Engineering Applications of Artificial Intelligence 30 (2014) 18–29

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

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

Combining learning in model space fault diagnosis with data validation/reconstruction: Application to the Barcelona water network

J. Quevedo [a,*], H. Chen [d], M.À. Cugueró [a], P. Tino [c], V. Puig [b], D. García [a], R. Sarrate [a], X. Yao [c]

[a] Advanced Control Systems (SAC), Automatic Control Department (ESAII), UPC, Rambla Sant Nebridi, 1008222 Terrassa, Spain
[b] Advanced Control Systems (SAC), Automatic Control Department (ESAII), UPC, Institut de Robòtica i Informàtica Industrial (IRI), CSIC-UPC, Pau Gargallo, 508028 Barcelona, Spain
[c] School of Computer Science, The University of Birmingham Edgbaston Birmingham B15 2TT, United Kingdom
[d] UBRI, University of Science and Technology of China Hefei, Anhui, China, 230027

ARTICLE INFO

Article history:
Received 20 July 2013
Received in revised form
3 November 2013
Accepted 13 January 2014
Available online 14 February 2014

Keywords:
Learning in model space
Sensor data validation/reconstruction
Time series
Fault diagnosis
Reservoir computing

ABSTRACT

In this paper, an integrated data validation/reconstruction and fault diagnosis approach is proposed for critical infrastructure systems. The proposed methodology is implemented in a two-stage approach. In the first stage, sensor communication faults are detected and corrected, in order to facilitate a reliable dataset to perform system fault diagnosis in the second stage. On the one hand, sensor validation and reconstruction are based on the combined use of spatial and time series models. Spatial models take advantage of the (mass-balance) relation between different variables in the system, whilst time series models take advantage of the temporal redundancy of the measured variables by means of Holt-Winters time series models. On the other hand, fault diagnosis is based on the learning-in-model-space approach that is implemented by fitting a series of models using a series of signal segments selected with a sliding window. In this way, each signal segment can be represented by one model. To rigorously measure the 'distance' between models, the distance in the model space is defined. The deterministic reservoir computing approach is used to approximate a model with the input–output dynamics that exploits spatial–temporal correlations existing in the original data. Finally, the proposed approach is successfully applied to the Barcelona water network.

1. Introduction

Critical infrastructure systems (CIS), including water, gas or electricity networks, are complex large-scale systems, geographically distributed and decentralized with a hierarchical structure, requiring highly sophisticated supervisory and real-time control (RTC) schemes to ensure high performance achievement and maintenance when conditions are non-favourable (Schütze et al., 2004; Marinaki and Papageorgiou, 2005) due to e.g. sensor and actuator malfunctions (faults). Each sub-system composing the CIS is constituted of a large number of elements with time-varying behaviour, having many different operating modes and subject to changes due to operational constraints. To deal with this problem, the use of an on-line fault diagnosis system able to detect such faults and correct them by activating different kinds of techniques e.g. data validation/reconstruction of sensor faults is desirable. This will prevent the RTC from being stopped every time that a fault appears, which is one of the main reasons why global RTC is not widely applied in the world (Schütze et al., 2004). Furthermore, the fault diagnosis process intends to identify which fault is causing the monitored events, including e.g. hardware and software faults.

Generally, two main strategies are available in the literature when addressing the fault diagnosis problem, which are hardware redundancy (preferred in critical systems) based on the use of extra sensors and actuators, and analytical redundancy, based on the use of software sensors or models combining information gathered by the sensor measurements or using other actuators to compensate the faulty ones. Nevertheless, the use of hardware redundancy in large-scale systems is very expensive and increases the number of maintenance and calibration operations, which calls for the use of combined hardware and analytical redundancy approaches in CIS (Carrozza et al., 2008). The capability to detect and isolate faults in these systems is important to keep their integrity safe. This problem has been targeted by numerous researchers from many different points of view, as overviewed in the compilation of techniques included in Venkatasubramanian et al. (2003a–c), and more recently in Ding (2008).

In this paper, an innovative framework that investigates the fault diagnosis problems in the model space instead of the data/signal space is developed. This fault diagnosis framework is integrated with a data validation/reconstruction methodology introduced in Quevedo et al. (2010a). 'Learning in model space'

* Corresponding author.
E-mail address: joseba.quevedo@upc.edu (J. Quevedo).

0952-1976/$ - see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.engappai.2014.01.008

(Chen et al., 2014) is implemented by fitting a series of 'approximated' models using a series of signal segments selected with a sliding window, and then apply the learning techniques to discriminate and isolate fault models from healthy models. Reservoir computing is chosen as an example to approximate the signal segments. Dynamic reservoirs of reservoir models have been shown to be 'generic' in the sense that they are able to represent a wide variety of dynamical features of the input driven signals, so that given a task at hand only the linear readout on top of reservoir needs to be retrained (Lukoševičius and Jaeger, 2009). Hence, in the formulation the underlying dynamic reservoir will be the *same* throughout the signal – the differences in the signal characteristics at different times will be captured solely by the linear readout models and will be quantified in the function space of readout models.

Here it is assumed that, for some sufficiently long initial period, the system is in a 'normal/healthy' regime so that when a fault occurs the readout models characterizing the fault will be sufficiently 'distinct' from the normal ones. A variety of novelty/anomaly detection techniques can be used for the purposes of detection of deviations from the 'normal' regime. In this contribution, support vector machines (SVMs) in the readout model space are used, so new faults occurring will be captured by the algorithm proposed operating in the readout model space.

The contributions of this paper are listed as follows:

- First, data validation and reconstruction techniques are integrated with learning in the model space for effective fault diagnosis.
- Second, SVMs are used in the model space for fault detection/isolation.

Finally, the proposed methodologies are applied to the Barcelona water network as a case study in this paper.

2. Data validation/reconstruction approach

In systems like CIS, a telecontrol system is acquiring, storing and validating data gathered from different kinds of sensors every given sampling time to accurately real-time monitor the whole system. In this process, problems in the communication system, e.g. between sensors and data loggers or in the telecontrol system itself, are frequent and produce data loss which may be of great concern in order to have valid historic records. When this is occurring, lost data should be replaced by a set of forecasted data which should be a representative of the data lost. Another common problem in CIS is caused by the unreliable sensors, which may be affected by e.g. offset, drift, freezing in the measurements (Kanakoudis and Tolikas, 2001; Kanakoudis and Tsitsifli, 2011; Tsitsifli et al., 2011). These unreliable data should also be detected and replaced by forecasted data, since it may be used for system management tasks e.g. maintenance, planning, investment plans, billing, security and operational control (Quevedo et al., 2010a) and system fault detection and isolation (Fig. 1).

Different types of data validation methods with distinct degrees of complexity may be considered according to the available system knowledge. Generally, two types of methods can be considered, one for elementary 'low-level' signal based methods and another for 'high-level' model-based methods. The first class use simple heuristics and limited statistical information from the sensors (Burnell, 2003; Jörgensen et al., 1998) and is typically based on checking either signal values or variations, whilst the second class uses models for consistency-checking of the sensor data (Tsang, 2003). Here, the first class of data validation methods has been used to deal with sensor communication faults.

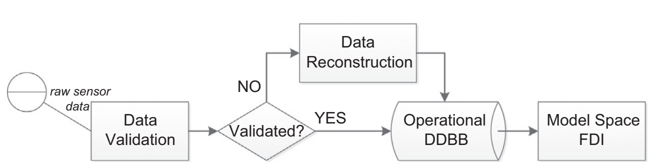

Fig. 1. Raw data validation/reconstruction and system FDI approach.

2.1. Data validation process

The data validation process is inspirited in the Spanish AENOR-UNE norm 500540 (Quevedo et al., 2010a). The methodology applies a set of consecutive validation tests to a given dataset (Fig. 2), to finally assign a certain quality level depending on the tests passed.

In a system like the one considered here, and in telecontrolled systems in general, one of the most common faults occurring is sensor communication fault. This type of fault is related with level zero of the sensor validation methodology in Quevedo et al. (2010a). This level checks whether the data is properly recorded, assuming that data acquisition systems sample data at a certain fixed rate. Hence, this level allows detecting problems in the data acquisition or communication system.

Here, communication faults are considered as the faults affecting the sensor of the telecontrol system, and data validation and reconstruction procedures are used as a prefilter to estimate the missing data when this type of faults is occurring.

2.2. Data reconstruction process

The output of the data validation process (Fig. 2) is used to identify the invalidated data that should be reconstructed. Spatial and time series (TS) models (Levels 4 and 5 in Fig. 2) are used for this purpose, depending on the performance of each model.

On the one hand, spatial models (SM) take advantage of the relation between different variables emplaced in the system. For example, in hydraulic systems, this relation is generally obtained from the mass balance model of the element relating the different measured variables involved, which states that the incoming and outcoming flows in a tank subsystem must be equal

$$\hat{x}_{SM}(k) = x(k-1) + \Delta t(q_{in}(k-1) - q_{out}(k-1)) \quad (1)$$

where $\hat{x}_{SM}$ is the spatial model tank volume estimation, x is the measured tank volume, q_{in} is the incoming tank flow, q_{out} is the outcoming tank flow and Δt is the sampling time. From this equation, the volume estimation for a particular tank subsystem may be stated. Estimation of other variables (e.g. $\hat{q}_{in}$, $\hat{q}_{out}$) may be obtained from algebraic manipulation of the latter.

However, real elements include uncertainty (due to e.g. unexpected behaviour of the plant, inaccuracy of the model) which may lead to the non-satisfaction of the mass balance in the element considered. Hence, consistency of the data collected by a certain sensor with its spatial model (Quevedo et al., 2010b), i.e. the correlation between data coming from spatially related sensors, may be maintained. For example, the data of the flow meters located in different points of the same pipe in a transport water network allows for checking the reliability of the sensor set and performing the corresponding correction, e.g. by using a linear regression model of input–output measured data in the pipe (which is ideally the identity function). In the case of the tank level estimation (1), this correction is introduced as

$$\hat{x}_c(k) = a\hat{x}_{SM}(k) + b \quad (2)$$

where $\hat{x}_c$ is the corrected estimation of the volume using regression model parameters $[a, b]$ obtained with the training dataset