

Using Gated Experts in Fault Diagnosis and Prognosis

Hamid Berenji¹, Yan Wang², David Vengerov¹, Reza Langari³, and Mo Jamshidi²

¹Intelligent Inference Systems Corp, MS:566-109, NASA Research Park, Moffett Field, CA 94035

²ECElectrical and Computer Engineering Department and Autonomous Control Engineering (ACE) Center, University of New Mexico, Albuquerque, NM 87131

³Texas A&M University, College Station, TX 77842

Abstract—Three individual experts have been developed based on Extended Auto Associative Neural Networks (E-AANN), Kohonen Self Organizing Maps (KSOM), and the Radial Basis Function based Clustering (RBFC) algorithms. An integrated method is proposed later to combine the set of individual experts managed by a Gated Experts algorithm, which assigns the experts based on their best performance regions. We have used a Matlab Simulink model of a chiller system and applied the individual experts and the integrated method to detect and recover sensor errors. It has been shown that the integrated method gets better performance in diagnostics and prognostics compared with each individual expert.

I. INTRODUCTION

The current research focuses on development of a library of routines and algorithms for fault diagnosis including approaches like Kohonen Self-Organizing Neural Networks, Extended Auto Associative Neural Networks, and Radial Basis Function based Clustering to train individual experts and then integrate the operations of these experts using the Gated Experts algorithm. In other words, our Gated Experts architecture manages the performances of our individual experts and assigns individual experts to regions where it performs the best. Our selected chiller system (at Texas A&M University) demonstrates many features of general dynamic systems. Most of our developments are performed using the Matlab environment and its SIMULINK simulator.

The algorithms that are being developed in this project are advanced pattern recognition techniques that are critically needed to separate nominal and faulty input-output component data vectors in a complex high-dimensional space. The prognostic tasks in component health monitoring require recognizing early signs of system malfunctioning and recommending whether a maintenance procedure should be carried out. A prognostic system needs to separate nominal component behavior from faulty ones even in the cases that those behaviors are similar.

II. DIAGNOSTIC ALGORITHMS

A. Extended AANN (E-AANN) for Sensor Diagnostics

Auto Associative Neural Networks (AANNs) have been used extensively in the recent past for sensor fault detection and identification [1-4]. The rationale for the use of AANNs in this mode is their capacity to provide a robust identity mapping between the input and the output of the network, which could be exploited in sensor fault detection. Consider the AANN shown in Fig. 1. In principle, this five-layer network maps its input $\{X_i, i = 1, 2, \dots, m\}$ to its output, $\{Y_i, i = 1, 2, \dots, m\}$ in such a manner that $\{Y_i = X_i, i = 1, 2, \dots, m\}$. When small perturbations in a given input to the network, say X_i , occur, for example as a result of sensor drift or other incipient faults in the sensor measuring X_i , the network will produce Y_i as a close approximation of the true X_i . The difference between X_i and Y_i can be used as an indicator of potential failure of the sensor that produces the corresponding reading.

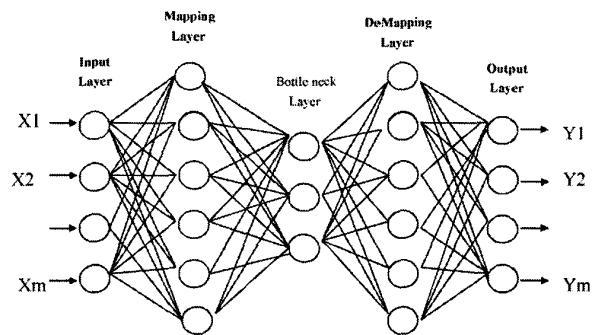


Fig. 1. AANN structure

In spite of their reported use in the literature [2] AANNs are often difficult to put into practice. This problem stems from the fact that AANNs generally rely on a strong correlation among multiple sensors in the given sensor set. Correlation among sensor readings, however, can cause undesirable variation of AANN outputs in presence of perturbations in its inputs, reducing the possibility of

isolating the one or two faulty sensors that may be involved. For instance, perturbation in, say X_i , can cause perturbations in every one of Y_j s, thereby undermining the ability of AANN to reproduce the true value of X_j s as intended. So an extension to the AANN (E-AANN) was developed by Dr. Langari and his students. It has been verified that E-AANN has the ability to recover the true value of X_j s in presence of input perturbations [5].

B. Kohonen Self-Organizing Maps

The application of Kohonen Self-Organizing Maps (KSOM) [6] to the fault prognostics problem is described in [7]. KSOM is formed in an unsupervised learning process and defines a mapping from the input data space \mathfrak{R}^n onto a regular two-dimensional array of nodes in Fig. 2. In the chiller system application, a KSOM input is a vector combining both inputs and outputs of the chiller system. Every node i is defined by a prototype vector $m_i \in \mathfrak{R}^n$. Each input vector $x \in \mathfrak{R}^n$ is compared with every m_i and the best match m_b is selected. The input is then mapped onto the corresponding location on the grid.

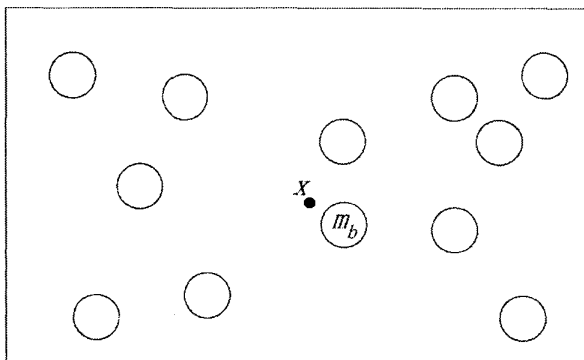


Fig. 2. KSOM structure

The optimal locations of the prototype vectors in the input space are determined using the following training procedure. At each step t of the learning process, a data sample $x(t) \in \mathfrak{R}^n$ is presented to the KSOM. Next, the node b with a prototype vector m_b closest to $x(t)$ in terms of Euclidean distance is determined. After that, the prototype vectors m_i of all KSOM nodes are updated using the following equation:

$$m_i(t+1) = m_i(t) + h_{ib}(t)[x(t) - m_i(t)] \quad (1)$$

where $h_{ib}(t)$ is a decreasing function of distance between prototype vectors m_i and m_b , which determines how much the node i is updated when unit b is the best-matching one. During repeated application of the above updating rule, the prototype vectors of nearby KSOM nodes become similar, an early malfunction of the chiller system can be observed as a gradual shift from a best-matching prototype vector in the middle of the KSOM grid toward some prototype vector near the boundary.

As described above, KSOM can be used to visualize the evolution of input-output component data in two dimensions by observing how the best matching prototype m_b changes over time. More importantly, a KSOM allows detecting atypical input-output data points by monitoring the distance between each input vector and its closest prototype vector. Monitoring this measure of “normality” together with component inputs can be used for making the maintenance decisions as described in [7].

C. Radial Basis Function based Clustering (RBFC)

The Radial Basis Function (RBF) rulebase is identified by our clustering algorithm. We will consider a specific case of a rulebase with n inputs and 1 output. The generalization to m outputs is described in [8]. The inputs to the rulebase are assumed to be normalized to fall within the range [0, 1].

The approach to identifying an RBF rulebase from data begins by creating an RBF rule with the Gaussian center coinciding with the first data point. When the next data point is encountered, the parameters of the first rule are adapted to account for both data points. If the error on the second data point is still too large, then a second rule is created centered at the second data point. The process continues until all data points have been considered. This is a very fast one-pass algorithm, which still gives very good results as our experiments indicate. For more information about the clustering algorithm, please see [8].

D. Gated Experts

The Gated Experts (GE) architecture [9] in Fig. 3 was developed as a method for adaptively combining predictions of multiple experts (Expert 1..., Expert K in Fig. 3) operating in an environment with changing hidden regimes. The predictions are combined using a gate block (Gating Network in Fig. 3), which dynamically assigns probabilities to the forecast of each expert being correct based on how close the current regime in the data fits the area of expertise for that expert.

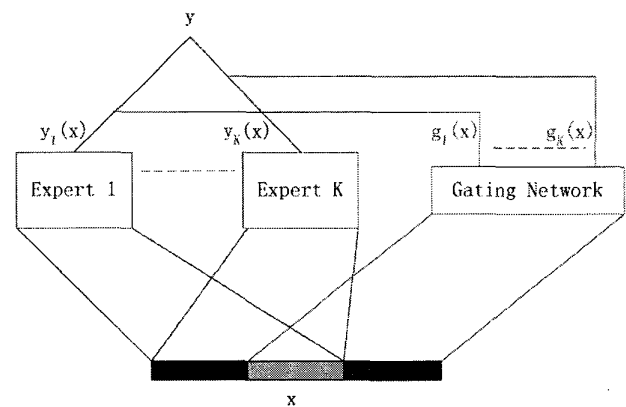


Fig. 3. GE structure

The training process for the GE architecture uses the expectation-maximization (EM) algorithm, which combines both supervised and unsupervised learning. The supervised component in experts learns to predict the conditional mean for the next observed value, and the unsupervised component in the gate learns to discover hidden regimes and assign the probabilities to experts' forecasts accordingly. The unsupervised component is also present in experts in the form of a variance parameter, which each expert adjusts to match the variance of the data for which it was found most responsible by the gate. This idea is similar to that of kernel smoothing, where the widths of kernels adjust to the local curvature of the data.

E. Integrated Method for Fault Diagnostics and Prognostics (IFDP)

The proposed approach IFDP to be evaluated in the context of the current effort is integrating all those prognostic units mentioned earlier (E-AANN, KSOM, RBFC) using the Gated Expert (GE) framework. This approach is based on the premise that each of the functional prognostic units likely functions best in a *limited* operating regime of the system or with respect to *certain types* of failure modes. With this in mind, the GE network will be used as a mechanism to determine the optimal operating regimes of each of the prognostic functional units thus proving the user with the most reliable prognosis or diagnosis of pending or current failures at any given time. This module will systematically summarize the prediction generated by each of the functional prognostic units in view of their performance as captured by the aforementioned adaptive mechanism.

III. CHILLER MODEL AND DATA GENERATION

A chiller model was used as a system to test the developed diagnostic approach using synthetic data. The model was previously developed as a part of a Masters thesis. The detailed specifications of the model can be found in [10]. In this section, we explain the model's inputs/outputs, the process of data generation, and the method of inducing noise to the system.

A. Model Inputs/Outputs

The model simulates the performance of a reciprocating, vapor compression cycle chiller (Fig. 4). The inputs and outputs are shown in Fig. 5.

The inputs to the chiller model are:

- 1) Temperature of water into condenser: $T_{in,c}$
- 2) Temperature of water into evaporator: $T_{in,e}$
- 3) Mass flow rate of water in evaporator: M_{we}

The outputs to the model are:

- 1) Temperature of water exiting the condenser: $T_{out,c}$
- 2) Temperature of water exiting the evaporator: $T_{out,e}$
- 3) Coefficient of the chiller performance: COP

- 4) Power consumed by compressor: W
- 5) Efficiency of Compressor: η_{comp}

The input ranges are¹: $T_{in,c} \approx 289 \rightarrow 309$ degK, $T_{in,e} \approx 275 \rightarrow 288$ degK, and $M_{we} \approx 0.22 \rightarrow 0.64$ Kg/sec.

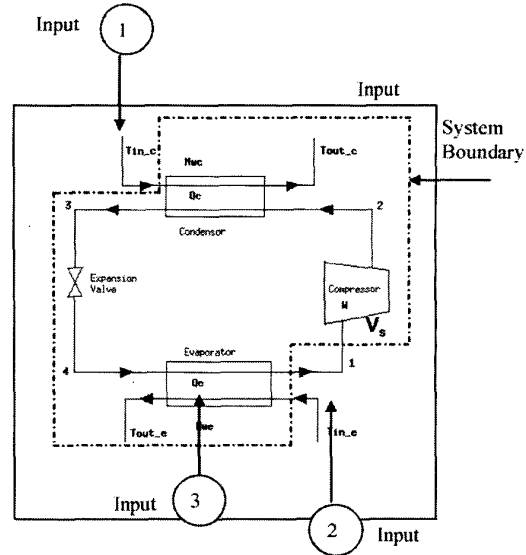


Fig. 4. System boundary of the chiller model, from [9]

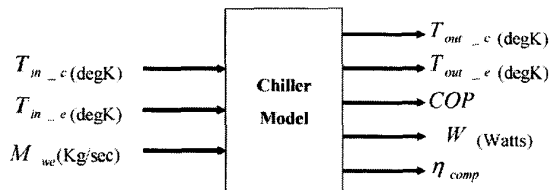


Fig. 5. Chiller Model Inputs/Outputs

B. Data Generation

From 1000 samples with 2% noise generated by using the model, 300 samples were randomly chosen as the test set and the remainder (700 samples) were used as the training set. The training set is used to train the network and the test set is used to test the trained network.

IV. PROGNOSTIC ALGORITHMS ON THE CHILLER

We applied a drift error in the sensor M_{we} at sample = 100 as a fault. The problem is to use the mentioned algorithms on the chiller model system to detect the sensor error and reconstruct its correct value. All the inputs and outputs of the chiller model are fed into the algorithm inputs. The outputs of the algorithms are diagnostic results, which are the sensor errors in this case.

Fig. 6 shows the diagnostic result from the E-AANN. The

¹ These ranges have been defined by designers in [10]

M_{we} output is increasing above zero about from sample = 120 and other variable outputs remain zero with some bias in the whole process. The sensor error is detected in the M_{we} and starts from sample = 120, a little later than the real starting point sample = 100.

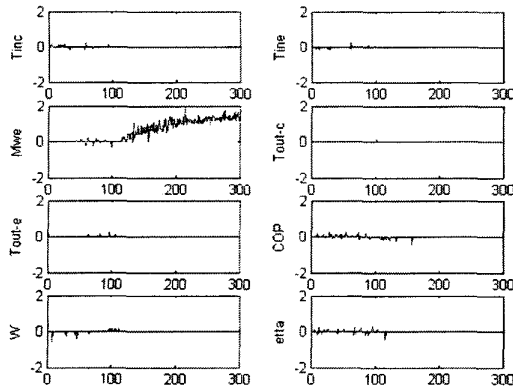


Fig. 6. E-AANN diagnostic outputs

Similar results can be seen in Fig. 7 when using KSOM. Except the sensor M_{we} error there also exists error in the sensor W from the diagnostic outputs. In this case, KSOM is not able of isolating sensor errors, i.e., if one of the sensors is wrong it can affect other sensor outputs from KSOM so that not only the wrong sensor is diagnosed as an error in it but also other sensors are diagnosed as errors in them. On the other hand, the fault in M_{we} ends about sample = 280, earlier than the real case.

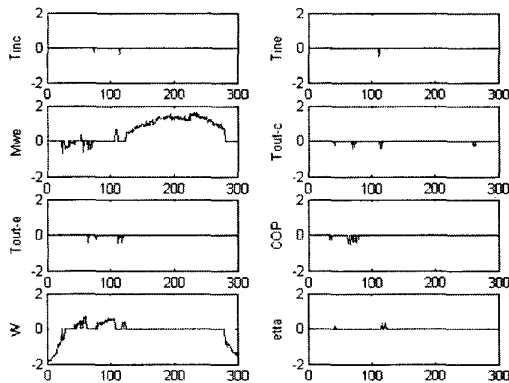


Fig. 7. KSOM diagnostic outputs

Also, RBFC is used to detect the sensor fault and the diagnostic result is shown in Fig. 8. It detects the fault in the sensor M_{we} starting about sample = 100.

Until now we have been caring about the sensor error detection. In this case, E-AANN and RBFC can correctly determine which sensor is wrong but KSOM cannot. Except

to diagnose the error existence, we also care about a more difficult task, the error reconstruction, which can keep the system working normally even though there exists some faults in it. In this step we will illustrate the effectiveness of our proposed approach.

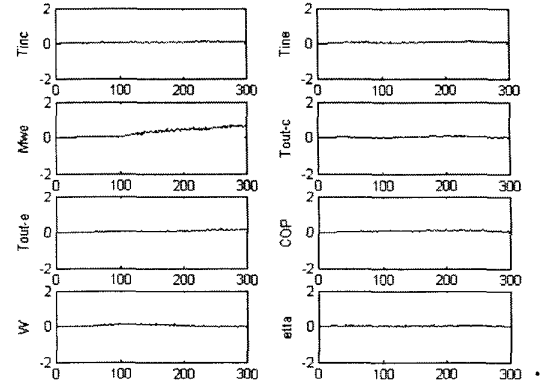


Fig. 8. RBFC diagnostic outputs

In IFDP we apply the gated experts to combine all the three algorithms and take advantage of each algorithm at every sample point. The gates for the three experts, E-AANN, SOM, RBFC, are shown in Fig. 9. The diagnostic output of the sensor M_{we} from IFDP is compared to that from the individual experts in Fig. 10. The diagnostic sensor error from the algorithms is plotted in blue and the real sensor error is plotted in red. The diagnostic error from E-AANN is matching the real error well except the initial period, sample 0-120. RBFC is working well in this period but there is some gap during sample 120-300. KSOM is the worst to recover the sensor error in the whole process. IFDP combines all the three algorithms (in fact, here only two) together using the gates experts and automatically determines the best diagnostic region for each expert and activates them in those regions so gets the best result in the whole process.

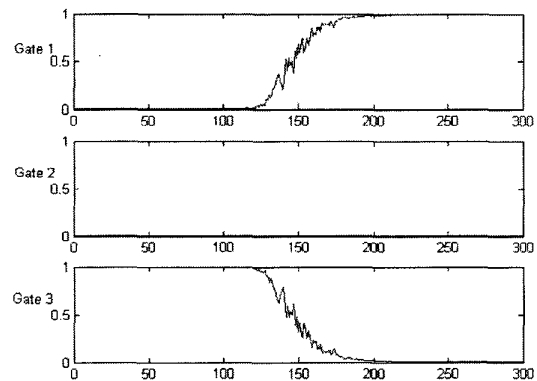


Fig. 9. Gate outputs

In Fig. 9 the gate outputs control the activation of the experts. When the gate output is 1 the corresponding expert is fully activated and when it is 0 it is fully suppressed. Some value between 0 and 1 is activating the expert in some degree. We can see the expert E-AANN is activated during sample 120-300 and expert RFBC is activated during sample 0-200. However, the expert KSOM is not activated for the whole process because it is not working well to reconstruct the sensor value and the combination between E-AANN and RFBC is the best choice picked by GE to reconstruct the sensor value. This can also be seen in Fig. 10 as illustrated above.

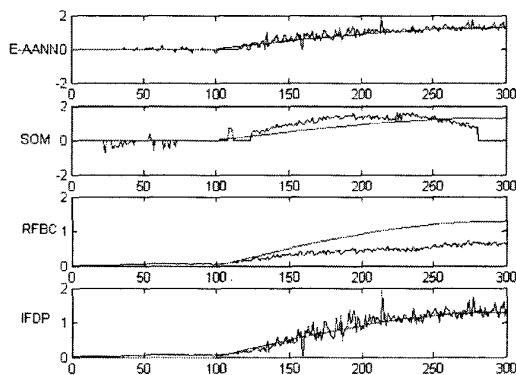


Fig. 10. Error reconstruction comparison

V. CONCLUSION

Due to the expected availability of a huge number of sensors on today's dynamic systems, our approach for fault diagnostics and prognostics must be capable of intelligent data reduction in such a way that no important data is lost and all the crucial data be used for smart prognosis with minimum false alarms. An example of such an integrated system was presented in this paper. The integrated method was tested on the chiller system to detect and recover the sensor error. It was shown that the proposed method could obtain better performance in the reconstruction compared with the individual experts. It is expected that a library of these strong methods which are under development at IIS Corp and University of New Mexico will significantly benefit the fault diagnosis and prognosis programs.

REFERENCES

[1] M. A. Kramer, "Autoassociative Neural Networks", *Computers in Chemical Engineering*, vol. 16, no 4, pp. 313-328, 1992.
 [2] J. W. Hines and R. E. Uhrig, "Use of Autoassociative Neural Networks for Signal Validation", *Journal of Intelligent and Robotic Systems*, vol. 21, no 2, pp. 143-154, Feb. 1998.

[3] J. W. Hines, D. J. Wrest, and R. E. Uhrig, "Plant wide sensor calibration monitoring," *Proceedings of the 1996 IEEE International Symposium on Intelligent Control*, Dearborn, MI, pp. 378-383, 1996.
 [4] J. W. Hines, A. V. Gribok, I. Attieh, and R. E. Uhrig, "Improved methods for on-line sensor calibration verification", *Proceedings of the 8th International Conference on Nuclear Engineering*, Baltimore, MD, 2000.
 [5] M. Najafi, "Use of Auto Associate Neural Networks for Sensor Diagnostics", *Masters Thesis, Texas A&M University, Mechanical Engineering*, 2003.
 [6] T. Kohonen, *Self-Organizing Maps*, Hidelberg: Springer-Verlag, 1995.
 [7] T. J. Böhme, N. Valentin, C. S. Cox and T. Denoeux, "Comparison of autoassociative neural networks and Kohonen maps for signal failure detection and reconstruction", *Intelligent Engineering Systems through Artificial Neural Networks 9 (Proc. of ANNIE'99)*, St. Louis, Missouri, pp. 637-644, 1999.
 [8] H. Berenji, J. Ametha, and D. Vengerov, "Inductive Learning for Fault Diagnosis", *12th IEEE International Conference on Fuzzy Systems (FUZZ-IEEE '03)*, vol. 1, pp.726 – 731, May 2003.
 [9] A. S. Weigend, M. Mangeas, and A. N. Srivastava, "Nonlinear gated experts for time series: Discovering regimes and avoiding overfitting", *International Journal of Neural Systems*, vol. 6, no. 4, pp. 373-399, Dec. 1995.
 [10] R. Prabhu, "A Neuro Computational Approach to Chiller Fault Identification and Isolation", *Masters Thesis, Texas A&M University, Mechanical Engineering*, 2002.