Pergamon

Computers chem. Engng Vol. 19, Suppl., pp. S747–S752, 1995

0098-1354/95 $9.50 + 0.00

0098-1354(95)00105-0

DYNAMIC PROCESS DIAGNOSIS VIA INTEGRATED NEURAL NETWORKS

Chii-Shang Tsai and Chuei-Tin Chang

Department of Chemical Engineering, National Cheng Kung University,
Tainan, Taiwan 70101, Republic of China

ABSTRACT

A generic scheme of integrated artificial neural networks has been studied in this work for the purpose of fault detection and diagnosis in dynamic systems with varying inputs. Two general types of neural models, i.e. the feedforward networks (FFNs) and the recurrent networks, were integrated in the proposed fault monitoring system. To demonstrate the utility of the proposed methodologies, extensive experimental studies have been carried out on a pilot plant which simulates the operation of a semi-batch storage system. It can be observed that the predictions of the normal system behavior are very accurate and, also, the diagnostic conclusions obtained with the integrated networks are highly reliable.

KEYWORDS

Fault detection; fault diagnosis; neural networks; dynamic process; residual.

THE FRAMEWORK OF INTEGRATED NEURAL NETWORK

Initial studies on artificial neural networks (ANNs) for fault detection and diagnosis suggested that the FFNs could monitor the continuous chemical processes adequately (Hoskins and Himmelblau, 1988). Later, application studies were carried out using steady-state data from practical systems (Venkatasubramanian and Chan, 1989; Himmelblau, 1992). Although the results reported in these studies are in general quite satisfactory, their methods are applicable only to a limited number of realistic situations. This is due to the facts that, after the inception of faults, the system may go through a long period of continuous change before reaching a new steady state and, furthermore, steady-state data may not even be available if the system inputs are varying with time. Thus, an on-line fault monitoring system for dynamic processes should consist of two distinct functions, i.e. fault detection and diagnosis. A framework of the integrated neural network (see Fig. 1) has been developed in this work to perform these two tasks separately. A brief description of its components is presented in the sequel.

In this work, the "plant" refers to a chemical process operated in an unsteady or batch mode. The outputs $\mathbf{y}(t)$ from the plant are obviously corresponding to the sensor measurements and the inputs $\mathbf{u}(t)$ can be regarded as signals sent to the actuators or as other external parameters affecting the process conditions of the system. The inputs can always be measured on-line. However, in many batch operations, the actuator inputs are usually manipulated with a programmable logic controller according to a recipe. Thus, their target values at any time during operation should be available in advance and can be used directly as inputs to the neural networks.

In this fault monitoring system, the *residual generator* is the most critical component. The residuals, $\Delta\mathbf{y} = \mathbf{y} - \hat{\mathbf{y}}$, are essentially measures of the discrepancy between the observed

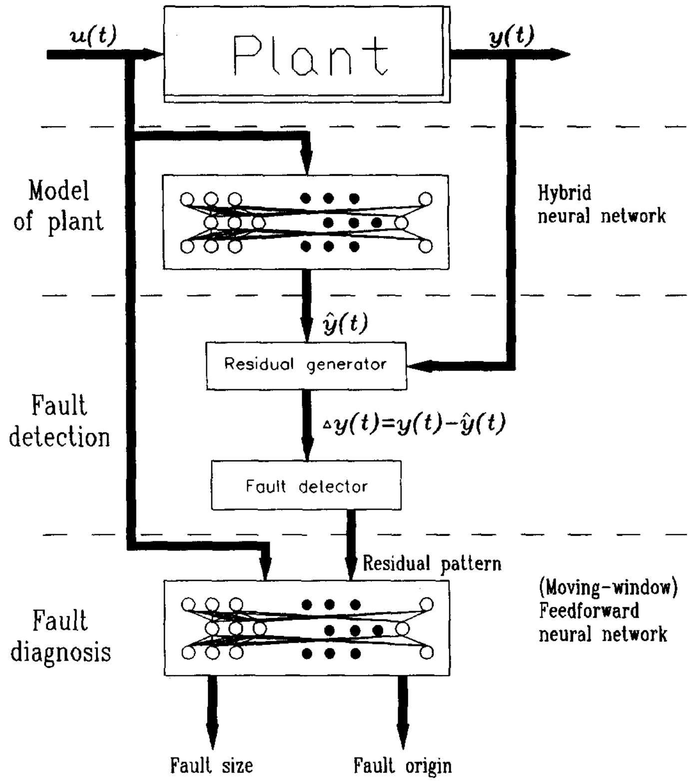

Fig. 1. The framework of integrated neural networks.

system behavior and that should result under normal conditions (Kramer and Mah, 1993). In previous publications, transient data corresponding to the latter were usually computed with the aid of mathematical models. However, since it is sometimes difficult to derive realistic models from basic principles of chemical engineering, a hybrid neural network has been developed for the same purpose. The residuals are then processed in two other components of the system. First, simple threshold tests are performed in a fault detector to identify abnormal conditions on-line. Next, the task of fault diagnosis is carried out with a moving-window feedforward network which maps the patterns of residuals to fault origins.

THE HYBRID NETWORK

Theoretically, the long-term dynamic behavior of a process system can be successfully predicted with an external recurrent network (ERN) if enough data are given and, also, the convergence criteria in training are satisfied (Su and McAvoy, 1992). However, this approach is questionable if applied blindly. First of all, the training process may be extremely difficult to converge. Secondly, even if convergence is achieved, poor prediction may still be possible in the generalization stage. This is caused by mistakenly identifying an unsuitable local minimum in the iterative training process.

In this work, physical insights of the systems were used as an aid for conjecturing the configurations of the neural network. It has to be realized that, in most chemical engineering applications, at least a qualitative description of the system is available. In addition to the cause-and-effect relations amongst input and output variables, knowledge about the dynamic characteristics of the system can sometimes be obtained. If the response times of different output variables with respect to the inputs are known to vary widely, the *qualitative* model of the plant can be expressed in the following form:

$$\frac{d\mathbf{x}_1}{dt} = \mathbf{f}_1(\mathbf{x}_1, \mathbf{x}_2, \mathbf{u}, t) \qquad \mathbf{x}_2 = \mathbf{f}_2(t, \mathbf{x}_1, \mathbf{u}) \tag{1a,1b}$$

$$\mathbf{y}_1 = \mathbf{h}_1(\mathbf{x}_1) \qquad \mathbf{y}_2 = \mathbf{h}_2(\mathbf{x}_2) \tag{2a,2b}$$

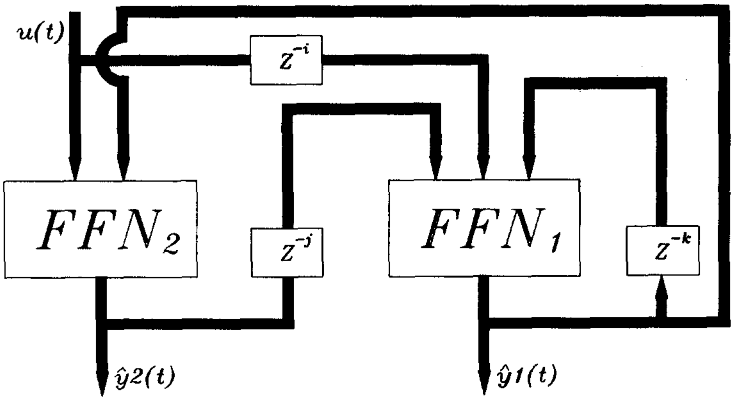

Fig. 2 The configuration of a generalized hybrid network.

where t is the time, $\mathbf{x}_1$ and $\mathbf{x}_2$ vectors of state variables, and $\mathbf{u}$ a vector of known inputs. Eqs. (2a) and (2b) simply indicate that the outputs $\mathbf{y}_1$ and $\mathbf{y}_2$ are functions of the state variables $\mathbf{x}_1$ and $\mathbf{x}_2$ respectively. Since an ERN is not designed to handle the coupled relations implied in Eqs. (1a) and (1b), it is necessary to introduce a change in the corresponding network structure.

In our study, the above information was incorporated in a generalized hybrid network (Fig. 2) to facilitate the construction of proper models for batch processes. Essentially, this network is built with two feed forward networks. The first one (FFN_1) maps the estimated output values and the input values at current and previous time steps to the estimates of $\mathbf{y}_1$ at present time, i.e.

$$\hat{\mathbf{y}}_1(t) = \mathcal{F}_1\left[\hat{\mathbf{y}}_1(t-1), \hat{\mathbf{y}}_1(t-2), \cdots; \hat{\mathbf{y}}_2(t), \hat{\mathbf{y}}_2(t-1), \ldots; \mathbf{u}(t), \mathbf{u}(t-1), \ldots\right] \tag{3}$$

In other words, it is a network similar to the FFN embedded in an ERN. The second network FFN_2 is used to describe the functional relations represented by Eqs. (1b) and (2b), i.e.

$$\hat{\mathbf{y}}_2(t) = \mathcal{F}_2\left[\hat{\mathbf{y}}_1(t), \mathbf{u}(t)\right] \tag{4}$$

Notice that the outputs of FFN_2 are the variables in $\hat{\mathbf{y}}_2(t)$. The inputs to FFN_2 are limited to $\mathbf{u}$ and $\hat{\mathbf{y}}_1$ at current time t.

THE MOVING-WINDOW FEEDFORWARD NETWORK

After detection of a fault, a feedforward network was used in this study to map the corresponding residual pattern to its origin. Since residuals associated with abnormal operating conditions may vary with time, it is important to include a *profile trend* as one of the inputs to FFN. Intuitively, this task can be accomplished with a moving-window FFN similar to the ones reported in previous publications, e.g. Vaidyabanathan and Venkatasubramanian (1990).

Several methods of feeding this FFN were investigated in this work. In general, we have found that it is better to pre-process the raw time-series data associated with Eqs. (1a) and (2a) and use the extracted features as inputs. At each time step t, the raw data in a window of size I, i.e. $\mathbf{\Delta y}_1(t - i\Delta t)$ $(i = 0, 1, \ldots, I-1)$, can be decomposed into the time-averaged values in $\mathbf{\Delta \bar{y}}_1(t)$ and profile trend in $\mathbf{\Delta \tilde{y}}_1(t - i\Delta t)$ $(i = 0, 1, \ldots, I-1)$ and then fed to the network. Specifically, these inputs can be calculated with the following equations:

$$\mathbf{\Delta \bar{y}}_1(t) = \frac{1}{I}\sum_{i=0}^{I-1} \mathbf{\Delta y}_1(t-i) \tag{5}$$

$$\mathbf{\Delta \tilde{y}}_1(t - i\Delta t) = \mathbf{\Delta y}_1(t - i\Delta t) - \mathbf{\Delta \bar{y}}_1(t) \tag{6}$$

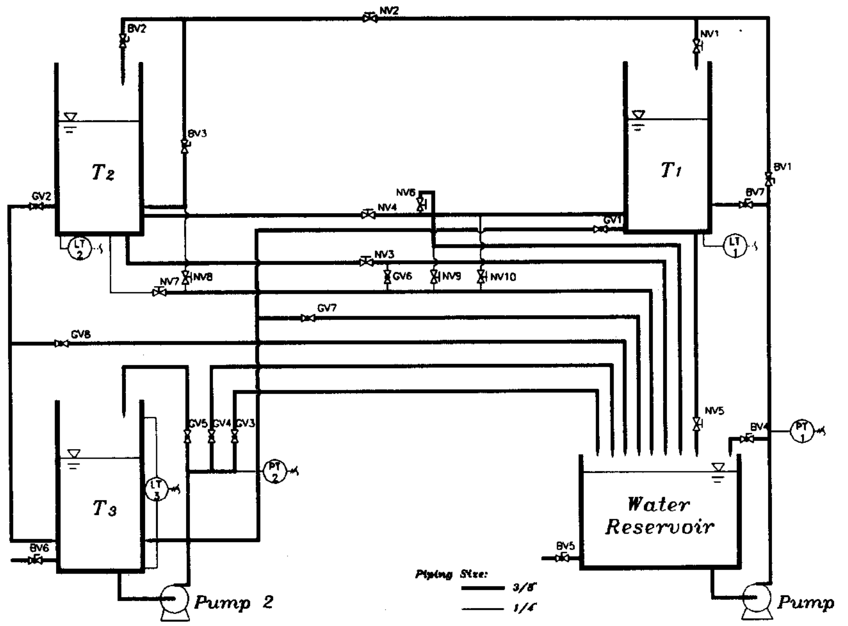

Fig. 3 The simplified P&ID of the pilot plant.

On the other hand, the raw time-series data of $\mathbf{\Delta y_2}(t - j\Delta t)$ $(j = 0, 1, \ldots, J-1)$ and $\mathbf{u}(t - k\Delta t)$ $(k = 0, 1, \ldots, K-1)$ are still utilized in the network at each time t.

THE PILOT PLANT

One of the most celebrated advantages of artificial neural network is its ability to capture the complex input/output relations of a variety of physical systems directly from realistic data without the need to derive mathematical models. To verify the feasibility and effectiveness of this approach, a pilot plant (shown in Fig. 3) has been built in this study for producing the training and testing data.

It can be seen from Fig. 3 that this experimental setup consists of a water reservoir, three open tanks (T_1, T_2 and T_3) and the connecting pipelines. The reservoir's height is 70 cm and the diameter 75 cm. The three tanks are identical in dimensions. The height and bottom diameter of each equal 123 cm and 50 cm respectively. The water reservoir and tank T_3 are on the ground. The other two tanks, i.e. T_1 and T_2, are placed on a bench with a height of 77 cm. The 3/8″ connecting pipelines are marked in Fig. 3 with thick solid lines, others 1/4″ pipelines are indicated by thinner solid lines. In order to regulate the interconnecting flows and to alter system configuration, needle valves, i.e. NV_1, NV_2, ..., NV_7, ball valves, i.e. BV_1, BV_2, ..., BV_7, and globe valve, i.e. GV_1, GV_2, ..., GV_8, are installed on these pipelines. In addition, branch pipelines which made of 1/4″ coppor tubes are attached to the 3/8″ pipelines. By opening the normally-closed needle valves, e.g. NV_8, NV_9 and NV_10, on these branches, small leaks in the pipelines can be simulated.

In our experimental setup, a number of process variables can be measured on-line. The discharge heads of the above two pumps are monitored with two identical pressure transmitters PT_1 and PT_2. The corresponding measurement range is from 0 to 15 psig. The levels in T_1 and T_2 are measured via differential-pressure type level transmitters, i.e. LT_1 and LT_2. Their ranges are from 0 to 2.5 psig. The level in the third tank T_3 is detected with an external-displacer type level transmitter LT_3.

EXPERIMENTAL RESULTS

A total of five systems have been studied in this work. Due to space limitation, only a sample of the results concerning a simple two-tank system (see Fig. 4) can be reported in this paper. Notice that there are three variables that can be measured on-line, i.e. the

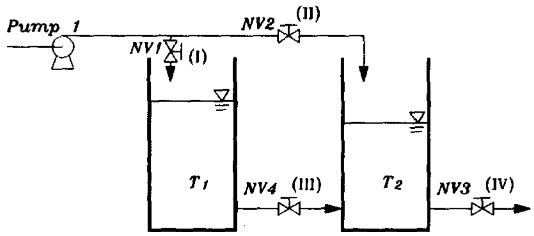

Fig. 4 The flow diagram of a two-tank system.

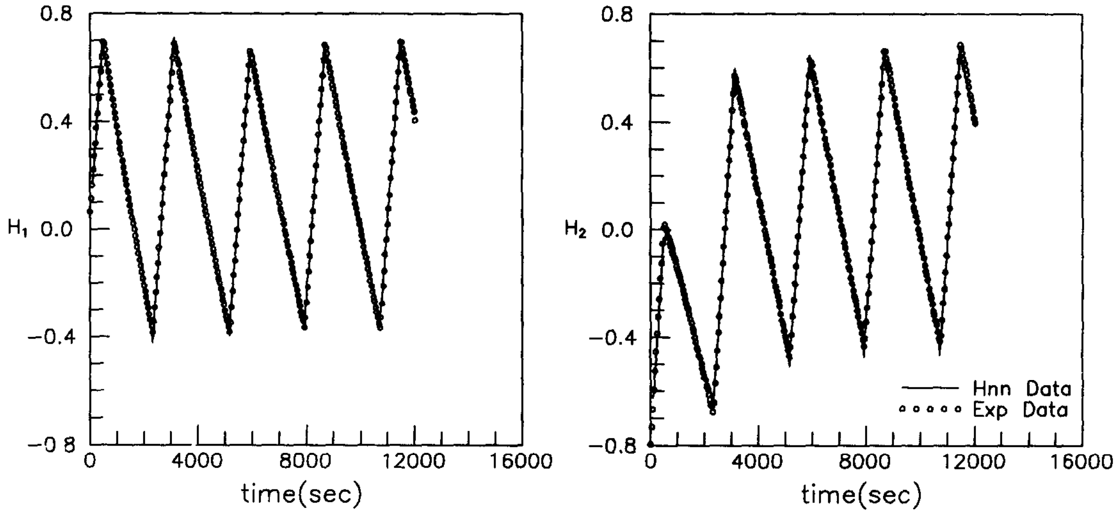

Fig. 5 Long-term prediction of the liquid levels in T_1 and T_2.

outlet pressure of pump No. 1 (p) and the height of liquid level in T_1 and T_2 (h_1 and h_2) respectively. The sampling interval in all experiments was 40 seconds.

The specific hybrid neural network can be constructed according to Fig. 2. The corresponding supervised learning process was terminated when the objective function reached a value less than 0.02. In most cases, the performance of the trained network was quite satisfactory. An example of the generalization tests is shown in Fig. 5. One can see that the predicted values of the normalized liquid levels H_1 and H_2 match the *untrained* transient data very closely for a period of more than three hours.

The standard FFN configuration has been adopted for fault diagnosis. The outputs of this network were related to six abnormal events that can be simulated experimentally, i.e. partial blockage in pipeline I, II and III, leakage in each of the two tanks and leakage in pipeline IV. The residuals associated with the three on-line measurements, i.e. ΔH_1, ΔH_2 and ΔP, and the ON/OFF status of pump No. 1, i.e. U, at consecutive sampling intervals were decomposed and fed into the input layer of the network. A uniform window size of 8 was adopted for these four types of inputs. The training process was ended when the objective value was less than 0.025. Experimental tests showed that diagnosis corresponding to all fault origins were fairly accurate. The results corresponding to one of the simulated events, i.e. a leak develops in T_2, are presented in Fig. 6. One can observe that, despite the fact that the symptoms of this event are very similar to those of partial blockage in the inlet pipeline to T_2, the proposed network is still able to identify the correct cause of abnormal system behavior within a short period of time.

CONCLUSIONS

Integrated neural-network-based techniques have been developed in this study for fault detection and diagnosis in dynamic processes. The principal difference between this paper and the published works which also address the problem of fault identification from transient data (Watanabe *et al.*, 1989; Vaidyanathan and Venkatasubramanian, 1990) is that the external inputs to the given process are no longer assumed to be constant. As a result, it is necessary to first generate residuals associated with the on-line measurements and then

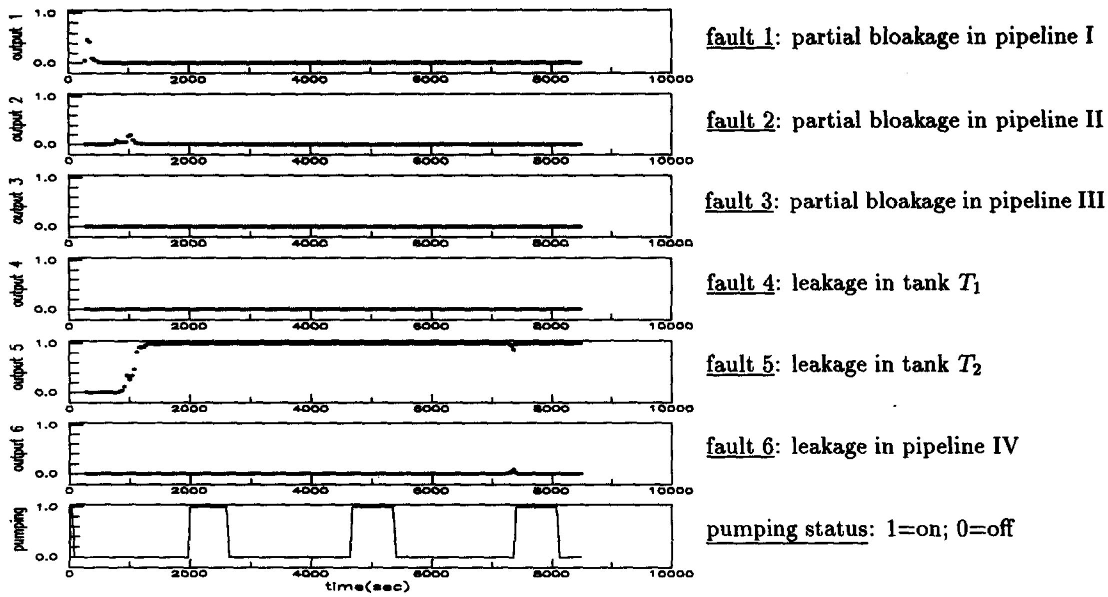

Fig. 6 Results of fault diagnosis. (1 = fault identified; 0 = otherwise)

locate the fault origins accordingly. This turns out to be the major point of departure from the previous approaches.

Extensive experimental studies have also been carried out in this work. Based on the test data we have produced so far concerning the hybrid networks, one can conclude that the accuracy demonstrated in their predictions definitely meets the need for a reliable fault detection scheme. Also, from the diagnostic results obtained in experiments, it can be observed that the moving-window FFN is in general able to correctly identify the residual pattern caused by a single fault origin.

REFERENCES

Himmelblau, D. M. (1992). Use of artificial neural networks to monitor faults and for troubleshooting in the process industries. In: *IFAC Symposium: On-line Fault Detection and Supervision in the Chemical Industries* (P. S. Dhurjati, ed.), pp. 144-149. Pergamon Press, Oxford.

Hoskins, J. C. and D. M. Himmelblau (1988). Artificial neural networks models of knowledge representation in chemical engineering. *Comput. & Chem. Engng.*, 12, 881-890.

Kramer, M. A. and R. S. H. Mah (1993). Model-based monitoring. In: *Proceedings of the Second Conference on Foundations of Computer-Aided Process Operations* (D. W. T. Rippon, J. C. Hale and J. F. Davis, ed.), pp. 45-68. CACHE, Austin.

Su, H. T. and T. J. McAvoy (1992). Long-term predictions of chemical processes using recurrent neural networks: a parallel training approach. *Ind. Eng. Chem. Res.*, 31, 1338-1352.

Vaidyanathan, R. and V. Venkatasubramanian (1990). Process fault detection and diagnosis using neural networks: II. dynamic processes. *AIChE Annual Meeting*, November. Chicago, IL.

Venkatasubramanian, V. and K. Chan (1989). A neural network methodology for process fault diagnosis. *AIChE J.*, 35, 1993-2002.

Watanabe, K., I. Matsuura, M. Abe, M. Kubota and D. M. Himmelblau (1989). Incipient fault diagnosis of chemical process via artificial neural networks. *AIChE J.*, 35, 1803-1811.

Watanabe, K., S. Hirota and L. Hou (1994). Diagnosis of multiple simultaneous faults via hierarchical artificial neural networks. *AIChE J.*, 40, 839-848.