Automatica 97 (2018) 143–149

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Residual selection for fault detection and isolation using convex optimization☆

Daniel Jung[a,b,*], Erik Frisk[a]

[a] Linköping University, Linköping, Sweden
[b] The Ohio State University, Columbus, OH, USA

ARTICLE INFO

Article history:
Received 19 July 2017
Received in revised form 10 April 2018
Accepted 25 June 2018

Keywords:
Fault detection and isolation
Feature selection
Model-based diagnosis
Convex optimization
Computer-aided design tools

ABSTRACT

In model-based diagnosis there are often more candidate residual generators than what is needed and residual selection is therefore an important step in the design of model-based diagnosis systems. The availability of computer-aided tools for automatic generation of residual generators have made it easier to generate a large set of candidate residual generators for fault detection and isolation. Fault detection performance varies significantly between different candidates due to the impact of model uncertainties and measurement noise. Thus, to achieve satisfactory fault detection and isolation performance, these factors must be taken into consideration when formulating the residual selection problem. Here, a convex optimization problem is formulated as a residual selection approach, utilizing both structural information about the different residuals and training data from different fault scenarios. The optimal solution corresponds to a minimal set of residual generators with guaranteed performance. Measurement data and residual generators from an internal combustion engine test-bed is used as a case study to illustrate the usefulness of the proposed method.

1. Introduction

A model-based diagnosis system is typically based on a set of residual generators, sometimes referred to as monitors, to detect if faults have occurred or not (Blanke, Kinnaert, Lunze, Staroswiecki, & Schröder, 2006). Each residual generator is designed to monitor a specific part of the system and then, based on which residuals that trigger, a set of diagnosis candidates (fault hypotheses) can be computed (Cordier et al., 2004).

There are two main motivational observations for this work. First, the number of possible residual generator candidates in general grows exponentially with the degree of redundancy of the model (Krysander, Åslund, & Nyberg, 2008). This means that in many cases there are significantly more candidates possible than what is needed to detect and isolate the faults. A second observation is that in realistic scenarios all candidate residual generators do not perform equally well, mainly due to the inherent uncertainties in the model and measurement noise. Fig. 1 shows a typical situation with a set of residuals that are all sensitive to the same fault. In an ideal case, all residuals in the plot should react in the gray regions, but clearly the detection performance varies and some has no clear reaction at all, making them less useful for this particular fault. Thus, selecting an appropriate subset of residual generators is a key step in the design process to ensure that satisfactory detection and isolation performance can be achieved at low computational cost.

Even though residual selection is important to achieve satisfactory fault detection and isolation performance, it has received relatively little attention compared to other steps in the model-based diagnosis system design, e.g., sensor selection (Bhushan & Rengaswamy, 2000; Krysander & Frisk, 2008; Nejjari, Sarrate, & Rosich, 2010) and residual generator design (Basseville, 1997; Frank & Ding, 1997; Venkatasubramanian, Rengaswamy, Yin, & Kavuri, 2003). In previous works, for example Nejjari et al. (2010), Perelman, Abbas, Koutsoukos, and Amin (2016) and Svärd, Nyberg, and Frisk (2013), the residual generators are assumed ideal when formulating the residual selection problem. Residual selection by optimization has been proposed in Nejjari et al. (2010) using a Binary Integer Linear Programming approach, in Svärd et al. (2013) using a greedy heuristic, and in adaptive on-line solutions in Chanthery, Travé-Massuyès, and Indra (2016); Krysander, Heintz, Roll, and Frisk (2010), also here assuming ideal performance. A main limitation with these methods is that quantitative residual performance is not taken into consideration in the residual selection process, i.e., assuming that the detection performance of all residuals in Fig. 1 are equal which is clearly not the case.

☆ The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Angelo Alessandri under the direction of Editor Thomas Parisini.

* Corresponding author at: Linköping University, Linköping, Sweden.
E-mail addresses: daniel.jung@liu.se (D. Jung), erik.frisk@liu.se (E. Frisk).

https://doi.org/10.1016/j.automatica.2018.08.006

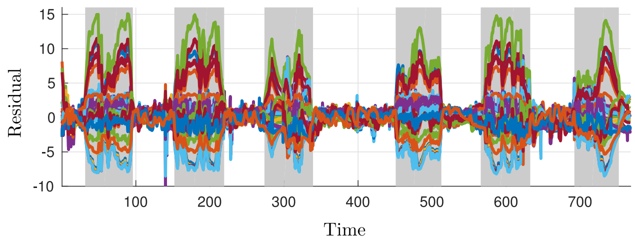

Fig. 1. A comparison of residuals sensitive to the same fault but with different detection performance. The gray-shaded intervals indicate where the fault is active.

The main property to consider in the selection process is robustness in the detector with respect to model uncertainties and noise. One approach would be to model noise and model uncertainty using, e.g., probabilistic methods, see for example Eriksson, Frisk, and Krysander (2013) and Wheeler (2011). However, in general this is difficult unless uncertainties are well modeled by stationary random processes. The approach adopted here is to let measured data model the uncertainties and the effects of different faults.

Residual selection is closely related to the feature selection problem in machine learning (Chandrashekar & Sahin, 2014; Friedman, Hastie, & Tibshirani, 2009; Guyon & Elisseeff, 2003). Different feature selection algorithms for data-driven fault diagnosis have been proposed, for example Jegadeeshwaran and Sugumaran (2015) and Jiang, Yan, and Huang (2016). Performance of feature selection algorithms depends on the quality of available training data (Tidriri, Chatti, Verron, & Tiplica, 2016). Collecting representative data from different faults is time-consuming, costly, and often infeasible since it is not known exactly how different faults manifest. This means that available data from different faults is often limited and not representative of all fault scenarios (Sankavaram, Kodali, Pattipati, & Singh, 2015) and then a data-driven classifier trained on this data is not expected to achieve reliable performance (Tidriri et al., 2016) for new fault manifestations and sizes.

In Jung and Sundström (2017), a residual selection algorithm is proposed which uses information from both models and training data. The residual selection problem is there solved as a set of separate optimization problems, one for each requirement. This univariate approach is clearly suboptimal and a main contribution here is that all performance requirements are solved simultaneously in one optimization problem. This means smaller solution sets since the residual selection algorithm can identify residuals that fulfill multiple requirements and utilizes residual correlations.

A main contribution here is the formulation of a residual selection problem, combining model-based and data-driven methods, as a convex optimization problem, which can be solved efficiently using general-purpose solvers. A key contribution is the re-formulation of the inherently multi-objective problem as a single optimization problem that finds a *set* of residual generators given all performance requirements. It is assumed that training data is available from all relevant fault modes and, most importantly, it is also assumed that data is limited and *not* representative of all realizations of each fault. A main contribution of this work is systematic utilization of the analytic model in the data-driven feature selection process, alleviating the fundamental problem of limited training data from different fault scenarios. The proposed residual selection algorithm can handle both single-fault and multiple-fault isolation. To illustrate the proposed algorithm, it is applied to a real industrial use-case with data from an internal combustion engine.

2. Model-based diagnosis

Before defining the residual selection problem, a summary of some model-based diagnosis notions needed is given in this section. Structural properties of residual generators are defined which will be used to formulate the fault isolability constraints in the residual selection problem. An ideal residual generator is defined as

Table 1
Fault signature matrix of residual set $\mathcal{R}^*$.

Residual	f_{Waf}	f_{pim}	f_{pic}	f_{Tic}
r_2	X	X		
r_{19}	X		X	
r_{26}		X		X
r_{27}		X	X	
r_{29}	X			
r_{30}				X

Definition 1 (*Ideal Residual Generator*). An ideal residual generator $r_k(z)$ for a given system is a function of sensor and actuator data z where a fault-free system implies that the residual output $r_k(z) = 0$.

An ideal residual generator $r_k(z)$ is said to be *sensitive* to a fault f_i if there exists a realization of the fault that implies that the residual output $r_k(z) \neq 0$ (Svärd et al., 2013). Information about which set of faults each residual is sensitive to can be summarized in a Fault Signature Matrix (FSM). An example is shown in Table 1 where a mark at position (k, l) means that residual r_k is sensitive to fault f_l. A fault f_l is said to be *decoupled* in r_k if the residual is not sensitive to that fault.

Instead of discussing single-faults and multiple-faults, the term *fault-mode* is used to describe the system state. A fault mode $\mathcal{F}_i \subseteq \mathcal{F}$ describes which faults that are present in the system and the no-fault case $\mathcal{F}_i = \emptyset$ is denoted NF. Based on fault modes, the following definition of fault detectability and isolability will be used to formulate the residual selection problem (Svärd et al., 2013).

Definition 2 (*Fault Detectability and Isolability*). Let $\mathcal{R} \subseteq \mathcal{R}_{all}$ denote a set of residual generators. A fault mode $\mathcal{F}_i$ is detectable in $\mathcal{R}$ if there exists a residual $r_k \in \mathcal{R}$ that is sensitive to at least one fault $f_i \in \mathcal{F}_i$. A fault mode $\mathcal{F}_i$ is isolable from another fault mode $\mathcal{F}_j$ if there exists a residual $r_k \in \mathcal{R}$ that is sensitive to at least one fault $f_i \in \mathcal{F}_i$ but not any fault $f_j \in \mathcal{F}_j$.

To determine if any of the residuals has deviated from its nominal behavior, different test quantities are used, such as thresholded residuals or cumulative sum (CUSUM) tests (Page, 1954).

3. Problem formulation

A first thing to observe is that for a given model there can be many possible residual generators. In general, the number of candidates grows exponentially with the degree of model redundancy (Krysander et al., 2008). To illustrate this, consider the small example

$$x = g(u), \;\; y_i = x, \quad i = 1, \ldots, n$$

where u is a known control input and there are n measurements of the unknown variable x. With $n = 1$ there is only one possible residual generator, i.e., $r = y_1 - g(u)$, but with an increasing n the number of possibilities increases. It is straightforward to realize that the number of residual generators based on a minimal number of equations is given by

$$|\{\text{minimal residual generators}\}| = \binom{n+1}{2}$$

since any pair of two equations, from the set of $n+1$ equations, can be used to compute a residual. This simple observation generalizes to more general models (Krysander et al., 2008).