COMBINING INTERVAL AND QUALITATIVE REASONING FOR FAULT DIAGNOSIS

G. Calderón-Espinoza * **J. Armengol** * **M.Á. Sainz** *
P. Herrero *

* *Institut d'Informàtica i Aplicacions. Universitat de Girona. Campus Montilivi. E-17071 Girona, Spain.*
(gcalder,armengol,pherrero)@eia.udg.es, sainz@ima.udg.es

Abstract: A methodology is proposed to diagnose faults from a detection system that allows to express uncertainty with intervals in parameters and measurements. A fault is detected by comparing current measurements with computed external and internal approximations of behavior. The results are applied to diagnose faults. A qualitative data set is obtained in order to create fault rules. A fault library is built from these rules to make fault diagnosis. Diagnosis is given by a set of the most possible faults with similar behavior.

Keywords: Fault detection, Fault diagnosis, Intervals, Uncertainty, Uncertain Dynamic System.

1. INTRODUCTION

A fault is something that changes the behavior of a technological system such that the system does no longer satisfy its purpose (Blanke *et al.*, 2003). In order to avoid production deteriorations or damage to machines and humans, faults have to be found as quickly as possible and decisions that stop the propagations of their effects have to be made.

To guarantee safety and correct operation of systems, the improvements of methods used in supervision, monitoring and control are necessary. Fault detection and diagnosis are tasks included in these methods. Detection consists on finding an abnormal behavior in the process and diagnosis consists on determinating the type, size as well as location of the fault.

One approach to diagnosis is the called Model-based diagnosis (MBD). MBD systems reason starting from a model. The model represents the behavior of the system to be diagnosed. If the behavior of the observed situation is different from the estimation carried out by the model to the same situation, the system concludes that there is a fault. A later analysis about differences tries to identify which component is the cause of the fault (Gertler, 1998).

Mathematical models can be used to calculate the reference behavior in order to be compared with available system's measurements. However, an accurate and complete model of a physical system is almost never available. Usually the parameters of a system may vary with time in an uncertain manner, and the characteristics of disturbances and noise are unknown so that they cannot be modelled.

An alternative approach is to represent uncertainty in models with intervals where structure of the models is known but parameters and sensors' noise may be given as imprecise numerical intervals. The present work starts from detection based on interval models. SQualTrack (Armengol *et al.*, 2003) is a robust fault detection tool that monitors the behavior of a real process in order to detect internal faults which are present in the process. Simulation is based on modal interval arithmetic, which produces overbounded and underbounded envelopes for the supervised system. It uses error-bounded estimations and time windows to deal with the computation problem, reducing the computational effort and improving the fault detection results. The present work describes a methodology to complement the fault detection system in order to develop a support system for making decisions in fault diagnosis.

The paper is organized in 5 sections. This section introduces fault detection and diagnosis tasks and briefly explains the problem of uncertainty. Section 2 describes the dynamical models and the detection system. In Section 3 the proposed methodology to make diagnosis from detection with interval models is described. Experimental results of the proposed methodology with an example of a DC motor are shown in Section 4 and finally conclusions and future work are presented in Section 5.

2. INTERVAL DYNAMIC MODELS

As it was said in the introduction one way to detect faults is by comparing the real system behavior with the predicted one obtained from a model. Continuous-time systems are usually described by differential equations. Usually the input, state and output variables are sampled time-signals defined over a time variable k, which belongs to a discrete set. All signals are assumed to be sampled synchronously at a fixed sampling period. It is for this reason that discrete models are used. Then a fault is detected when the predicted behavior from the model is different from the corresponding measurement:

$$y(t) \neq \hat{y}(t) \tag{1}$$

Most of the times this equation will be true and consequently the constant detection of faults happens. One reason is because of in industrial monitoring of processes, the uncertainty is often present due to sensors and signal noises, imprecise knowledge about model parameters or because the parameters may vary with time. With Intervals it is possible to have uncertain and less precise models but more accurate.

These models are used to determine, for the measured sequence U, the sequence of the model output $\hat{Y}(t)$. The consistency of the system with the model can be checked at every time t by determining the difference:

$$r(t) = y(t) - \hat{y}(t) \tag{2}$$

which is called residual. When there is no fault the value is close to zero.

The reference behavior for fault detection in the present work is obtained by simulation of interval models which consider uncertainty in model parameters and sensor measurements. One example of this kind of models is equation 3. The equation is a n-th order SISO (Single Input, Single Output) system represented by a difference equation where u are inputs, T the sampling time, and a or b are parameters of the system, in this case they are intervals:

$$y_t = \sum_{i=1}^{m+1} a_i y_{t-iT} + \sum_{j=1}^{p+1} b_j u_{t-jT} \tag{3}$$

Equation 3 is the starting point of our approach. It defines an imprecise model of the supervised system with uncertain parameters that are independent. The simulation of a real-valued model produces a trajectory for each output variable which is a curve representing the evolution of the variable of the system across time: $y_r(t)$. In the case of an interval model, there is a set of models indeed where a set of curves (a band) represents the evolution of each variable (Armengol *et al.*, 2003). The limits of the band are:

$$Y_r(t) = [min(y_r(t)), max(y_r(t))] \tag{4}$$

2.1 Modal interval simulator

To compute the band limits is necessary to compute the range of a function in a parameter space at each simulation step, which is a task related to global optimization and usually needs an important computation effort. With SQualTrack (Armengol, 1999; Armengol *et al.*, 2003) similar results can be obtained at a lower cost by calculating external estimations $Y_{rex}(t)$ to the range of the function at each iteration. After infinite iterations it would calculate the exact range, but it stops when the estimation is sufficiently close to detect the fault, thus saving much computational effort when a fault is detected. However if there is not a fault and the simulator never stops this drawback can be overcome by using an internal estimation $Y_{rin}(t)$. If the measurement is inside this envelope the fault, if it exists, will never be detected so the algorithm must stop.

The simultaneous use of internal and external estimations obtain the same fault detection results than $Y_r(t)$ but with a much lower computation effort. Both estimations form an error-bounded estimation because although $Y_r(t)$ is not known, it is known that $Y_{rin}(t) \subseteq Y_r(t) \subseteq Y_{rex}(t)$.

Three zones are defined by error bounded estimations depicted in figure 1. The simulator guaran-

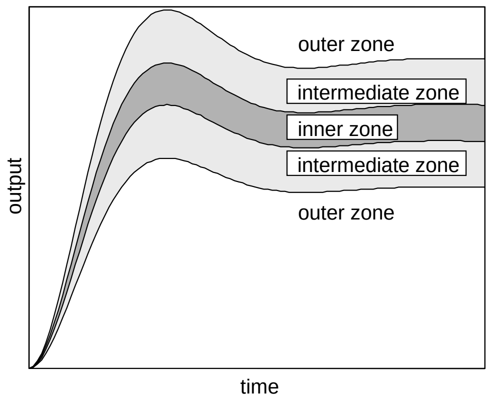

Fig. 1. Three zones defined by error-bounded envelopes

tees that a fault exits when the measurement is out of the overbounded envelope (outer zone) by eliminating in this way false alarms. However if the measurement is in the intermediate zone there can be a fault and not be detected (missed alarm). This is due to that overbounded envelope includes values that do not belong to the values space of the system represented by interval models. Another reason is the dynamics of the system so more time may be needed for detecting a fault. If the measurement is in the inner zone, the algorithm stops because it is not possible to detect faults.

To computed the error bounded-estimations in a time instant, the simulator uses the values of the interval measurement from the past instant t-window. In case of multi-incidences in the function and using algorithms based on Interval Arithmetic, overbounded results are obtained. However these multi-incidences are taken into account by using an extension of Interval Analysis called Modal Interval Analysis (SIGLA/X, 1999). In this way spurious solutions are reduced.

3. METHODOLOGY

The fault detection system uses the proposed methodology in order to diagnose faults. With SQualTrack the symptoms of faults caused by changes in process parameters are obtained. The models proposed to be introduced in SQualTrack are based on structural analysis that considers the links between variables and parameters in the model.

In the structural analysis (Blanke *et al.*, 2003) the model of the system is seen as a set of constraints that are applied to a set of variables. There is a subset formed by known variables (sensors and control variables). The set of constraints is given by component models of the system which form the elementary analytical relations (EAR) between the values of variables from physical laws.

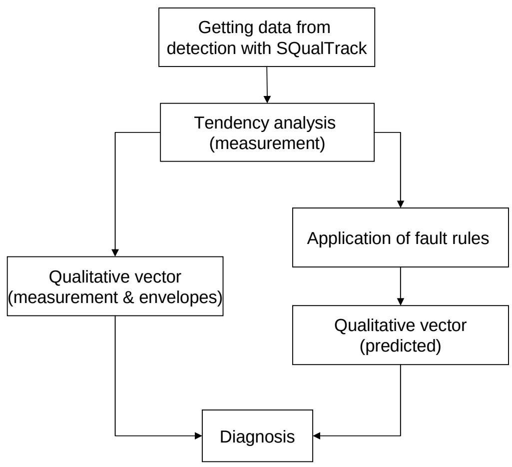

Fig. 2. Steps for Diagnosis

A constraint that applies only to known variables and parameters constitutes an analytical redundancy relation (ARR). The structure of the model is a matrix where a number 1 represents that the constraint applies to the variable and 0 otherwise. In this way the fault signature matrix is built. The constraints can be evaluated from only observed variables and they can be used in fault detection and diagnosis (Cordier *et al.*, 2000).

In the present approach a graphical interface and the functions to analyze the symptoms were developed on MATLAB. Two stages are considered; off-line to implement and configure the system and fault diagnosis after detection. On one hand off-line tasks are the generation of ARRs, the simulation and off-line detection of faults in order to have representative data from different faults and the creation of rules by including qualitative knowledge from fault detection. The rules will be used to build automatically a fault library according to the fault scenery. On the other hand the diagnosis tasks consists on six steps depicted in figure 2 and described in the following subsections. Some of these tasks are similar to those carried out in the off-line stage but with different data.

3.1 Getting data from detection

First step consist on getting the data from detection results obtained by SQualTrack. Data are included in a file with 9 columns in the following order: 1) step/time 2) Inferior measurement 3) Superior measurement 4) Inferior underbounded 5) Superior underbounded 6) Inferior overbounded 7) Superior overbounded 8) Fault/no fault 9) Window where fault was detected.

3.2 Tendency analysis

This step consists on making an analysis of monotony to the measurement data. The increas-

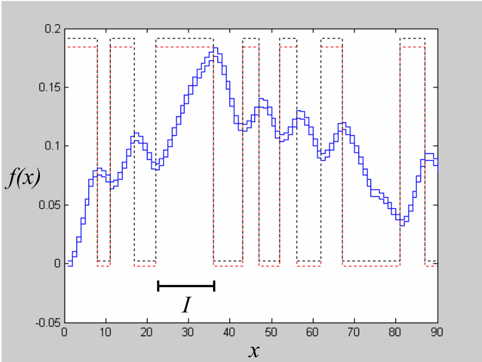

Fig. 3. Monotony analysis from measurement. $f(x)$ is the measurement, x is the step time and I is the interval of monotony.

ing behavior corresponds to the up-motion while the decreasing behavior corresponds to the down-motion. These two function properties are closely related to the behavior from the derivative of the function (when it is differentiable). In figure 3, the behaviors can be described in the following way: vertical axis $f(x)$ is increasing on the time interval I if for any $x_1, x_2 \in I$

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} > 0, \tag{5}$$

and $f(x)$ is decreasing on I if for any $x_1, x_2 \in I$

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} < 0 \tag{6}$$

The analysis is based on measurement values from the sampling time. The task consists on putting labels to the measurement values during time intervals. On one hand if the value of the measurement in next time step is bigger than present value, the label is ”Increasing”. On the other hand if the value is smaller, the label is ”Decreasing”. All data vector is analyzed in this way and the result is a qualitative vector that includes a behavior description of the current measurement. There is a third label called ”Constant” where the value remains constant. Figure 3 shows the intervals of monotony from a signal. Vertical axis is the measurement to be compared with the simulated one, horizontal axis refers to the step time. If the signal is increasing it is indicated in the graph by the maximum value. If the signal is decreasing it is indicated in the graph by the minimum value.

3.3 Qualitative vector from detection state

On this step a qualitative vector describes the relation of the current measurement with regard to the envelopes simulated by SQualTrack. The measurements on each time step are labelled. If the current measurements values are bigger than the superior overbounded, the label is ”Plus”. On the other hand if the measurements are smaller than the inferior overbounded the label is ”Minus”. In case of no fault has been detected (according to the properties of detection module) the label is ”Uncertain Plus” if at least one limit from the measurement envelope is bigger than the superior overbounded and the label will be ”Uncertain Minus” in any other case. They are uncertain since it is unknown if a fault exists or not.

3.4 Fault rules and fault library

As a suitable tool for treatment of heuristic knowledge (especially in the diagnostic domain), specific rules are applied in order to set up logical interactions between observed symptoms (effects) and faults (causes) to structure the knowledge in a problem-adapted manner. An example of a single rule is described as

$$\mathit{If} < \mathit{condition} > \mathit{THEN} < \mathit{conclusion} >$$

where condition part contains facts (symptoms) with AND and OR connectives and the conclusion part represents an event as a logical cause of these facts. On the present work a vector is generated by using rules of faulty behaviors. Each one of faulty behaviors to be detected are described by means of rules. An example of one simple rule is the following:

IF measurement is ”Growing” THEN ”Plus”
(Measurements are bigger than superior overbounded)

Tendency information and qualitative vectors describing faulty behaviors are used as training set to create a set of rules. The set of rules, used for classification and training, includes all kind of faults to be detected. The data are used with an algorithm called C4.5 introduced by Quinlan *(http://www.cse.unsw.edu.au/∼quinlan/)* for inducing Classification Models, also called Decision Trees, from data. A set of records is given. Each record has the same structure, consisting of a number of attribute/value pairs. One of these attributes represents the category of the record.

The aim is to determine a decision tree that will predict correctly the value of a non-category attribute. Usually the category attribute takes only the values between true or false. In this case the values depend on the number of differentiable faults.

3.5 Fault diagnosis

Diagnosis of the detected fault is formed by the steps depicted in figure 2. Data from detection are saved in a file. A monotony analysis is carried out to the current measurement (vector including

”Growing” or ”Falling” labels). After this a qualitative vector is obtained (”Plus” or ”Minus” labels depending on the behavior from the measurement with regard to the computed envelopes). At the same time another qualitative vector is obtained which predicts the behaviors of different faults by using the fault rules generated by the classification algorithm C4.5. Finally a comparison among qualitative current-vector and qualitative predicted vectors is made in order to find the most similar fault from the library. The idea of similarity is made by the definition of Hamming distance. The Hamming distance between these two vectors is the number of labels that disagree. The diagnosis is given by a list of the most possible faults in order of possibility being the most possible the one with the smaller distance.

4. EXPERIMENTAL RESULTS

The performance of the diagnosis system has been evaluated by using simulated supervised systems such as electrical circuits, linear systems and in this case the example of a DC motor is presented.

The elementary models of components are obtained from physical laws. Equations 7 and 8 correspond to the electrical subsystem and equations 9 and 10 to the mechanical subsystem:

$$L\frac{di_a}{dt} + Ri_a(t) + e_b(t) = e_a(t) \quad (7)$$

$$e_b(t) = K_1\omega(t) \quad (8)$$

$$J\frac{d\omega}{dt} + B\omega(t) = T(t) \quad (9)$$

$$T(t) = K_2 i_a(t) \quad (10)$$

Where imprecise parameters are: Resistance (R), inductance (L), rotor inertia (J) and viscose friction coefficient (B). Also back-EMF (e_b), rotor torque (T), torque constant (K_2) and for an ideal DC motor it equals the back-EMF constant (K_1) determine the system.

Table 1. Fault signature

	J	B	R	L
ARR1	1	1	0	0
ARR2	0	0	1	1

The fault signature matrix (Table 1) is obtained from structural analysis. For this reason it only includes known variables: Current (i_a), voltage supply (e_a) and angular speed (w). The fault signature indicates that, analytical redundancy relation number one just involves to the components J and B corresponding to the mechanical subsystem (equation 11) and equation 12 corresponds to the electrical subsystem. On this case is naturally separated but in more complex cases the ARRs can discriminate or put at the end of the list components which are not involved in an ARR where a fault has been detected.

$$w(t+1) = w(t)(1 - \frac{B\Delta t}{J}) + \frac{\Delta t}{J}(K_2 i_a(t)) \quad (11)$$

$$i(t+1) = i(t)(1 - \frac{R\Delta t}{L}) + \frac{\Delta t}{L}(e_a(t) - K_1 w(t)) \quad (12)$$

Several faults were introduce to diagnose in a DC motor for example increasing the resistance in the circuit simulating and increased stator temperature. Another case is to increase friction, this could be due to missing lubrication and beginning corrosion (see Table 2).

Table 2. Faults in the system

Fault	Kind of fault	Name	Value
1	normal	nofault	nominal
2	J $>$ normal	J-bigger	J = 0.1
3	J $<$ normal	J-smaller	J = 0.005
4	B $>$ normal	B-bigger	B = 0.15
5	B $<$ normal	B-smaller	B = 0.05
6	L $>$ normal	L-bigger	L = 0.7
7	L $<$ normal	L-smaller	L = 0.1
8	R $>$ normal	R-bigger	R = 1.5
9	R $<$ normal	R-smaller	R = 0.5

Figure 4 shows the results of the fault detection software SQualTrack, this information is saved in a file to be analyzed by the diagnosis stage. Data are obtained from a faulty behavior corresponding to the fault 2 where the parameter J is bigger than the considered as normal. The first graph (starting from the upper) shows the envelopes for the output variable and the corresponding measurements. The second graph shows red bars with value of 1 when a fault is detected. Third graph indicates the window length that has been used at the corresponding time step. Finally, the lower graph simply shows the current measurement.

Diagnosis system starts to deal with data applying the steps explained and shown in figure 2. The diagnosis support system gives information consisting on Hamming distances from similarity between the current behavior and the faulty predicted behaviors. The smaller distance is the most similar fault. Table 3 present the diagnosis corresponding to the results from a fault where J is bigger than the normal value. Column 1 indicates the Hamming distance of 200 samples which determines the order of possibility depending on the similarity and column 2 indicates the kind of fault.

Table 3. Diagnosis

Similarity distance	Kind of fault
18	Jbigger
78	Bbigger
102	Bsmaller
180	Jsmaller
199	nofault

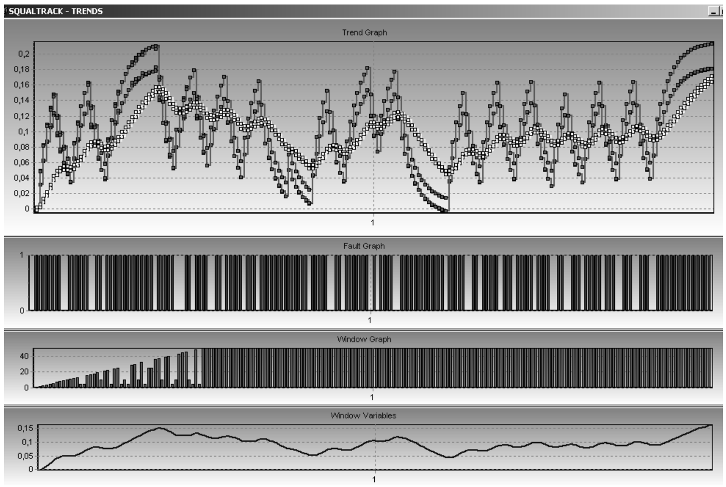

Fig. 4. SQualTrack fault detection. Case 2: J-bigger. Upper graph shows envelopes of measurement and simulation. Following graph shows bars with 1 when a fault is detected. Next graph indicates last window length applied at the specific time. Lower graph shows the analyzed measurement.

5. CONCLUSIONS AND FUTURE WORK

A mathematical interval model from a dynamical process has been simulated with SQualTrack which takes a decision about the presence of a fault. The elimination of false alarms makes it a useful tool for fault diagnosis. The structural analysis carries out the task of isolation through specific equations. Heuristic knowledge from these symptoms is used by making an automatic and adaptive fault library.

A methodology of fault diagnosis based on detection with interval models and based on a fault library created with an algorithm of rules (C4.5) is explained. The advantage of this methodology is that the rules can be applied to the same system with different parameters. Once the rules have been defined, fault library is build in an automatic way based on the current measurements. As the rules include qualitative knowledge from measurements with regard to the output, the rules can be applied to the system with different input signals (frequency and magnitude).

Combination of approaches has been successful. Interval models and structural analysis, techniques from FDI with classification algorithms and qualitative reasoning from AI have been complemented for making a more complete fault diagnosis system. The fault diagnosis capability was demonstrated with simulated data.

Future work consists on improving the rules in order to give quantities of intensity and size of the fault depending on how well a variable belongs to a specific fault. Generation of additional symptoms is another future work. Further research will concern fault diagnosis with real data and more complex systems.

ACKNOWLEDGEMENTS

This work has been funded by Mexican National Council for Science and Technology (CONACyT), the European Union (European Regional Development Fund) and the Spanish government (Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica, Ministerio de Ciencia y Tecnología) through the co-ordinated research project DPI2003-07146-C02-02.

REFERENCES

Armengol, J. (1999). Application of Modal Interval Analysis to the simulation of the behavior of dynamic systems with uncertain parameters. PhD. Universitat de Girona.

Armengol, J., J. Vehí, M.Á. Sainz and P. Herrero (2003). Fault detection in a pilot plant using interval models and multiple sliding time windows. In: *SAFEPROCESS'03*. Washington D.C.

Blanke, M., M. Kinnaert, J. Lunze and M. Staroswiecki (2003). *Diagnosis and Fault-Tolerant Control*. Springer-Verlag. Berlin, Germany.

Cordier, M-O, P. Dague, M. Dumas, F. Lévy, J. Montmain, M. Staroswiecki and L.Travé-Massuyès (2000). A comparative analysis of AI and control theory approaches to model-based diagnosis. In: *ECAI2000*. Germany. pp. 136–140.

Gertler, J. (1998). *Fault detection and diagnosis in engineering systems*. Marcel Dekker. New York.

SIGLA/X (1999). Modal intervals. Basic tutorial. In: *MISC99*. Universitat de Girona. Spain. pp. 157–227.