Mechanical Systems and Signal Processing 45 (2014) 170–192

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

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

Data-driven and adaptive statistical residual evaluation for fault detection with an automotive application

Carl Svärd [a,*], Mattias Nyberg [a], Erik Frisk [b], Mattias Krysander [b]

[a] Systems Development, Scania CV AB, SE-15187 Södertälje, Sweden
[b] Department of Electrical Engineering, Linköping University, SE-58183 Linköping, Sweden

ARTICLE INFO

Article history:
Received 10 May 2012
Received in revised form
19 February 2013
Accepted 5 November 2013
Available online 28 November 2013

Keywords:
Fault diagnosis
Fault detection
Residual evaluation
Statistical hypothesis testing
Automotive diesel engine

ABSTRACT

An important step in model-based fault detection is residual evaluation, where residuals are evaluated with the aim to detect changes in their behavior caused by faults. To handle residuals subject to time-varying uncertainties and disturbances, which indeed are present in practice, a novel statistical residual evaluation approach is presented. The main contribution is to base the residual evaluation on an explicit comparison of the probability distribution of the residual, estimated online using current data, with a no-fault residual distribution. The no-fault distribution is based on a set of a priori known no-fault residual distributions, and is continuously adapted to the current situation. As a second contribution, a method is proposed for estimating the required set of no-fault residual distributions off-line from no-fault training data. The proposed residual evaluation approach is evaluated with measurement data on a residual for fault detection in the gas-flow system of a Scania truck diesel engine. Results show that small faults can be reliably detected with the proposed approach in cases where regular methods fail.

1. Introduction

Fault diagnosis is becoming more and more important with the increasing demand for dependable technical systems, driven mostly by economical, environmental, and safety incentives. One example is automotive systems, where good fault diagnosis is essential in order to meet customer demands regarding up-time, efficient repair and maintenance, and also to fulfill on-board diagnosis (OBD) legislative regulations.

Model-based fault diagnosis typically comprises fault detection and isolation [13], and the fault detection part contains the essential steps: residual generation and residual evaluation. In the first step, a model of the system is used together with measurements to generate residuals. In the second step, the residuals are evaluated with the aim to detect changes in the residual behavior caused by faults in the system. This work concerns the second step, residual evaluation.

Ideally, residuals are signals that are zero when no faults are present in the system, and non-zero otherwise. Due to the presence of uncertainties and disturbances, caused by for instance modeling errors, measurement noise, and unmodeled phenomena, residuals typically however deviate from zero even in the no-fault case. Moreover, due to changes in the operating mode of the system, the magnitude of these uncertainties and disturbances is time-varying, causing the behavior of residuals to be non-stationary. An illustration is given by Fig. 1, where a residual for fault detection in the gas-flow system

* Corresponding author. Tel.: +46707702384.
E-mail addresses: carl.svard@scania.com, carl.svard@gmail.com (C. Svärd), mattias.nyberg@scania.com (M. Nyberg), frisk@isy.liu.se (E. Frisk), matkr@isy.liu.se (M. Krysander).

0888-3270/$ - see front matter
http://dx.doi.org/10.1016/j.ymssp.2013.11.002

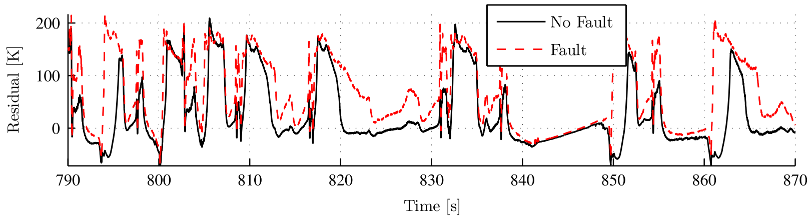

Fig. 1. A residual for fault detection in the gas-flow system of a heavy-duty truck diesel engine in the no-fault (solid) and fault (dashed) cases.

of a truck diesel engine is shown. Clearly, the residual is not zero in the no-fault case, and it is obvious that the residual exhibits non-stationary features. It can also be noted that the difference between the residual in the no-fault and fault cases is time-varying. Nevertheless, the fact that there is a difference implies that the present fault is potentially detectable. Note also that the residual does not exhibit any periodic time-variations.

There are two main approaches [22] for residual evaluation, statistical [60,27,7,50,3,14,59,30,48] and norm-based [23,25,26,53,18,62,63,33,4,36,1]. Statistical approaches exploit the framework of statistical hypothesis testing in order to detect changes in some parameter of the probability distribution of the residual, typically by means of likelihood ratio testing [28]. In norm-based approaches, residual evaluation is typically done by adaptive or constant thresholding of some norm of the residual.

Apparently, when encountering a residual as the one depicted in Fig. 1, neither statistical-based approaches assuming stationary probability distributions, nor norm-based approaches using constant thresholds, would be successful. A potential solution is to consider adaptive thresholds [19,24], and use a priori knowledge, either qualitative [33,62,31,23] or quantitative [53,25,47], to derive non-constant thresholds to take the time-varying uncertainties and disturbances into account. Furthermore since the residual depicted in Fig. 1 is non-periodic, diagnosis approaches for machines working in non-stationary operating conditions [16,52,43,51] are not applicable.

This paper instead proposes an adaptive statistical residual evaluation method, which exploits quantitative a priori knowledge in the form of data.

The main contribution is to base the residual evaluation on an explicit comparison of the probability distribution of the residual, estimated on-line using current data, with a no-fault residual distribution. The no-fault distribution is based on a set of a priori known no-fault distributions and to handle changes in the operating mode of the system, and thus time-varying residual features, it is continuously adapted to the current operating mode of the system. The comparison is done in the framework of statistical hypothesis testing by application of the Generalized Likelihood Ratio (GLR). As a second contribution, a method is proposed for estimating the required set of no-fault residual distributions off-line from no-fault training data. Thus, using the method for distribution estimation, the overall residual evaluation method becomes fully data-driven and neither assumptions regarding the properties of the probability distribution of the residual, nor the properties of the faults to be detected, are made.

The paper is organized as follows. Section 2 discusses and formalizes the problem setup and the residual evaluation problem is formulated in the framework of statistical hypothesis testing. In Section 3, the GLR is utilized to design a preliminary test statistic for the residual evaluation hypotheses, and the emerging likelihood maximization problems are considered. In Section 4, the preliminary test statistic is improved in terms of required computational effort, and a residual evaluation algorithm suitable for implementation in an online environment is given. Section 5 presents an off-line algorithm for learning no-fault residual distributions from no-fault training data. In Section 6 the proposed residual evaluation approach is applied to a residual for fault detection in the gas-flow system of a real Scania truck diesel engine. Finally, Section 7 concludes the paper.

2. Problem formulation

The residual evaluation problem, as considered in this work, is formally stated in this section.

2.1. Prerequisites

A *residual*, r, is considered to be the output from a *residual generator*, taking measurements from a *system* as input. Typically, the measurements consist of the input and output of the system. The system is considered to be subject to faults, and the intention is to detect if any fault is present in the system by monitoring the behavior of the residual. Note that if a set of residuals sensitive to different faults is used, faults can also be isolated, see for example [13].

The system typically operates in a number of different *operating modes*, and normal operation usually involves several of these modes. For an example, consider a heavy-duty truck diesel engine, for which a residual is shown in Fig. 1. Naturally, this system is designed to operate in a number of different operating modes typically characterized by engine torque, engine speed, ambient temperature, ambient pressure, etc.