A NOVEL MODULAR NONLINEAR NETWORK FOR FAULT DIAGNOSIS AND SUPERVISED PATTERN CLASSIFICATION

B. Bhushan[*] and J. A. Romagnoli[**]

** Department of Chemical Engineering*
The University of Sydney NSW, 2006, Australia
Email: bharat@chem.eng.usyd.edu.au
*** Department of Chemical Engineering, Louisiana State University,*
Baton Rouge, LA 70803
Email: jose@lsu.edu

Abstract: A novel modular network is proposed in this work for supervised pattern classification. The parameters of the hidden layer are determined using polygonal line algorithm. No further training of the network is required. Firstly, an abnormality is detected and responsible sensors identified using polygonal line based radial basis function network algorithm. Furthermore, the proposed strategy is applied for fault diagnosis. A continuous pilot plant is selected as the case study to show the efficiency of the proposed strategy. The result shows that, the proposed framework is a promising direction towards fault detection and diagnosis in real time, non-linear systems.

Keywords**:** Fault detection, Fault identification, Fault diagnosis, Nonlinear PCA, Polygonal lines, Modular network

1. INTRODUCTION

The advent of faster and more reliable computer systems has revolutionized the manner in which industrial processes are monitored and controlled. Once thought of as just data logging and storage units, these computer systems now perform sophisticated computer-based control strategies, and real-time simulation and optimization. These advances have resulted in the generation of a large amount of process data, yet the task of interpreting and analysing these data is daunting.

Fault detection and diagnosis is the primary module for any process monitoring framework. Principal component analysis (PCA) and projection to latent structure (PLS) are one of the most used multivariate statistical process control (MSPC) techniques. The major drawback of this method is that, it assumes linear correlation between data which is not always true in case of process data that are generally nonlinearly correlated. However, the philosophy behind these approaches is to reduce the dimensionality of the problem by forming a new set of latent variable to obtain an enhanced understanding of the process behaviour.

Many methodologies have been proposed for nonlinear principal component analysis (NLPCA). Kramer (1991) proposed a NLPCA based on five layer auto associative neural networks. Dong and McAvoy (1996) proposed NLPCA based on principal curves and neural networks. The principal curve method was used to calculate the associated score and corrected data point for each original data point. But, since principal curve method does not produce a nonlinear principal component in the sense of principal loading, Dong and McAvoy (1996) developed an alternative approach based on multi layer perceptron to model the calculated data. Two three layer neural networks were trained separately to map the data to lower dimensional feature space and remapping the data back to the sample space. The number of hidden layer nodes was decided using cross validation scheme. A methodology based on "well – defined" architecture of radial basis function (RBF) network and polygonal line (PL) has been suggested by Bhushan and Romagnoli (2005) for dimensionality reduction and fault detection. The data of the normal operating region is used to fit the polygonal lines and the output generated has been used for determining the architecture of the network and to train the model. Online data are projected to the RBF-PL model and an abnormality is indicated whenever the prediction is significantly different from the projected measurements. Furthermore, the measured variables that makes significant contribution towards the deviation in the model prediction is identified. However, this information is insufficient for the operator to find the root cause, since the operator needs to infer the root cause which is difficult in case of process with large number of variables.

Vedam and Venkatasubramanian (1999) proposed an integrated approach based on PCA and signed digraphs (SDG) for fault detection and diagnosis. Fault detection is performed using PCA. Whenever an abnormality is detected, the contribution of measured variables is presented as an input to the SDG to perform fault diagnosis.

Leonard and Kramer (1992) suggested a decomposition strategy based on modular neural network approach for solving large scale fault diagnosis problems. Though, RBF networks are many time faster than similar back propagation networks (BPN); it still required large computational resources. Two decompositions are proposed for this work: decomposition in time, reducing the dimensionality of the input space; and decomposition among the fault classes, reducing the size of the training set for each subnet.

In this work, a novel modular network is suggested to accomplish the diagnosis task. The advantage of this methodology over others is that it uses the same PL algorithm of fault detection to decide the architecture and related parameters of each module of the network. Furthermore, there is no additional training required of the network and hence it is computationally very less expansive. The output of the proposed network can be the partial belongingness of the input pattern to more than one fault classes and the strength of the fault.

The remaining part of the paper is organized as follows. In section 2, a brief introduction of RBF-PL methodology for fault detection and identification is given. The proposed network for classification is explained in section 3. Section 4 contains the results and discussion on the application of the entire strategy to a real time pilot plant environment. Finally, section 5 contains the conclusion and future direction of the work.

2. FAULT DETECTION AND IDENTIFICATION

Polygonal line algorithm proposed by Verbeek, *et al*., 2002 is used for fitting the data. Each data point is projected orthogonally onto the PL. Thus for each data points there are corresponding lengths $t_1, t_2, \ldots, t_n$ along the curve where n is the number of data points in d-dimensional space.

In analogy to PCA, this length represents the non-linear scores of the data points. Thus the sample vector can be represented as

$$X = f_1\,(t(X)) + E_1 \tag{1}$$

where t is the non-linear principal component score and $\boldsymbol{E_1}$ is the residual vector.

The next non-linear component score can be found by projecting the data points of $\boldsymbol{E_1}$ on the PL constructed using $\boldsymbol{E_1}$. These steps are repeated until all the information is extracted. It is found that the first few nonlinear principal components explain most of the variance of the dataset. Though, this method is quite effective in reducing the data dimensionality, it is to be noted that $\boldsymbol{f}$ has no parametric form, and it is quite cumbersome and memory expensive to use it for online application. A RBF network is trained to model the relationship. This network is further used for fault detection and identification. A non-parametric approach based on kernel density estimation (KDE) is used to determine the confidence limit. The detail of this methodology can be found in Bhushan *et. al.* (2005).

3. SUPERVISED CLASSIFICATION AND FAULT DIAGNOSIS

It should be noted here that each segment of the PL in the sample space represents a localised region around which the data is concentrated. We propose that each of these regions in the input space can be represented by a multidimensional Gaussian function (Figure 1).

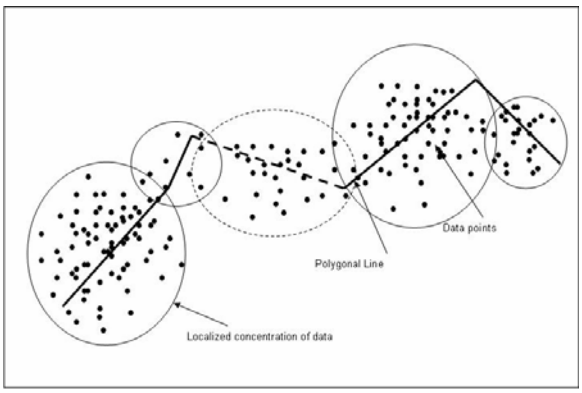


Fig. 1: Schematic representation of the region covered by each segment of PL.

The Gaussian function with equal spread in all the directions is defined as:

$$\varphi(x;c,\sigma) = exp\left(-\frac{\|x-c\|^2}{2\sigma^2}\right) \qquad (2)$$

where $\| \boldsymbol{x} - \boldsymbol{c} \|$ is the Euclidean distance of $\boldsymbol{x} = (x_1, x_2, \ldots, x_d)$ from the vector centre $\boldsymbol{c} = (c_1, c_2 \ldots c_d)$ and σ is the spread. When the spread of the data points is not uniform, a multidimensional Gaussian function takes the form

$$\varphi^i(x, c^i, \sigma^i) = exp\left(-\left[\frac{(x_1 - c_1^i)^2}{2\sigma_1^{i\,2}} + \ldots + \frac{(x_d - c_d^i)^2}{2\sigma_d^{i\,2}}\right]\right) \qquad (3)$$

where c^i is the centre of the region and is defined as the mean of the data contained in the region i and σ^i represents the standard deviation of the dataset in the region.

Since it is the supervised classification, the class of each training data set is known in advance. The training data set is grouped according to its class. Each group is presented to the PL algorithm. The number of segments required to fit the polygonal line is found out. However, segments which have just been used to construct the PL and do not contain any data points are neglected.

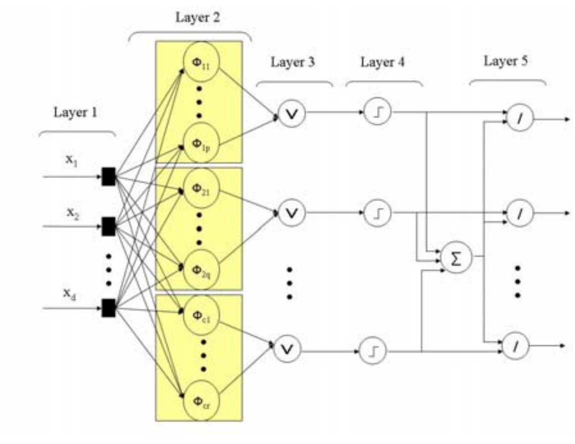


Fig. 2: Proposed modular fault diagnosis network.

Figure 2 shows the proposed modular fault diagnostic system. The system contains five layers. Nodes at layer one are input nodes representing the input variables and the last layer is the output nodes. The number of nodes in the output layer is same as the number of fault classes. A node at layer two represents a region in the domain of a specific fault, in other words, it is one of the segment of the PL which fitted that fault class. The nodes of a fault class constitute one module in layer two and there will be as many modules as the number the fault classes. The number of nodes at layer three is the same as the fault classes and each node of this layer is linked to the nodes of only one module of layer two. The nodes in this layer calculate the maximum strength of the input data in that particular class. Each node in layer four is linked with only one node of layer three and decides whether the strength of the belongingness is strong enough to be considered as a fault or not. Finally, the nodes in layer five decide the contribution of each fault class to the abnormality.

The function of a node in each layer of the proposed network is described in detail next.

Layer 1: The input vector is presented to the nodes of this layer. The number of nodes is same as the dimension of the input vector and each element is linked to one of the node of this layer. The nodes in this layer transmit the input data to the next layer without any change.

The output from this layer y_j^1 is defined as

$$y_j^1 = x_j^1, \; j = 1, \ldots, d \qquad (4)$$

Where y_j^1 denotes the output of node j in layer one and x_j^1 denotes the input to node j at layer one.

Layer 2: The output of each node of layer one is presented to each node of this layer. This is one of the most important layers of this network. As mentioned earlier, the parameters of each node are determined using the data of that region.

$$c_{j,c}^k = \frac{\sum_{i=1}^{I_k} x_{j,c}^i}{I_k} \qquad c = 1,2,\ldots,C; \; j = 1,2,\ldots,d \qquad (5)$$

where $c_{j,c}^k$ is the centre of the region k in class c, I_k represents the data points in region k and C is the number of classes or modules. The spread of k^{th} node in class c is defined as

$$\varsigma_{j,c}^k = \delta.\sigma_{j,c}^k \quad c = 1,2,\ldots,C; \; j = 1,2,\ldots,d \qquad (6)$$

where $\sigma_{j,c}^k$ is the standard deviation of the input vectors contained in region k of class c in j^{th} dimension and δ is a user defined parameter which ensures the optimum receptive field covered by each region.

A Gaussian membership function is constructed with $c_{j,c}^{k}$ and $\varsigma_{j,c}^{k}$ as the centre and the width respectively. Each node in this layer is represented by one such function and all the nodes generated by using segments of the PL of a group constitute a module. Therefore, we have as many modules as the number of classes. The output from this module defines the belongingness of a input vector to a particular region.

$$y_c^{2,k}(x,c_c^k,\varsigma_c^k) = exp\left(-\left[\frac{(y_1^1 - c_{1,c}^k)^2}{2\,\varsigma_{1,c}^{k\;2}} + \ldots + \frac{(y_d^1 - c_{d,c}^k)^2}{2\,\varsigma_{d,c}^{k\;2}}\right]\right) \tag{7}$$

where $y_c^{2,k}$ is the output from node representing region k of class c in layer two.

Layer 3: The output from each module is fed to not more than one node of this layer. Hence the number of nodes is same as the number of fault classes. The output of the node from this layer represents the strength of the belongingness of the input data in a particular class. An input vector may belong to more than one region of the same class; however, the belongingness of the vector in a class is dictated by the maximum membership value. Therefore, the output of each node from this layer is defined as:

$$y_c^3 = max(y_c^{2,1}, y_c^{2,2}, \ldots, y_c^{2,k}) \tag{8}$$

where *k* represents the number of regions in class *c*.

Layer 4: The nodes in this layer decide whether the output from previous layer is strong enough to assign the data in to a particular class or not. This task is accomplished by a function defined as follows:

$$y_c^4 = \begin{cases} y_c^3 & if \;\; y_c^3 \geq \lambda \\ 0 & if \;\; y_c^3 < \lambda \end{cases} \tag{9}$$

where λ is a user defined parameter.

Layer 5: This layer is the decision making layer. It gives an idea to the operator which fault is more severe if there are multiple faults. The output from this layer is defined as:

$$y_c^5 = \frac{y_c^4}{\sum_{k=1}^{C} y_k^4} \tag{10}$$

It should be noted that once the network is build there is no further requirement of the training since there is no weight adjustment required.

4. APPLICATION TO PILOT PLANT ENVIORNMENT

To test the overall strategy in real time, a general-purpose pilot plant facility is used. The process contains two CSTRs, a mixer, a feed tank and a number of heat exchangers.

Each CSTR consists of a reaction vessel, a steam jacket, a cooling coil and a stirrer. Material from the feed tank is heated before being fed to the first reactor and the mixer. The effluent from the first reactor is then mixed with the material in the mixer before being fed to the second reactor. The effluent from the second reactor is, fed back to the feed tank and the cycle continues. The pilot plant is well instrumented to provide many possible control scenarios and configurations.

Nine variables [Fin (feed flow rate in), Tin (temperature of feed in), Tc,in (temperature of cooling water in), Ts,in (temperature of steam in), Lvl (level of the reactor), Fout (feed flow rate out), Tout (temperature of feed out), Tc,out (temperature of cooling water out), Ts,out (temperature of condensate)] related to the first CSTR are considered for this study. Once the plant reached its normal operating condition, 100 training data points at 5 second interval were collected. All variables in the training data set were normalized in order to give equal weights to each. The training data set were exposed to PL algorithm (kmax = 24) and the nonlinear scores were found by projecting the data point onto the polygonal line. The residual was calculated and was exposed again to PL algorithm and so on. The PL algorithm fitted the training dataset into nine segments yielding a RBF mapping network with nine input nodes, nine hidden layer nodes and two output nodes. The centre and spread of the hidden layer nodes were calculated using the centre and standard deviation of the segments and hence only weights of the output layer were to be calculated. GA with 100 generations was used to first get near an optimum solution followed by BFGS Quasi-Newton algorithm for faster convergence. The total time taken for training was 12.80 seconds. The demapping layer with two input nodes, nine hidden layer nodes and nine output layer nodes was trained using the similar strategy, however all the parameters were trained and hence the training time was 35.46 seconds.

KDE is used for finding the two warning limits at 95% and 99%. 95% and 99% warning limit was found to be 9.19592 and 10.8476 respectively. SPE along with these warning limits were constructed to facilitate fault detection. Any violation to the 95% warning limit, fault identification algorithm is triggered to identify the measurements responsible for out of control signal followed by the fault diagnosis model to decide the root cause.

For this paper, two different fault scenario were planned. Firstly, process condition under normal operation was achieved that was similar to when data were collected to train the network. Secondly, to simulate a process upset scenario, the feed flow rate was increased from 0.6 l/min to 1.0 l/min. This change affected many other variables including the reactor level and the effluent flow rate. For simulating the single sensor failure condition, a random bias of mean 6 (30% of the actual) and standard deviation 2 was added to the feed temperature. 100 values at 5 second interval for all these conditions were captured.

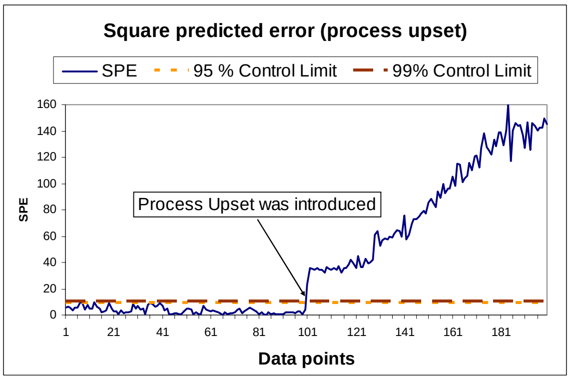
The data of these three conditions (normal, process upset and sensor failure) were fed to the PL algorithm which fitted it into 9, 22 and 12 regions respectively. Therefore, the structure of the fault diagnosis network was 9-43-3-3-3. The value of δ and λ used are 1.5 and 0.05 respectively. The result of the fault detection and diagnosis are as follows:

1) Process Change (Flow rate increased from 0.6 l/min to 1.0 l/min)

Figure 3(a) and 3(b) shows the SPE and contribution plot for process upset respectively. It should be noted that as soon as feed flow rate was increased, SPE crossed both the warning limits and was well above this condition throughout the period this condition prevailed.

However, the contribution plot shows that though the contribution by feed flow rate was high in the beginning, in the later part, prime contribution was due to the level measurement which is readily expected for this process. Also, this change in the process condition affected the effluent flow rate. From the process point of view, these three variables are related to each other and they are also being reflected in the result.

(a)



(b)

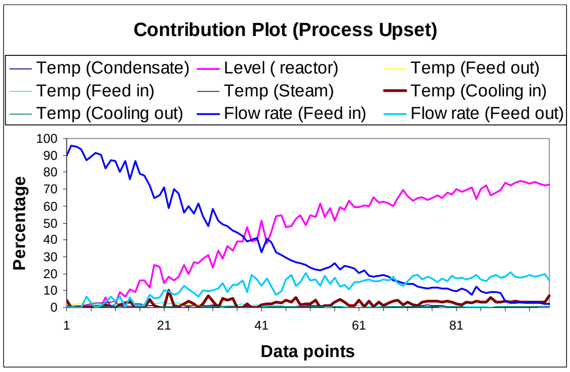


Figure 3: (a) Square predicted error plot (b) contribution plot in case of process upset.

Table 1: Results of the fault diagnosis in case of process upset

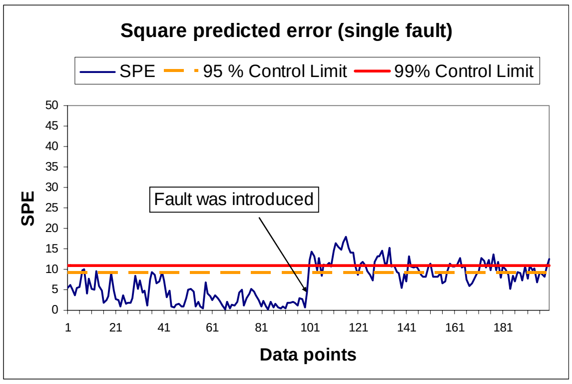
Normal	Process Upset	Sensor Failure
0.00	1.00	0.00
0.00	1.00	0.00
0.00	1.00	0.00
0.00	1.00	0.00
0.00	1.00	0.00
0.00	1.00	0.00
0.00	1.00	0.00
0.00	1.00	0.00
0.00	1.00	0.00
0.00	1.00	0.00

Table 1 shows the result of the proposed fault diagnostic system. The result for first ten points is only shown however it detected correctly for all the testing data. In the result, 0 indicates that this condition is not prevailing in the process whereas 1 indicates that according to the knowledge of the network this condition is 100% present.

2) Sensor failure (a random noise of 30% of the actual with std. dev of 2 was added to feed temperature)

In case of sensor failure, SPE is above the warning limits in most cases (Fig. 4(a)), though in some cases the magnitude of SPE is not very high because of the presence of noise in the measurements the normal feed temperature is quite close to the value in case of the faulty sensor (maximum normal feed temperature: 28.83 0C, minimum feed temperature in case of sensor fault: 31.13 0C). The contribution plot (Fig. 4(b)) clearly identify feed temperature sensor as the sensor which has the highest contribution in fault.

(a)



(b)

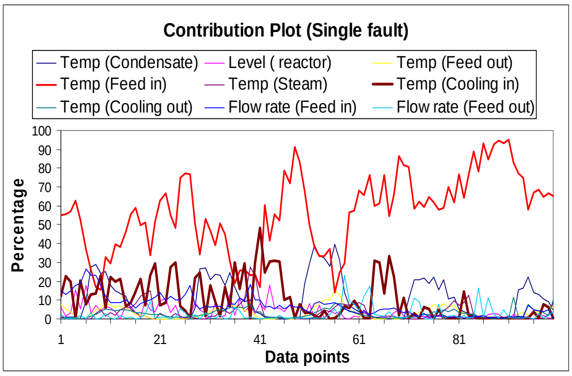


Figure 4: (a) Square predicted error plot (b) contribution plot in case of single fault.

The result of the fault diagnostic module is shown in table 2. It should be noted that a module for this fault was already included in the network; hence it did diagnose all the conditions correctly.

Table 2: Results of the fault diagnosis in case of Sensor failure

Normal	Process Upset	Sensor Failure
0.00	0.00	1.00
0.00	0.00	1.00
0.00	0.00	1.00
0.00	0.00	1.00
0.00	0.00	1.00
0.00	0.00	1.00
0.00	0.00	1.00
0.00	0.00	1.00
0.00	0.00	1.00
0.00	0.00	1.00

5. CONCLUSION AND FUTURE WORK

In this work, a novel modular network is proposed for supervised classification. The key feature of this network is its simplicity and less computational complexity. First, the training data were separated into different classes, and each set was fed to the PL algorithm for finding the optimum number of regions in that class. Each region was used to find out the parameters of the network. There was no further training required. This work is integrated with non linear PCA based on RBF and PL for simultaneous fault detection and diagnosis. The proposed methodology is used for monitoring the condition of a continuous pilot plant. Two different types if abnormalities were simulated to test the capability of the framework. The results show that the fault was detected and diagnosed in both of the cases correctly.

However, the model needs to be validated for controller faults and multiple faults which will be a part of future work.

6. REFERENCES

Dong, D., and T.J. McAvoy (1996). Nonlinear Principal Component Analysis-Based on Principal Curves and Neural Networks. *Computers Chem.. Engng.*, **20**, 65-78.

Bhushan, B., J A Romagnoli and D. Wang (2005). Fault Detection Using Radial Basis Function Network and Polygonal Line. *16th IFAC world congress*, Prague, Czech Republic.

Kramer, M. A. (1991). Nonlinear Principal Component Analysis Using Autoassociative Neural Networks. *AIChE Journal*, **37**, 233-243.

Leonard, J. A. and M. A. Kramer (1993). A Decomposition Approach to Solving Large-Scale Fault Diagnosis Problems with Modular Neural Networks. *IFAC Symposia Series*,1, 237-242.

Looney, C.G. (1996). *Pattern Recognition Using Neural Networks*. Oxford University Press, Oxford.

Verbeek, J. J.,, N. Vlassis and B. Krose (2002), A k-segments algorithm for finding principal curves. *Sample Recognitions Letters*, **23**, 1009-1017.

Venkatasubramanian, V., R. Rengaswamy, S. N. Kavuri and K. Yin (2003), A Review of Process Fault Detection and Diagnosis Part III: Process History based Methods. *Computers and Chemical Engineering,* 27, 327-346.

Vedam, H. and V. Venkatasubramanian (1999), PCA-SDG based Process Monitoring and Fault Diagnosis. Control Engineering Practice, 7, 903-917.