Scientia Iranica, Vol. 10, No. 3, pp 300–310

SCIENTIA
IRANICA

Multiple Simultaneous Fault Diagnosis via Hierarchical and Single Artificial Neural Networks

R. Eslamloueyan[1], M. Shahrokhi* and R. Bozorgmehri[1]

Process Fault Diagnosis (PFD) involves interpreting the current status of the plant given sensor readings and process knowledge. There has been considerable work done in this area with a variety of approaches being proposed for PFD. Neural networks have been used to solve PFD problems in chemical processes, as they are well suited for recognizing multi-dimensional nonlinear patterns. In this work, the use of Hierarchical Artificial Neural Networks (HANN) in diagnosing the multi-faults of a chemical process are discussed and compared with that of Single Artificial Neural Networks (SANN). The lower efficiency of HANN, in comparison to SANN, in PFD is elaborated and analyzed. Also, the concept of a multi-level selection switch is presented and developed to improve the performance of hierarchical artificial neural networks. Simulation results indicate that application of multi-level selection switches increases the performance of the hierarchical artificial neural networks considerably.

INTRODUCTION

Fault detection and diagnosis is an important task in process engineering. A fault might originate from outside the plant battery limit (e.g. fluctuation in flow rate, temperature or pressure of the utility streams), or from inside the plant itself (e.g. equipment failure, catalyst deactivation or malfunction of instrument and control valves). Although based on the HAZOP (Hazard and Operability) studies, a plant is normally equipped with the required interlock and protection systems, however, these facilities are usually activated only when the symptoms of the faults are so dangerous that the plant, or at least part of it, should be shut down. It is evident that the emergency shut down of a plant is economically undesirable. On the other hand, even if emergency shut down does not occur, the yield and product quality of a plant operating in an abnormal situation (i.e. process variables deviate significantly from their nominal values) would be low.

Industrial statistics now estimate the economic impact due to emergency shut down and abnormal situations to be around $20 billion a year in the US petrochemical industries alone [1]. If the value of losses caused by abnormal situations of all the process plants in the world were estimated, the importance of abnormal situation management would be more clarified.

The first step in abnormal situation management is to detect the process of the abnormal condition and diagnose the faults creating this situation (i.e. process fault detection and diagnosis). Today's process plants are very complex with a lot of measured variables for plant monitoring and, hence, Process Fault Diagnosis (PFD) in such plants is a very difficult task, even for an experienced operator. Therefore, designing an intelligent real time system for PFD has received considerable attention both from industry and academia, due to the economic and safety impact involved [2].

Although there are various methods for PFD in open literature, based on the form of process knowledge used by these methods, they can be classified as process model-based and process history-based techniques. The models used in the former category can be a quantitative deep model [3] or a qualitative causal model, such as a signed digraph [4]. The major restriction of quantitative model-based methods is the fact that finding a rigorous model for a plant consisting of so many unit operations is either very expensive or impossible. On the other hand, the qualitative

1. *Department of Chemical Engineering, Sharif University of Technology, P.O. Box 11365-9465, Tehran, I.R. Iran.*
*. *Corresponding Author, Department of Chemical Engineering, Sharif University of Technology, P.O. Box 11365-9465, Tehran, I.R. Iran.*

model-based methods, such as signed digraph and fault tree techniques, suffer from the generation of a large number of hypotheses leading to a poor resolution. Furthermore, due to intensive computations, the causal qualitative model-based algorithms are not suitable for on-line process fault diagnosis. However, it should be noted that these methods usually result in a complete set of possible fault candidates and, because of mimicking human reasoning, they posses the power to explain how the fault originated and propagated.

The process history-based methods make use of the large amount of process data obtained from recorded measured variables of the plant during abnormal and normal situations. This category consists of techniques like expert systems and statistical and neural network methods. Expert systems are knowledge-based techniques which are able to represent existing expert knowledge, accommodate existing data bases and accumulate new knowledge, leading to practical and justifiable decisions. The knowledge-base can be obtained from heuristics, expert testimony (shallow knowledge) and/or structural behavioral and mathematical models (deep knowledge) of the process. Several recent papers describing applications of expert systems in process fault diagnosis are available [5-12]. Multivariate statistical techniques, based on Principal Component Analysis (PCA), Partial Least Squares (PLS) and Fisher Discriminant Analysis (FDA), have been employed for fault detection and diagnosis [13-18].

In this work, a Hierarchical Artificial Neural Network (HANN) and a Single Artificial Neural Network (SANN) for multiple-fault diagnosis are discussed and their performances are compared. The use of neural networks in process fault diagnosis has received considerable attention over the last few years. Venkatasubramanian et al. proposed a SANN for process fault diagnosis and compared its capabilities with a knowledge-based approach [19]. The potential of their method has been shown through a case study. According to their studies, the SANN was able to diagnose both the faults it was trained upon and novel fault combinations not included in the training data. Furthermore, the SANN could also handle incomplete and uncertain data. Watanabe et al. showed that a two-stage HANN could discriminate between the causes of faults and, also, identify the degree of fault deterioration [20]. In the first stage, the fault is diagnosed by a three-layer network and depending on the type of fault detected in the first stage, the corresponding network in the second stage is triggered to determine fault deteriorations. In another attempt, Fan et al. enhanced the representation capability of a conventional back-propagation network for diagnosing faults with various levels of deterioration, by adding to the input layer a number of functional units (sine and cosine functions) [21]. Watanabe et al. developed a HANN for diagnosing multiple simultaneous faults of a chemical process [22]. In their approach, a set of neural networks was used in a hierarchical manner. These networks were arranged in two stages. The first stage, which consisted of one neural network, was used to discriminate between normal and single fault patterns. The second stage, whose role was distinguishing various combinations of binary faults, consisted of a set of conventional feed forward networks. The number of networks in the second stage was the same as the number of outputs of the network in the first stage. They claimed that the proposed HANN could be trained easier and faster than the other neural net-based alternatives. Some other neuromorphic approaches have been suggested for process fault diagnosis [23-30].

In this paper, the structure and performance of HANN is critically reviewed. It is shown that a single conventional artificial neural network outperforms a HANN, provided that it is trained with an effective and powerful training algorithm. Also, by using a multi-level selection switch, the performance of the HANN is improved significantly. The paper is organized as follows. Firstly, the proposed two-stage and three-stage hierarchical artificial neural networks are presented and their training procedure is briefly described. Secondly, the proposed single artificial neural network and the training technique are explained. Thirdly, the simulated process used for generating the training and validating fault patterns is described. Finally, the trained hierarchical and single artificial neural networks are applied to diagnose the process faults and the advantages and disadvantages of these methods are discussed in detail. Also, the concept of a multi-level selection switch is described and applied to the two- and three-stage hierarchical artificial neural networks.

HANN STRUCTURE AND TRAINING METHOD

The HANN structure used here for evaluation is a two-stage HANN proposed by Watanabe et al. [22]. The first stage consists of a conventional backpropagation network which receives the process measurements and discriminates normal, single and multiple fault patterns. If normal condition is detected in the first stage then the output of this stage will be considered as the HANN output, otherwise, based on faults detected in the first stage, the corresponding networks in the second stage are activated. Since the neural networks in the second stage discriminate single fault and binary fault combinations, the total number of neural networks in the second stage is equal to the number of single faults. For multiple faults, the fault set diagnosed by HANN is obtained through

performing an "OR" operation on the output of the second stage networks. The HANN structure is shown in Figure 1, where R/B is a real to binary number converter which receives a real number from zero to one. If the input value to R/B is greater than, or equal to, a predetermined threshold (in this case 0.5), its output will be 1, otherwise it will be set to zero. In this figure, NSS (Network Selection Switch) specifies those networks in the second stage which should be triggered.

The number of neurons in the input and output layers of each SANN are equal to the number of input measurements and number of patterns that the SANN is supposed to classify, respectively.

The main objective of this HANN structure is to divide, systematically, a large number of patterns into smaller subsets, in order to facilitate the training of the networks and classification of the faults.

The HANN structure is suitable for classifying single, double and, at most, triple-fault patterns. To accommodate the triple-fault patterns in the training

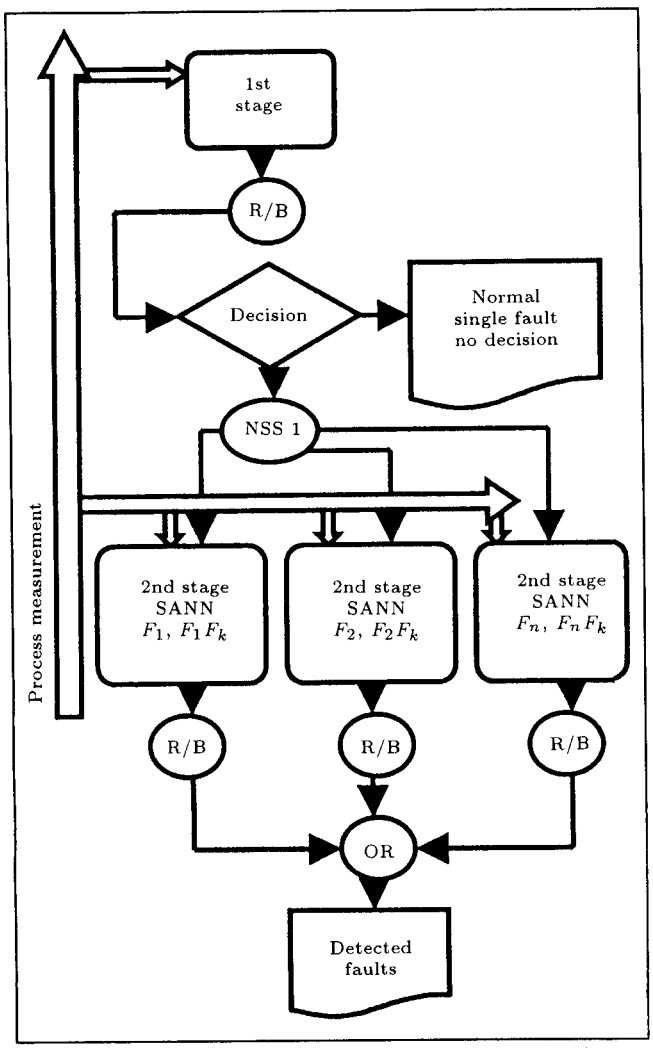

Figure 1. Schematic diagram of a two-stage HANN.

set of the HANN, based on the design concept of the two-stage HANN, a three-stage HANN has been developed, as shown in Figure 2. The first and second stages of the three-stage HANN are similar to the two-stage HANN. The third stage consists of a set of neural networks, where each network is responsible for discriminating a given double fault and its corresponding triple faults. If the result of the 'OR' operation after the second stage networks is a single or double fault then these faults will be considered as the final HANN's output and the third stage networks will not be triggered. On the other hand, when the second stage 'OR' operation generates more than two possible faults, all binary combinations of these faults are generated and the network selection switch 'NSS2' triggers third-stage corresponding networks. For example, if the fault set created by a second-stage 'OR' operation consists of the faults F_i, F_j and F_k then the third-stage networks corresponding to F_iF_j, F_iF_k and F_jF_k will be activated. In this case, the final result of the HANN is the fault set obtained from the third-stage 'OR' operation.

Each network in two-stage and three-stage hierarchical neural networks is trained independently, based on its own training patterns. All networks are conventional backpropagation networks with log-sigmoid neurons.

The training algorithm used here is called variable learning-rate backpropagation with momentum. When the momentum filter is added to the parameter changes, one obtains the following equations for the training algorithm:

$$\Delta W^m_{k+1} = \gamma \Delta W^m_k - (1-\gamma)\alpha s^m (a^{m-1})^T, \tag{1}$$

$$a^m = f^m_{(W^m a^{m-1} + b^m)}, \tag{2}$$

$$\Delta b^m_{k+1} = \gamma \Delta b^m_k - (1-\gamma)\alpha s^m,$$

$$\text{for } m = 1, \cdots, M, \tag{3}$$

where:

ΔW^m_k	the weight matrix update of layer 'm' at iteration 'k',
Δb^m_k	the bias vector update of layer 'm' at iteration 'k',
α	learning rate of the training algorithm,
γ	momentum of the training algorithm,
a^m	output vector from layer 'm' having transfer function 'f^m',
s^m	the sensitivity vector of mth layer.

The sensitivity vector is determined by the following

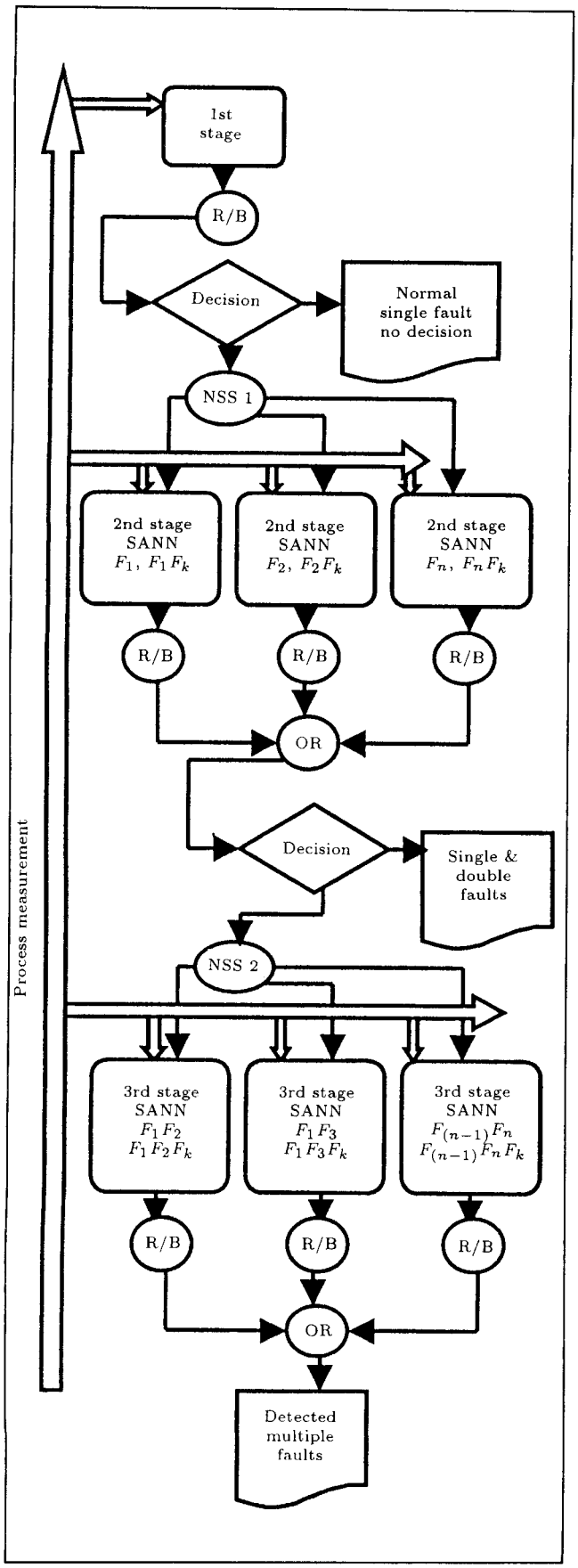

Figure 2. Schematic diagram of the proposed three-stage HANN.

equations:

$$\dot{F}^m_{(n^m)} = \begin{bmatrix} \dot{f}^m_{(n^m_1)} & 0 & \cdots & 0 \\ 0 & \dot{f}^m_{(n^m_2)} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & \dot{f}^m_{(n^m_{C^m})} \end{bmatrix}, \tag{4}$$

$$\dot{f}^m_{(n^m_j)} = \frac{\partial f^m_{(n^m_j)}}{\partial n^m_j}, \tag{5}$$

$$n^m_j = \sum_{i=1}^{s^{m-1}} W^m_{j,i}\, a^{m-1}_i + b^m_j, \tag{6}$$

$$s^M = -2\dot{F}^M_{(n^M)}(t-a), \tag{7}$$

$$s^m = \dot{F}^m_{(n^m)}(W^{m+1})^T s^{m+1},$$

$$\text{for } m = M-1, \cdots, 2, 1, \tag{8}$$

where:

C^m	total number of neurons in layer 'm',
n^m_j	net input to layer m,
f^m	transfer function of the neurons in the layer 'm',
M	total number of layers in the network,
$a^0 = P$	the input vector to first layer,
$a^M = a$	the output vector from the last layer,
t	target output vector.

There are different approaches for changing the learning rate. The approach used here is a straightforward batching procedure [31].

SANN STRUCTURE AND TRAINING METHOD

The SANN structure used in this paper for fault diagnosis is similar to the first stage network of the HANN described previously. Two structurally similar single artificial neural networks were trained separately, through applying a different set of learning patterns. The first SANN, which is called, henceforth, SANN1, has been trained by a set of normal, single and double-fault patterns. The data set used to train the second SANN (called SANN2) consists of normal, single and double, as well as triple fault patterns.

The synaptic weights and biases of SANN1 were determined by the method of variable learning-rate backpropagation with momentum. However, since SANN2 cannot be trained into a reasonably small error by this method, a specific type of conjugate gradient technique is used for its training. This technique, which is called Scaled Conjugate Gradient (SCG) algorithm, was proposed by Moller to avoid the

time consuming line search procedure of the conjugate gradient method [32]. Although the SCG algorithm is a relatively complex method, its basic idea is to combine the model trust region approach used in the conjugate Levenberg Marquardt algorithm, with the conjugate gradient approach.

SIMULATED PROCESS

The process considered here for simulation is the same process used by Watanabe et al. [22]. In this process, heptane is catalytically converted to toluene in a reactor heated by circulating steam. The reaction is endothermic and the required heat is supplied by an external electrical heater installed in the inlet line of circulating steam to the reactor. The schematic diagram of the process is shown in Figure 3. The following reaction takes place in the reactor:

$$C_7H_{16(g)} \rightarrow C_7H_{16(g)} + 4H_{2(g)}. \quad (9)$$

The mathematical model and steady state conditions used in the simulation of this plant have been presented in Appendix 1. As supposed by Watanabe et al. [22], it is assumed that the plant is operating in near normal

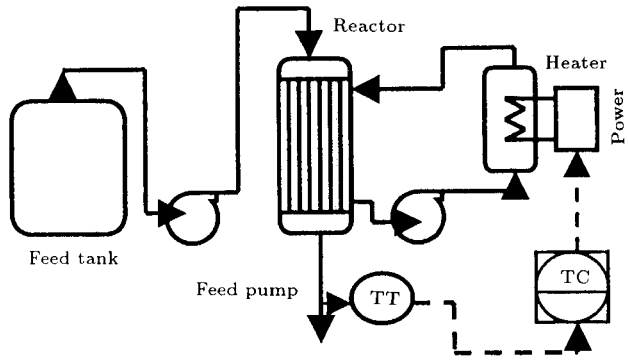

Figure 3. Heptane to toluene reaction unit.

operating conditions and any process faults evolved at an early stage tend to be strongly suppressed and alleviated by the PI controller. Therefore, no catastrophic or abrupt event occurs during plant operation. Assuming the controller is always on a non-faulty operating mode, the following seven possible causes of faults are considered for the plant. These are the same faults considered by Watanabe et al. [22]:

- F_1 0.1% decrease in the controller set point, u_0,
- F_2 2.0% decrease in the frequency factor of the rate equation, k_0,
- F_3 0.1% decrease in the activation energy, E_a,
- F_4 2.0% decrease in the heater gain, K,
- F_5 2.0% decrease in the overall heat transfer coefficient of the reactor, U,
- F_6 1.0% decrease in the feed concentration, $C^i_{C_7H_{16}}$,
- F_7 1.0% decrease in the feed temperature, T_i.

The following seven variables are measured for the process fault diagnosis [22]:

- T_h outlet temperature of electrical heater,
- T product temperature at the reactor outlet,
- S_i integrator output in the PI controller,
- $C_{C_7H_{16}}$ outlet concentration of heptane,
- $C_{C_7H_{16}}$ outlet concentration of toluene,
- $C^i_{C_7H_{16}}$ heptane concentration in the feed,
- T_i reactor feed temperature.

As mentioned before, the faults are not strong enough to push the process far from the linear operating region. Hence, based on the superposition principle, the pattern of measured variables resulting from multiple simultaneous faults can be produced by adding the patterns of the corresponding single faults separately. Table 1 presents normal and single-fault patterns obtained from the plant simulation.

For example, if $\Delta y_i, \Delta y_j$ and Δy_k are the deviation vectors of measured variables, resulting from single

Table 1. Normal and single-fault patterns from plant simulation.

Fault	Measured Input Pattern						
	ΔT_h °C	ΔT °C	Δs_i mV	$\Delta C_{C_7H_{16}}$ Mole/lit	$\Delta C_{C_7H_8}$ Mole/lit	$\Delta C^i_{C_7H_{16}}$ Mole/lit	ΔT_i °C
Normal	0	0	0	0	0	0	0
F_1	-1.85	-0.74	-1.19	5.55	-5.55	0	0
F_2	-0.83	0	-0.83	5.04	-5.04	0	0
F_3	0.90	0	0.90	-5.55	5.55	0	0
F_4	0.00	0	4.55	0.00	0.00	0	0
F_5	3.04	0	3.04	0.00	0.00	0	0
F_6	-0.86	0	-0.86	-4.76	-5.23	-10	0
F_7	0.43	0	0.43	0.00	0.00	0	-3

faults, F_i, F_j and F_k, then, the pattern of simultaneous three faults, $F_iF_jF_k$, is given by:

$$\Delta y_{ijk} = \Delta y_i + \Delta y_j + \Delta y_k. \quad (10)$$

As long as the plant operating condition is not far from the normal operating region, the accuracy of the above approximation for generating multiple-fault patterns is acceptable. It should be noted that noise signals were added to each fault pattern to produce the necessary amount of learning data for training each neural network. Thirty neurons have been considered here for the hidden layer, which is the same as the number of hidden layer neurons used by Watanabe et al. [22]. Experience in training SANN and the neural networks of HANN showed that having thirty neurons for the hidden layer is relatively good. In the next section, the training results and abilities of each diagnostic method are discussed. Also, the main reasons for HANN's shortcomings are explained and a new method is proposed to overcome some of them.

TRAINING RESULTS

All aforementioned neural networks have been trained by using noisy learning data. Test patterns are comprised of normal, single-fault, double-fault, triple-fault and quadruple-fault patterns. It should be noted that although the multiple-fault patterns, for training purposes, have been derived based on the superposition principle, the multiple-fault patterns, used in validating fault diagnosis techniques, have been obtained from the plant simulation. For instance, Tables 2 and 3 show patterns of double and multiple faults with higher degrees of deterioration, respectively.

Table 4 shows information corresponding to the training of various fault diagnosis techniques. The error trajectories during the training of SANN1 and SANN2 are shown in Figures 4 and 5, respectively.

Table 2. Double-fault patterns from plant simulation.

Fault	Measured Input Pattern						
	ΔT_h °C	ΔT °C	Δs_i mV	$\Delta C_{C_7H_{16}}$ Mole/lit	$\Delta C_{C_7H_8}$ Mole/lit	$\Delta C^i_{C_7H_{16}}$ Mole/lit	ΔT_i °C
F_1F_2	-2.68	-0.74	-2.01	10.6	-10.6	0	0
F_1F_3	-0.94	-0.74	-0.28	0.00	0.00	0	0
F_1F_4	-1.85	-0.74	3.35	5.55	-5.55	0	0
F_1F_5	1.17	-0.74	1.84	5.55	-5.55	0	0
F_1F_6	-2.70	-0.74	-2.03	0.74	-10.73	-10	0
F_1F_7	-1.42	-0.74	-0.75	5.55	-5.55	0	-3
F_2F_3	0.08	0.00	0.08	-0.51	0.51	0	0
F_2F_4	-0.83	0.00	3.71	5.04	-5.04	0	0
F_2F_5	2.20	0.00	2.20	5.04	-5.04	0	0
F_2F_6	-1.68	0.00	-1.68	0.23	-10.22	-10	0
F_2F_7	-0.40	0.00	-0.40	5.04	-5.04	0	-3
F_3F_4	0.90	0.00	5.48	-5.55	5.55	0	0
F_3F_5	3.97	0.00	3.97	-5.55	5.55	0	0
F_3F_6	0.04	0.00	0.04	-10.25	0.26	-10	0
F_3F_7	1.34	0.00	1.34	-5.55	5.55	0	-3
F_4F_5	3.04	0.00	7.66	0.00	0.00	0	0
F_4F_6	-0.86	0.00	3.68	-4.76	-5.23	-10	0
F_4F_7	0.43	0.00	4.99	0.00	0.00	0	-3
F_5F_6	2.17	0.00	2.17	-4.76	-5.23	-10	0
F_5F_7	3.48	0.00	3.48	0.00	0.00	0	-3
F_6F_7	-0.43	0.00	-0.43	-4.76	-5.23	-10	-3

DIAGNOSIS RESULTS

The validation fault patterns were introduced to the diagnostic schemes, SANN1, SANN2, two-stage HANN and three-stage HANN. The performances of these diagnostic schemes have been presented in Table 5.

Table 3. Patterns of multiple faults combined with different deterioration degree.

Fault	Measured Input Pattern						
	ΔT_h °C	ΔT °C	Δs_i mV	$\Delta C_{C_7H_{16}}$ Mole/lit	$\Delta C_{C_7H_8}$ Mole/lit	$\Delta C^i_{C_7H_{16}}$ Mole/lit	ΔT_i °C
F_1(1%)*, F_2(0.5%)	-2.70	-0.75	-2.03	10.65	-10.65	0.00	0.00
F_1(1%), F_3(0.5%)	-0.96	-0.75	-0.29	0.02	-0.02	0.00	0.00
F_2(0.5%), F_5(1%)	2.23	0.00	2.23	5.05	-5.05	0.00	0.00
F_1(0.5%), F_2(1%), F_5(0.5%)	0.32	-0.74	0.99	10.65	-10.65	0.00	0.00
F_1(0.5%), F_3(0.5%), F_4(0.5%)	0.95	-0.74	4.29	-0.01	0.01	0.00	0.00

* % parentheses is deterioration degree applied to previously defined faults.

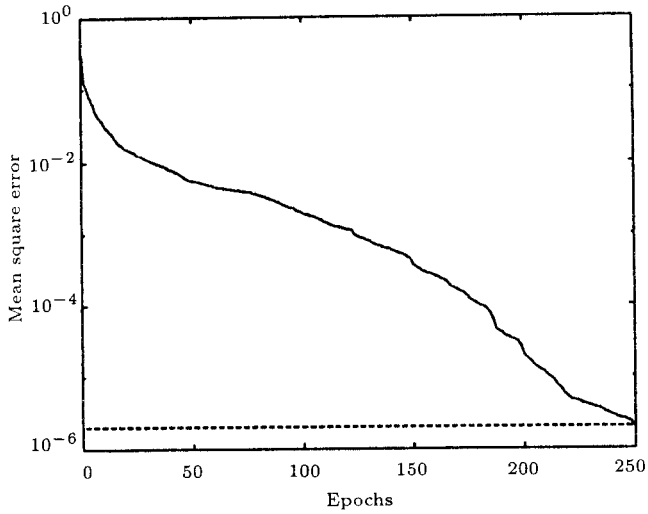

Figure 4. Trajectory of error convergence for SANN1.

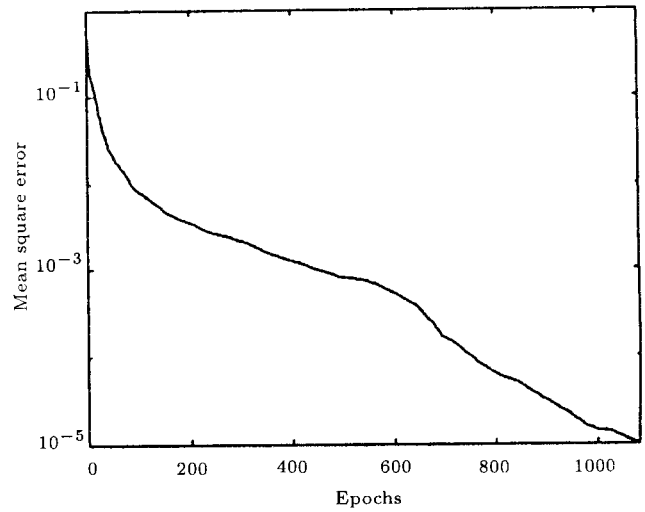

Figure 5. Trajectory of error convergence for SANN2.

Table 4. Some training characteristics of the diagnostic methods.

Diagnostic Technique	Number of Training Patterns	Training Method	MSE[1]	Epochs
SANN1	1 normal, 7 single-fault, 21 double-fault	Scaled conjugate gradient	2×10^{-6}	510
SANN2	1 normal, 7 single-fault, 21 double-fault and 35 triple-fault	Scaled conjugate gradient	10^{-5}	685
2-stage HANN	1st stage: 1 normal, 7 single-fault 2nd stage: 7 $(F_i, F_iF_j, i\#j)$ for each SANN	VRBPM[2]	10^{-5}	280-1299
3-stage HANN	1st stage: 1 normal, 7 single-fault 2nd stage: 7 $(F_i, F_iF_j, i\#j)$ for each SANN 3rd stage: 6 $(F_iF_j, F_iF_jF_k, k\#i, j)$	VRBPM	10^{-5}	19-1299

1- Mean squared error.

2- Variable-learning rate backpropagation with momentum.

Table 5. Responses of different schemes to validation fault patterns.

Fault Pattern	Response			
	Two-Stage HANN	Three-Stage HANN	SANN1	SANN2
Normal and Single Faults	All	All	All	All
Double Faults	17(2/2), 1(1/2), 3(0/2)	17(2/2), 1(1/2), 3(0/2)	All	All
Triple Faults	9(3/3), 16(2/3), 9(1/3), 1(0/3)	9(3/3), 16(2/3), 9(1/3), 1(0/3)	19(3/3), 7(2/3), 9(1/3)	All
Quadruple Faults	1(4/4), 3(3/4), 29(2/4), 2(0/4)	1(4/4), 3(3/4), 29(2/4), 2(0/4)	3(4/4), 11(3/4), 21(2/4)	20(4/4), 5(3/4), 10(2/4)
Higher Degree Double and Triple Faults	3(2/2), 2(1/3)	3(2/2), 2(1/3)	3(2/2), 1(2/3), 1(1/3)	All

Note: Numbers in parentheses represent the number of faults detected over total number of faults.

As can be seen from this table, all schemes have been able to diagnose normal and single-fault patterns. For double faults, SANN1 and SANN2 yield completely correct diagnoses. However, two-stage and three-stage hierarchical neural networks equally result in seventeen full diagnoses, one partial diagnosis and no detection for three cases. The correct diagnoses of SANN1 and SANN2 in the case of double faults is not unusual, because double fault patterns have been considered in their training set. The failure of two and three-stage hierarchical neural networks can be explained as follows: The partial and "no detection" results of the hierarchical neural networks in diagnosing double faults are due to the first stage network, which cannot, appropriately, activate the second stage networks. With respect to triple-faults, SANN2 exhibits excellent performance and all faults are detected correctly. Both SANN1 and the hierarchical neural networks have not been trained by triple-fault patterns, but, as Table 5 shows, the performance of SANN1 in diagnosing triple faults is superior to hierarchical neural networks. The main reason for this discrepancy can be found in the deficiency of the network selection switch (NSS1) in the first stage of the HANN. Indeed, as stated before, the first stage uses a rigid threshold value (here 0.5) to select the associated network in the next stage, which, in some cases, results in the wrong selection of the correct network or no selection at all. This is why the addition of a new stage to the two-stage HANN is futile. Regarding the quadruple fault diagnosis, the performance of SANN2 is still much better than other schemes and more than 57% of faulty situations have been fully diagnosed by this scheme.

The extrapolation capability of SANN2 to the validation data, which have not been considered in its training pattern set, is relatively good. The last row of Table 5 indicates the response of each method to the multiple faults with higher degrees of deterioration. In this case, the diagnosis of SANN2 is perfect while other schemes show some mistakes. It should be noted that despite the partial detection result shown for two-stage HANN in the last row of Table 5, Watanabe et al. presented a full detection response for this method in diagnosing higher degree multiple faults [22]. However, this discrepancy has no significant effect on the fact that the performance of SANN2 is better than other schemes.

In order to improve the performance of the hierarchical neural networks, the concept of a Multi-Level Selection Switch (MLSS) is proposed. To evaluate the power of MLSS in comparison to previous network selection switches NSS1 and NSS2, a Three-Level Selection Switch (TLSS), with two threshold values and a few general decision rules, is developed. Based on whether the output neuron values of the first stage network are less than 0.33 and greater than 0.67, in the range of 0.33-0.67, neural networks of the second stage are classified impossible, probable and most probable, respectively. Then, the following rules are utilized for activating the second stage networks:

- If the output of the first neuron in the output layer of the first-stage network is "most probable", then the output of the two-stage HANN will be a normal case and the second stage networks will not be triggered;
- Otherwise, "probable" and "most probable" networks of the second stage are triggered.

The output neurons of the triggered networks in the second stage are classified in the above manner as "impossible", "probable" and "most probable". Then, by applying the following decision rule, the process faults are diagnosed:

> The "probable" and "most probable" output neurons of the "most probable" second stage networks and, also, the "most probable" output neurons of the "probable" second stage networks are considered as faults.

The above three-level selection switch was used for a two-stage HANN. In the same way, a TLSS was developed and implemented for a three-stage HANN. The simulation results for the modified two- and three-stage hierarchical neural networks have been presented in Table 6.

Figure 6 indicates the overall performances of different schemes used in this work. In this figure, percentages of partial, full and no detection for each diagnostic scheme over the validation patterns have been demonstrated. According to Figure 6, the best performance belongs to SANN2 and, after that, the

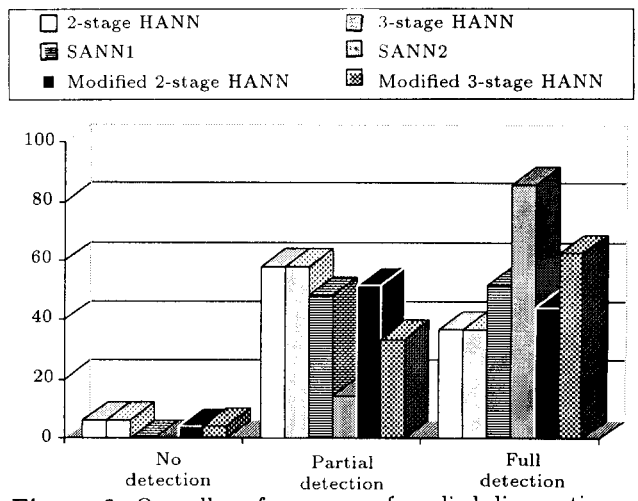

Figure 6. Overall performances of applied diagnostic schemes.

Table 6. Performance of modified two- and three-stage HANN.

Fault Patterns	Diagnostic Scheme	
	Modified Two-Stage HANN	Modified Three-Stage HANN
Normal and Single Faults	All	All
Double Faults	17(2/2), 1(3/2), 1(1/2), 2(0/2)	17(2/2), 1(3/2), 1(1/2), 2(0/2)
Triple Faults	15(3/3), 12(2/3), 8(1/3)	26(3/3), 9(1/3)
Quadruple Faults	2(4/4), 6(3/4), 25(2/4), 2(0/4)	10(4/4), 6(3/4), 17(2/4), 2(0/4)
Higher Degree Double and Triple Faults	3(2/2), 2(1/3)	3(2/2), 2(1/3)

Note: Numbers in parentheses represent the number of faults detected over total number of faults.

modified three-stage HANN performs better than other schemes.

CONCLUSIONS

In this paper, the performance of a two-stage HANN structure proposed for diagnosing multiple faults in the literature has been evaluated. In order to improve the performance of the two-stage HANN, particularly in diagnosing triple and quadruple concurrent faults, it has been extended to a three-stage structure. Simulation results indicate that no significant improvement is achieved. To alleviate this shortcoming, a Multi-Level Selection Switch (MLSS) is introduced, which improves the performance of the hierarchical neural network structures considerably. In addition, two single artificial neural networks (SANN1 and SANN2), which have the same structures but different training, are proposed and their performances are compared with multi-stage hierarchical neural networks. Among the schemes trained by single, double and triple fault patterns, the best performance belongs to SANN2. It should be noted that this improvement is obtained at the expense of more computational efforts for training. Results also indicate that the performance of SANN1 is superior to the two-stage and modified two-stage HANNs' performances, although they are trained with the same training patterns.

REFERENCES

1. Dash, S. and Venkatasubramanian, V. "Challenges in the industrial applications of fault diagnostic systems", *Comput. Chem. Eng.*, **40**, pp 785-791 (2000).
2. Nimo, I. "Adequately address abnormal situation operations", *Chem. Eng. Progress*, **91**(9), pp 36-45 (1995).
3. Frank, P.N. "Fault diagnosis in faulty systems using analytical and knowledge based redundancy- A survey and some new results", *Automatica*, **26**(3), pp 459-474 (1990).
4. Vedam, H. and Venkatasubramanian, V. "Signed digraph based multiple fault diagnosis", *Comput. and Chem. Eng.*, **21**(Supp.), pp 5655-5660 (1997).
5. Leung, D. and Romagnoli, J. "Dynamic probabilistic model-based expert system for fault diagnosis", *Computers Chem. Eng.*, **24**, pp 2473-2492 (2000).
6. Young Eo, S., Chang, T.S., Shin, D. and Yoon, E.S. "Cooperative problem solving in diagnostic agents for chemical processes", *Computers Chem. Eng.*, **24**, pp 729-734 (2000).
7. Ozyurt, B., Sunol, A.K., Camurdan, M.C., Mogili, P. and Hall, L.O. "Chemical plant fault diagnosis through a hybrid symbolic-connectionist machine learning approach", *Computers Chem. Eng.*, **22**(1-2), pp 299-321 (1998).
8. Prasad, P.R., Davis, J.F., Jirapinyo, Y., Josephson, J.R. and Bhalodia, M. "Structuring diagnostic knowledge for large scale process systems", *Computers Chem. Eng.*, **22**(121), pp 1897-1905 (1998).
9. Mylaraswamy, D. and Venkatasubramanian, V. "A hybrid framework for large scale process fault diagnosis", *Computers Chem. Eng.*, **21**(Supplementary), pp S935-S940 (1997).
10. Wang, X.Z., Lu, M.L. and MacGreavy, C. "Learning dynamic fault models based on a fuzzy set covering method", *Computers Chem. Eng.*, **21**(6), pp 621-630 (1997).
11. Becraft, W.R. and Lee, P.L. "An integrated neural network/expert system approach for fault diagnosis", *Computers Chem. Eng.*, **17**(10), pp 1001-1014 (1993).
12. Kramer, M.A. "Malfunction diagnosis using quantitative models with non-boolean reasoning in expert systems", *AIChE Journal*, **33**(1), pp 130-140 (1987).

13. Raic, A. and Cinar, A. "Statistical process monitoring and disturbance diagnosis in multivariable continuous processes", *AIChE Journal*, **42**(4), pp 995-1009 (1996).
14. Lin, W., Qian, U. and Li, X. "Nonlinear dynamic principal component analysis for on-line process monitoring and diagnosis", *Computers Chem. Eng.*, **24**, pp 423-429 (2000).
15. MacGregor, J.F., Jaeckle, C., Kiparissides, C. and Koutoudi, M. "Process monitoring and diagnosis by multiblock PLS methods", *AIChE Journal*, **40**, pp 826-839 (1994).
16. Yon, S. and MacGregor, J.F. "Statistical and causal model-based approaches for fault detection and isolation", *AIChE Journal*, **46**(9), pp 1813-1824 (2000).
17. Duda, R.O. and Hart, P.E., *Pattern Classification and Scene Analysis*, John Wiley & Sons, New York (1973).
18. Russel, E.L. and Braatz, R.D. "Fault isolation in industrial processes using fisher discriminant analysis", In *Foundations of Computer-Aided Process Operations, AIChE*, Pekney, J.F. and Blau, G.E., Eds., New York, pp 380-385 (1998).
19. Venkatasubramanian, V. and Chan, K. "A neural network methodology for process fault diagnosis", *AIChE Journal*, **35**(12), pp 1993-2002 (1989).
20. Watanabe, K., Matsuura, I., Abe, M., Kubota, M. and Himmelblau, D.M. "Incipient fault diagnosis of chemical processes via artificial neural networks", *AIChE Journal*, **35**(11), pp 1803-1812 (1989).
21. Fan, J.Y., Nikolaou, M. and White, R.E. "An approach to fault diagnosis of chemical processes via neural networks", *AIChE Journal*, **93**(1), pp 82-88 (1993).
22. Watanabe, K., Hirota, S. and Hou, L., "Diagnosis of multiple simultaneous fault via hierarchical artificial neural networks", *AIChE Journal*, **40**(5), pp 839-848 (1994).
23. Sorsa, T. and Koivo, H.N. "Application of artificial neural networks in process fault diagnosis", *Automatica*, **21**(4), pp 843-849 (1993).
24. Farell, A.E. and Roat, S.D. "Framework for enhancing fault diagnosis capabilities of artificial neural networks", *Computers Chem. Eng.*, **18**(7), pp 613-635 (1994).
25. Tsai, C.S. and Chang, C.T. "Dynamic process diagnosis via integrated neural networks", *Computers Chem. Eng.*, **19**(Supplementary), pp S747-S752 (1995).
26. Vedam, H. and Venkatasubramanian, V. "A wavelet theory based trend analysis system for process monitoring and diagnosis", *Proceeding of the American Control Conference*, New Mexico, pp 309-313 (June 1997).
27. Zhao, J., Chen, B. and Shen, J. "Multidimensional nonorthogonal wavelet-sigmoid basis function neural network for dynamic process fault diagnosis", *Computers Chem. Eng.*, **23**, pp 83-92 (1998).
28. Chen, B.H., Wang, X.Z., Yang, S.H. and MacGreavy, C. "Application of wavelets and neural networks to diagnostic system development- 1. feature extraction", *Computers Chem. Eng.*, **23**, pp 899-906 (1999).
29. Ruiz, D., Nougues, J.M., Calderon, Z., Espuna, A. and Puigjaner, L. "Neural network based framework for fault diagnosis in batch chemical plants", *Computers Chem Eng.*, **24**, pp 777-784 (2000).
30. Rengaswamy, R. and Venkatasubramanian, V. "A fast training neural network and its updation for incipient fault detection and diagnosis", *Computers Chem. Eng.*, **24**, pp 431-437 (2000).
31. Hagan, M.T., Demuth, H.B. and Beal, M., *Neural Network Design*, PWS Publishing Company (1996).
32. Moller, M.F. "A scaled conjugate gradient algorithm for fast supervised learning", *Neural Networks*, **6**, pp 525-533 (1993).

APPENDIX 1

State Variables

T	reaction temperature, 740.0 K
$C_{C_7H_8}$	outlet concentration of $C_{C_7H_8}$, 524.0 mol/m^3
C_{H_2}	outlet concentration of C_{H_2}, 2097.0 mol/m^3
$C_{C_7H_{16}}$	outlet concentration of $C_{C_7H_{16}}$, 476.0 mol/m^3
T_h	heater outlet temperature, 889.0 K
S_i	output of the integrator in PI controller, 223.0 mV

Intermediate Variable

S_h	driving signal for the heater, 222.0 mV

Process Inputs

T_i	temperature of the reaction inlet stream, 300.0 K
$C^i_{C_7H_{16}}$	inlet concentration of C_7H_{16}, 1000.0 mol/m^3
cT	inlet temperature of heating steam, 0.9×740.0 K
c	efficiency of heater, 0.9

Reactor

k_0	frequency, 5.01×10^8 h^{-1}
E_a	activation energy, 1.369×10^5 J/mol
R	gas constant, 8.319 J/mol.K
ΔH	heat of reaction: $2.2026 \times 10^5 + 6.2044 \times 10T$ $-5.536 \times 10^{-2}T^2 - 1.15 \times 10^{-6}T^3$ $+3.1469 \times 10^{-7}T^4$, J/mol
C_p	specific heat, 490.7 J/mol.K
ρ	density, 593.0 mol/m^3

a area of heat exchange, 10.0 m^3
U overall heat transfer coefficient, 6.05×10^5 $J/m^2.h.K$
q inlet and outlet volumetric flow rate, 3.0 m^3/h
V effective reactor volume, 30.0 m^3

Heater

τ V'/q', heater time constant, 0.2 h
K $a'h'k'_h/(\rho' C'_p q')$, heater gain, 1.0 K/mV

PI Controller

K_c proportional gain, 20.0
T_i^* integrator coefficient, 0.3 h
$K_{mv/T}$ gain of temperature to mV transducer, 1.0 mV/K
u_0 set point, 740.0 mV

Equations

Reactor energy balance:

$$k(T) = k_0 e^{(-E_a/RT)},$$

$$\frac{dT}{dt} = \frac{q}{V}(T_i - T) - \frac{\Delta H}{\rho C_p} k(T) C_{C_7H_{16}} + \frac{aU}{\rho C_p V}(T_h - T), \qquad T(0) = T_0.$$

Reactor mass balances:

$$\frac{dC_{C_7H_8}}{dt} = -\frac{q}{V} C_{C_7H_8} + k(T) C_{C_7H_{16}},$$

$$C_{C_7H_8}(0) = C^0_{C_7H_8},$$

$$\frac{dC_{H_2}}{dt} = -\frac{q}{V} C_{H_2} + 4k(T) C_{C_7H_{16}},$$

$$C_{H_2}(0) = C^0_{H_2},$$

$$\frac{dC_{C_7H_{16}}}{dt} = -\frac{q}{V}(C_{C_7H_{16}} - C^i_{C_7H_{16}}) - k(T) C_{C_7H_{16}},$$

$$C_{C_7H_{16}}(0) = C^0_{C_7H_{16}}.$$

Heater energy balance:

$$\frac{dT_h}{dt} = \frac{1}{\tau}(cT - T_h) + \frac{K}{\tau} s_h, T_h(0) = T^0_h,$$

$$\frac{ds_i}{dt} = \frac{K_c}{T_i^*}(u_c - K_{mv/T} T),$$

$$s_i(0) = s^0_i,$$

$$s_h = K_c(u_c - K_{mv/T} T) + s_i.$$