A Brief Review of Model-Based Fault Diagnosis

Gordana Janevska[1]

Abstract – **This paper outlines the basic concept of fault detection and isolation (FDI), i.e. fault diagnosis in dynamic systems based on analytical process models. It gives a brief review of the most important approaches in literature.**

Keywords – **fault diagnosis, analytical redundancy**

I. Introduction

Within the last two decades there has been increasing interest in the field of fault diagnosis both within the academic community and in industry. The increasing complexity of automatic control systems and their use in safety critical areas such as flight control and chemical and power plants has helped to fuel this interest. An undetected fault in such a system could often result in dire consequences and the vast array of data which complex system generate can make in difficult, if not impossible, for operators to assess the location and nature of a fault. Automated fault diagnosis has an important role both in systems such as flight control, where automatic reconfiguration of sensors and control systems can be carried out, and with process management, where the role of a fault diagnosis system is more one of information processing and filtering with the final decision and choice of action being performed by a human operator.

Away from the safety critical areas, fault diagnosis is also attracting interest from those wishing to improve productivity and reduce plant downtime. The knowledge of the state of a plant can be used to schedule maintenance and allow reconfiguration or operating point changes to be carried out more effectively to increase the efficiency of a plant's operation. Such a scheme is often termed condition monitoring and can result in significant decreases in the running costs of plants.

The increase in the need for fault diagnosis systems has been matched with advanced in computer technology which facilitates the analysis of large amounts of data. Methods that would have been computationally infeasible only a decade age can now be applied in real time.

II. The Fault Diagnosis Problem

Diagnosis is a procedure to detect and locate faulty components in a dynamic process. Faults and failures in complex automated control systems are, in general, unavoidable facts and they require quick detection, location and identification. A diagnosis scheme is of importance in, for example, nuclear plants, aeroplanes, automotive engines. This is due to increasing demand for higher performance, higher safety and reliability. Different fault detection and isolation techniques have been developed over the recent years.

A general diagnosis procedure for a dynamic system consists of several tasks. In literature the following steps are suggested.

- *Fault detection*: Detect when a fault has occurred. That is often done with a suitable comparison, for example in parameter estimation, the estimated physical parameters are compared to their nominal values;
- *Fault isolation*: Isolate the fault. Primarily to determine the faults origin but also the fault's type, size and time.

These two tasks are commonly referred to as FDI (*fault detection and isolation*), which sometimes is referred to as *diagnosis* and the other way around.

The system to be diagnosed often includes a control loop, which further complicates the problem. A control loop tends to hide or mask a faulty component or sensor making it even more important, in a controlled system, to detect faults. The control loop can also damp the system's signals making it necessary to excite the signals from the system.

We speak of *faults* and *failures* in diagnosis. In diagnosis literature there is a distinction between the two and the definition can be written as:

Definition 1. *A failure suggests a complete breakdown of a process component while a fault is thought of as an unexpected component change that might be serious or tolerable.*

Fault diagnosis and fault detection is not a new problem and before model based fault diagnosis, they were accomplished e.g. by introducing hardware redundancy in the process. A critical component was then duplicated, triplicate (TMR) or even quadrupled and a majority decision rule was then used. Hardware redundancy methods are fast and easy to implement but they have several drawbacks

- Extra hardware can be very expensive
- It introduces more complexity in the system
- The extra hardware is space consuming which can be of great importance, e.g. in a space shuttle. Also the components weight sometimes has to be considered.

Instead of using hardware redundancy, analytical redundancy can be utilized to reduce, or even avoid, the need for hardware redundancy. Analytical redundancy is in principle the relationships that exist between process variables and measured output signals. If an output signal is measured, there is information about all variables that influences the output signal in the measurement. If the relationships are known, by quantitative or qualitative knowledge, this

[1]Gordana Janevska is with the Faculty of Technical Sciences, I.L.Ribar bb, 7000 Bitola, R. Macedonia E-mail: gordana.janevska@uklo.edu.mk

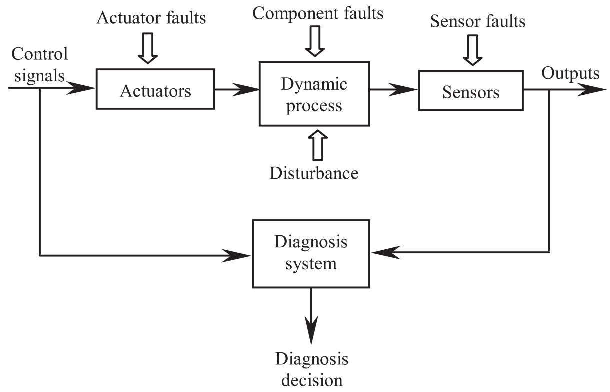

Fig. 1. Structure of a diagnosis system

information can be extracted and the extracted information from different measurements can be checked for consistency against each other.

There are different types of analytical redundancy. Instead of measuring several outputs, the different output measurements at different times can be compared. If the relationship between time series of outputs and inputs are known, from this relationship fault information can be extracted. This kind of analytical redundancy is called temporal redundancy.

The faults acting upon a system can be divided into three types of faults.

- *Sensor (Instrument) faults*: Faults acting on the sensors
- *Actuator faults*: Faults acting on the actuators
- *Component (System) faults*: fault acting on the system or the process we wish to diagnose.

A general FDI scheme based on analytical redundancy can be illustrated as in Fig. 1, an algorithm with measurements and control signals as inputs and fault detection as output.

It is unrealistic to assume that all signals acting on the process can be measured; therefore an important property of an algorithm is how it reacts on these unknown inputs. It is also unrealistic to assume a perfect model; the modelling errors can be seen as unknown inputs. An algorithm that continues to work satisfactory even when unknown inputs vary is called *robust*. Some approaches give the possibility to achieve disturbance decoupling, i.e. make the isolation decision independent of unmeasured disturbances.

III. Advantages of Model Based Diagnosis

This paper outlines the model based diagnosis, i.e. the procedure of diagnosis based on the mathematical model of the system. Why is there need for a mathematical model to achieve diagnosis? It is easy to imagine a scheme where important entities of the dynamic process is measured and tested against predefined limits. The model based approach instead performs consistency checks of the process against a model of the process. There are several important advantages with the model based approach.

- Outputs are compared to their expected value on the basis of process state, therefore the thresholds can be set much tighter and the probability to identify faults in an early stage is increased dramatically.
- A single fault in the process often propagates to several outputs and therefore causes more than one limit check to fire. This makes it hard to isolate faults without a mathematical model.
- With a mathematical model of the process the FDI scheme can be made insensitive to unmeasured disturbances, and also feasible in a much wider operating range.
- It might be possible to perform the diagnostic task without installing extra sensors, i.e. the sensors available for e.g. control might suffice.

There is of course a price to pay for these advantages in increased complexity in the diagnosis scheme and a need for a mathematical model.

IV. Quantitative Approaches to Diagnosis

In quantitative approaches the diagnosis procedure is explicitly parted into two stages, the residual *generation* stage and the residual *evaluation* stage, as illustrated in Fig. 2.

The residual evaluation can in its simplest form be a threshold test on the residual, i.e. a test if $|r(t)| > Threshold$. More generally the residual evaluation stage consists of a change detection test and a logic inference system to decide what caused change. A change here represents a change in normal behavior of the residual.

The residual generation approaches can be divided into three subgroups, *limit & trend checking*, *signal analysis* and *process model based*.

- **Limit & trend checking** – This approach is the simplest imaginable, testing sensor outputs against predefined limits and/or trends. This approach needs no mathematical model and therefore it is simple to use, but it is hard to achieve high performance diagnosis.
- **Signal analysis** – These approaches analyze signals, i.e. sensor outputs, to achieve diagnosis. The analysis can be made in the frequency domain, or by using a signal

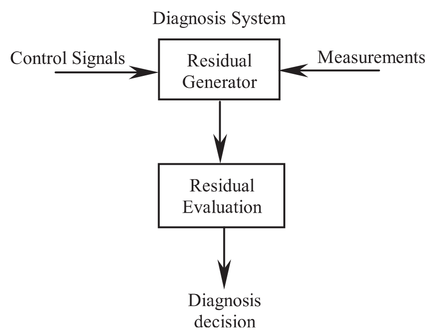

Fig. 2. Two stage diagnosis system

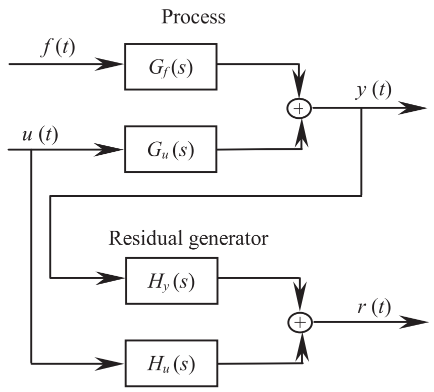

Fig. 3. General structure of a linear residual generator

model in the time domain. If fault influence is known to be greater than the input influence in well known frequency bands, a time-frequency distribution method can be used.

- **Process model based residual generation** – These methods are based on a process model. The process model based approaches are further parted into two groups, *parameter estimation* and *geometric approaches*.

The approaches mention here generate residuals which can be defined as:

Definition 2. *A residual (or parity vector) $r(t)$ is a scalar or vector that is 0 or small in the fault free case and ≠0 when a fault occurs.*

The residual is a vector in the parity space. This definition implies that the residual $r(t)$ has to be independent of, or at least insensitive to, system states and unmeasured disturbances.

A general structure of a linear residual generator can be described as in Fig. 3. The transfer function from the fault $f(t)$ to the residual $r(t)$ then becomes

$$r(s) = H_y(s)G_f(s)f(s) = G_{rf}(s)f(s) \quad (1)$$

To be able to detect the i:th fault the i:th column of the response matrix $[G_{rf}(s)]_i$ has to be nonzero, i.e.

Definition 3. ***Detectability*** *The i:th fault is detectable in the residual if $[G_{rf}(s)]_i \neq 0$*

This condition is however not enough in some practical situations. This leads to another definition,

Definition 4. ***Strong detectability*** *The i:th fault is said to be strongly detectable if and only if $[G_{rf}(s)]_i \neq 0$*

Note that in Definition 4, the frequency $\omega = 0$ is made particularly important. Which frequency that is particularly important depends on which type of faults that are interesting. There are three different types of temporal fault behavior:

- Abrupt, step faults
- Incipient (developing) faults
- Intermittent faults

V. Isolation Strategies

In the case of the strongly detectable residuals, the literature describes two general methods for isolation,

- Structured residuals
- Fixed direction residuals

The idea behind **structured residuals** is that a vector valued of residuals is designed making each element in the residual insensitive to different faults or subset of faults whilst remaining sensitive to the remaining faults, i.e. if three faults should be isolate then a three dimension residual should be designed with components $r_1(t)$, $r_2(t)$ and $r_3(t)$ insensitive to one fault each. Then if component $r_1(t)$ and $r_3(t)$ fire it can be assumed that fault 2 has occurred.

The idea with fixed direction residuals is the basis of the f*ault detection filter* (FDF) where the residual vector get a specific direction depending on the fault that is acting upon the system.

VI. Robustness

One problem, as was noted earlier, is that unmeasurable signals often act upon the system plus the influence by modelling errors. This makes it hard to keep the false alarm rate at an appropriate level. This problem is called the robustness problem and a diagnostic algorithm that continues to work satisfactory, even when subjected to modelling errors and disturbances, is called robust.

Since the ideal situation never occurs in a real application, the robustness aspect is one of the most important issues when designing a diagnosis system. The methods to tackle the robustness problem can be divided into two categories

- Robust residual generation, active robustness
- Robust residual evaluation, passive robustness

Robust residual generation methods strive to make the residuals insensitive or even invariant to model uncertainty and disturbances, and still retain the sensitivity towards faults. There are two different types of disturbances, structured and unstructured disturbances. If it is ”known” exactly how a disturbance signal influences the process it is called structured uncertainty and this high degree of disturbance knowledge is enough to actively reduce or even eliminate the disturbance influence on the residual. However if no knowledge of the disturbance is known, no active robustness can be achieved.

However, it is possible to increase robustness in the fault evaluation stage, i.e. in the threshold selection step, for example by using adaptive threshold levels or statistical decoupling. This is called passive robustness. It is not likely that one method can solve the entire robustness problem; a likely solution is one where disturbance decoupling is used side by side with passive robustness.

VII. Model Structure

To proceed in the analysis of residual generation approaches, an analytical model is needed. A state representation of the model is given with the following equation:

$$\begin{aligned}\dot{x}(t) &= f(x(t), u(t)) \\ y(t) &= h(x(t), u(t))\end{aligned} \tag{2}$$

The linear (time-continuous) state representation

$$\begin{aligned}\dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t)\end{aligned} \tag{3}$$

As was noted earlier, there are three general types of faults: **sensor (instrument) faults**, **actuator faults** and **component (system) faults**. There are also uncertainties about the model or unmeasured inputs to the process. If these uncertainties are structured, i.e. it is known how they enter the system dynamics, this information can be incorporated into the model.

In the linear case and model uncertainties are supposed structured, the complete model becomes

$$\begin{aligned}\dot{x}(t) &= Ax(t) + B(u(t) + f_a(t)) + Hf_c(t) + Ed(t) \\ y(t) &= Cx(t) + Du(t) + f_s(t)\end{aligned} \tag{4}$$

where $f_a(t)$ denotes actuator faults, $f_c(t)$ component faults, $f_s(t)$ sensor faults and $d(t)$ disturbances acting on the system. H and E is called the distribution matrices for $f_c(t)$ and $d(t)$.

VIII. Parameter Estimation

Process model based residual generators could be parted into two approaches: parameter estimation and geometric approaches.

A parameter estimation method is based on estimating important parameters in a process, e.g. frictional coefficients, volumes or masses, and compares them with nominal values. The typical parameter estimation diagnosis method can be outlined with three steps

- **Data processing**, with the help of the model and measured output data, model parameters can be estimated
- **Fault detection**, which includes a comparison between the estimated parameters and the nominal values
- **Fault classification**, in the case of fault presence, isolation of the fault source is the final stage in a parameter estimation method.

IX. Parity Space Approaches

Geometric approaches to residual generation are called parity space approaches because they generate residuals that are vectors in the parity space. The methods can be divided into open- and closed-loop approaches. In an open-loop approach there are, as the name suggests, no feedback from previously calculated residuals.

The idea behind closed-loop approaches, i.e. observer base approaches, is to use a state estimator as a residual generator. There are a number of approaches suggested in literature like

- State observers
- Fault detection filter
- Unknown Input Observers
  - By parity equations
  - By Kronecker canonical form
  - By eigenstructure assignment of observer

Note that these are methods to design the residual generator. Several of these designs may result in the same residual generator in the end.

X. Summary of Approaches in Literature

To summarize the relationships between the different diagnosis methods a tree-structured is presented in Fig. 4. The different residual generation methods are related as in Fig. 5. All these methods have their advantages and disadvantages and it is likely that in a complete diagnosis application several of these methods will be used.

The presentation done here is in no way complete as there exist numerous of approaches, e.g. the neural network approach.

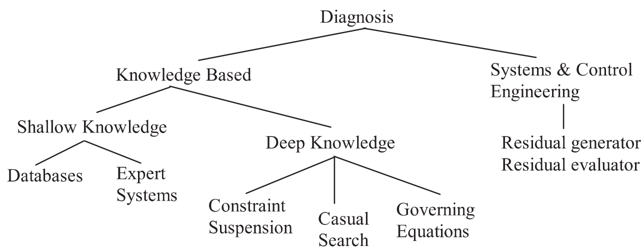

Fig. 4. Categorization of FDI methods

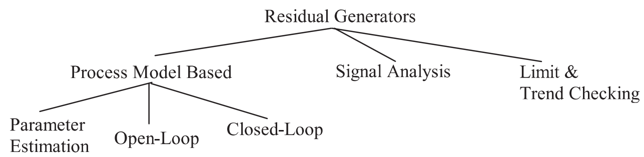

Fig. 5. Categorization of residual generation methods

References

[1] P.M. Frank, "Fault Diagnosis in dynaic system using analytical and knowledge based redundancy – A survey and some new results", *Automatica*, Vol.26, (3), pp.459-474, 1990.

[2] P.M. Frank, "Application of fuzzy logic to process supervision and fault diagnosis",*IFAC Symp. SAFEPROCESS '94*, pp.531-538, Espo, Finland, 1994.

[3] P.M. Frank, "Analytical and qualitative model-based fault diagnosis – a survey and some new results",*European Journal of Control*, no. 2, pp.6-28, 1996.

[4] E. Frisk, *Model-based fault diagnosis applied to a SI-engine*, Master's thesis, Reg.nr: LiTH-ISY-EX-1679, Linköping University, 1996.

[5] M. Nyberg, *Model Based Fault Diagnosis: Methods, Theory and Automotive Engine Applications*, Phd. thesis 591, Department of Electrical Engineering, Linköping University, Linköping, Sweden, 1999.

[6] R.J. Patton, ”Fault-Tolerant Control Systems: the 1997 situation”, *IFAC Symposium on Fault Detection Supervision and Safety for Technical Processes*, Vol. 3, pp.1033-1054, Kingston Upon Hull, UK, 1997.

[7] R.J. Patton, F. J. Uppal & C. J. Lopez-Toribio, ”Soft Computing Approaches to Fault Diagnosis for Dynamic Systems: A Survey”, *4th IFAC Symposium on Fault Detection Supervision and Safety for Technical Processes*, Vol. 1, pp.298-311, Budapest, 2000.