Fault detection in high redundancy actuation using an interacting multiple-model approach

This item was submitted to Loughborough University's Institutional Repository by the/an author.


Additional Information:

- This is a conference paper, it was presented at SafeProcess 2009 [© IFAC]. The definitive version is available from: http://www.ifac-papersonline.net/

Metadata Record: https://dspace.lboro.ac.uk/2134/4979

Version: Accepted for publication

Publisher: © IFAC

Please cite the published version.
This item was submitted to Loughborough’s Institutional Repository (https://dspace.lboro.ac.uk/) by the author and is made available under the following Creative Commons Licence conditions.

For the full text of this licence, please go to: http://creativecommons.org/licenses/by-nc-nd/2.5/
Fault Detection in High Redundancy
Actuation using an Interacting
Multiple-Model Approach

J. Davies∗H. Tsunashima**R.M. Goodall∗†R. Dixon∗T. Steffen∗

**Department of Mechanical Engineering, College of Industrial Technology, Nihon University, 1-2-1 Izumi-cho, Narashino-shi, Chiba 275-8575, Japan

Abstract: The High Redundancy Actuation (HRA) project investigates the use of a relatively high number of small actuation elements, assembled in series and parallel in order to form a single actuator which has intrinsic fault tolerance. Both passive and active methods of fault tolerant control are being considered for use with the HRA. In either approach, some form of health monitoring is required to indicate the requirement for reconfiguration in the latter case and the need for maintenance in the former. This paper presents a method of detecting faults in a HRA using an Interacting multiple-model (IMM) algorithm.

Keywords: Fault detection, Fault diagnosis, Kalman filters, Multiple model

1. INTRODUCTION

1.1 Traditional Approaches to Fault Tolerant Actuation

In automated processes, faults in hardware or software will often produce undesired reactions. These faults can result in failures, where the system as a whole does not complete an expected action, possibly causing damage to the plant, its environment, or people in the vicinity of that plant [Blanke et al., 2001]. Fault Tolerant Control (FTC) aims to prevent failures and their consequences by providing adequate system performance in the presence of faults.

The majority of FTC research to date has concentrated on sensor faults. Significant advances have been made in this area, but most of these strategies are not applicable to actuator faults. This is attributable to the fundamental differences between actuators and sensors. Sensors deal with information, and measurements may be processed or replicated analytically to provide fault tolerance. Actuators, however, must deal with energy conversion, and as a result actuator redundancy is essential to keep the system in control and bring it to the desired state in the presence of actuator faults [Patton, 1991].

The common solution for fault tolerant actuation in critical systems involves straightforward parallel replication of actuators. Each redundant actuator must be capable of performing the task alone and possibly override the other faulty actuators. This solution is over-engineered, reducing the efficiency of the system i.e. in triplex systems 200% more capability, cost and weight than required is introduced to ensure a certain level of reliability.

1.2 High Redundancy Actuation

High Redundancy Actuation (HRA) is a novel, state-of-the-art approach to actuator fault tolerance that aims to reduce the over-engineering incurred by traditional approaches. The HRA concept is inspired by musculature, where the tissue is composed of many individual cells, each of which provides a minute contribution to the overall contraction of the muscle. These characteristics allows the muscle, as a whole, to be highly resilient to individual cell damage.

This principle of co-operation in large numbers of low capability modules can be used in fault tolerant actuation to provide intrinsic fault tolerance. The HRA uses a high number of small actuator elements, assembled in parallel and series to form one high redundancy actuator (see Figure 1). Faults in elements will affect the maximum capability, but through control techniques, the required performance can be maintained. This concept allows the same level of reliability to be attained in exchange for less over-dimensioning.

Figure 1. High Redundancy Actuation.

The HRA is an important new approach within the overall area of fault tolerant control. When applicable, it can provide actuators that operate at the desired level of performance in the presence of multiple faults in actuator elements, and gracefully degrade after the designed level of fault tolerance has been exceeded.

1.3 High Redundancy Actuation and Fault Detection

The project thus far has investigated two methods of controlling the HRA: robust control (passive fault tolerance) and reconfigured control (active fault tolerance) [Dixon et al., 2009, Davies et al., 2008a, Steffen et al., 2008]. Both of these approaches, to
different extents, require some form of fault detection (FD). In the latter case, a clear indication of the HRA’s remaining capability, and thus it’s fault state is required in order to reconfigure the control laws appropriately. In passive control, the controller is static and thus not reliant on the fault state. However, health monitoring of the system is still required to indicate to a user the remaining capability of the HRA or indicate requirement for maintenance if fault levels approach the performance limits.

1.4 Overview

This paper presents an approach to fault detection for a HRA using an Interacting Multiple-Model (IMM) methods. Section 2 describes the modelling of HRA that uses electromagnetic actuation technology. The IMM algorithm is outlined in Section 3, and simulation results of its application to parallel, serial and mixed configuration elements are discussed in Section 4. Finally, conclusions are made in Section 5 and future work is considered.

2. HRA MODELLING

This paper assumes that the underlying technology of the HRA is electromagnetic, moving coil actuation, which is similar to a voice-coil in operation. Many other technologies are possible and indeed, the next stage of the project aims to address which technology will be best suited to manufacturing HRAs with large numbers of elements. However, many technologies will lead to a model with a similar structure to that presented here.

2.1 Single Element

The full order modelling of a moving coil actuator and of HRA configurations using these actuation elements is presented in [Davies et al., 2008b]. However, this paper will utilise a simplified version of this model.

A moving coil actuator typically comprises a coil wound round the centre pole of a magnetic assembly that produces a uniform magnetic field perpendicular to the current conducted in the coil. On providing a voltage, a current flows in the coil (inversely proportional to the input resistance \( R_{in} \)), which generates a force known as the Lorentz force:

\[
F = BNLI = kI = \frac{k}{R_{in}}v
\]

Where \( B \) is the magnetic flux density, \( N \) is the number of turns and \( I \) is the length of the conductor. These are all constant and thus may be combined to form a single force constant, \( k \). This force causes the coil, and the rod which is mounted to it, to move. The movement of the coil in the field generates a counter-electromotive force which can be expressed as below:

\[
E = BNlv = kv
\]

Where the derivative \( \dot{x} \) is the perpendicular component of the velocity of the wire relative to the flux lines i.e. the velocity of the coil.

The force produced by the electrical/magnetic part of the system acts upon the mechanical part which consists of the moving mass of the element and any stiffness and damping. Hence, using Newton’s second law of motion, the following second order model for the actuator can be derived:

\[
m\ddot{x} = \frac{k}{R_{in}}v - \frac{k^2 + R_{ad}d}{R_m}x - rx, \tag{3}
\]

where \( v \) is the input voltage, \( m \) is the moving mass, \( d \) is the damping factor, and \( r \) is the stiffness. Choosing \( x \) and \( \dot{x} \) as states leads to the following state space model:

\[
\begin{bmatrix}
\dot{x} \\
\ddot{x}
\end{bmatrix}
= \begin{bmatrix}
\Sigma_{i=1}^{n} K_i & -\sum_{i=1}^{n} r_i \\
m & 0
\end{bmatrix}
\begin{bmatrix}
x \\
\dot{x}
\end{bmatrix}
+ \begin{bmatrix}
\Sigma_{i=1}^{n} K_i v(n) \\
m
\end{bmatrix}
\tag{4}
\]

2.2 Parallel Elements

When elements are arranged in parallel (Figure 2), their forces act on a combined mass. Hence, assuming a common input voltage, the model for \( n \) parallel elements is:

\[
\begin{bmatrix}
\dot{x} \\
\ddot{x}
\end{bmatrix}
= \begin{bmatrix}
\Sigma_{i=1}^{n} K_i & -\Sigma_{i=1}^{n} r_i \\
m & 0
\end{bmatrix}
\begin{bmatrix}
x \\
\dot{x}
\end{bmatrix}
+ \begin{bmatrix}
\Sigma_{i=1}^{n} K_i v(n) \\
m
\end{bmatrix}
\tag{5}
\]

where:

\[
K_i = \frac{k^2 + R_{ad(i)}d_i}{R_{m(i)}} \text{ and } Kin_i = \frac{k_i}{R_{m(i)}}
\]

2.3 Serial Elements

If a number of elements \( n \) are arranged serially (Figure 3) then the system contains \( n \) moving masses. Forces produced by the elements act not only on their moving mass, but also counter-act upon the preceding moving mass. The model of three elements in series is then:

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2 \\
\dot{x}_3
\end{bmatrix}
= \begin{bmatrix}
a_{13} & a_{14} & a_{15} & a_{16} & 0 & 0 \\
a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\
0 & 0 & a_{31} & a_{32} & a_{33} & a_{34} \\
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3
\end{bmatrix}
+ \begin{bmatrix}
b_1 \\
b_2 \\
b_3
\end{bmatrix}
\tag{6}
\]

where:
\[
a_{11} = \frac{K_{i-1}}{m_i}, \quad a_{12} = \frac{r_{i-1}}{m_i}, \quad a_{13} = -\frac{K_i + K_{i+1}}{m_i}, \quad a_{14} = -\frac{r_i + r_{i+1}}{m_i}
\]
\[
a_{i1} = \frac{K_{i+1}}{m_i}, \quad a_{i2} = \frac{r_{i+1}}{m_i}, \quad b_i = \frac{K_i n_j - K_{i+1} n_j}{m_i}
\]

Models of higher numbers of serial elements follow this model’s structure. Also, models of mixed configuration arrangements which are necessary for creating HRAs can be constructed using these basic equations.

2.4 Element faults

Three fault types are considered within this paper; overheating, loose faults and lock-up faults. Overheating of an actuation element may be represented as an increase in the resistance i.e. an increase in \(R_m\).

A loose fault is where the actuation element loses the ability to translate force between its end points. Hence, a loose fault in a parallel assembly will reduce the force exerted on the mass. In serially connected elements, this fault is terminal as it effectively fails the whole serial branch. Thus, it is only useful to consider loose faults where elements are arranged in parallel.

A lock-up fault is where an element loses the ability to change the length between its two end points. This may occur if the coil of an actuation element is deformed and touches the magnet. This fixes the mass with respect to the reference point, and consequently the relative position and the speed are constant. In serially connected elements, this adds the mass of the locked element to the preceding element, and removes the mechanical states of that element from the system model. In parallel arrangements, this fault locks the whole assembly from end-to-end. Therefore, this fault type will only be considered where there are serial elements.

3. INTERACTING MULTIPLE-MODEL APPROACH

Conventional multiple-model estimation methods use a bank of filters, each of which is based on a model of the system when it is in a particular mode. The outputs of these filters are combined with a probabilistically weighted sum to achieve an overall state estimate.

However, there is no interaction between the filters, and as such the approach is not suited to situations where the parameters or structure of the system changes [Zhang and Jiang, 2001]. Nonetheless, non-interacting methods of multiple-model estimation have been applied to FD applications, where sudden parameter and structural changes to the system occur using ad hoc solutions [Menke and Maybeck, 1995, Napolitano and Swaim, 1991].

The Interacting multiple-model (IMM) method, developed in the field of tracking [Blom and Bar-Shalom, 1988, Bar-Shalom et al., 2001] deals with these issues. In the IMM approach, the initial estimate at the beginning of each iteration is a mixture of recent estimates from the filters. As a result the accuracy of estimation is increased and dependency on the previous mode history is introduced. This increases its suitability to detecting faults and thus it has been applied within this field [Mehra et al., 1998, Zhang and Jiang, 2001, Hayashi et al., 2006, Hashimoto et al., 2007, Hayashi et al., 2008].
white Gaussian with zero mean. The Kalman filter algorithms can then be expressed as:

\[
\hat{x}^j_{(t|t-1)} = F^j(\hat{x}^j_{(t-1)}) + D^j u_{(t-1)} \\
\hat{x}^j_{(t|0)} = \hat{x}^j_{(t-1)} + K^j \left[ y - (H^j(\hat{x}^j_{(t-1)}) + L^j u_{(t-1)}) \right]
\]

(14) \hspace{2cm} (15)

\[
K^j = P^j_{(t|t-1)} H^j T + R^j_{(t-1)}^{-1}
\]

(16) \hspace{2cm} (17)

\[
p^j_{(t|t-1)} = F^j(\hat{x}^j_{(t-1)}) + E^j T + G^j \left[ \lambda^j_{(t-1)} Q^j_{(t-1)} G^j T \right]
\]

(18) \hspace{2cm} (19)

Mode Probability Calculation The mode probability, \( p_{j(t)} \) (for mode \( j \) at time \( t \)) is then updated based on the likelihood function \( \Lambda \) for each mode filter:

\[
p_{j(t)} = \frac{\Lambda_{j(t)} e^j}{\sum_{i=1}^{m} \Lambda_{i(t)} e^i}
\]

(20)

\[
\Lambda_{j(t)} = \left| 2 \pi \sigma_{j(t)} \right|^{-\frac{1}{2}} \exp \left[ -\frac{1}{2} \left( \hat{x}_j(t) - \left( \hat{x}_j_{(t-1)} + L^j u_{(t)}) \right) \right)^T \sigma_{j(t)}^{-1} \left( \hat{x}_j(t) - \left( \hat{x}_j_{(t-1)} + L^j u_{(t)}) \right) \right]
\]

(21) \hspace{2cm} (22)

The mode probabilities give a time-varying estimate on the likelihood of the system state being one of the model-based modes and thus are used in the indication of fault type for FD applications. The probabilities are smoothed using a moving average window.

Combination of Estimates Finally, the combined state estimate \( \hat{x}^j_{(t)} \) and covariance \( P_{(t)} \) are derived by weighting the estimated state and the mixed covariance for each mode with the mode probabilities:

\[
\hat{x}^j_{(t)} = \sum_{j=1}^{m} \hat{x}^j_{(t)} \rho_{j(t)}
\]

(23) \hspace{2cm} (24)

\[
P_{(t)} = \sum_{j=1}^{m} \rho_{j(t)} \left[ \rho_{j(t)} \left[ \hat{x}^j_{(t)} - \hat{x}^j_{(t-1)} \right] \left[ \hat{x}^j_{(t)} - \hat{x}^j_{(t-1)} \right]^T \right]
\]

4. SIMULATION EXAMPLES

This section discusses the simulation results of the IMM approach when applied to first purely parallel and serial configurations to illustrate that overheating, loose and lock-up faults can be diagnosed in these structures. The diagnosis of faults in a mixed configuration of parallel and serial elements will then be considered briefly.

In each case the simulation is set-up as shown in Figure 5. The elements receive a shared input from a classical controller, designed for good transient characteristics and frequency margins from voltage input to load position. The system is given a sine wave input reference with an amplitude that uses its full range of travel (±15mm). The known input and measured output is passed to the IMM algorithm which produces mode probabilities and a mixed state estimate.

Figure 5. IMM Simulation

Figure 6. Mode probabilities for parallel elements

4.1 Parallel Elements

IMM FD is applied to three parallel elements here. The IMM estimator is designed based on the following modes:

- Mode 1: Nominal system
- Mode 2: Overheating, a resistance increase of 20%
- Mode 3: Overheating, a resistance increase of 50%
- Mode 4: 1 Loose element
- Mode 5: 2 Loose elements

The transition matrix \( p_{ij} \) is set such that the probability of no transition from the current state is 0.999 and \( 2.5 \times 10^{-5} \) for transitions to the other modes.

The measured output is the position of the load. A very small value of covariance is used for the noise on the measured position (\( 5 \times 10^{-12} \) m), as the glass encoder used has an rms noise value of \( 1 \mu \)m. The plant noise covariance \( Q \) is set at \( 1 \times 10^{-5} \) V, as this gives a noise level in the order of mV.

Simulation Results The simulation results shown in Figure 6 are the mode probabilities produced from the IMM algorithm for the parallel elements system with changing fault state. At \( t = 0 \), the system is nominal, after which at 5s intervals the system fault state is changed from mode 2 through to mode 5.

It can be seen that during each fault state, the correct mode is diagnosed with a high probability after approximately 0.5s except mode 4. The probability of mode 4 takes longer to rise due to its similarity to the nominal state.

In a realistically scaled HRA, the levels of parallel redundancy (used in conjunction with serial redundancy) will be higher e.g. 10 or more parallel elements. A greater similarity between the nominal system and small proportions of loose faults will exist. This may make the clear diagnosis of low numbers of loose faults more difficult. However, if the behaviour of a HRA with a very low proportion of loose elements is sufficiently near the nominal behaviour, then detection of these faults at this fault level is not crucial, as the health status of the HRA will be high and control reconfiguration will not be necessary.
4.2 Serial Elements

Three serial elements with overheating and lock-up faults are used in this example. As the location of a lock-up fault in the system will result in slightly different fault models, more modes are needed to diagnose lock-ups within serial elements. Therefore, the IMM estimator in this instance uses the following nine modes:

- Mode 1: Nominal system
- Mode 2: Overheating, a resistance increase of 20%
- Mode 3: Overheating, a resistance increase of 50%
- Mode 4: Element 1 lock-up
- Mode 5: Element 2 lock-up
- Mode 6: Element 3 lock-up
- Mode 7: Elements 1 and 2 lock-up
- Mode 8: Elements 1 and 3 lock-up
- Mode 9: Elements 2 and 3 lock-up

The transition matrix \( P \) is set such that the probability of no transition from the current state is 0.999 and \( 1.25 \times 10^{-5} \) for transitions to the other modes. Relative positions were used as the measured quantities in the simulation. Relative measurements were chosen over absolute as the HRA rig in development will have position encoders on each element. The noise covariance for each sensor is the same as that used in the parallel example. The plant noise covariance is also the same as that used in the parallel case.

Estimation in the presence of actuator lock-ups presents a special issue, particularly when actuators are arranged in series. The fault model for a serial assembly of \( n \) actuation elements with one locked element will effectively be a model for \( n-1 \) elements with one mass augmented with the locked element's mass (if it is not the ground connected mass). When actuators lock at a non-zero point along their travel, this unknown position is not included in the fault model and thus position estimation and correct mode identification (without velocity information) becomes difficult.

One solution to this problem is to include in the fault model a high damping factor in the faulty element's dynamics. This will incorporate the locked position into the estimation resulting in a more accurate overall estimation and accurate mode identification. This approach is used within this simulation.

Simulation Results The resulting mode probabilities for an example fault profile simulation are shown in Figure 7. The mode probabilities in the example are typical of all fault profiles. The correct mode is clearly indicated in each time period. More fluctuation of the mode probabilities is present during nominal conditions and overheating in comparison with the parallel element results. This may be explained by the increased number of sensors in the system. In this case three sensors are used, each measuring a smaller quantity than the one measurement in the parallel case. However, the same noise covariance is present on each sensor. Thus there is more noise present in the system. These fluctuations are less prevalent when actuation elements lock, as this fault mode is more removed from the nominal behaviour.

4.3 Parallel in Series HRA

Having illustrated that it is possible to diagnose overheating, loose faults and lock-up faults in purely parallel or serial arrangements of elements, the application of IMM FD to a system that contains both parallel and serial elements is briefly considered. A 3 × 3 Parallel in Series (PS) system (Figure 8) is used as an example.

This configuration has relatively high intrinsic tolerance to loose faults. Loose faults in the parallel branches will have little effect on the system until there are loose faults in every branch i.e. one loose fault in every parallel branch is equivalent to one loose branch in a purely parallel system. Hence, at least 2 loose faults (but at maximum 4 if they are divided equally between two branches) can occur before a reduction in force capability is observed.

The system has less tolerance to lock-up faults, however. A locked element will lock a whole parallel branch, reducing its travel capability by a third and thus the same fault tolerance is achieved as in a purely serial arrangement.

Many more mode filters are required to cover all the possible fault combinations within this system. As before, 3 modes are required to diagnose nominal conditions and two levels of overheating; 2 for diagnosing 2 reductions in force capability (i.e. loose faults within the system); and 6 modes for diagnosing travel capability reductions (lock-up faults). However, if we were to consider occasions where both force and travel capabilities are reduced, as would be necessary for a HRA, then the required number of modes rises to 23. Considering that this is a quite low level of redundancy, then this number is high. In a higher order, more realistically dimensioned HRA containing, for example 10 × 10 elements, then using the current approach to mode allocation, 2286 modes would be needed to diagnose all the fault combinations for up to 50% reduction in force and travel capabilities.

Simulation Results The higher number of modes in this example does not affect the diagnosis quality. An example simulation for the 3 × 3 PS system is shown in Figure 9. The system is nominal for the first 5s period, followed by a lock-up in element 1 at t=5. Loose faults resulting in a 1/3 loss of force capability are injected at t=10, and element 2 locks at t=15. Finally, another loose fault in the remaining unlocked branch occurs at t=20. In each case, the correct mode is diagnosed with short detection delays. The state in 15s-20s is more difficult
to diagnose as it is similar to no loose faults and 1/3 force capability, but the correct mode is still clearly indicated. The higher number of modes, however, does affect the required simulation run-time for the IMM algorithm.

5. CONCLUSIONS

This paper discussed the utilisation of IMM techniques to achieve fault detection in a HRA. Simulation results of the IMM method applied to parallel elements, serial elements and a small mixed configuration HRA were presented. These results suggested that, using a comprehensive set of mode filters, it is possible to detect overheating faults; location independent loose element faults; and location specific lock-up faults. However, the required number of modes for detection in low level redundancy HRA is relatively large, and for more realistically dimensioned HRA (100+ elements) the required number becomes much greater. This may make real-time diagnosis unfeasible.

However, in HRA applications, the location of the locked element is not of interest. Only the actuator’s remaining travel or force capability is required to give an indication of health, or reconfigure global control laws. Hence, if a simplified model of the system is used with the IMM algorithm, where each mode filter represents a level of capability, the number of required modes would be limited dramatically. This approach will be the focus of the next stage of work on this specific area of the HRA project’s research. Also, the development of a 4 × 4 experimental rig for the HRA is underway and application of these fault detection techniques to further assess their feasibility is planned.

This approach to fault detection is by no means the only one that can be taken to meet the requirements of this application. Indeed, the project aims to examine other fault detection and health monitoring methods in the future. A comparison may then be made between the fault detection types to further assess the effectiveness and feasibility of using this IMM approach with the HRA.

ACKNOWLEDGEMENTS

This project is a cooperation of the Control Systems group at Loughborough University, the Systems Engineering and Innovation Centre (SEIC), and the actuator supplier SMAC UK Ltd.. The project is funded by the UK’s Engineering and Physical Sciences Research Council (EPSRC) under reference EP/D078350/1.

REFERENCES


