Modified fault diagnosis system

Pavel Trnka[1,*]

[1] *CTU in Prague, Faculty of mechanical engineering, Department of instrumentation and control engineering, Technická 4, 166 07 Praha 6, Czech republic*

Abstract

Paper presents our system for the fault detection and identification (FDI) and a comprehensive overview of the changes and improvements made by author. FDI system is based on a Markov model. The main presented enhancements include the introduction of a dynamic model for transients between failures and application of empirical modal decomposition algorithm to observed data in order to find closer relations to the individual faults.

Key words: FDI; fault diagnostics; Markov model; Markov chain; empirical mode decomposition, dynamic system

Introduction

Modern trends in the industry are directed towards increasing automation and thus the autonomy of technological processes. The aim is to reduce potentially hazardous human factor and to move the operator role from the executive towards the decision field. Larger autonomy requires a more advanced system for fault detection, isolation and identification. It is necessary to ensure the safety, security and environmental protection. Early detection, identification and eventual removal of the fault can help to avoid a complete breakdown of the system and thus to the irreparable damage or to loss of human life.

The role of fault diagnosis is generally to detect as quickly as possible each deviation from the regular system behavior, while it should minimize the number of wrong decisions. Once the fault is detected, the actions aimed to minimize its consequences usually follow. Therefore, it is important to properly locate and describe the fault discovered. The monitoring system used to detect faults, to locate them and to determine their severity is called the Fault Diagnosis System. The task of fault diagnosis consists of detection, isolation and identification. Fault detection is based on simple decision, if the system is in fault condition. This does not necessarily mean a complete collapse of the system, but generally speaking, any sufficiently significant deviation from the desired function. Fault isolation is task to accurately locate the source(s) of failure, e.g. a faulty sensor, a structural change in the system, improper human intervention to the controls, etc. Fault identification aims to specify the kind and extent of the fault.

Markov chains are special stochastic models, which are characterized by the so-called Markov property. Markov property says that the probability of the current state of some process does not depend on its entire previous history but only on its immediately preceding state. The main advantage of Markov chain models is the possibility to work with a strongly nonlinear systems and relatively easy identification. The disadvantage, especially when working in real time, is a large volume of data processed and also large dimensions of transition matrix. Also a discrete nature of the Markov model may be a slight constriction.

Hilbert-Huang transformation (HHT) belongs among time-frequency signal analysis methods. It consists in the decomposition of sample signal into components using empirical mode decomposition (EMD) and the subsequent application of the Hilbert transform (HT) in order to find the instant frequencies and instant magnitudes of the oscillating signal. For a reasonable usage of HT, the signal spectrum must be as narrow as possible and the mean value of signal must be close to zero. EMD decomposes signal into components (intrinsic modal Intrinsic functions - IMF) that satisfies above conditions very well. IMFs represent modes of signal oscillation and those satisfy above conditions very well. Especially for Markovian FDI system, it is also important that the IMF can be used to design a good regression vector.

1. Markov fault diagnosis system

Markov FDI model is a special case of Markov model. The regression vector (RV) consists of input and output quantities of monitored system. State of monitored system predicted by Markov FDI is understood as its fault condition, i.e. it estimates, which fault state actually appeared. (From this perspective, the fault-free state or nominal behavior is only one of known fault states.)

For the identification of the Markov model it is required to have a sufficiently large amount of data for each expected failure as well as for the fault-free state of the monitored system. This dataset is used to create the statistics of dependencies between the instances of RV and the appropriate fault condition of the monitored system. The structure of the regression vector is crucial for the quality of the FDI.

Output from our Markov FDI system is probability distribution of known fault states. Then the most probable fault state is easy to estimate. E.g. the nominal behavior is represented by number 0, first fault state by 1 etc. It is also possible to distinguish an unknown new state. Typi-

[*] Contact to author: Pavel.Trnka@fs.cvut.cz

cally (with well designated RV structure), one fault state is significantly more probable than all the others. If the fault condition probability distribution is flat (i.e. actual probabilities of all known fault states are almost same), it seems to be an unknown fault state.
Markov FDI system works in two stages:

• ***Learning stage.*** At this stage, the Markov model with the chosen structure of RV generates statistics based on training dataset obtained from monitored system. This dataset must be appropriately prepared; it must cover all possible fault (and fault-free) states that might occur. During this stage the system must have available an additional information about the current fault state. The result is a transition matrix between known values (instances) of regression vector and individual fault states (fault condition of system).

• ***Diagnostic stage.*** During this stage the FDI system has no information about actual fault condition of monitored system but just the opposite FDI must estimate it. The monitored quantities are continuously transformed into the RV instance, which is compared to known RVs with trained statistics. Then it outputs the actual probability distribution of the known fault states. Then the most probable fault is selected and then passed on for further use.

Both stages can alternate. When a new unknown fault occurs (fault-free behavior of the system does not respond, nor any of the known fault states), the learning can be applied retroactively using continuously recorded behavior of monitored system and the new fault can be added to the statistics.

2. Dimensionality reduction

The biggest problem of Markov model based methods is a huge dimensionality of the statistics or the transition matrix. In our system, the method of Approximation of Markov chains based prediction (AMCP) (see [1]) is used to reduce size of transition matrix. If the new instance of regression vector (RV) is obtained (i.e. unknown combination of values of measured quantities), then the similar known RV instances (neighborhoods) are found, then their relevancy (weight) is computed based on their distance from actual RV and finally the statistics for new RV is computed from weighted statistics of neighborhoods. Due this method, much less amount of RV must be stored and thus the transition matrix is significantly smaller while maintaining the same quality of FDI.

3. Dynamics of state transitions

The recognition of faults with fast dynamics can be fairly faster, if we use a somewhat different categorization of fault states than that described above.
Data series used as source for generating of statistics in the Markov FDI model must be relatively long. It primarily contains steady data when the system persists in some fault state. Then individual parameters of the system remain in a relatively narrow range of its values with only small variations. However, the data series includes also very significant, although relatively very short transitions between steady states. Therefore the dataset assigned to the fault state is quite dynamically unbalanced, because in addition to the relatively small amount of data gathered during transition occurs a large amount of data gathered in a steady state or in its vicinity.
The dynamic unbalance of the source data while learning introduces a bias into the statistics and thus the entire Markov FDI model. However, the shape and dynamics of the transition are very important for the correct FDI. Transients are often only fault indicators, because after settling the RV returns almost to the original state.
Redefined set of fault states serves as a solution to the above problem (see [2]).

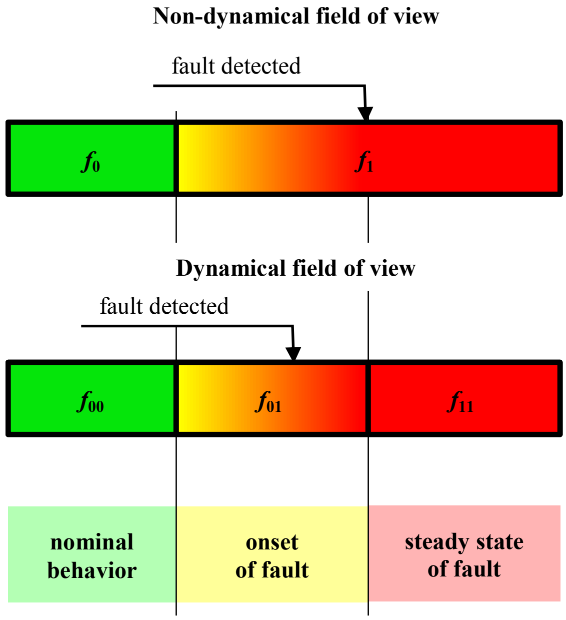

Fig. 1. *The dynamic nature of the transitions between faults.*

This change preserves the shape of the regression vector and the range and resolution of the measured quantities, but it changed the categorization of fault states. Each fault state is divided into a steady part and one or more transient parts. The steady part replaces the original fault state and each of the transient parts is associated to a new fault state. From the viewpoint of the Markov FDI model the set of fault states and thus the statistics and the transition matrix changes, but the algorithm itself stays unchanged.
Consider that individual faults does not combine each other because the failure always occurs after long interval of nominal behavior and the system returns back to the nominal behavior after removal of the fault source. Then each fault state disintegrates into exactly two fault sub-states (one “steady” sub-state and one “transition from failure-free” sub-state) and only the failure-free state breaks up into $n + 1$ sub-states, where n is the number of transients from individual faults back into the nominal behavior.
The taking into account of the dynamic nature of the monitored system is especially useful in case of rapid start of fault. Experiments show that the time required to detecting the fault shortened by up to half.

4. Enhanced logic module

In order to take into account the relation of steady and transient sub-states to the common state, the diagnostic system has been improved by the enhanced logic module (ELM). In fact, the ELM is a simple rule-based system that is able to solve primarily two tasks:

• ***Pooling of states.*** It enables to cover multiple sub-states by one common state. In case of realization of the dynamic approach of FDI this mechanism enables to link the transient and steady fault sub-states, so that we get back only one general fault state for each alarm. However, because the link between sub-states is only qualitative but not quantitative, the importance transient sub-state stays high.

• ***Chaining of states.*** Known information about dynamic system properties enables usually to derive simple rules which can significantly reduce the amount of possible transitions between its states. This applies for the set of fault states (and sub-states), too. E.g. consider steady states A,B,C and transition states A→B,B→A. Then the sequence of states A, A→B, B or A, A→B, B→A can occur with high probability, while the sequence A, A→B, C is almost impossible. The set of rules that defines the transient probabilities between known fault sub-states is not used directly to estimate the current fault state but it gives rule to decide in case of ambiguous outcome from the Markov model.

4. EMD based regression vector

As stated above, the structure of the regression vector has a major impact to the quality of fault detection and identification. Interesting possibilities offers a decomposition of measured quantities using empirical mode decomposition algorithm (EMD), if we build a regression vector only from the obtained intrinsic modal functions (IMF). Each IMF represents some signal oscillation mode which is associated to some component of dynamic behavior of the monitored signal. This can be used to obtain a more detailed representation of the dynamics of the system as well as to create a RV structure which reflects better the differences between the various fault states.

The EMD is carried out in recursive sifting algorithm in which the spurious trends are gradually removed from the signal until the residue satisfies the IMF definition with sufficient precision (i.e. envelopes of local extremes are symmetrical around zero mean value). Then this IMF is separated from the original signal and a next round of sifting algorithm applies on the residual signal. Sifting algorithm is repeated until all possible IMFs are found. Final residue that is not possible to decompose represents the overall trend of the original signal within the entire dataset.

In order to take advantage of EMD in fault diagnosis, it has been necessary to select its online variant capable of the decomposition continuously in real time. Problematics of online EMD deal articles [3], [4].

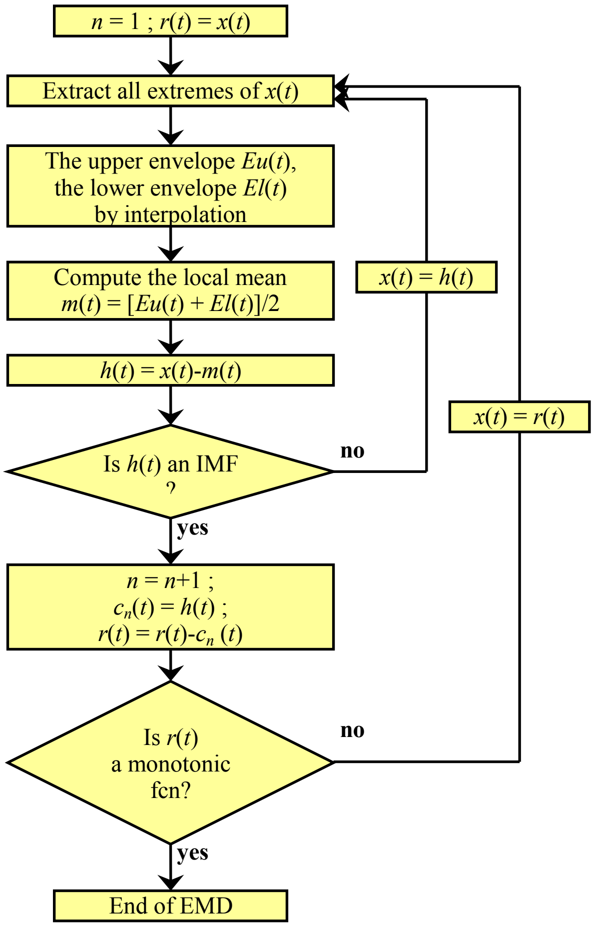

Fig. 2. *The principle of the sifting algorithm of EMD.*

5. Conclusion

The Markov Fault Detection and Identification system is based on the Markov model with the transition matrix dimensionality reduction algorithm. The empirical mode decomposition has been applied to the observed data in order to obtain more detailed regression vector structure. The model of the diagnostic system has been enhanced by the dynamical field of view on the data obtained from a monitored system and the dynamical character of transitions between its fault states.

Also the enhanced logic module has been included into FDI system thus it is possible to use extra information about dependencies between fault sub-states or about some properties of monitored system.

Acknowledgement

This work was supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS16/210/OHK2/3T/12.

Abbreviations

AMCP	Approximation of MC-based Prediction
ELM	Enhanced Logic Module
EMD	Empirical Mode Decomposition
FDI	Fault Detection and Identification
HT	Hilbert Transform
HHT	Hilbert-Huang Transform
IMF	Intrinsic Modal Function
MC	Markov Chain
RV	Regression Vector

Sources

[1] GARAJAYEWA, Guncha: *Bayesian approach to real-time fault detection and isolation with supervised training, PhD. dissertation work.* Praha: ČVUT v Praze, 2005, vedoucí disertační práce prof. Ing. Milan Hofreiter, CSc.

[2] HOFREITER, Milan - TRNKA, Pavel: *Fault Diagnosis for Nonlinear Stochastic Dynamic Systems.* In: Proceedings of The 9th International Conference Process Control 2010, Pardubice: Universita Pardubice, s. C009b - 1-9. ISBN 978-80-7399-951-3.

[3] TRNKA, Pavel - HOFREITER, Milan: *The Empirical Mode Decomposition in Real-Time.* In: Proceedings of the 18th International Conference on Process Control. Bratislava: Slovenská technická univerzita, 2011, s. 284-289. ISBN 978-80-227-3517-9.

[4] TRNKA, Pavel - HOFREITER, Milan: *Empirická modální dekompozice environmentálních časových řad v reálném čase.* In: Automatizácia a riadenie v teórii a praxi 2011. Košice: Technická univerzita v Košiciach, 2011, s. 52-1-9. ISBN 978-80-553-0606-3.