less than $10km$ [1]. Because of the low density

[1] Hydrologists estimate that, for the considered problem of rainfall, if the distance between two sensors is more than 10km, the rainfall

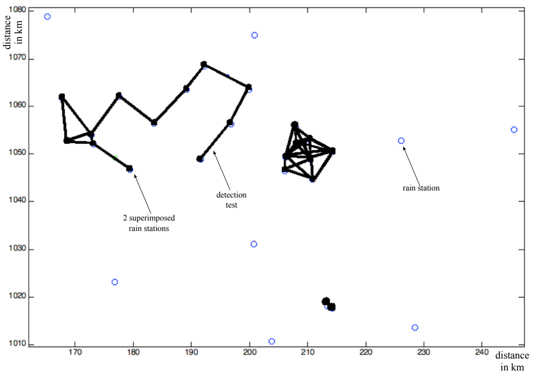

Fig. 2. Tests created in the basin of Oueme

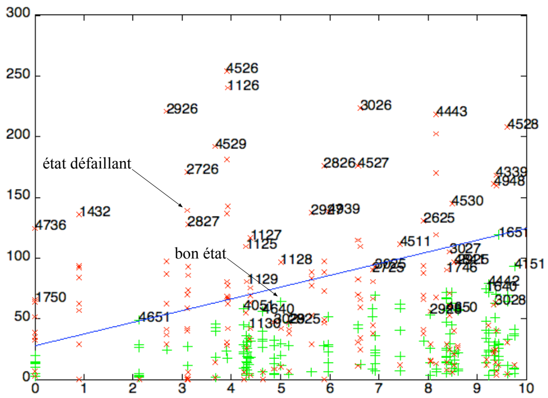

Fig. 3. Detection threshold

of the sensor network, only 25 of 46 rain gauges can be tested. Figure (2) shows all the tests that may be performed.

Thank to a table given by hydrologists which identifies the state of each sensor for each period of two weeks, decision thresholds for the correlation level is established. Figure (3) shows that the threshold was chosen as a linear function of the distance which minimizes the number of non-detection while strictly prohibiting false alarms that are critical for a diagnostic. The 'x' indicate failing false states and the '+' indicate normal states. Note that many faulty states may be confused with normal states.

Reiter's algorithm of diagnostic analysis has been applied. For the studied eight months of 2002 [2] , 13 to 32 possible diagnoses are obtained for the two weeks periods: each diagnosis has at least 4 sensors which are simultaneously faulty. For example, for the month August 2002, 16 diagnoses have been obtained:

```
diagnosis #0 (Formal:100.0%,
         Contextual:88.64%, A priori:0.01%)
Component d643 (gangamou) is faulty
 and component d626 (dapefougou) is faulty
 and component d614 (adiangdia) is faulty
 and component d647 (parakou_2) is faulty
```

amounts received by theses two sensors are independent of each other.

[2] from March to October because out of this period, it does not rain on this basin

```
...
------------------------------
diagnosis #15 (Formal:100.0%,
         Contextual:72.73%, A priori:1.00E-4%)
Component d639 (kolokonde) is faulty
 and component d611 (donga) is faulty
 and component d644 (gountia) is faulty
 and component d645 (koko-sika) is faulty
 and component d632 (adiangdia_oues) is faulty
 and component d636 (parakou) is faulty
```

It is very complex for an expert to exploit these results: how to select one of these diagnoses? A computer-aided iterative process is described in section 4.

Another problem arises because false alarms have to be avoided in Reiter's approach. In case of uncertainty, it is better to consider that the behavior is normal rather than asserting a faulty behavior. But this problem leads to very conservative diagnoses. A fuzzy reasoning is proposed in section 3. Section 5 presents an application of these results to the problem of diagnosing a grid of rain gauges.

3. FUZZY REASONING

Considering the impact of a false alarm, it can be difficult to adjust the detection threshold of a test. In fact, to prevent the false alarm happening, adjustment can become very pessimistic and leads to non-detection problem. To tackle this problem, Touaf and Ploix (2004a,b); Touaf (2005) aim at transforming the crisp logic reasoning of the formal diagnosis analysis to fuzzy logic of Zadeh (1975). The proposed approach avoids the introduction of necessity measurement as suggested in Cayrac et al. (1996) and preserves the result of crisp logic when the degrees of membership become certain $\{0 \text{ or } 1\}$. It relies on the fuzzy logic proposed by Yager (1986).

Constraints modeling behavior are sometimes not sufficient to model systems to be diagnosed. Indeed, in some applications, a constraint does not model the behavior whatever the system state is: it is necessary to add validity constraints that determine when the behavioral constraint applies.

Let $\mathcal{K}(V) = 0$ be a behavioral constraint that has to be satisfied when a component c is in mode m and when $\mathcal{K}'(V) \diamond 0$ where $\diamond$ stands for a comparator such as $<$, $\leq$, $>$, $\geq$ or possibly $=$. Behavioral and validity constraints can therefore be modeled by:

$\forall V \in dom(V)$, $\mathcal{K}(V) = 0$	$\forall V \in dom(V)$, $\mathcal{K}'(V) \diamond 0$	$m(c)$
$false$	$false$	$true$
$false$	$true$	$false$
$true$	$false$	$true$
$true$	$true$	$true$

This table is summarized by the following proposition that has to be satisfied:

$$\forall V \in dom(V),$$
$$m(c) \leftrightarrow ((\mathcal{K}(V) = 0) \vee \neg(\mathcal{K}'(V) \diamond 0)) \quad (1)$$

Let $\mathcal{D} \subset dom(V)$. Equation (1) becomes:

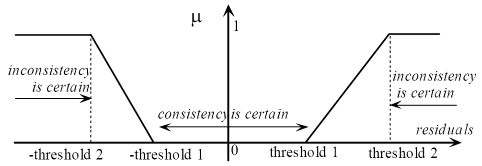

Fig. 4. Interpretation of residues in a fuzzy uncertain

$$\forall V \in \mathcal{D},$$
$$m(c) \rightarrow ((\mathcal{K}(V) = 0) \vee \neg(\mathcal{K}'(V) \diamond 0)) \quad (2)$$

For the sake of simplicity, it is assumed that the constraint $\mathcal{K}'(V) \diamond 0$ contains the same variables, or a subset of the same variables, that those present in $\mathcal{K}(V) = 0$.

3.1 Fuzzification of symptoms

Let $\mathcal{K}(V)$ be a set of behavioral constraints which depict partially or totally the behavior of items, and $\mathcal{K}'(V)$ is a set of constraint of validity which depict whether the elementary behavioral constraints are suitable. The main idea is that if the validity constraints are not satisfies, the system is considered as being in an indeterminate state. It is represented by the following relationship when an exoneration assumption is assumed:

$$\bigwedge_i ok(item_i) \leftarrow (\mathcal{K}(V) = 0) \wedge (\mathcal{K}'(V) = 0) \quad (3)$$

It yields a constraint corresponding to uncertainty in the case of invalidity:

$$\neg\,(\mathcal{K}'(V) = 0) \equiv \text{uncertainty of} \bigwedge_i ok(item_i) \quad (4)$$

and therefore by a constraint corresponding to:

$$\bigwedge_i ok(item_i) \rightarrow (\mathcal{K}(V) = 0) \vee \neg\,(\mathcal{K}'(V) = 0) \quad (5)$$

Therefore, two sets of constraints must be set. In both cases, with the fuzzy approach, it is considered that there are more than two possible values for satisfaction degree (*satisfies* or *non-satisfied*. The truths of test and validity constraints belong to [0 1]. The value '1' means surely satisfied and the value 0 means surely unsatisfied. This is illustrated by figure (4). During the test phase, it exists two membership functions $\mu_{\mathcal{K}} = \mu(\mathcal{K}(V) = 0)$ and $\mu_{\mathcal{K}'} = \mu(\mathcal{K}'(V) = 0)$. By using the fuzzification operator $\mu(A \vee B) = min(1, \mu(A) + \mu(B))$, where A and B are propositions, (5) is transformed[3] into:

$$min \begin{pmatrix} 1,\ 1 - \mu\left(\bigwedge_i ok(item_i)\right) + \\ \mu\left(\mathcal{K}(V) = 0\right) + \\ \left(1 - \mu\left(\mathcal{K}'(V) = 0\right), 1\right) \end{pmatrix} = 1$$

By integrating the previous notions:

$$min \begin{pmatrix} 1,\ 2 - \mu\left(\bigwedge_i ok(item_i)\right) + \\ \mu_{\mathcal{K}} - \mu_{\mathcal{K}'} \end{pmatrix} = 1$$

In order to satisfy this relation, it must verify:

$$\mu\left(\bigwedge_i ok(item_i)\right) \leq 1 + \mu_{\mathcal{K}} - \mu_{\mathcal{K}'} \quad (6)$$

[3] Let's recall that $A \rightarrow B$ is equivalent to $\neg A \vee B$.

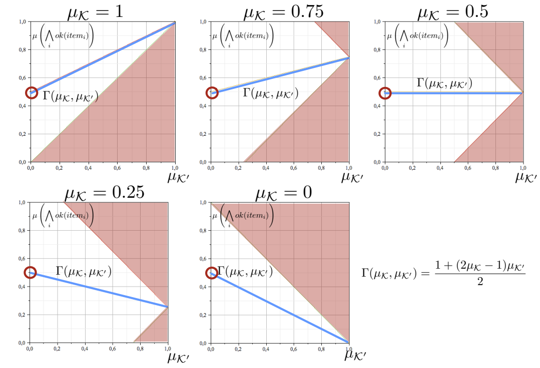

Fig. 5. Fuzzification function for fuzzy tests

The relation (4) leads to:

$$(\mu_{\mathcal{K}'} = 0) \equiv \left(\mu\left(\bigwedge_i ok(item_i)\right) = 0.5\right) \quad (7)$$

The relation (3), which is valid only with the exoneration assumption, is transformed into:

$$min\left(1, 2 - \mu_{\mathcal{K}} - \mu_{\mathcal{K}'} + \mu\left(\bigwedge_i ok(item_i)\right)\right) = 1$$

In order to satisfy this relation, it must verify:

$$\mu\left(\bigwedge_i ok(item_i)\right) \geq \mu_{\mathcal{K}} + \mu_{\mathcal{K}'} - 1 \quad (8)$$

The proposed problem amounts to find a fuzzification function that allows to infer the truth degree of test $\bigwedge_i ok(item_i)$ created from the membership degree of satisfaction of constraints $(\mathcal{K}(V) = 0)$ and validity $(\mathcal{K}'(V) = 0)$. This function must satisfy the constraints (6) and (7) as shown in figure (5). It corresponds to the hatched area. Moreover, if the fuzzification function can verify (4), it should be used with exoneration assumption. We have shown that the following function is suitable:

$$\Gamma(\mu_{\mathcal{K}}, \mu_{\mathcal{K}'}) = \frac{1 + (2\mu_{\mathcal{K}} - 1)\mu_{\mathcal{K}'}}{2} \quad (9)$$

For each test, the fuzzification function makes it possible to evaluate the truth of symptoms depending on the satisfaction degree of behavior constraints and validity constraints: $\mu\left(\bigwedge_i ok(item_i)\right) = \Gamma(\mu_{\mathcal{K}}, \mu_{\mathcal{K}'})$.

In case of malfunctioning, instead of evaluating $\mu\left(\bigwedge_i ok(item_i)\right)$, $\mu\left(\bigvee_i cfm(item_i)\right) = 1 - \mu\left(\bigwedge_i ok(item_i)\right)$ is evaluated. For the model of normal functioning, it is obtained: $\neg ok(item) = cfm(item)$. It leads to:

$$\mu\left(\bigvee_i cfm(item_i)\right) = \quad (10)$$
$$\mu\left(\neg\,(\mathcal{K}(V) = 0) \wedge (\mathcal{K}'(V) = 0)\right) =$$
$$1 - \Gamma(\mu_{\mathcal{K}}, \mu_{\mathcal{K}'})$$

Let's now imagine that two sets of observations are available. In this case, since constraints depend on data flow,

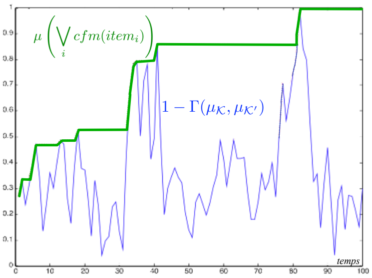

Fig. 6. The degree of alarm in a fuzzy context

two set of constraints are given by:($\mathcal{K}_1(V)=0, \mathcal{K}'_1(V)=0$) and ($\mathcal{K}_2(V)=0, \mathcal{K}'_2(V)=0$). The relation (5) becomes:

$$\begin{gathered}\bigwedge_i ok(item_i) \rightarrow \\ ((\mathcal{K}_1(V)=0) \vee \neg(\mathcal{K}'_1(V)=0)) \\ \wedge((\mathcal{K}_2(V)=0) \vee \neg(\mathcal{K}'_2(V)=0))\end{gathered}$$

In case of fault, it comes:

$$\begin{gathered}\neg((\mathcal{K}_1(V)=0) \vee \neg(\mathcal{K}'_1(V)=0)) \;\vee \\ \neg((\mathcal{K}_2(V)=0) \vee \neg(\mathcal{K}'_2(V)=0)) \\ \rightarrow \bigvee_i cfm(item_i)\end{gathered}$$

Then, using the fuzzification operator *max* for $\vee$ to prevent the saturation problem yields:

$$\mu\left(\bigvee_i cfm(item_i)\right) = max\begin{pmatrix}1-\Gamma(\mu_{\mathcal{K}_1},\mu_{\mathcal{K}'_1}),\\ 1-\Gamma(\mu_{\mathcal{K}_2},\mu_{\mathcal{K}'_2})\end{pmatrix}$$

This result generalizes to order n and shows that, without exoneration assumption, the degree of alarm $\mu\left(\bigvee_i cfm(item_i)\right)$ is either constant or increasing. Figure (6) illustrates this result on an example where new observations are recorded in a time series.

3.2 Fuzzy diagnosis reasoning

During detection phase, the truth degree $\mu_{test} = \mu\left(\bigwedge_i ok(item_i)\right)$ is used to evaluate each test. The diagnostic analysis phase can start. But does the formal diagnostic analysis can be transformed from crisp logical into fuzzy logic? Several situations can occur. Let $\mathbb{T}$, the set of tests that can be divided into

$\mathbb{T} = \mathbb{T}^{satisfied} \cup \mathbb{T}^{unsatisfied} \cup \mathbb{T}^{dubious}$ with $\mathbb{T}^{satisfied} = \{test; \mu(test)=1\}$, $\mathbb{T}^{unsatisfied} = \{test; \mu(test)=0\}$ and $\mathbb{T}^{dubious} = \{test; \mu(test) \in]0,1[\}$.

It is considered that an uncertain test can either be satisfied or unsatisfied. The resolution principle consists in examining all the possible combinations of the considered uncertain tests and, for each case, to compute diagnoses. Among this set of diagnoses, only minimal diagnoses are considered because they are generators of other diagnoses. Let's consider the 3 different situations.

The first one corresponds to the case where $\mathbb{T} \equiv \mathbb{T}_{satisfied}$ is satisfied. This situation is similar to the crisp logic: there is no reason to calculate diagnoses because no abnormalities were detected.

The second situation is the case where $\mathbb{T}^{unsatisfied} \neq \emptyset$. The diagnosis should then be able to explain not only normal tests but also uncertain tests $\mathbb{T}^{dubious}$ that correspond to possible faulty behaviors. Let $\mathbb{D}^{unsatisfied} = \{D_i^{unsatisfied}\}$ be a set of diagnoses explaining the set of tests $\mathbb{T}^{unsatisfied} \neq \emptyset$. To explain the combination of tests $\mathbb{T}^{dubious}$, each diagnosis $D_i^{unsatisfied}$ must be completed by a additional mode *cfm*. In other words, the diagnoses resulting from this completion are obviously non-minimal because it contains $D_i^{unsatisfied}$. Therefore, in the situation where $\mathbb{T}^{unsatisfied} \neq \emptyset$, the minimal diagnoses are deducted only from the set of tests $\mathbb{T}^{unsatisfied}$: $\{D_i\} = \{D_i^{unsatisfied}\}$ or it is rewritten as $\mathbb{D} = \mathbb{D}^{unsatisfied}$. Hence, diagnoses can be calculated but it remains to evaluate their truth degree. Diagnosis are induced by the test $\mathbb{T}^{unsatisfied}$ whose dissatisfaction degree is given by: $1-\mu(test)=1$. Since the dissatisfaction degrees of each tests, which leads to diagnoses $\mathbb{D}^{unsatisfied}$, is 1. It is deduce that:

$$\forall D_i \in \mathbb{D}^{unsatisfied}, \mu(D_i)=1 \tag{11}$$

Thus, if it exists tests with $\mu(test)=0$, diagnoses should be calculated only from tests verifying $\mu(test)=0$ and their truth degree is 1.

The last situation is the case where $\mathbb{T}^{unsatisfied} = \emptyset$ and $\mathbb{T}^{dubious} \neq \emptyset$. By considering all the possible combinations for the tests included in $\mathbb{T}^{dubious}$, it appears that minimal diagnoses corresponds to the case where only one uncertain test is unsatisfied [4]. Hence, diagnoses correspond to simple modes *cfm*: $\mathbb{D}^{dubious} = \bigcup_{test \in \mathbb{T}^{dubious}} \bigcup_{mode_i \in Expl(test)} mode_i$. To evaluate the truth degree of this diagnosis, the truth degree of the factors that lead to this diagnosis has to be evaluated first. Let's consider a diagnosis $cfm(item_i)$ that belongs to the set of diagnoses $\mathbb{D}^{dubious}$ because some tests contain it. By using *max* as a fuzzification operator for $\vee$, its truth degree is:

$$\begin{gathered}\mu(cfm(item_i)) = \ldots \\ \ldots max_{test \in \mathbb{T}^{dubious}}\left(1-\Gamma(\mu_{\mathcal{K}_{test}},\mu_{\mathcal{K}'_{test}})\right)\end{gathered} \tag{12}$$

The function Γ is used when the exoneration assumption is taken into account. It is therefore possible to use it to evaluate a distance between the effective signature defined by $\forall test_i \in \mathbb{T}, (\sigma^*_{\mathbb{T}})_i = 1-\Gamma(\mu_{\mathcal{K}_{test_i}},\mu_{\mathcal{K}'_{test_i}})$ [5], and theoretical signature of a diagnosis. Let D_j be a diagnosis. The theoretical signature $\sigma_{\mathbb{T}}(D_j)$ is given by:

$$\forall test_i \in \mathbb{T}, \begin{cases}(\sigma_{\mathbb{T}}(D_j))_i = 0, \\ \quad \text{if } Modes(D_j) \cap Expl(test_i) = \emptyset \\ (\sigma_{\mathbb{T}}(D_j))_i = 1, \\ \quad \text{if } Modes(D_j) \cap Expl(test_i) \neq \emptyset\end{cases} \tag{13}$$

The distance between the two signatures can be written:

$$distance_{\mathbb{T}}(D_j) = \frac{\sum_{test \in \mathbb{T}} \mid \sigma^*_{\mathbb{T}} - \sigma_{\mathbb{T}}(D_j)) \mid}{card(\mathbb{T})} \tag{14}$$

[4] There is no diagnosis if all tests are satisfied.

[5] if there are several sets of observations, $\sigma^*_{\mathbb{T}}$ corresponds to the maximum of $1-\Gamma(\mu_{\mathcal{K}_{test_i}},\mu_{\mathcal{K}'_{test_i}})$ for all sets of observation

An detailed example of fuzzy logic reasoning is presented in Touaf and Ploix (2004b).

4. COMPUTER AIDED INTERACTIVE DIAGNOSIS PROCESS

4.1 Directing the diagnosis with implicit knowledge

Since it was not reasonable to present all calculated diagnoses to expert, a solution to accompany the expert to establish a diagnosis has been preferred. The main idea is that only part of the expert knowledge can be formalized in the form of tests as described before. On the one hand, there is a tool-aided diagnosis with mathematical models and reasoning tool that allow tackling the complexity without difficulty, and on the other hand, expert with tacit knowledge that usually allows determining whether a sensor is failed or not by looking at his hyetograph and those of its neighbored sensors while the tool-aided diagnosis cannot detect.

An *interactive diagnosis matrix* has been designed. It is represented by figure (7). Let $\mathbb{T}$ be a set of valid tests. To simplify some notations, the satisfaction degree of a test $test_i = (\mathcal{K}_i, \mathcal{K}'_i) \in \mathbb{T}$, denoted by $\Gamma(\mu_{\mathcal{K}_i}, \mu_{\mathcal{K}'_i})$ in the formula (9) is now denoted by: $\mu(test_i)$. Then, for each test, $test_i \in \mathbb{T}$:

$$\mu\left(\bigwedge mode_i \in Modes(test_i)\right) = \mu(test_i)$$

The interactive diagnosis matrix relies on the test results $\mu(test_i)$ with $test_i \in \mathbb{T}$. Let's start by examining the zone (A) of the figure (7). (B) shows the list of items constituting the diagnosis in progress D_j established at iteration i by the expert. Button (H) allows removing the last item added to the diagnosis in progress D_j to direct towards a new explanation.

Zone (C) presents indicators which characterizes the relevance of the diagnosis in progress. Paragraph (3.2) points out that if it exists at least one test $test_i$ that is totally unsatisfied ($\mu(test_i) = 0$), then diagnoses is computed as it is done in crisp logic, only from the test $\{test_i \in \mathbb{T}; \mu(test_i) = 0\}$, and the obtained accuracy degree of diagnoses, denoted *FORMAL* or *F* in the interactive diagnosis matrix of diagnosis, are equal to 1 (see Eq. (11)). If there is no totally unsatisfied test, diagnoses will be given by each mode appearing in the explanations of uncertain tests ($\mathbb{T}_{uncertain} = \{test_i; \mu(t_i) \in]0, 1[$). In this case, the indicator *FORMAL* of each diagnosis $cfm(item)$ is $\mu(cfm(item)) = max_{test_i \in \mathbb{T}_{uncertain}}(1 - \mu(test_i))$ (See Eq. (12)). For a diagnosis in progress D_i, two cases are considered:

- It exists tests which are totally unsatisfied. Indicator *FORMAL* is set at 100% if the diagnosis in progress contains modes that can explain all the unsatisfied tests. In the contrary case, the indicator *FORMAL* will be set at 0%: the evaluated certain anomalies have not been explained yet under a formal viewpoint. *explication* or *E* in column (E) means, if it is marked, that the items of the corresponding line will explain all the totally unsatisfied test not yet explained.
- It exists only one uncertain test. In this case, each fault mode implied in uncertain test considered at a diagnosis. So it exists a *or logic* among fault mode of a diagnosis in progress. It is transformed into a fuzzy logic by using a fuzzification operator that does not saturate: $max_{cfm(item)}\mu(cfm(item))$, where $\mu(cfm(item))$ is calculated as presented before. The boxes (E) have no meaning in this case: they are unmarked.

Indicator *CONTEXTUAL*, denoted *C*, evaluates the testing results and diagnosis in progress. It has been shown that the distance in $\mathbb{T}$ between the effective signature and the theoretical signature associated with a diagnosis D_j (see eq. (14)) is:

$$distance_{\mathbb{T}}(D_j) = \frac{\sum_{test \in \mathbb{T}} \mid \sigma^*_{\mathbb{T}} - \sigma_{\mathbb{T}}(D_j)) \mid}{card(\mathbb{T})}$$

To transform into a scale where 100% is preferable, indicator *(*CONTEXTUAL) is defined by: $contextual_{\mathbb{T}}(D_j) = 1 - distance_{\mathbb{T}}(D_j)$.

Indicator *A PRIORI*, denoted also by *A*, evaluate priori reliability thank to probabilities of occurrence of fault modes: each probability $p(cfm(item_i))$ is defined for item $item_i$. For a diagnosis D_j, indicator *A PRIORI* is: $\Pi_{cfm(item_i) \in D_j} p(cfm(item_i))$. The function *log* is used in the zone (F) to make easy the reading.

Two cases must be distinguished for the construction of the suggestions of items appearing in column (D) of the matrix:

- It exists tests which are totally unsatisfied. The set of test $\mathbb{T}^{unsatisfied}$ is not empty. Each test must be explained. Let $\mathbb{T}^{unsatisfied}_{explained}(D_j) \in \mathbb{T}^{unsatisfied}$ be a set of tests of $\mathbb{T}^{unsatisfied}$ that verifies: $\{test_i \in \mathbb{T}^{unsatisfied}; Expl(test_i) \cap D_j \neq \emptyset\}$, and its complement $\mathbb{T}^{unsatisfied}_{non\ explained}(D_j)$ in $\mathbb{T}^{unsatisfied}$. D_j must be completed to empty $\mathbb{T}^{unsatisfied}_{non\ explained}(D_j)$. Candidates items displayed in column (D) are given by: $\bigcup_{test_i \in \mathbb{T}^{unsatisfied}_{non\ explained}(D_j)} Expl(test_i)$.
- It exists only uncertain tests. In this case, only modes of uncertain tests are presented ($\mu(test) < 50\%$), except for the modes that already belong to D_j: $\{\bigcup_{\{test_i} \in \mathbb{T}; \mu(test_i) < 50\%\} Expl(test_i) \backslash D_j\}$. If the box *all* is marked, all modes of all tests are presented: $\{\bigcup_{\{test_i} \in \mathbb{T}\} Expl(test_i) \backslash D_j\}$.

Zones (F) and (G) correspond to indicators (C) and to suggestions (D) that would be obtained by adding the item in the zone of the diagnosis in progress.

The button *Exoneration* allows to permanently remove some items of suggestions (column (D)) for the construction of a diagnosis.

Finally, to help experts, the top button of each column (E) and (F) can be clicked to change the classification of the suggested items.

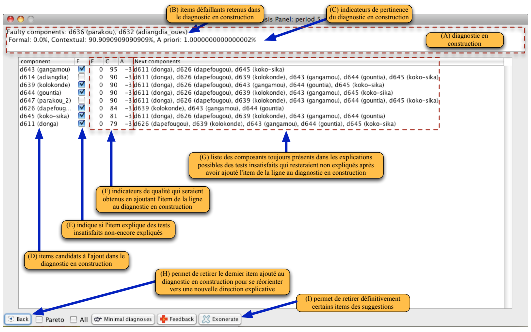

Fig. 7. Interactive diagnosis matrix

4.2 feedback analysis

The problem is to identify the tests, which leads to wrong diagnoses. Let assume that the testing results of valid tests are known and effective diagnosis, denoted by D^*, found by an operator is also known. The following algorithm has been proposed to solve the problem.

D^* **: a set of modes corresponding to the actual state of the system**
tests **: a set of tests done on the system**
Pour *test* $\in$ *tests* **faire**
  Si (*test* is valid and unsatisfied) **Alors**
    Si $(Expl(test) \cap D^* = \emptyset)$ **Alors**
      [false alarm criticizes: test has to be reviewed]
    Fin Si
  Fin Si
  Si (*test* is valid and satisfied) **Alors**
    Si $(Expl(test) \cap D^* \neq \emptyset)$ **Alors**
      [test has not detected the fault]
    Fin Si
  Fin Si
Fin Pour

Let's note that a false alarm is critical because it is unacceptable that a test infer an abnormality it does not exist. Hence, test has to be tuned by redesigning detection thresholds or intervals modeling uncertainties in the necessary case.

5. APPLICATION

5.1 Interactive process

Let's examine the use of this interaction matrix on a scenario to search the failed rain gauges during the month of August 2002. The expert discovered the figure (8). Since the indicator *CONTEXTUAL* is more important for d643 (d643 explains many symptoms), the expert selects naturally this diagnosis. The position of the sensor lights up in blue on the screen of the software *Hydrodiag* and a window with rainfall hyetographs of all sensors located within a area of 10km radius appears (9)). Obviously, the sensor d643 is failed. The expert continues this explanatory direction.

Then, four sensors appear on top with comparable scores. The expert chooses randomly the first (see figure 10): the D614. By looking at hyetograph of the sensor and its neighbored sensor, he concludes that the sensor D632 is faulty and not D614. He goes back by clicking on the button *Back* (H) and redirects to D632. Then, he obtains the matrix shown in figure (11).

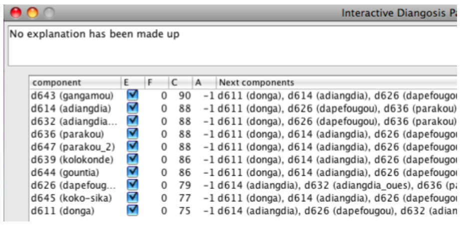

Fig. 8. Diagnosis analysis for the month of August 2002 - Step 1

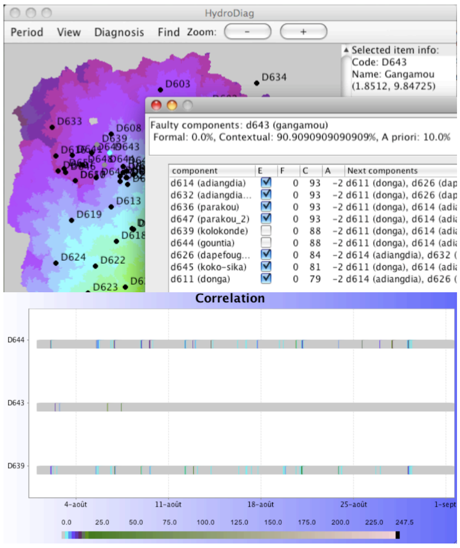

Fig. 9. Diagnosis analysis for the month of August 2002 - Step 2

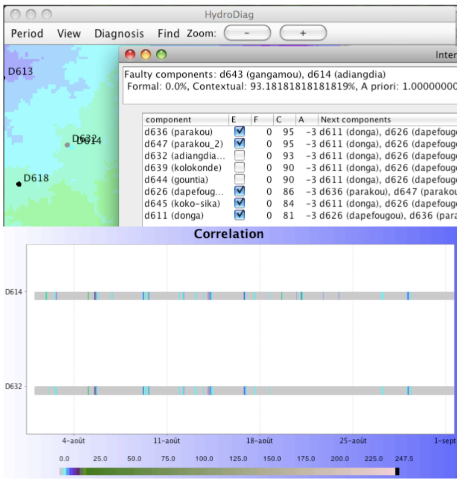

Fig. 10. Diagnosis analysis for the month of August 2002 - Step 3

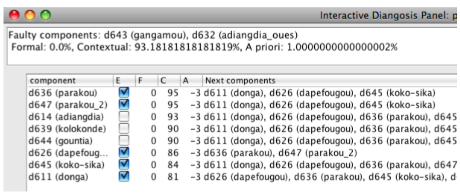

Fig. 11. Diagnosis analysis for the month of August 2002 - Step 4

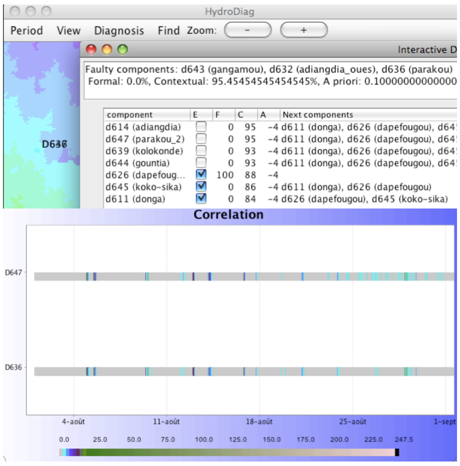

Fig. 12. Diagnosis analysis for the month of August 2002 - Step 5

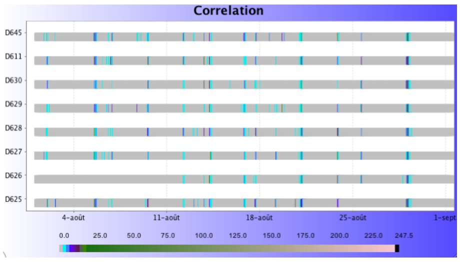

Fig. 13. Diagnosis analysis for the month of August 2002 - Step 6

Expert chooses d636. The results confirm that this direction is good (see figure 12).

It appears that the sensor D626 can explain the remaining tests that are totally unsatisfied. The expert chooses and verifies the explanatory direction on hyetograph (see figure 13). This explanatory direction appears to be good and all totally unsatisfied tests are explained. Hence, the fault gauges are: d643 (gangamou), d632 (adiangdia-oues), d636 (parakou), d626 (dapefougou). In comparison with the list of 16 possible diagnoses with 4 or more simultaneous faults represented before, the interests of an interactive approach is obvious. The found diagnosis corresponds to the fourth one in the list of the possible calculated diagnoses.

Thanks to this approach, non-expert hydrologists can find out the faulty rain gauges: the faulty sensors have been found without having been reported by expert hydrologists.

6. CONCLUSION

This paper presents a method for diagnosis in the context of human-machine cooperation. It corresponds to a computer-aided diagnosis system that solves the problem of determining the faults in a network of rain gauge sensor. For this purpose, a set of detection tests is established by using the average correlation level of each couple of sensors. Then, the fuzzy reasoning is integrated. It transforms the crisp logic reasoning of the formal diagnosis analysis to fuzzy logic. This method allows tackling the impact of false alarms in diagnosis analysis. Together with the proposed approach, a software has been designed: it consists in an interactive diagnosis matrix that accompany the expert to establish diagnosis. By interacting with computer during the diagnosis process, expert, with its tacit knowledge, can determine whether a sensor is faulty or not thank to hyetograph of each cluster of sensors.

REFERENCES

Blanke, M., Kinnaert, M., Lunze, J., and Staroswiecki, M. (2006). *Diagnosis and Fault-tolerant Control*. Springer-Verlag.

Cayrac, D., Dubois, D., and Prade, H. (1996). Handling uncertainty with possibility theory and fuzzy sets in satellite fault diagnosis application. *IEEE transactions on Fuzzy Systems*, 4(3), 251–269.

de Kleer, J. and Williams, B.C. (1992). Diagnosis with behavioral modes. *Readings in model-based diagnosis*, 124–130.

Frisk, E. and Krysander, M. (2007). Sensor placement for maximum fault isolability. In *The 18th International Workshop on Principles of Diagnosis, DX-07*. USA.

Nyberg, M. and Krysander, M. (2003). Combining ai, fdi, and statistical hypothesis-testing in a framework for diagnosis. In *IFAC Safeprocess'03*. Washington, U.S.A.

Ploix, S. and Chazot, P. (2006). Iterative expert driven fault diagnosis based on structural modeling. In *IFAC Symposium SAFEPROCESS'2006*.

Ploix, S., Desinde, M., and Touaf, S. (2005). Automatic design of detection tests in complex dynamic systems. In *16th IFAC World Congress*. Prague, Czech republic. 4-8 july.

Ploix, S., Touaf, S., and Flaus, J.M. (2003). A logical framework for isolation in fault diagnosis. In *SAFEPROCESS'2003*. Washington D.C., U.S.A.

Reiter, R. (1987). A theory of diagnosis from first principles. *Artificial Intelligence*, 32, 57–95.

Touaf, S. (2005). *Diagnostic logique des systèmes complexes dynamiques dans un contexte multi-agent*. Ph.D. thesis, Grenoble Institute of Technology (INPG).

Touaf, S. and Ploix, S. (2004a). Diagnosis for large scale distributed industrial plants - application to an hydraulic looper. In *11th IFAC Symposium on Automation in Mining, Mineral and Metal processing, MMM04*. Nancy, France.

Touaf, S. and Ploix, S. (2004b). Soundly managing uncertain decisions in diagnostic analysis. In *15th International Workshop on Principles of Diagnosis (DX2004)*. Carcasonne, France.

Yager, R.R. (1986). A characterization of the extension principle. *Fuzzy Sets and Systems*, 18, 205–217.

Yassine, A., Ploix, S., and Flaus, J.M. (2008). A method for sensor placements taking into account diagnosability criteria. *Applied Mathematics and Computer Science*, 18(4).

Zadeh, L.A. (1975). The concept of a linguistic variable and its application to approximate reasoning. *Information Sciences*, 8, 199–249.