AQTR 2004 (THETA 14)
2004 IEEE-TTTC - International Conference on Automation, Quality and Testing, Robotics
May 13 – 15, 2004, Cluj-Napoca, Romania

Fault Detection and Fault-tolerant Control - Basic Ideas, State of the Art and Perspectives -

Paul M. Frank

Universität Duisburg-Essen, Fakultät Ingenieurwissenschaften/AKS
Bismarckstrasse 81, 47048 Duisburg, Germany

Keywords: Fault detection, fault-tolerant control, diagnosis, safety, observers.

1 Basic Ideas

All real systems in nature – physical, biological and engineering systems – can malfunction and fail due to faults in their components. The chances for failures are increasing with the system's complexity. The complexity of engineering systems is permanently growing due to the growing system size and degree of automation, and accordingly increasing is the danger of faults and, at the same time, accordingly aggravating are the effects of system failures on man and environment. Therefore, increased attention has to be paid to the reliability, safety and fault tolerance in the design and operation of modern automated engineering systems. But obviously, compared to the high standard of perfection that nature has developed with the "self-healing" and "self-repairing" mechanisms in its complex biological organisms, the fault management in engineering systems is far behind their technological standards and is still in its infancy.

In technical automatic control systems, defects may happen in sensors, actuators, the framework and components of the plant, or within the hardware or software of the control equipment. Component faults can cause a failure of the whole system. This effect can be diminished but also be amplified by the closed loop. The closed loop may also hide an incipient fault from being observed until a situation has been reached in which a failing of the whole system has become unavoidable. Even making the closed loop *robust* or *reliable* (by using *robust* or *reliable* design algorithms), can not solve the problem in full. It may, in the presence of faults, ensure to retain stability of the closed loop and to continue its mission with the desired or a tolerable degraded performance, but when the faulty part continues to miss-function, this may cause damage to man and environment due to the consequences of the faults (i.e., leakages in gas tanks or in oil pipes etc.). So, robust and reliable control using available hard- or software redundancy may be efficient ways to complete the mission of a control system, but it can not guarantee environmental compatibility or safety of the whole system. A realistic fault management has to provide dependability which includes both reliability and safety. Dependability is a fundamental requirement in industrial automation, it can only be attained by *fault-tolerant control* (FTC).

Traditionally, the basic approaches to FTC are divided into two main categories: *passive* and *active* fault-tolerant control. Typical for *passive* FTC is that a fixed compensator is designed that maintains at least the stability of the control system when a fault occurs. Clearly, like robust and reliable control, this approach concentrates on the performance of the control system and does not include the safety and environmental aspects. Some current theoretical approaches must be considered highly problematic from a practical point of view. In contrast, a*ctive* fault-tolerant control makes use of a fault diagnosis mechanism, which detects and isolates a fault, and whenever a fault occurs, a supervisory system takes action to modify the parameters and/or the structure of the system. The key issue of an active FTC system is to prevent local faults from developing into a system failure that can end the mission of the system and/or cause safety hazards by the faulty devises for man and environment. Because of the increasing importance in industrial automation, FTC (both active and passive) has also become an emerging research topic of modern control theory.

Note that automation of *safety-critical* systems, where no failure can be tolerated, requires redundant hardware to accomplish a system that is not affected by any single failure. *Fail-operational* systems are made insensitive to any single component fault. *Fail-safe* systems perform a controlled shut-down to a safe state when a measurement indicates a critical fault. *Robust* and *passive fault-tolerant* control ensure stability and pre-assigned performance of the control loop in the presence of faults within a specified range. In contrast, active *fault-tolerant* control monitors on-line the system behaviour and diagnoses critical faults in the components, and after detection of the faults it causes appropriate remedial actions in order to prevent faults from developing into a failure. Depending on the special purpose of application, the overall FTC system has to keep the plant availability despite the faults, possibly with degraded performance, and, if the faults cause damage or endanger man or environment, handles them by system reconfiguration, e.g., shutting down the faulty devices and substituting their function.

The basic schematic arrangement of the active FTC framework has four basic functional units: the plant, i.e., the controlled object including actuators and sensors, the controller with all control devices, configuration commander, and the supervisory system which performs the *fault diagnosis* (FD).

It can be seen that the implementation of a fault tolerant control framework requires multi-level automation, i.e., the control level needs to be complemented by at least one further level, the supervisory level (sometimes this is organised in several higher levels of information or knowledge processing and logical decision making). The purpose of the supervision level is to observe the process in order to maintain the desired plant availability and avoid damages and accidents. Traditional approaches to the accomplishment of this task are:

- monitoring, i. e., checking of operating conditions, system states and measurable signals (magnitudes, tolerances, limit values, trends),
- giving alarms and instructions to human operators to take proper actions,
- setting reference inputs and tolerances,
- providing automatic protection of the process.

In advanced fault-tolerant control systems, fault diagnosis has become a key issue of the tasks of the supervision level. Note that the traditional non-model-based FD methods can only cope with the FD problem in an incomplete manner. Their advantage is their simplicity and reliability and that they do not need detailed knowledge of the system, which is often not available or too expensive. Their crux with respect to fault tolerance is, however, that only relatively large changes (sudden and long drifting faults) are detectable and that reliable results are only available in steady state operation of the system. Also, early detection of small abrupt and incipient faults and a full and systematic fault diagnosis are not possible.

These deficiencies can be overcome with more sophisticated methods of fault diagnosis, where the model-based methodology plays a fundamental role. The model-based approaches make use of dynamic models of the system under consideration and, if good models are available, are thus capable of detecting small faults, performing high-quality fault diagnosis by determining time, size and cause of a fault, and are applicable in case of dynamic system operation. At the occurrence of faults they detect the faults by generating discrete event signals from which an automatic reconfiguration commander can be triggered in order to do fault accommodation. But not only that model-based fault diagnosis is an essential ingredient of fault-tolerant control, it is also a basic tool for off-line tasks such as condition-based maintenance and repair, which is carried out according to the information obtained by condition monitoring of the system. Hence model-based fault diagnosis is an important issue in all kinds of advanced engineering systems.

In this paper, we outline the fundamentals, review the state of the art and try some perspectives of technical fault diagnosis and active fault-tolerant control with focus on the advanced model-based approaches to fault detection and isolation (FDI). Some critical comments are made to current developments in the theory of passive FTC.

2 State of the Art

The main contributions to the theory and practice of *fault detection* and *fault tolerant control* were made in the last three decades by introducing the model-based methodology. Up to the 1970s it was common practice in the engineering community to use model-free (signal-based) approaches - analytical or heuristic - , mostly conducted by the human operator. In the early 70s, the control engineering community gave a fresh impetus. Stimulated by the achievements of the theory of mathematical modelling, signal processing, observers and Kalman filters, control engineers introduced the *analytical redundancy* approach based on the use of dynamic mathematical models.

The fundament of the model-based approach to FDI was laid by the pioneering works of Beard (1971) and Jones (1973), who introduced the concept of the fault detection filter, an observer-based approach for FDI in linear systems. This was the beginning of a tremendous development, and model-based FDI became a special discipline of control theory. The observer-based approach was soon followed by the parity space approach (an extension of the idea of plausibility check) and the parameter identification

approach. The further development is characterised by various extensions of the observer-based approach towards banks of observers ("observer schemes"), unknown input observers, non-linear and adaptive observers and structural analysis. Because of the usual practical difficulties resulting from modelling uncertainties with analytical models, the robustness of analytical model-based FDI methods has become an important issue, both in practice and research, which has, among others, resulted in the framework of robust diagnostic observers.

In parallel to the activities in the field of analytical redundancy, there have been increasing efforts to use qualitative, data-based and knowledge-based models. An early important impetus came from medicine, followed by the artificial intelligence and computer science communities according to the growing significance of the computer in FD. Already in the early 70s, medical scientists introduced the concept of a diagnosis expert system (MYCIN) making use of medical expert knowledge. Whereas technical expert systems had no great future due to the well-known difficulties with knowledge acquisition, more encouraging results were obtained in the 80s and early 90s with qualitative modelling techniques, fuzzy logic, neural networks and computational intelligence. Currently, the focus of research and development is on residual generation using structural modelling, non-linear observers, adaptive observers, fuzzy and neural observers, knowledge observers, intelligent residual evaluation using intelligent agents, and the integration of different approaches with due regard of the specific conditions and demands in industrial applications.

3 Perspectives

The problem of unforeseen changes in engineering systems due to disturbances and faults has long been recognized in the control community and has become an important issue in research and development of control systems in recent years. The outcome is known under several terms such as robust control, reliable control and fault-tolerant (or fault-adaptive) control. There are clear differences between the three approaches; at present, the definitions are as follows: Robust control aims at continued system operation under *continuous* changes, reliable control aims at continued system operation under *abrupt* changes, and fault-tolerant control includes the safety aspect, i. e., it aims at continued system operation providing safety with respect to both failure of the control system *and* faults in the components. The latter is often overseen by control theoreticians. Note that fault-tolerance is demanding more than robustness.

A great problem at present is that the relevance of fault diagnosis and fault-tolerance is not adequately recognized in practice, and the research in this field suffers from an according lack of acceptance and support from industry. There are several reasons for this. The most important one is that there is an inherent difference between the role of control and the role of diagnosis. Control is *function-critical*, i. e., control is necessary to make the system work, without control there is no operation. In contrast, diagnosis is *malfunction-critical*, i. e., it is only needed if the system does *not* work which is an abnormal case; in other words, normally the system works without the need of diagnosis (and fault-tolerance). Since an FDI or FTC system is expected to never be needed, nobody wants to pay for it, the more that industry invests huge amounts of money and

makes great efforts to *avoid* faults in the system. But this situation will drastically change in the near future because of the following reasons.

When looking to what nature has achieved with its biological systems and organisms, we realize that it has not only been extremely productive and successful in building systems of greatest capability and sophistication creating an overwhelming spectrum of functionality, nature has also managed with fascinating perfection the task of self-repair and self-healing, i. e., it has equipped all its products with a perfect *fault protection system*. When we break a bone, we may just put it into a plaster dressing and wait for 6 weeks and it will be cured, but when we break an axis of a wheel in our car, we may also put it in a plaster dressing and wait as long as we whish and nothing will happen. This illustrates that technology is still in its infancy. In other words, technical systems are yet incomplete, what they lack is the fault protection mechanism. The stately engineering success in the technological development achieved so far is only the *first* step of the whole solution, the *second* step, the equipment with an efficient fault protection mechanism, is still missing. This deficit is more and more being recognized by our society in that there are increasing demands for the safety of man and environment with respect to the technological development.

As a result, one of the great technological challenges of the 21st century will be to step in its second phase, namely to cure this deficit and develop both an efficient theory and industrial practice to equip engineering systems with the missing fault protection and self-healing mechanisms. Most of the existing engineering artefacts will then have to be re-examined with respect to their safety and reliability properties, and many of them will even have to be re-designed in order to make them acceptable for the demands of our society. In the view of this, fault-tolerance and self-repair properties turn out to become key features in research and industrial applications in the years to come. In other words, there is a perspective that FDI and FTC are among the disciplines in automatic control that will have a gigantic future.

4 Bibliography

Beard R. V. (1971). Failure accommodation in linear systems through self-reorganization. *Man Vehicle Lab., M.I.T., Cambridge, Mass., Report no. MTV-71-1*.

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

Blanke M., Izadi-Zamanabadi R., Bøgh S. A. and Lunau C. P. (1997). Fault Tolerant Control - A Holistic View. *Control Engineering Practice* **5**, pp 693-702

Chen J., Patton R.J. (1998). Robust Model-based Fault Diagnosis for Dynamic Systems. *Kluwer Academic Publishers*.

Frank P.M., Keller L. (1984). Entdeckung von Instrumentenfehlanzeigen mittels Zustandsschätzung in technischen Regelungssystemen. *VDI Fortschrittsberichte, Reihe **8***, Nr. 80. VDI Düsseldorf.

Frank P.M. (1987). Advanced fault detection and isolation schemes using non-linear and robust observers (Invited Survey Paper). *Proceedings of the 10th IFAC World Congress Munich*, pp. 63-68.

Frank P.M. (1990). Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy, *Automatica* **26**, pp. 459-474.

Frank P.M., Ding X. (1994). Frequency domain approach to optimally robust residual generation and evaluation for model-based fault diagnosis. *Automatica* **30**(4), pp. 789-804.

Frank P.M., (1994). Enhancement of robustness in observer-based fault detection. *International Journal of Control* **59** (4) pp. 955-981.

Frank P.M. (1996). Analytical and qualitative model-based fault diagnosis – A survey and some new results, *European Journal of Control* **2** (1), pp.6-23.

Gertler J. (1998). Fault Detection and Diagnosis in Engineering Systems. *Marcel Dekker*.

Isermann R. (1984). Process fault detection based on modelling and estimation methods: A survey. *Automatica* **20**, pp. 387-404.

Isermann, R. (1997). Supervision, fault-detection and fault-diagnosis methods – an introduction. *Control Engineering Applications* **5**, pp. 639-652.

Isermann, R., Ballé, P. (1997). Trends in the application of model-based fault detection and diagnosis of technical processes. *Control Engineering Applications* **5**, pp. 709-719.

Jones, H. L. (1973). Failure detection in linear systems. PhD Dissertation, *Dept. Aero. and Astro., M.I.T. Cambridge, Mass. USA*

Mahmoud M., Jiang J. and Zhang Y (2003). Active Fault Tolerant Control Systems – Stochastic Analysis and Synthesis. *Springer Lecture Notes in Control and Information Sciences* **287**.

Mangoubi R. S. (1998). Robust Estimation and Failure Detection – a Concise Treatment. *Springer Verlag*.

Patton R.J., Frank P.M., Clark R.N., eds. (1989). Fault Diagnosis in Dynamic Systems, Theory and Application. *Prentice Hall*.

Patton R.J., Frank P.M., Clark R.N., eds. (2000). Issues of Fault Diagnosis for Dynamic Systems. *Springer*.

Willsky A.S.(1976). A survey of design methods for failure detection in dynamic systems. *Automatica* **12**, pp. 601-611.