AJCHE 2010, Vol. 10, No. 2, 1 – 9

A Two-Step Fault Detection and Diagnosis Framework for Chemical Processes

Lau Chee Kiong[1]
Che Rosmani[1]
Che Hassan[1]
Mohd. Azlan Hussain[1]
[1]Department of Chemical Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia.
** E-mail: chee_kiong5161@yahoo.com.my*

An effective process monitoring system serves as an early warning system for influences affecting the chemical plant and helps plant operator to devise remedial actions to mitigate the adverse effects. However, the design of such system presents challenges such as complex cause-effect correlations, imprecise process model and novelty identifiability. In this work, a two-step fault detection and diagnosis framework is presented. This framework utilizes boundary models developed from mass and energy balances for each section of the chemical plant. The fault detection step consists of a fuzzy inference system (FIS) to analyze the balances and identify the faulty section if the balances deviate from the normal boundary. Then, multiple adaptive neuro-fuzzy inference system (ANFIS) classifiers are constructed to diagnose the exact root causes of bad performance. The combination of boundary models and FIS provides fault isolation of the faulty plant section even when novel faults have occurred. Utilization of multiple ANFIS classifiers reduces the complexity of the networks and improves the proficiency of the process monitoring system. The proposed scheme is applied on a model of a large scale industrial process.

Keyword: Fault detection, Fault diagnosis, Boundary models, FIS, ANFIS.

INTRODUCTION

Effective detection of abnormal operating condition and identification of malfunction components are one of the main concerns of the process engineer to improve process operation, increase plant throughput, reduce process downtime and comply with increasingly stringent environmental rules and safety regulations.

Various fault detection and diagnosis methods have been proposed and studied extensively. These methods can be classified into three major categories: qualitative model-based, quantitative model-based and data driven based. Quantitative model-based methods such as observer-based method and parameter estimation utilize mathematical models constructed from first principles for process monitoring. However, the effectiveness of this approach depends on the precision of the mathematic models constructed. Qualitative model-based methods such as signed digraphs and fault trees analysis employ cause-effect reasoning to describe system behavior. However, the qualitative model-based approach is restricted to systems with a relatively small number of variables or states because the knowledge based creation task can be time consuming.

Data driven methods find patterns or compute meaningful statistics from the process historical data, for example, principal component analysis (PCA) and neural networks. The applications of neural networks and other artificial intelligence techniques in the field of engineering as well as process fault detection and diagnosis have been extensively studied [1]-[7]. The advantages of using artificial intelligence include adaptation, generalization and effective handling of uncertainty.

To improve the process monitoring of larger scale system, several methods have been proposed to reduce the computation loads and memory requirements. Power and Bahri [8] proposed a two-step supervisory framework utilizing neural networks for fault diagnosis and Petri net for fault detection of Bayer process. Watanabe et al. [9] discussed and applied hierarchical neural networks (HANN) for multiple simultaneous fault diagnosis of a chemical reactor. Ozurt and Kandel [10] proposed a hybrid diagnostic methodology based on hierarchical perceptron-elliptical neural networks structure and expert system for a hydrocarbon chlorination plant. Eslamloueyan [11] proposed a HANN for isolating the faults of Tennessee Eastman (TE) process through fuzzy C-mean clustering of the fault patterns.

This work focuses on the use of fuzzy inference system (FIS) for fault detection and the use of adaptive neuro-fuzzy inference system (ANFIS) for fault diagnosis of a dynamical process. Also, multi-scale principal component analysis (MSPCA) is incorporated into the diagnosis module to improve the process monitoring.

The paper is organized as follows. In Section 2, a brief description of FIS, MSPCA and ANFIS are described followed by the description of the fault detection and diagnosis framework. A case study is presented in Section 3 for the proposed method to detect and diagnose the faults in Tennessee Eastman (TE) process. The diagnosis results and the comparison with multivariate statistical fault detection using PCA are discussed in Section 4. The last section contains some concluding remarks.

FAULT DETECTION AND DIAGNOSIS

Fuzzy inference system

A FIS formulates the mapping from a given an input to an output using fuzzy logic through the use of membership functions, logical operations and *If-Then* rules as depicted in Fig. 1.

There are five parts of the fuzzy inference process. Initially, the inputs are converted into fuzzy domain and the antecedent for each rule is evaluated through logical operation. Then, the weight of the rule is determined and the membership function of the consequent is determined through implication method. The output of each rule are combined and

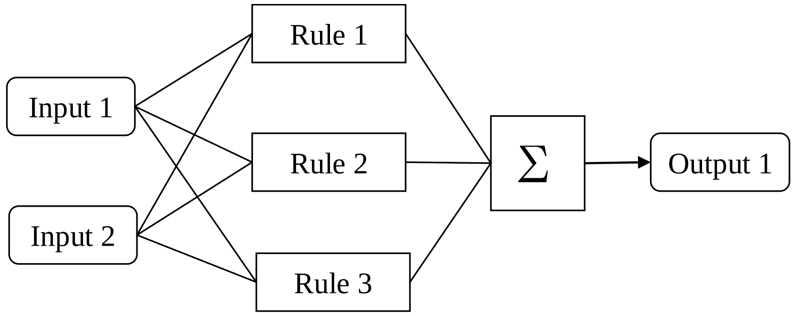

Figure 1. A basic structure of FIS with two inputs, three rules and one output.

aggregated. Finally, the output is defuzzied and the fuzzy set is transformed into a crisp value.

The use of FIS enables the integration of plant operator's experience and knowledge and improves the transparency of the diagnostic system. Detailed description of FIS can be found in [12].

Feature extraction using MSPCA

The data from all practical processes are multiscale in nature, for instance, the variables are measured at different sampling rates. Also, the events can occur at different locations in the plant and with different localization of time and frequency. As a result, conventional feature extraction method such as principal component analysis (PCA) which assumes the data are in single scale is not efficient in capturing the slow and fable changes in the measured variables.

The methodology of MSPCA can be summarized into following steps:

(i) The measurement data are decomposed into wavelet coefficients.
(ii) The PCA of the wavelet coefficients are calculated and the coefficients with significant events are retained.
(iii) The retained coefficients are converted back into time domain.
(iv) The PCA loading vectors of the filtered data are calculated.

Then, the observation data are projected into lower dimensional score and residual space.

MSPCA combines the advantages of PCA to capture the correlation across the measured variables with that of wavelet analysis to extract autocorrelation within the measured variables along the time axis. Detailed formulation of MSPCA has been discussed by Bakshi [13].

Adaptive neuro-fuzzy inference system

An ANFIS is a fuzzy inference system enhanced with learning, generalization and adaptively capabilities. A Sugeno-type fuzzy inference system [12] is created through fuzzy clustering of the data. By identifying the grouping of the data into a collection of *If-Then* rules, the characteristics of a nonlinear system can be concisely represented. Typically, a fuzzy rule in a Sugeno fuzzy model has the format

$$\text{If } x \text{ is } A \text{ and } y \text{ is } B \text{ then } z = f(x, y). \qquad (1)$$

where A and B are fuzzy sets in the antecedent; $z = f(x, y)$ is a crisp function in the consequent. For a first-order Sugeno fuzzy model, $f(x, y)$ is a first-order

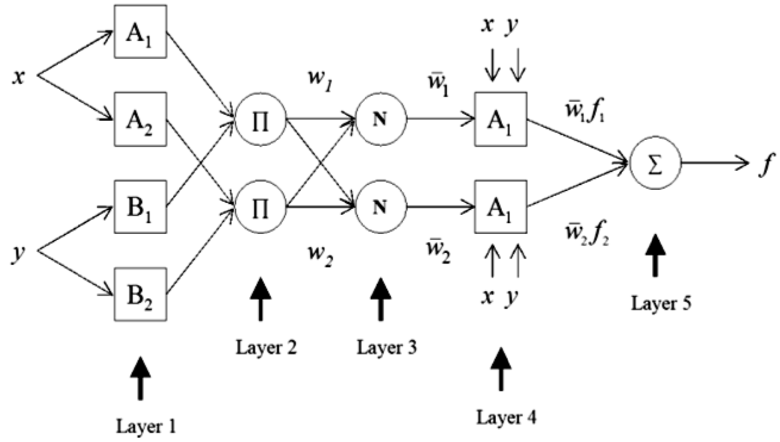

Figure 2. A typical structure of ANFIS with two inputs and one output.

polynomial. For a zero-order Sugeno fuzzy model, the output, *f*, is a constant. To facilitate the learning of a Sugeno fuzzy model, an adaptive network that can search optimal parameters systematically is incorporated. The resulting network structure, as shown in Fig. 2, is termed as adaptive neuro-fuzzy system (ANFIS).

The function of each layer in ANFIS is summarized as follows:

(i) Layer 1 – convert the crisp input to domain through a fuzzy membership function.

(ii) Layer 2 – generate the firing strength of a rule through multiplication.

(iii) 3 – calculate the i^{th} firing strength to the total firing strength.

(iv) ayer 4 – compute the contribution of the i^{th} rules.

(v) Layer 5 – calculate the weighted average output from each rule.

Detailed information of ANFIS has been covered in [12].

Fault detection and diagnosis framework

The two-step fault detection and diagnosis framework is presented in Fig. 3. A FIS fault detection module analyzes the mass and energy balances of the process and detects the faults based on the dynamic process data. The rule of FIS can be constructed from the process Failure Mode and Effect Analysis (FMEA), the plant operator troubleshooting knowledge and past historical data.

The diagnosis is then directed to ANFIS classifier. The MSPCA projects the observation data into a lower dimensional space and divides the data into score and residual space. Normal and faulty condition data are used to train ANFIS to learn the cause-effect correlations of the process. The ANFIS will monitor the subspaces to determine whether a fault has occurred. This two-step fault detection and diagnosis framework enables the isolation of the faulty section and quick diagnosis of the fault.

CASE STUDY

In this section, the description of the TE process and the process faults are described. Then, the detection and diagnosis results using the proposed

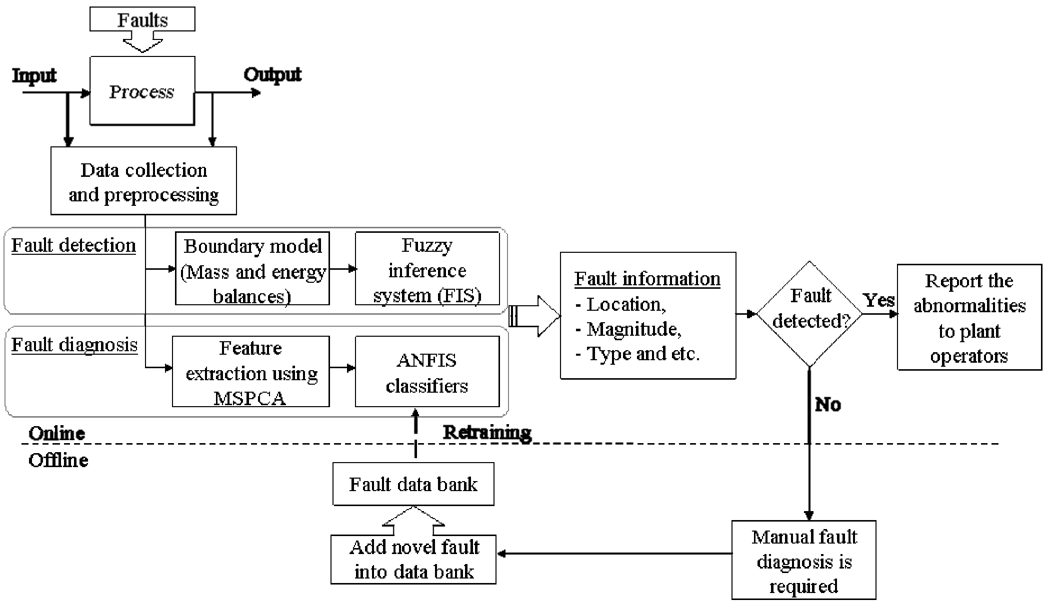

Figure 3. The two-step fault detection and diagnosis framework.

framework are illustrated. Also, the diagnosis results will be compared with the process monitoring using multivariate statistical PCA method.

Tennessee Eastman (TE) process

The process flowsheet of TE process is illustrated in Fig. 4. The process consists of five main units: a reactor, condenser, compressor, separator and stripper. The reactant A, C D and E are fed to the reactor where the liquid product G and H are formed through the following exothermic reactions:

A (g) + C (g) + D (g) → G (l),
A (g) + C (g) + E (g) → H (l),
A (g) + E (g) → F (l),
3D (g) → 2F (g),

The process has 22 continuous process measurement variables, 19 composition measurements and 12 manipulated variables. All the process measurements consist of Gaussian noise. At a particular time instant, an observation vector is given by:

$$X = [XMEAS(1)...XMEAS(22)\ XMV(1)...XMV(11)] \quad (2)$$

where the XMEAS and XMV stand for measured and manipulated variables respectively. Detailed of the process description can be found in [13].

Process fault

There are 21 preprogrammed faults in the TE process. Out of the 21 faults, only five faults are selected for evaluating the effectiveness of the proposed framework:
Fault 5 - Step change in condenser cooling water temperature.
Fault 12 - Random variation in condenser cooling water inlet temperature.
Fault 13 - Slow drift in reaction kinetics.
Fault 14 - Sticking of reactor cooling water valve.
Fault 15 - Sticking of condenser cooling water valve.

The simulation time of the TE process for each fault was 48 hours. The simulation

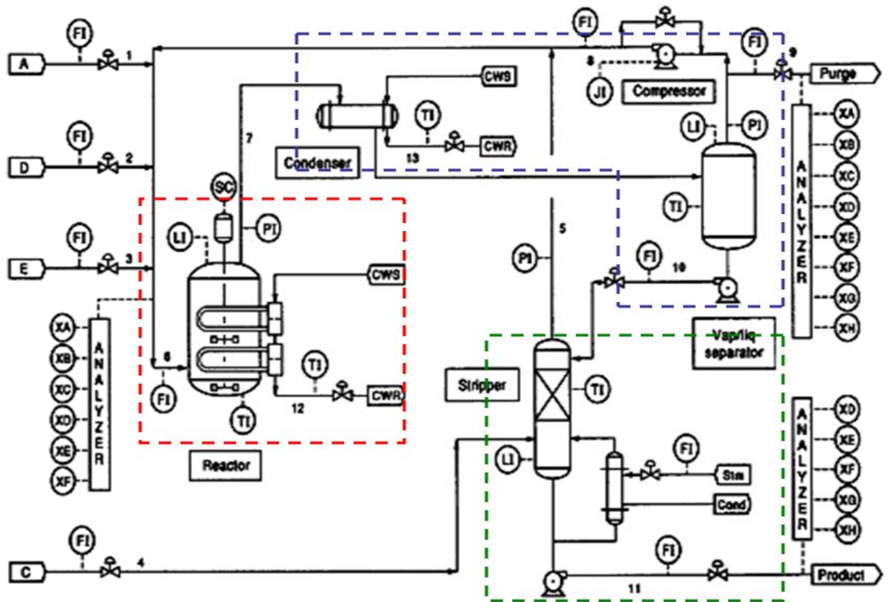

Figure 4. Reaction, Separation and Stripping section of the TE process

started with no fault, and the fault was introduced after 8 hours into the run. The total number of observations generated was 960 and each sample was taken for every 3 minutes.

RESULTS AND DISCUSSIONS

Below are diagnostic results for Fault 12 and 13 obtained from the diagnosis module of the proposed scheme.

Fault 12 is an intermittent fault which involves variation in the cooling water inlet temperature. The extracted features are shown in Fig. 5. Also, it exhibits different fault patterns even at the same initial operating point or mode. Residual space is utilized to diagnose Fault 12 as it can capture random variation in the characteristic properties more effective compared to the diagnosis using score space. The ANFIS classifier corresponding to Fault 12 is triggered at around 8 hours and consistently gives value above the threshold = 0.145 consistently.

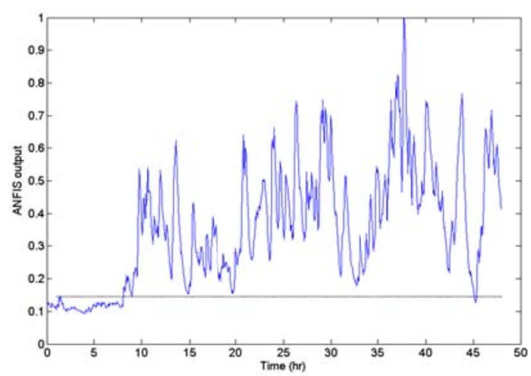

Figure 5. Diagnosis results using the residual space for Fault 12

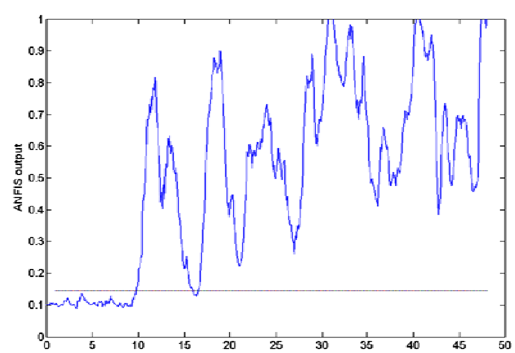

Figure 6. Diagnosis results using the score space for Fault 13

However, the misdetection rates for Fault 15 are relatively high. This is because the changes in the measured variables are less than 3% from the normal condition [15]. As the changes are insignificant and lack of distinctive information, the diagnosis

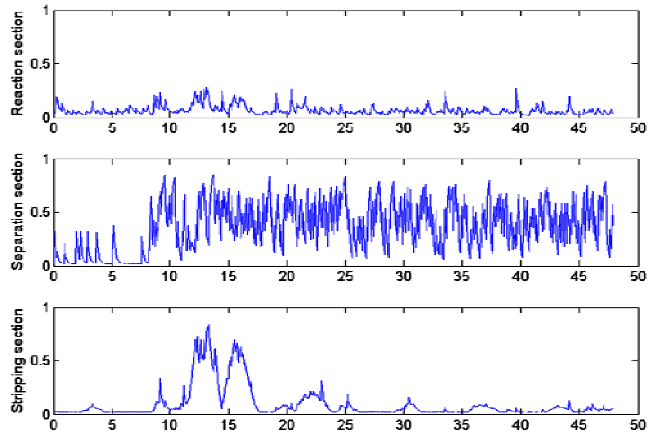

Figure 7. The FIS detection module outputs for fault detection of Fault 5

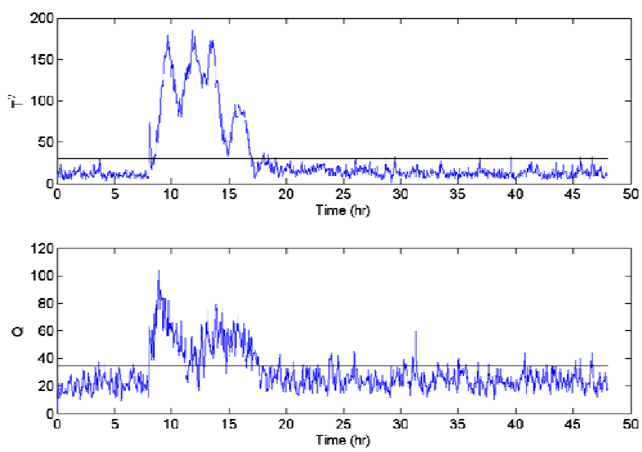

Fig.ure 8. The PCA multivariate statistics for fault detection for Fault 5

Table 1. Misdetection rates for faulty cases in the TE process

Fault	Misdetection rates		
	Propose d method	PCA-T^2	PCA-Q
5	0	0.956	0.038
12	0.031	0.029	0.025
13	0.076	0.060	0.045
14	0.001	0.158	0
15	0.779	0.988	0.973

module of the proposed framework unable to discriminate the fault from normal condition and thus, resulting high misdetection rates. However, the results are acceptable as the proposed method successfully cover the diagnosis for nearly all the faults.

The result of fault detection module and the comparison with multivariate statistic PCA are illustrated in Fig. 7 and Fig. 8 respectively. The fault is originated from the condenser which is in the Separation section. The detection module of the proposed framework identifies the Separation and Stripping section are faulty after the fault is introduced at the 8-th hour. However, the third output (Stripping section) retreats to below 0.1 after 20 hours as the plant reaches its new steady state and only the second output (Separation section) continuous follows the fault. The PCA statistics only can detect the fault between the 8 to 17 hours. The fault detection module can detect the fault more consistently compare to the PCA statistics.

CONCLUSION

A two-step fault detection and diagnosis framework was proposed for a relatively complex chemical plant. The fault detection module utilizes a FIS to analyze the mass and energy balances developed around the plant sections. The membership functions of the inputs and the rules were developed from the operating range for normal and faulty conditions. When a fault has occurred, the mass and energy balances will be violated and FIS will detect and isolate the faulty section. The diagnosis module consists of MSPCA feature extractor and multiple ANFIS classifiers. Each of the classifier is dedicated to one specific fault. The subspace (score or residual space) which is most sensitive to the fault will be monitored by the ANFIS classifier. This significantly improves the proficiency of the process monitoring. The classifier will analyze the extracted features and determined which fault has occurred. The use of multiple ANFIS classifiers reduces computation load and complexity of the network structure.

REFERENCES

1) M. A. Hussain, Review of the applications of neural networks in process control – simulation and online implementation, J Artif. Intell. Eng. **13**(1) (1999) 55-68.
2) M. A. Hussain, C. R. Che Hassan, K. S. Loh and K. W. Mah: Application of artificial intelligence technique in process fault diagnosis, J. Eng. Sci. and Tech. **2**(3) (2007) 260-270.
3) C. K. Lau, Y. S. Heng, M. A. Hussain and M. I. Mohamad Nor: Fault diagnosis of the polypropylene production process (UNIPOL PP) using ANFIS, ISA T. (2010) doi:10.1016/j.isatra.2010.06/007.
4) S. H. Jamal, M. A. Hussain, M. K. Aroua and D. Yaakop: Artificial neural network models applied to chemical engineering: a review, Trends Chem. Eng. **10** (2006) 1-15.
5) J. Y. Fan, M. Nikolaou and R. E. White: An approach to fault diagnosis of

chemical process via neural networks, AIChE J. **39** (1993) 82-88.

6) Y. C. Kok, C. P. Lim and K. L. Weng: Fault diagnosis in a power generation plant using a neural fuzzy system with rule extraction. In: V. Palade, C.C. Bocanialam and L. Jain (Eds): Computational Intelligence in Fault Diagnosis (2001) Springer.
7) K. Watanabe, I. Matsuura, M. Abe, M. Kubota and D. M. Himmelblau: Incipient fault diagnosis of chemical processes via artificial neural networks, AIChE J. **35**(11) (1989) 1803-1989.
8) Y. Power, P. A. Bahri: A two-step supervisory fault diagnosis framework, Comput. Chem. Eng. **28** (2004) 2131-2140.
9) K. Watanabe, S. Hirota, L. Hou and D. M. Himmelblau: Diagnosis of multiple simulataneous faults via hierarchical artificial neural networks, AIChE J. **40**(5) (1994) 839-848.
10) B. Ozyurt and A. Kandel: A hybrid hierarchical neural network-fuzzy expert system approach to chemical process fault diagnosis, Fuzzy Sets Syst. **83** (1996) 11-25.
11) R. Eslamloueyan: Designing a hierarchical neural network based on fuzzy clustering for fault diagnosis of the Tennessee-Eastman process, Appl. Soft Comput. J. (2010), doi:10.1016/ j.asoc.2010.04.012.
12) J.-S R. Jang and C.-T. Sun, Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence (1997), Prentice Hall.
13) B. R. Bakshi: Multiscale PCA with application to multivariate statistical process monitoring, AIChE J. **44**(7) (1998) 1596-1610.
14) J. J. Downs and E. F. Vogel: A plant-wide industrial process control problem, Comput. Chem. Eng. **17** (1993) 245-255.
15) L. H. Chiang, E. L. Russell and R. D. Braatz: Fault Detection and Diagnosis in Industrial Systems (2006) Springer.