Research Paper Engineering

Identification of Incipient Faults by Employing SOM after Model Order Reduction

* Tarun Chopra ** Dr. Jayashri Vajpai

* Associate Prof., Dept. of Electrical Engg. ,Govt. Engg. College Bikaner

** Associate Prof, Dept. of Electrical Engg., J.N.V.University, Jodhpur

ABSTRACT

The Self Organizing Map (SOM) plays a versatile role in providing an initial organization of the data and is an epistemological tool for acquiring an understanding of the semantics of data and for generating hypothesis about the associated faults. An important aspect of the SOM based architecture is that use of SOM as constructive learning based epistemic tool has encouraged the researcher to take up of more complex issues. The understanding gained about the data with the help of SOM has been further supported by ANFIS based model order reduction in the proposed methodology. This has paved the way for more improvement in the classification results of highly overlapping faults.

Keywords : DAMADICS Benchmark Process Control System, Fault Diagnosis, SOM, Model Order Reduction

Introduction

SOM can be used to quickly create a qualitative overview of the data. It maps nonlinear statistical relationships among different variables of a high dimensional input data on a low dimensional network, preserving most of the topographic relationships from the input space[1-3]. It aggregates clusters of input information from the raw data, and projects them on a much simpler two or three dimensional network, thereby contributing to relatively comprehensible visualizations

Keeping in view the high level of complexity, there is a need for reduced order model for fault diagnosis involving incipient fault data. Model order reduction tries to quickly capture the essential features of a model /structure. When the process of model order reduction is stopped, all necessary properties of the original model should be captured with sufficient precision.

Reduced order models can describe the behavior of a system accurately, without the disadvantage of unnecessary detail. These models are computationally less demanding, being smaller than the original model. Their behavior is comparable to the behavior of the original model. This makes the models interpretable and efficient.

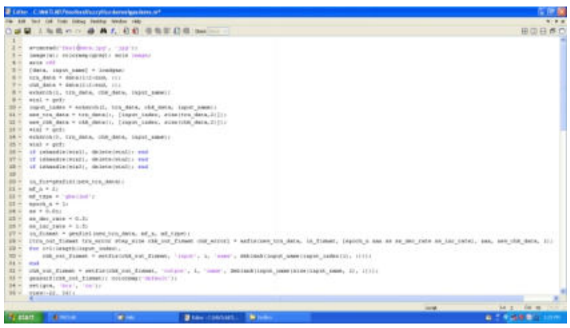
Proposed Methodology

ANFIS based modeling [4] is chosen for model order reduction in this work because it is a quick and straight forward method. ANFIS is a network structure that facilitates systematic computation of gradient vectors, i.e., the derivative of the output error with respect to each modifiable parameter. It employs an efficient hybrid learning method that combines least squares method and gradient descent.

Least squares method leads to fast training and gradient descent serves to slowly change the underlying membership function that generates the basis functions for the least squares method. ANFIS can generate the satisfactory results right after the first epoch of training i.e. after first application of least squares method. Since least squares method is computationally efficient, ANFIS models for various combinations of inputs is constructed and trained with single application of least squares method. The model with best performance is chosen to proceed for further training.

Here, ANFIS Exhaustive search function has been used for model order reduction. The MATLAB Program used for this purpose is shown in Figure 1.

Figure 1: Screen shot of MATLAB Program



To select the best input attribute, 'exhsrch' constructs four ANFIS, each with a single input attribute (Measured Parameter). After executing exhsrch (1, trn_data, chk_data, input_name), following results have been obtained as shown in Table 1 and Figure 2.

ISSN - 2249-555X

Table 1: Four ANFIS models, each with 1 input selected from 4 candidates.

	Measured Parameter	Training Error	Checking Error
ANFIS model 1	CV	0.0173	0.0208
ANFIS model 2	P1	0.2224	0.2522
ANFIS model 3	P2	0.1301	0.2595
ANFIS model 4	T	0.1524	0.3978

Obviously, 'CV' is the most influential input attribute and the training and checking errors of other measured parameters are comparable in size, which implies that there is no over fitting and more input variables can be selected.

Intuitively, one can simply select 'CV' and 'P1'or 'CV' and 'P2'directly. However, this will not necessarily lead to a two-ANFIS model with the minimal training error. To verify this, the exhsrch(2, trn_data, chk_data, input_name) can be used to select the best two inputs from all possible combinations.

Figure 2: RMS Error Values for Training and Testing Data when 1 input is selected

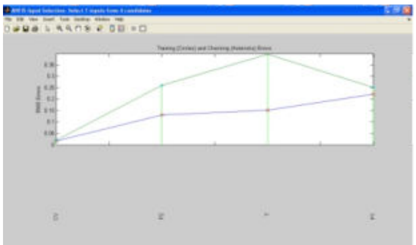
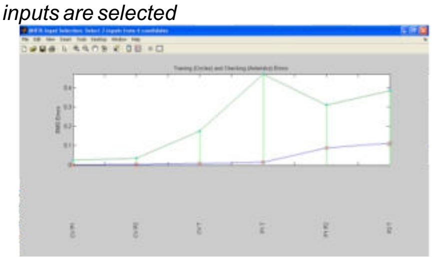


Table 2 and Figure 3 depict the results of selecting two inputs. 'CV' and 'P1' are selected as the best two input variables, which is quite reasonable.

Table 2 : Six ANFIS models, each with 2 inputs selected from 4 candidates.

	Measured Parameter	Training Error	Checking Error
ANFIS model 1	CV,P1	0.0003	0.0233
ANFIS model 2	CV,P2	0.0022	0.0334
ANFIS model 3	CV,T	0.0070	0.1755
ANFIS model 4	P1, P2	0.0886	0.3111
ANFIS model 5	P1, T	0.0131	0.4719
ANFIS model 6	P2, T	0.1130	0.3860

Figure 3: RMS Error Values for Training and Testing Data when 2 inputs are selected



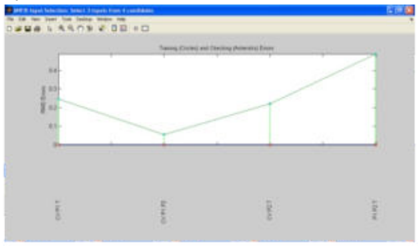
It can be inferred that 'CV,P1' are the most influential input attributes. Since, the training and checking errors of other measured parameters are comparable in size, there is no over fitting and more input variables can be selected.

Now, exhsrch is used to select three inputs. Table 3 and Figure 4 show the result of selecting three inputs, in which 'CV', 'P1', and 'P2' are selected as the best three input variables.

Table 3 : Four ANFIS models, each with 3 inputs selected from 4 candidates.

	Measured Parameter	Training Error	Checking Error
ANFIS model 1	CV,P1,P2	0.0001	0.0549
ANFIS model 2	CV,P1,T	0.0000	0.246
ANFIS model 3	CV,P2,T	0.0001	0.2215
ANFIS model 4	P1, P2,T	0.0015	0.4892

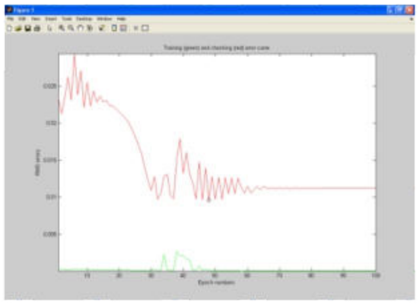
Figure 4: RMS Error Values for Training and Testing Data when 3 inputs are selected



The function exhsrch only trains each ANFIS for a single epoch in order to be able to find the right inputs shortly. The model with best performance with three input attributes namely 'CV,P1,P2' is chosen.

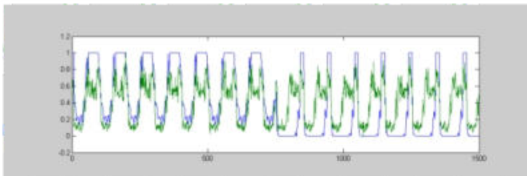
Now that the inputs are fixed, ANFIS can be trained further. The plot indicating error curves for 100 epochs of ANFIS training has been shown in Figure 5. The green curve (lower) shows the training errors and the red (upper) one shows the checking errors.

Figure 5: Training and Checking errors



The model accuracy is reflected in Figure 6.

Figure 6: Actual Output and Modeled Output 8



Results after Model Order Reduction

Following results after Map training step are obtained for the chosen data set corresponding to set of fault {F3, F4, F5, F6, F9}; when three inputs namely CV, P1 and P2 are considered.

- Map size = [8, 6] i.e., A two-dimensional SOM of 48 neurons (8 by 6), organized in a hexagonal neighborhood lattice.

Following results are obtained for the data set, after fine tuning phase:-

- Quantization error: 0.384
- Topographic error: 0.000

Histograms and Scatter Plots have been depicted in Figure7 .U-matrix, Component planes and Labels have been shown in Figure 8. Map Analysis by Visual Inspection may be done .

Figure 7: Histograms and Scatter Plots

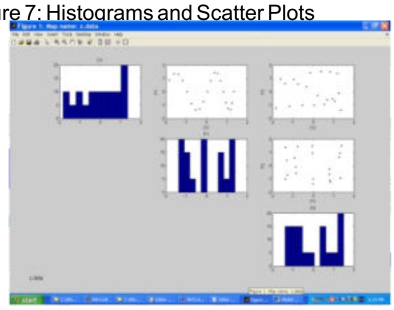


Figure 8: U-matrix and Component Planes

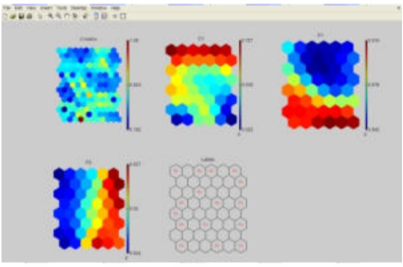
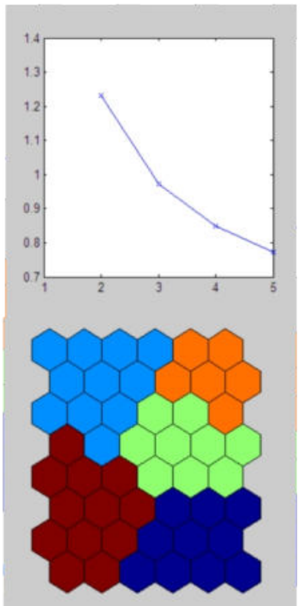


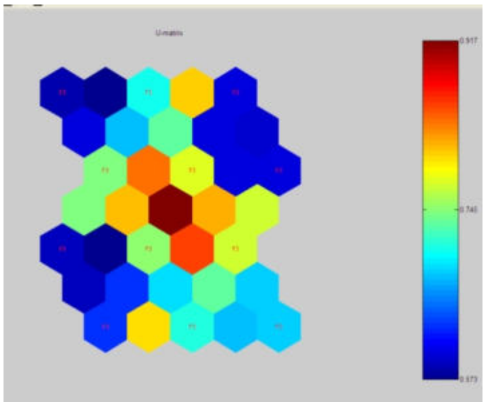
Figure 9 shows the Davies-Bouldin clustering index, which is minimized with best clustering

Figure 2: Assignment of apriori Probabilities



After performing Classification task as shown in Figure 10, Classification accuracy of 85.3% is obtained for the data set, which is about 6% more than the results obtained by using the techniques without model reduction.

Figure 10: SOM after Supervised Learning



Conclusion

The nature of data pertaining to incipient faults is generally overlapping with normal operating condition on one hand and abrupt fault condition on the other hand. Hence, it is essential to first understand the data that is being processed for obtaining better and meaningful results. The central task for gaining the necessary understanding is data exploration, which has been attempted by using SOM. This has resulted in the classification of incipient faults. The understanding gained about the data with the help of SOM has been further supported by ANFIS based model order reduction in the proposed methodology. This has paved the way for more improvement in the classification results of highly overlapping faults.

REFERENCES

Demartines P. and Hérault J.. “Curvilinear Component Analysis: A Self-Organizing Neural Network for Nonlinear Mapping of Data Sets”, IEEE Transactions on Neural Networks, January 1997, 8(1):148154. | Kohonen T., “Self-Organizing Maps”, 3rd ed. Springer, Berlin 2001. | Vesanto J.., “Neural network tool for data mining: SOM Toolbox”, In Proceedings of Symposium on Tool Environments and Development Methods for Intelligent Systems, Oulu, Finland, Oulunyliopistopaino, 2000, pages 184196. | Jang J.S.R., “ANFIS: Adaptive network based fuzzy inference system”, IEEE Transactions on Systems