242l An Hybrid System for Robust and Transparent Process Fault Diagnosis

Ignacio Yélamos, Estanislao Musulin, and Luis Puigjaner

Process fault diagnosis deals with the detection and isolation of abnormal events. It consists on interpreting the current status of the plant from sensors readings and process knowledge. Principal Components Analysis (PCA), one of the most popular multivariate statistical techniques, has been widely used for process monitoring, helping to reduce the number of false alarms by monitoring the variables correlation in addition to variables themselves. Besides, this method informs about the contribution of the different process variables to changes in the operation conditions, which is very useful in plant fault diagnosis. However, the interpretation and association of the contribution of the variables with the root causes of faults are not trivial and an expert system is required to ensure a reliable fault diagnosis. In this work, a fault diagnosis system (FDS) has been developed integrating PCA with a fuzzy logic knowledge based (FLKB) system. It makes use of PCA reduced models and the capabilities of FLS to behave in uncertainty environments in order to augment the robustness of the FDS.

The FLKB system is the core of this work and it has been thoroughly designed in order to achieve a robust, flexible and easy to use FDS. Additionally, transparency of FDS that is a characteristic very appreciated by plant operators has been faced, since it is seriously degraded by the high number of rules and linguistic values in each rule, that the complex FLKB systems result in. In order to obtain a transparent FDS, a simple off-line methodology to automatically extract rules from historical data has been designed. The association of variables with a process state is made through an if-then rule base.

The algorithm involves three main steps and it is depicted in figure 1: a) first step consists in an input space partitioning by means of trapezoidal membership functions. The selection of the membership functions parameters is of crucial importance to the later FDS resolution; b) second step makes an assignment of linguistic values for each input variable. The membership degree of each sample to the sets Low, High or Normal can be evaluated using the membership functions created in step one; c) third step consists on merging the extracted antecedents from second step, building a complete rule to identify the current state. For each considered state, one or more if-then rules are extracted. They have the following structure:

Rk : If x1 is Ak1 and x2 is Ak2 and ... xn is Akn Then Sj is Bkj

where xi represents inputs variables, Aki represents input fuzzy sets (Low, Normal or High), Sj is the current state and Bkj represents output fuzzy sets (Low or High). Additionally, two rules are added to perform fault detection.

To account for temporal changes in the variables trends during the faulty behavior of the process, the automatic rules extraction algorithm allows to define several sample zones. Then, a set of rules is extracted for each zone, the new rules that are not redundant are introduced in the rule-base. Obviously, the number and location of the sample zones influence the detection and diagnosis performance.

By means of this FLS automatic design, a flexible framework which is able of continuously updating when new abnormal situations are detected, is achieved. Furthermore, it allows to implement an optimization routine to improve the FDS robustness.The optimization procedure proposed involves two main actions:

1: the FDS performance is initially evaluated using different samples from historical data sets. Depending on the information gathered by each kind of sample from historical data, different rules sets are achieved. The most representative sample, the best FDS performance obtained. After this searching, the best sample zone is chosen.

2: then, a genetic algorithm to find the best parameters of fuzzy logic membership functions is implemented to improve the FDS resolution. The objective function proposed to drive this optimization, takes into consideration the diagnosis urgency. For this reason, a weight exponential function to put greater relevance to diagnosis at early moments after detection, is introduced in the overall objective function.

The methodology is validated in a simple case study (multicomponent distillation column) and in an industrial case study (The Tennessee Eastman benchmark).

As conclusions, a simple methodology to automatically extract rules from data is presented, taking special interest in maximizing classification transparency. Besides, a genetic algorithm-based optimization procedure is applied to optimize the overall system performance, maintaining the comprehensive structure. For both case studies, very good fault isolation has been obtained and noticeable fault detection improvements have been achieved if comparing with PCA results.

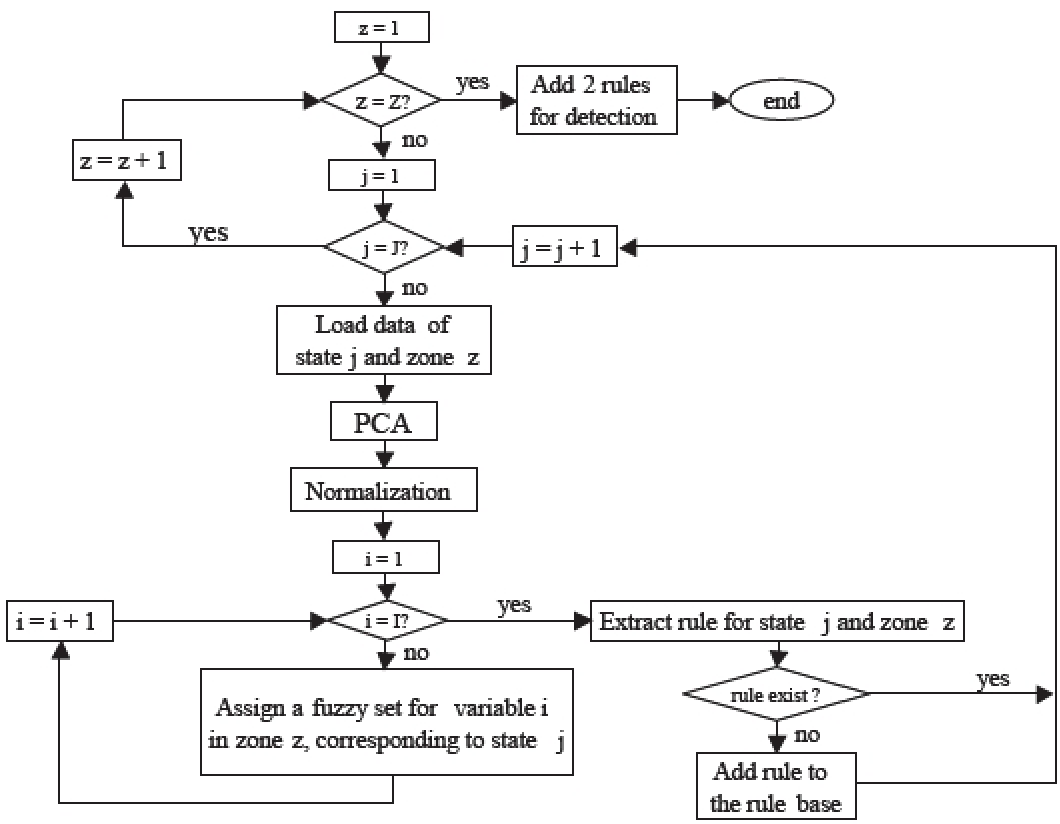

Fig. 1. Algorithm to extract rules from data; z: sample zone, j: process state, i: process variable