Computers in Industry 74 (2015) 75–94

Contents lists available at ScienceDirect

Computers in Industry

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

An overview on fault diagnosis and nature-inspired optimal control of industrial process applications

Radu-Emil Precup [a,*], Plamen Angelov [b,c], Bruno Sielly Jales Costa [d], Moamar Sayed-Mouchaweh [e]

[a] Politehnica University of Timisoara, Department of Automation and Applied Informatics, Bd. V. Parvan 2, 300223 Timisoara, Romania
[b] Lancaster University, Data Science Group, School of Computing and Communications, Lancaster LA1 4WA, United Kingdom
[c] Chair of Excellence, Carlos III University, Madrid, Spain
[d] Federal Institute of Rio Grande do Norte – IFRN, Campus Natal, Zona Norte, Natal, Brazil
[e] École des Mines de Douai, Mines-Douai, IA, 59500 Douai Cedex, France

ARTICLE INFO

Article history:
Received 31 December 2014
Received in revised form 27 February 2015
Accepted 4 March 2015
Available online 23 March 2015

Keywords:
Data-driven control
Data mining
Evolving soft computing techniques
Fault diagnosis
Nature-inspired optimization algorithms
Wind turbines

ABSTRACT

Fault detection, isolation and optimal control have long been applied to industry. These techniques have proven various successful theoretical results and industrial applications. Fault diagnosis is considered as the merge of fault detection (that indicates if there is a fault) and fault isolation (that determines where the fault is), and it has important effects on the operation of complex dynamical systems specific to modern industry applications such as industrial electronics, business management systems, energy, and public sectors. Since the resources are always limited in real-world industrial applications, the solutions to optimally use them under various constraints are of high actuality. In this context, the optimal tuning of linear and nonlinear controllers is a systematic way to meet the performance specifications expressed as optimization problems that target the minimization of integral- or sum-type objective functions, where the tuning parameters of the controllers are the vector variables of the objective functions. The nature-inspired optimization algorithms give efficient solutions to such optimization problems. This paper presents an overview on recent developments in machine learning, data mining and evolving soft computing techniques for fault diagnosis and on nature-inspired optimal control. The generic theory is discussed along with illustrative industrial process applications that include a real liquid level control application, wind turbines and a nonlinear servo system. New research challenges with strong industrial impact are highlighted.

1. Introduction

For four decades fault diagnosis, also referred to as Fault Detection and Isolation (FDI), has been a strong field of study, however, it has gained more attention in the later years. Considering the complexity of the current industrial applications, FDI is a very challenging issue nowadays. Studies have shown that the human operator is responsible for 70–90% of the accidents in industrial environments [1,2]. Therefore, FDI techniques are used to improve operational safety, preventing (or reducing) accidents and unscheduled stoppages.

The data-driven methods, or process history-based methods, are regularly used for FDI tasks, many times combined with model-based approaches. The main advantage of such methods is that they do not require much (or none at all) expertise of the operator/designer, since they are mainly based on the data collected from the process, which can be historical (off-line) or real-time (on-line). This is a very important feature, since data-driven methods can cope with the problem of data drift and other unpredicted disturbances.

The data drift is defined in [3] as a change in the learned structure that occurs over time and can lead to a drastic drop of classification accuracy. Many times a system is designed to work under determined circumstances and its behaviour can be inadequate when dealing with unpredictable changes. These variations can occur due to process faults, disturbances or even slow environmental alterations.

At this point it is important to distinguish the concepts of *adaptive* and *evolving* systems. The term "adaptive" usually refers to the conventional systems, known in control theory for working

* Corresponding author. Tel.: +40 256 403229; fax: +40 256 403214.
E-mail address: radu.precup@upt.ro (R.-E. Precup).

http://dx.doi.org/10.1016/j.compind.2015.03.001
0166-3615/

with linear parameter adaptation, e.g., the traditional Adaptive Neural Fuzzy Inference Systems (ANFIS). The term "evolving", on the other hand, is associated to the systems that are also able to perform a gradual change in its core [4], either updating (adding and removing rules) the rule base, in the case of fuzzy systems, or updating (adding and removing nodes and layers) the structure, in the case of neural networks.

The decision regarding the use of process history-based methods rely on the fact that model-based techniques are very restrictive in the sense that they require information that it is usually not available in practical applications. While quantitative model-based methods require accurate mathematical descriptors of the process, qualitative model-based techniques work with a qualitative database, usually built from knowledge from an operator or an expert system. These models are often not available and, even it they are, it is many times impracticable to obtain all information of relevant physical parameters of the system, not to mention that external parameters, such as unpredictable disturbances, model uncertainties and so on, are not considered.

Process history-based techniques, on the other hand, do not require any knowledge, either quantitative or qualitative, about the process. Instead, they use massive historical information collected from the process. This data is, then, transformed and presented as a priori information to the FDI system.

The task of FDI can be analyzed as a classification problem in both stages–detection and isolation. Fault detection is the task where it is possible to identify whether the system is working in a normal operating state or in a faulty mode. However, in this stage, vital information about the fault, such as physical location, length or intensity, are not provided to the operator [5]. The detection stage, thus, can be addressed by a general one-class classifier, able to distinguish if the current collected data samples belong to a determined class of data, e.g. "normal". Fault isolation, on the other hand, refers to determination of kind, location and time of detection of a fault, and follows the fault detection stage [6]. The challenge is to match each pattern of the symptom vector with one of the pre-assigned classes of faults, in the case of supervised approaches, and the fault-free case [7], as a multi-class classifier.

Once the fault diagnosis is guaranteed, a systematic way to meet the performance specifications of control systems in industrial applications is the optimization in terms of optimization problems with variables represented by the tuning parameters of the controllers. Nature-inspired algorithms can solve these optimization problems, and they ensure the optimal tuning of controllers in order to meet the performance specifications expressed by adequately defined objective functions and constraints. The constraints are due to technical and/or economical operating conditions of industrial process applications, and they include stability, sensitivity, robustness and fault diagnosis conditions.

The motivation for nature-inspired optimization algorithms (NIOAs) in the optimal control of industrial process applications concerns the ability of such algorithms to cope with non-convex or non-differentiable objective functions because of the process complexity, of the controllers' structures and eventually the controllers' nonlinearities, which can lead to multi-objective optimization problems. In addition, the complexity of the classical optimization algorithms is very high, and this requires enormous amount of computational work. Therefore, the NIOAs are appreciated because they are better in terms of efficiency and complexity than classical optimization algorithms.

As generally shown or with focus on certain classes of NIOAs in [8–13], these algorithms are based on biological, physical, and chemical phenomena of nature. NIOAs have the distinct ability of finding the global minimum (or maximum) of certain objective functions under specific conditions. In addition, the analytical expression of the objective functions depending on other design (or tuning) parameters may be difficult or even impossible to formulate. These are the reasons why the NIOAs applied to optimal control of industrial processes are justified and also challenging.

This paper addresses the following topics: the state-of-the-art on machine learning, data mining and evolving soft computing techniques for fault diagnosis is discussed in Section 2. The unsupervised and autonomous self-evolving fault diagnosis is treated in Section 3. An illustrative fault diagnosis application related to a real liquid level control system controlled by a multi-stage fuzzy controller using a pilot plant for industrial process control is included. The problem setting for fault diagnosis in wind turbines and a classification of the methods used with this regard is presented in Section 4. It discusses the interest, motivation and challenges related to the fault diagnosis of wind turbines. Machine learning and data mining techniques are next organized in the framework of a general scheme that achieves fault diagnosis of wind turbines. Then, the different steps of this scheme are detailed in order to emphasize the links between the methods and techniques that they use and how they answer the challenges related to the fault diagnosis of wind turbines. Section 5 discusses NIOAs in the optimal tuning of linear controllers for industrial process applications and gives an example concerning the position control of a nonlinear servo system. The nature-inspired optimal tuning of nonlinear controllers in industrial process applications is next treated in Section 6 with focus on fuzzy controllers, but neural network controllers and sliding mode controllers are also considered. The drawbacks and research challenges in fault diagnosis and nature-inspired optimal control are discussed in Section 7. The concluding remarks are highlighted in Section 8.

2. State-of-the-art on machine learning, data mining and evolving soft computing techniques for fault diagnosis

Many authors have, very recently, contributed to the fault diagnosis field of study, with extensive studies, compilations and throughout reviews. In this section, we are going to address a few important approaches to FDI using intelligent techniques, specially focusing on machine learning, data mining, clustering and evolving techniques applied to industrial problems.

Among the recent studies on the topic, the paper [14] proposes an incremental support vector data description and extreme learning machines are used to solve the problem of classification of faults when the number of classes is unknown and tend to increase over time. While the proposed Support Vector Machine (SVM) is used to quickly detect new failure modes, Extreme Learning Machine (ELM) is changed into an elastic structure whose output nodes can be added incrementally to cope with the new fault scenario. The algorithm is applied to a Diesel engine under eleven different fault conditions, however can be suitable to other mechanical equipment.

An application of ELM to real-time FDI with data pre-processing through wavelet packet transform and time-domain statistical features is suggested in [15]. The process of feature extraction is, then, performed by a kernel Principal Component Analysis (PCA) algorithm. As a case of study, a comparison between ELM and SVM on a fault detection problem is conducted, resulting in a considerable advantage for the ELM algorithm.

An automatic method for bearing FDI using the vibration signals as input is given in [16]. A one-class v-SVM is used to distinguish normal and faulty operation modes, where the model of normality is built from data extracted under normal conditions. Band-pass filters and Hilbert Transform are used to isolate the fault. An experimental study is then performed using two different data sets: real data from a laboratory test-to-failure experiment and data obtained from a fault-seeded bearing test. The results