FAULT DIAGNOSIS WITH OUTPUT OPTIMIZATION PROCEDURE

Alexandar Kamenov Ichtev

Faculty of Automatics, Technical University of Sofia, boulevard Kliment Ohridski 8, Sofia 1000, Bulgaria, phone: (+359 2) 965 2420, e-mail: ichtev@tu-sofia.bg

In this paper a method for fault diagnosis is presented. It is based on the optimization procedure and information obtained from bank of models. This is iterative method. It is capable of detecting single, as well as multiple, subsequent as well as simultaneously developing, modeled ones as well as partial faults in the system. The alternative optimization with output optimization is also disused. Experiments with a benchmark example – long bridge is carried out for demonstration purposes. The obtained results are discussed.

Keywords: Fault diagnosis, optimization procedure, bridge benchmark

1. INTRODUCTION

In the last decades, there is increasing demand on performance for systems working in different environments. In order to satisfy this demand, more and more sophisticated systems with a larger number of sensors, actuators and other components are being built. As a result, the probability of a fault is increasing. On the other hand there are increasing safety demands. In order to satisfy those demands for automated systems reliable methods for fault detection and isolation are required.

This article proposes a method that continuously monitors the outputs of a plant. The objective is to detect fault as soon as possible after its occurrence as well as to determine its exact location and size. The results from fault detection might be used for fault tolerant control systems. The knowledge of the exact fault location will also be beneficial for rapid repair of the plant. This allows extension of the time between scheduled maintenances, which significantly reduces operational costs of the plant.

2. MULTIPLE MODEL APPROACH

The investigated systems are subject to abrupt as well as gradual developing faults. One way of describing such systems is by modeling them as hybrid dynamic system, whose state may jump as well as vary continuously. The jumps are used to model random abrupt faults. The dynamics between jumps is used to describe the investigated system, in faults free working regime or for gradually developing faults.

The model of the hybrid system is represented with

$$x(k+1) = \mathbf{F}(m(k+1))\mathbf{x}(k) + \mathbf{G}(m(k+1))\mathbf{u}(k) + \mathbf{T}(m(k+1))\boldsymbol{\eta}(k)$$
$$\mathbf{y}(k) = \mathbf{H}(m(k+1))\mathbf{x}(k) + \boldsymbol{\upsilon}(k)$$

where $m(k+1)$ is the model used in moment $k+1$.

It is assumed that the known model adequately describes the plant as well as that that jumps are described as first order Markov chain with transition probabilities from one model to another:

$$P\left(m_i(k+1)/m_j(k)\right) = \pi_{i,j}(k) \quad \text{with} \quad \sum_j \pi_{i,j}(k) = 1$$

The transition probability matrix (π) is from great importance for the proper operation of the interactive multiple model (MM) algorithms [3]. However, the only way to determine this matrix is with the trial and error method.

The MM approach assumes that a set of N models can be set, which can approximate the hybrid system with the following N pairs of equations

$$\mathbf{x}(k+1) = \mathbf{F}_j\mathbf{x}(k) + \mathbf{G}_j\mathbf{u}(k) + \mathbf{T}_j\boldsymbol{\eta}(k)$$

$$\mathbf{y}(k) = \mathbf{H}_j\mathbf{x}(k) + \boldsymbol{\upsilon}(k) \qquad \text{for } j = 1, 2, \cdots, N$$

The set of all models will be referred to a model set M.

Each pair of equations corresponds to an operating mode of the system or presence of fault. If a mathematical description of the plant is known then the models with faults can be obtained directly from it. For example, *60%* partial fault of the sensor can be modeled by multiplying the corresponding row of matrix $\mathbf{H}$ with *0.4*. Total failure can be modeled if it is multiplied with *0*.

3. Fault detection and isolation

In the literature there are two types of methods for fault detection and isolation. The first one relies on information obtained form redundant sensors, actuators and component. This is known as hardware redundancy. The second type is based on additional information from computational methods. This is referred to as analytical redundancy [1], [2]. The first type requires use of redundant equipment. An important example for this approach is aerospace industry, where there are three (sometimes four) systems for the implementation of the same task. This leads to increased costs for implementation, maintenance and operation. Therefore, in this article the latter approach is adopted.

In the model based approach [1], [2], [3], [4], [5] for fault detection and isolation there are two main stages: (i) determination of the plant's models (can be performed off-line) and (ii) on-line residuals generation and evaluation. The residual is analytically calculated value that represents the difference between the measured values form the plant and analytically calculated ones. The residuals are determined from a model of the plant and actually measured values. They represent the difference between the current (real) situation with our expectation for the system's behavior.

4. Problem formulation

The problem for fault detection and isolation is usually solved by performing a hypothesis test in order to determine the current model [1], [2]. This is done by choosing one model from the model set M. Afterwards, it is assumed that this is the thru one. The hard decision can be serious drawback when a system is working with partial faults and fault tolerant control is used [4], [5]. One way to overcome this problem is to extend MM to linear differential inclusions. Let us consider the MM set M. Linear differential inclusions are defined as a set of all plants that are convex combination of the model in M:

$$M = \sum_{i=1}^{N} \mu_i M_i \quad \text{with} \quad \sum_{i=1}^{N} \mu_i = 1$$

The used criterion for convex combination is a weighted sum from the outputs:

$$\mathbf{y}(k) = \sum_{i=1}^{N} \mathbf{y}_i(k)\mu_i$$

where $\boldsymbol{\mu}$ is vector containing probabilities for each of the models. Each element of this vector represents the probability that the particular model is the true one at the given time instant. This vector will be referred to mode probability vector.

If the set M is known in advance, i.e. all possible faults are determined, then the fault detection and isolation task boils down to calculation of the mode probability vector for each time instant. If the probability that corresponds to the nominal model of the plant is close enough to one and the probabilities for the rest models are small enough (close to zero), then the system is fault free. Each significant difference from this situation indicates that there is a fault in the system, which is fault detection. The calculation of the mode probabilities also provides additional information. By tracking the most likely model it can be established which fault scenario is present in the system, which is fault isolation.

5. ALGORITHM

The base of the MM algorithms consists in the use of separate descriptions for each mode of the sustem. A model is set to represent this situation. Each model is evaluated for each time instant. The interactive MM algorithm consists from four main parts:

1) Mode dependant reinitialization. In this part the iteratively of the algorithm is given. It is from great importance, since in the no iterative algorithms it is assumed that the system mode did not changes, while in the fault detection this mode change is essential. The initial values for all models are chosen based on the estimated model form the previous time instant. This assessment is probabilistic sum from previous estimates (see step 4 from this algorithm). This is how the interaction between the models is achieved. In this algorithm there is no need of knowing in advance the π matrix, in contrast to the proposed in [3] algorithm.

2) Model evaluation: All models are evaluated for each time instant and information regarding their outputs is stored.

3) Mode probability calculation and fault detection and isolation. The most significant difference between this approach and the one described in [3] is precisely in this step. Here it is proposed to be used an optimization procedure, as the presented results below are used with quadratic programming. Similar results are obtained by using other optimization algorithms that can solve the problem with limits.

During the optimization two limits are set in respect to the mode probabilities of the models.

$$\sum_{i=1}^{N} \mu_i(k) = 1 \quad \text{and} \quad 0 \le \mu_i(k) \le 1 \quad \text{for} \quad i=1,2,\cdots,N$$

The minimization at each step is

$$\min_{\boldsymbol{\mu}} \left[(\mathbf{Y}_m - \mathbf{Y}_s\boldsymbol{\mu})^T (\mathbf{Y}_m - \mathbf{Y}_s\boldsymbol{\mu}) \right] = \min_{\boldsymbol{\mu}} \left[\boldsymbol{\mu}^T \mathbf{Y}_s^T \mathbf{Y}_s \boldsymbol{\mu} - 2\mathbf{Y}_s^T \mathbf{Y}_m \boldsymbol{\mu}) \right]$$

where $\mathbf{Y}_m$ is the vector containing the measured outputs from the system $\mathbf{Y}_s$ is a vector containing the estimated output for each of the N models and $\boldsymbol{\mu}$ I s the vector containing the probabilities for each of the models.

$$\mathbf{Y}_m(k) = \left[y_1(k) \quad y_2(k) \quad \cdots \quad y_r(k) \right]^T$$

$$\boldsymbol{\mu}_m(k) = \left[\mu_1(k) \quad \mu_2(k) \quad \cdots \quad \mu_n(k) \right]^T$$

$$\mathbf{Y}_s(k) = \begin{bmatrix} y_{11}(k) & y_{12}(k) & \cdots & y_{1N}(k) \\ y_{21}(k) & y_{22}(k) & \cdots & y_{2N}(k) \\ \vdots & \vdots & \ddots & \vdots \\ y_{r1}(k) & y_{r2}(k) & \cdots & y_{rN}(k) \end{bmatrix}$$

The optimization is perfumed over the outputs of the system instead over the states. This aims to accelerate the algorithm, by simulating the output instead of perfuming the state estimation. It must be mentioned that, for the fault detection and isolation of component faults, it is better to use the optimization based on the states [3], while for fault detection of the actuators and especially sensor fault it is more convenient to use optimization over the outputs.

In some cases, especially when there is abrupt change in the operation point and when the system is subjected to strong noise, it is possible that some discrepancies between the actual situation and calculated mode probabilities can occur. In such cases the optimization should be performed over a time window instead of just one time instant. This will slow down the fault detection procedure, but it will overcome the mentioned problems.

4) Estimate combination. In this last step of the algorithm the model and the initial conditions for the next time instant are determined. They are calculated as weighted sums from the states of all models. The probabilities from the mode probability vector are used.

6. EXAMPLE

As an illustrative example in this article a long bridge is considered. This plant would be good illustrative example, because for control purpose of such plant it is necessary to use a large number of sensors and actuators.

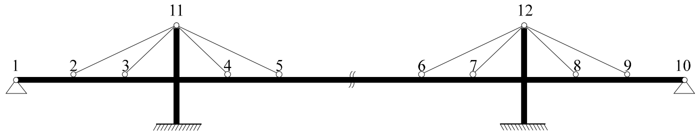

Fig.1 Principle diagram of the plant

Mathematical models. As mathematical model of the plant a second order matrix differential equation is assumed:

$$\mathbf{M}\frac{d^2\mathbf{Y}}{dt^2}+\mathbf{C}\frac{d\mathbf{Y}}{dt}+\mathbf{KY}=\mathbf{DV}+\mathbf{BU}$$

where **Y** is n-dimensional vector of displacements in the basic points of the plant, **V** is a vector of external forces, **U** is a vector of the controls, and **M**, **C**, **K**, **D** and **B** are respectively matrices of masses, damping, stiffness inputs and controls. By introducing a *2n*-dimensional state vector $\mathbf{X}^T=\begin{bmatrix}\mathbf{Y}^T & \dot{\mathbf{Y}}^T\end{bmatrix}$, the model yields the form

$$\dot{\mathbf{X}}=\begin{bmatrix}\mathbf{0} & \mathbf{I}\\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C}\end{bmatrix}\mathbf{X}+\begin{bmatrix}\mathbf{0}\\ \mathbf{M}^{-1}\mathbf{D}\end{bmatrix}\mathbf{V}+\begin{bmatrix}\mathbf{0}\\ \mathbf{M}^{-1}\mathbf{B}\end{bmatrix}\mathbf{U}$$

Experimental results. The nominal matrices in the plants model are

$\mathbf{M}=diag\begin{bmatrix}m_1 & m_2 & m_3 & m_4 & m_5\end{bmatrix}$,

$$\mathbf{C}=\begin{bmatrix}c_{1,1} & -c_{1,2} & 0 & 0 & 0\\ -c_{1,2} & c_{2,2} & -c_{2,3} & 0 & 0\\ 0 & -c_{2,3} & c_{3,3} & -c_{3,4} & 0\\ 0 & 0 & -c_{3,4} & c_{4,4} & -c_{4,5}\\ 0 & 0 & 0 & -c_{4,5} & c_{5,5}\end{bmatrix},\ \mathbf{K}=\begin{bmatrix}k_{1,1} & -k_{1,2} & 0 & 0 & 0\\ -k_{1,2} & k_{2,2} & -k_{2,3} & 0 & 0\\ 0 & -k_{2,3} & k_{3,3} & -k_{3,4} & 0\\ 0 & 0 & -k_{3,4} & k_{4,4} & -k_{4,5}\\ 0 & 0 & 0 & -k_{4,5} & k_{5,5}\end{bmatrix}$$

$$D=\begin{bmatrix}d_1 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & d_2\end{bmatrix}^T,\ \mathbf{B}=diag\begin{bmatrix}b_1 & b_2 & b_3 & b_4 & b_5\end{bmatrix}$$

Their parameters have the following numerical values:

$m_1=m_2=m_3=m_4=m_5=1$,

$k_{i,i}=2, i=1,2,3...,\ k_{i,i+1}=1, i=1,2,3...,\ c_{i,j}=0.01k_{i,j}$,

$d_1=d_2=b_1=b_2=b_3=b_4=b_5=1$

The experiment involves sensor fault simulations. There are three total faults: in sensors *1*, *2* and *5*, and three partial faults, which are 40% error in the measurements. During the first *30* times samples (*3* seconds) of the experiment the bridge and sensors are with no fault. In the next *30* time samples the system is under the influence of fault 1. Further, alternately fault free and next fault modes with the same frequency are simulated. In the last two faults scenarios are simulated non modeled partial failures. Thus, the method's ability to cope with failures that are convex combination of modeling one is verified. This is achieved without introducing additional models in the model set.

As can be seen from Figure 2 the algorithm correctly detects the true model. Excluding the first interval, which is transitory, it can be seen that the probability of the correct model is more than 90% (0,9). If we take a limit of 50% (or even 70%), i.e. be considered the most likely model as a true one, it can be concluded that the algorithm works perfectly, even in the first (transitional) interval.

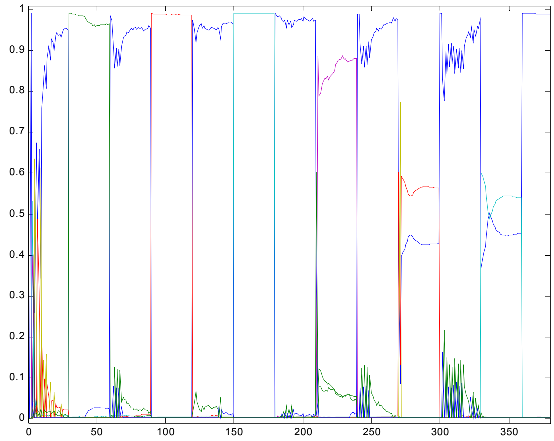

Fig.2 Mode probabilities for each of the models

From great interest are the last two fault intervals. It can be seen that the partial non modeled faults can be successfully detected, isolated and identified as well. This shows that the proposed method can evaluate each partial failure without adding additional models. This is the main advantage of this method with respect to the other methods in the literature. It is essential for control in the presence of partial faults. Faults tolerant systems are discussed in [4] [5].

7. Conclusion

This article proposes a new interactive method for fault detection and isolation. The output optimization allows faster evaluation then similar ones based on state optimization. Its main advantage is that it can detect and isolate partial faults, without introduction of additional faults. This advantage is demonstrated by the long bridge benchmark example. The methods advantages are important for fault tolerant control.

8. Acknowledgement

The research results presented in this paper are partially funded by Technical University – Sofia research project No. 102ni046-08.

9. References

[1] Gertler, Janos, Fault Detection and Diagnosis in engineering systems, Marcel Dekker, Inc., USA, 1998

[2] R. Isermann Fault-Diagnosis Systems, Springer 2006.

[3] Ichtev. A. "Fault tolerant control with the controlled bridge structures". VSU'2002, v.1, pp. I – 137-142, 2002 (In Bulgarian)

[4] A. Ichtev, J. Hellendoorn, R. Babuska, S. Mollov, Fault Tolerant Model Based Predictive Control Using Multiple Takagi-Sugeno Fuzzy Models, FUZZ-IEEE 2002, Hawaii, USA, Volume: 1, 346 -351

[5] A. Ichtev, P. Petkov, Fault Tolerant Control by Compensating Partial Faults, 6th WSEAS conference, Crete, Greece, July 7-14, 2002, 259-264