ADVANCED SUPERVISORY CONTROL FUNCTIONS BASED ON COMPUTATIONAL INTELLIGENCE

Dragan Kukolj, Miroslava Berko-Pusic

Faculty of Engineering, University of Novi Sad,
Trg Dositeja Obradovica 6, 21000 Novi Sad, Yugoslavia
E-mail: kukolj@uns.ns.ac.yu

Abstract: The paper describes a methodology for design of advanced features within the SCADA system, such as fault detection, estimation, and prediction of process variables. These features are developed using two computational intelligence techniques, namely artificial neural networks and Takagi-Sugeno fuzzy models. Development and testing of the additional functions is performed within a SCADA system for natural gas pipeline transmission network. The impact of a number of factors on accuracy of artificial neural networks and Takagi-Sugeno fuzzy models employed in the SCADA system is analysed.

Keywords: supervisory control, networks, decision making, artificial intelligence, neural networks, fuzzy models

1. INTRODUCTION

Complex distributed systems are characterised with a large number of measured variables. It is therefore of great importance that functions, which enable effective and reliable supervisory control of the system, are embedded within the SCADA (Supervisory Control And Data Acquisition) system. Such functions enable dispatchers to appropriately react when a system overload or a fault occurs. The features that appear to be the most important are: (1) estimation of selected state-space vector components – process variables; (2) prediction of process variables and/or consumption; (3) fault detection – early indication of undesired phenomena within the system, such as failures of the measuring equipment, errors in the signal transmission, etc.

With regard to the listed additional features, a number of differences relevant for the selection of the input variables of adequate functions can be stated. As far as the state estimator is concerned, the input variables should be the key measured variables or their deviation from reference values (Isermann, 1984). Prediction on the other hand requires as inputs values of the output in preceding time intervals, week-day indicators and/or weather conditions, etc. Finally, a diagnostic function looks at variation in states obtained by estimation and by measurement (Isermann, 1984; Kukolj and Berko, 1998).

Functions of this type are built using two methods of computational intelligence: artificial neural networks (ANN) and Takagi-Sugeno fuzzy models (TS models) (Jang, *et al.*, 1997). These computational intelligence methods enable creation of models for dynamic systems of high complexity. Both methods are capable of generalization and knowledge accumulation regarding the process, on the basis of representative samples used at the training stage. These samples must contain sufficient information about system behavior in varying operating conditions. A SCADA database, that contains information regarding all the changes in the system

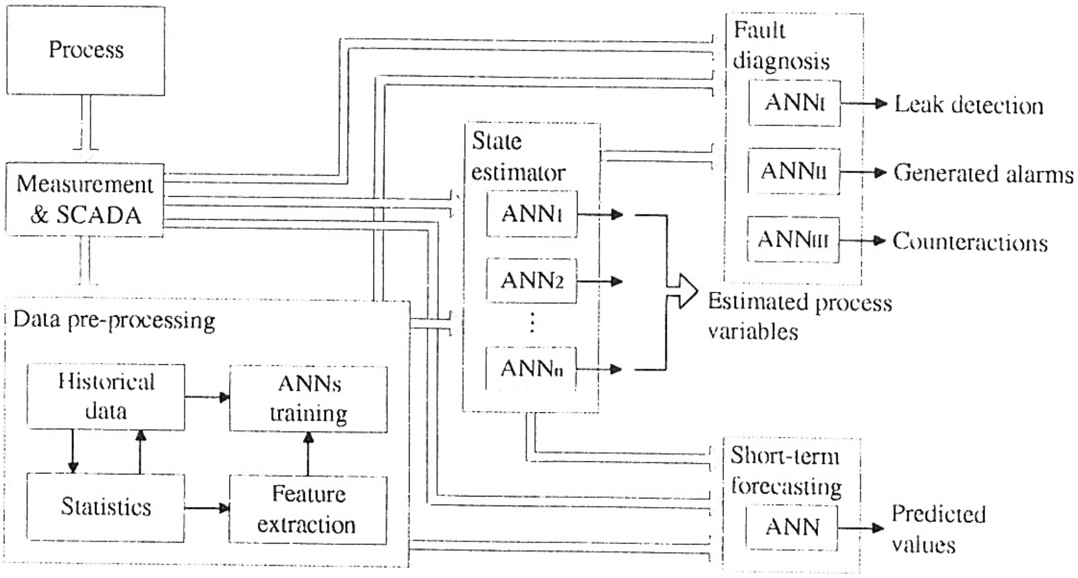

Fig. 1. Schematic representation of the ISO functions.

4. DESCRIPTION OF THE INTELLIGENT STATE OBSERVER STRUCTURE AND FUNCTIONS

State observer for the complete gas pipeline is composed of a number of modules comprising ANN or TS models. Each module utilizes 3 to 5 measured variables as inputs, while the output is an estimate or a prediction of a certain process variable. Figure 1 illustrates structure of the developed ISO with ANN type of models.

The requirements, imposed on the ISO, have necessitated fulfillment of a number of conditions related to the corresponding programme modules and internal database. Major characteristics of the ISO are listed in what follows:

(1) A connection with the SCADA database is established, with the aim of extraction and filtration of the required values of the process variables.

(2) An ISO internal database is created, that contains parameters of individual function modules (input-output quantities, time interval, sampling time, etc.). In addition, it contains parameters of corresponding ANN or TS models (training parameters and model configuration).

(3) Procedures for on-line execution of ANN or TS models for the purposes of estimation or prediction of a process variable are embedded.

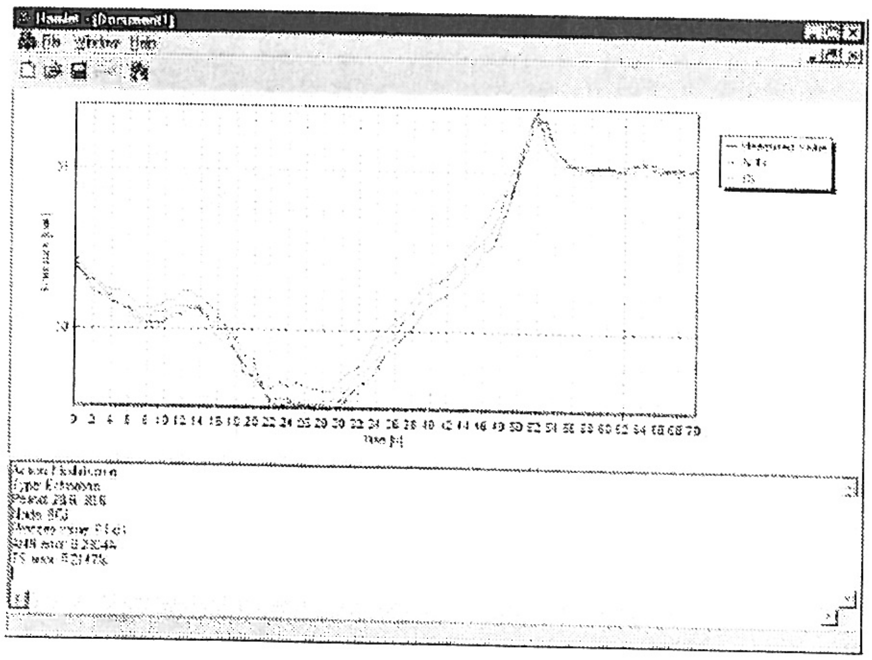

Fig. 2. Node Beèej – measured and estimated pressure.

in the past, is a prerequisite for successful application of these methods. A set of functions for prediction and estimation of selected process variables represents the system called Intelligent State Observer (ISO). ISO functions are developed for SCADA system of a natural gas pipeline transmission network.

2. SHORT OVERVIEW OF THE USED COMPUTATIONAL INTELLIGENCE TECHNIQUES

Feed-forward ANNs are used in this paper (Haykin, 1994). The important features of the ANN are its capability to perform non-linear mappings, fault tolerance capability and its ability to approximate a wide class of non-linear functions (Haykin, 1994; Narendra, 1996). Although a ANN can not incorporate any dynamics within it, a non-linear dynamic system is emulated by time-delayed inputs. The training of the ANNs is performed using the well-known error back-propagation method (Jang, *et al.*, 1997; Haykin, 1994).

The Takagi-Sugeno fuzzy model of the first order relies on product space clustering (Setnes, *et al.*, 1998). This means that when K clusters are formed, the corresponding TS fuzzy model contains K rules of the following form:

$$R_i\text{: If } x_1 \text{ is } A_{i1} \text{ and } x_2 \text{ is } A_{i2} \text{ and } \dots \text{ and } x_n \text{ is } A_{in} \text{ then } y_i = \mathbf{x}\mathbf{a}_i + b_i,\ i=1,2,\dots K \quad (1)$$

where R_i is a i-th rule, $x_1, x_2, \dots, x_n$ are input variables, $A_{i1}, A_{i2}, \dots, A_{in}$ are the fuzzy sets assigned to the input variables, y_i is the value of the output of the i-th rule, and $\mathbf{a}_i$ and b_i are parameters of the consequent function.

The modeling procedure consists of two phases. The first phase identifies the structure of the rule base. During this phase, partition of the input-output space by means of fuzzy c-means or Gustafson-Kessel clustering methods takes place (Jang, *et al.*, 1997; Setnes, *et al.*, 1998). The training data that contain N input-output samples $z_k = [x_k; y_k]^T$, $k = 1,\dots,N$, are used. Each resulting cluster represents a certain operating region of the system, where input-output data values are highly concentrated. These information clusters are interpreted as rules.

Once when the structure of the fuzzy model is defined, the second phase - parameter identification, is initiated. This phase enables calculation of the parameters that are present in the antecedent and consequent parts of the TS fuzzy rules. One-dimensional fuzzy sets in antecedent part of the i-th rule A_{ij} are obtained from multidimensional fuzzy clusters, on the basis of partition matrix, by point-wise projection onto the space of the input variable x_j. In order to determine parameters of the consequent function, it is necessary to calculate normalized firing strength of each rule for all the input samples. By forming the matrix composition of normalized firing strengths, it is possible to determine the consequent parameters using least squares method (Jang, *et al.*, 1997).

3. DESIGN OF THE FUNCTIONS BASED ON ANN AND TS FUZZY MODELS

Independently from type of control function under consideration, the design procedure is same. Successful input-output mapping depends upon external influences originating from input vector samples, and upon internal characteristics of the model used. External influences can be assessed by the amount of data, its representativeness, completeness, informativeness, etc. Internal group of influences consists of factors such as: number of hidden layers and number of neurons in them, a type of neuron's activation function, neurons connections, a process of learning, or type of membership functions, number of rules, clustering method and so on. Impact of those factors on model output can be expressed by functional dependence $y = f(x_1, x_2, \dots x_i, \dots x_m)$, where x_i is i-th influencing factor and m is total number of factors.

In order to assess the influence of external and internal factors on the model accuracy, three phase experiment is created. Quality of the information contained in samples is assessed through two different (external) factors: feature selection and data quantity. The cross-correlation is used to determine the existence and nature of the relations between features. Training data can be subjected to reduction by using different clustering techniques or unsupervised learning methods (Jang, *et al.*, 1997; Kukolj, *et al.*, 1997).

In second phase, 10-fold cross validation with Taguchi's method is used in order to analyze learning and generalization abilities of both model types (Peterson, *et al.*, 1995). With n-fold cross validation and an initial data set of size N, n test trials are carried out. Each trial employs N-N/n samples as the training set and the remaining N/n samples as the test set. In this experiment ten-fold cross validation is used, which means that initial data set η is partitioned into ten equal sets $\eta_1, \eta_2, \dots, \eta_{10}$. During k-th trial, the set $\eta - \eta_k$ is used for training and η_k is left for testing. The error measure used in this experiment is the mean square output error.

The last phase of the experiment consists of the analysis of variance. Analysis of variance determines statistical significance of the factors by using Fisher's test.

(4) Procedures for training – creation of ANN or TS models are integrated into the ISO.

(5) A module for continuous monitoring of selected estimates of process variables, that are subsequently compared to corresponding measured values obtained through SCADA system, is developed and constitutes a part of the ISO. This module serves the purpose of fault detection in various parts of the system.

The system is designed and tested using the data of the SCADA system of the natural gas pipeline transmission network "NIS-GAS" for the territory of Vojvodina (the northern province of Serbia). Pressure and flow of the gas are predicted and estimated. Data that correspond to different time intervals over a number of years are used for this purpose. All the tested models have shown high accuracy, with an error of less than 1% for pressures and less than 5% for flows. As an illustration of the ANN model application for the purpose of pressure estimation, Fig. 2 depicts comparative chart of the change of estimated and measured pressure in one node of the network.

Comparison of results of numerous tests indicates that there are not any substantial differences in performances between the ISO modules based on ANN and TS models. Figure 3 shows a comparison of the error for eight tests of ISO modules based on ANN and TS models.

5. CONCLUSION

The paper describes design of SCADA functions for quick and accurate estimation and prediction of the key process variables. The approach to the design utilizes artificial neural networks and Takagi-Sugeno fuzzy models. It is shown by experiments that, among many different factors, sampling time and the number of inputs have the greatest impact on the model error. This means that, by changing these parameters, errors of the function modules can be most easily reduced. The modules are very robust and computationally efficient. It is believed that upgrading of a SCADA system, by adding the described functions, could significantly improve the performance of a dispatcher assigned to control a complex process.

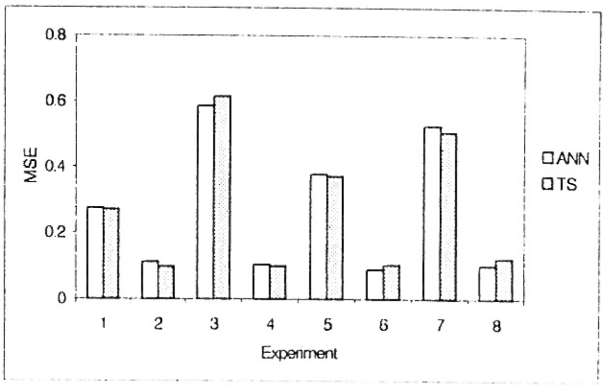

Fig. 3. Comparison of errors of ISO modules based on ANN and TS models.

REFERENCES

Haykin, S. (1994). *Neural Networks*. Macmillan Publ. Co., New York.

Isermann, R. (1984). Process Fault Detection Based on Modeling and Estimation Methods - A survey, *Automatica*, **20**, 387-405.

Jang, J.R., C. Sun, E. Mizutani (1997). *Neuro-Fuzzy and Soft Computing*, Prentice Hall, NJ.

Kukolj, D., D. Popovic, and M. Borota (1997). Applied Unsupervised Learning in Model Reduction of Linear Dynamic Systems, *Computers & Mathematics with Applications*, **33**,**3**, 95-103.

Kukolj, D., M. Berko (1998). The Experimental Design of the Diagnostic System Based on Unsupervised and Supervised Learning Methods, *Neural Network World*, **8**,**4**, 375-386.

Narendra, S.K. (1996). Neural networks for control: theory and practice, *Proc. of the IEEE*, **84**,**10**, 1385-1406.

Peterson, G.E., D.C. Clair, S.R. Aylward and W.E. Bond (1995). Using Taguchi's method of experimental design to control errors in layered perceptrons, *IEEE Trans. on Neural Networks*, **6**, 949-960.

Setnes, M., R. Babuška, H.B. Verburger (1998). Rule-Based Modeling: Precision and Transparency, *IEEE Transactions on Systems, Man, and Cybernetics - Part C*, **28**,**1**, 65-169.