Journal of Natural Gas Science and Engineering 29 (2016) 303–310

Contents lists available at ScienceDirect

Journal of Natural Gas Science and Engineering

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

Enhancing the noise tolerance of fault diagnosis system using the modified adaptive boosting algorithm

Mojtaba Cheki*, Hooshang Jazayeri-Rad, Pooria Karimi

Department of Instrumentation and Automation Engineering, Petroleum University of Technology, Ahvaz, Iran

ARTICLE INFO

Article history:
Received 13 March 2015
Received in revised form
10 December 2015
Accepted 16 December 2015
Available online 19 December 2015

Keywords:
Artificial neural network
Fault diagnosis
MadaBoost
AdaBoost
Noise

ABSTRACT

Recent investigations in the field of fault diagnosis have increasingly used ensemble learning techniques due to their exceptional performance in dealing with high-dimensionality, small sample size, and complex data structure. An ensemble of ANNs can enhance the generalizability and reliability of a single ANN system, through training the ANNs for a given assignment and incorporating the outcomes. One of the conventional and widely recognized ensemble methods is the adaptive boosting (AdaBoost). A major problem associated with this method is the fact that it is not strong enough and sometimes fails in noisy environments. In this paper, to eliminate this drawback, the weighting system of the typical AdaBoost algorithm was modified. This paper applies the new boostig (MadaBoost) algorithm for the fault diagnosis of a chemical plant by using an ensemble of neural networks, with the objective to improve the noise tolerance of fault diagnosis system using the modified adaptive boosting algorithm.

1. Introduction

Changes in the physical and external conditions of the process units and the control systems may lead to faults. According to Himmelblau (1978), a fault is expressed as a divergence from the allowable span of a specified variable. In fault diagnosis, the source of each fault is determined and this procedure is, in fact, inevitable for the countermeasure or removal of faults.

In machine learning, numerous techniques have been developed to provide supervised classifications. These methods, which have been successfully utilized for fault diagnosis purposes, include Kernel Independent Component Analysis (KICA), Support Vector Machine (SVM) (Zhang, 2008), and Principal Component Analysis (PCA) (Chen and Liao, 2002). From a variety of classification techniques, artificial neural networks (ANNs) have particularly received considerable attention in the past few years by many researchers. The characteristics of ANNs with regard to process fault diagnosis are quiet useful. For instance, they are capable of managing the nonlinear and also un-determined processes even without any mathematical model. In addition, with the aid of the training data, they can learn the diagnosis procedure.

In this study, new boosting algorithm (MadaBoost), which utilizes an ensemble of ANN is used for fault diagnosis. The proposed ANN ensemble is then compared to the AdaBoost and the single ANN order to evaluate the classification performance. In addition, it will be exhibited that the utilized method significantly improves the performance of the fault diagnosis system for noisy data.

The case study of the Tennessee–Eastman Process (TEP) is subsequently described in Section 2, which is used for fault generation in this work. In the next section, a brief description of the ensemble methods is presented and two boosting methods (i.e., AdaBoost and MadaBoost) are depicted. In Section 4, the performance of the single ANN, the AdaBoost ANN, and the MadaBoost ANN methods in fault diagnosis for the noise-free and noisy data conditions are examined. Finally, the conclusions of this study are provided in Section 5.

2. Case study: Tennessee–Eastman Process (TEP)

The “Tennessee–Eastman Process (TEP)” is a benchmark problem in process engineering, offered by the Eastman Chemical Company. It presents a real industrial process for the process control and monitoring techniques (Downs and Vogel, 1993). The TEP consists of five main operation units, including reactor, condenser, compressor, stripper and separator (see Fig. 1). As far as the TEP is unstable in open-loop, Chiang et al. (2004) carried out a simulation of a plant through utilizing a closed-loop control structure, which

* Corresponding author.
E-mail address: mojtaba.cheki@yahoo.com (M. Cheki).

http://dx.doi.org/10.1016/j.jngse.2015.12.029
1875-5100/

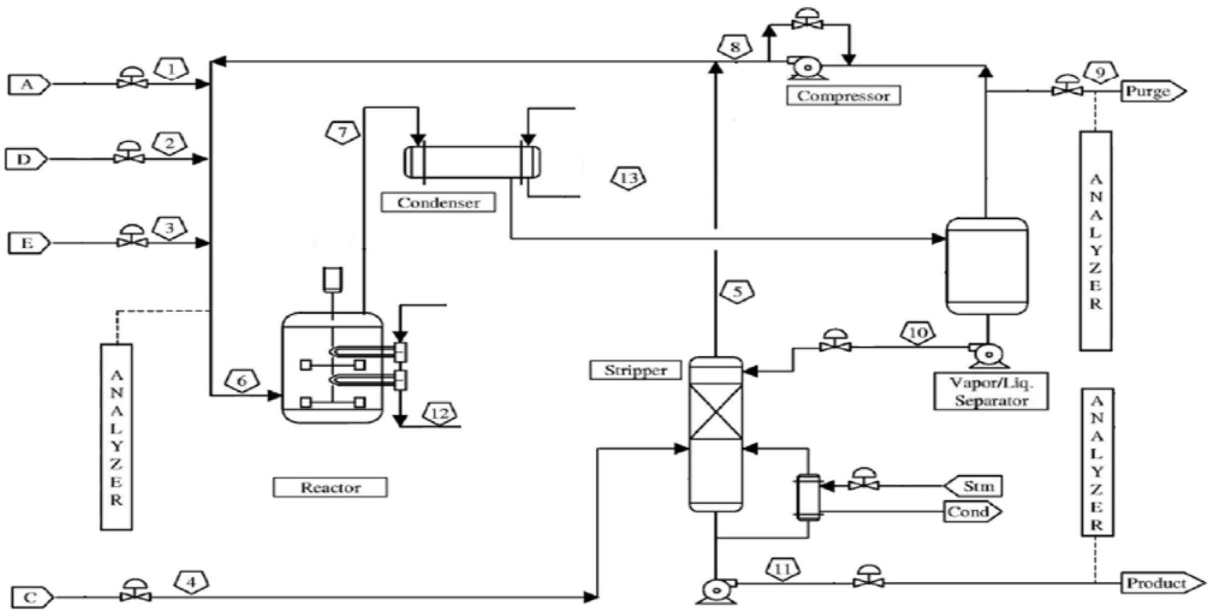

Fig. 1. The Tennessee–Eastman process (TEP).

was described by Lyman and Georgakis in 1995. For the present study, the generated simulation data, available in (Karimi and Jazayeri-Rad, 2014), have been used. The TEP included 41 measured variables and 12 manipulated variables of the control loops, where 52 out of the 53 variables were utilized for this work. The reactor agitator speed is fixed. The TEP simulation includes 21 pre-programmed faults. The first simulation run was generated without any fault (fault 0), whereas the remaining 21 runs were operated with various faults (faults 1 to 21). While 16 out of the whole set of the TEP faults are known, 5 faults are unknown, as shown in Table 1. For each type of fault, 480 training samples and 800 testing samples were generated after the introduction of the fault.

3. Ensemble methods

Ensemble learning has been broadly recognized as a very effective method and increasingly implemented for complex multiple learning algorithms to enhance the overall prediction accuracy (Dietterich, 2000b). Ensemble methods are capable of diminishing the problem of small sample sizes, via averaging and integrating the outputs of the multiple classification models and as a result, the possibility of over-fitting the training data will be reduced (Dietterich, 2000a). An ensemble-based method can be created through training various sub-models and incorporating the obtained outputs using the pre-specified methods. In the present work, in order to combine the independent decisions, the majority voting method has been utilized.

Table 1
The types of faults in the TEP.

Fault	Description	Type
1	A/C Feed ratio, B Composition constant (Stream 4)	Step
2	B Composition, A/C Ratio constant (Stream 4)	Step
3	D Feed temperature (Stream 2)	Step
4	Reactor cooling water inlet temperature	Step
5	Condenser cooling water inlet temperature	Step
6	A Feed loss (Stream 1)	Step
7	C Header pressure loss – reduced availability (Stream 4)	Step
8	A, B, C Feed composition (Stream 4)	Random
9	D Feed temperature (Stream 2)	Random
10	C Feed temperature (Stream 4)	Random
11	Reactor cooling water inlet temperature variation	Random
12	Condenser cooling water inlet temperature	Random
13	Reaction kinetics	Slow Drift
14	Reactor cooling water valve	Sticking
15	Condenser cooling water valve	Sticking
16	Unknown	Unknown
17	Unknown	Unknown
18	Unknown	Unknown
19	Unknown	Unknown
20	Unknown	Unknown
21	The valve for stream 4 was fixed at the steady-state position.	Constant Position