Journal of the Korean Nuclear Society
Volume 33, Number 5, pp.457 ~ 467, October, 2001

An Input Feature Selection Method Applied to Fuzzy Neural Networks for Signal Estimation

Man Gyun Na and Young Rok Sim
Chosun University
375 Seosuk-dong, Dong-gu, Kwangju 501-759, Korea

E-mail: magyna@chosun.ac.kr

(Received March 7, 2001)

Abstract

It is well known that the performance of a fuzzy neural network strongly depends on the input features selected for its training. In its applications to sensor signal estimation, there are a large number of input variables related with an output. As the number of input variables increases, the training time of fuzzy neural networks required increases exponentially. Thus, it is essential to reduce the number of inputs to a fuzzy neural network and to select the optimum number of mutually independent inputs that are able to clearly define the input-output mapping. In this work, principal component analysis (PCA), genetic algorithms (GA) and probability theory are combined to select new important input features. A proposed feature selection method is applied to the signal estimation of the steam generator water level, the hot-leg flowrate, the pressurizer water level and the pressurizer pressure sensors in pressurized water reactors and compared with other input feature selection methods.

Key Words : vitrification, dry active waste, dust generation ratio, off-gas, cold crucible melter

1. Introduction

In recent years, the general problem of selecting a salient feature set for fuzzy neural networks has been generating a great deal of interest. It is experienced that the performance of a neural network strongly depends on the input features selected for its training. Non-salient input features to a fuzzy neural network can have even negative results. In signal estimation applications, there are a large number of input signals related with an output and it would require a large amount of time to train a fuzzy neural network with all the input signals as the number of connection weights for neural networks and parameters for fuzzy inference would be extremely large. Also, by eliminating unimportant signals and signal parameters, the cost and time of collecting the data can be reduced. As the number of input features grows, the training time required grows exponentially. Thus, it is important to reduce the number of inputs to a fuzzy neural network and to select optimum number of mutually independent inputs that are able to clearly define the input-output mapping. The process of selecting these important input variables is usually known as feature selection. The feature selection criterion is the measure used to rank feature subsets. Depending on the problem encountered, various criteria can be used. If the task is to predict properties of the measurements like the sensor signal estimation, a criterion that evaluates the predictive ability of the selected features should be used.

Also, sensor signal monitoring may require an estimation of a certain variable of interest. This estimation can be used to aid the operators in controlling the nuclear plants, and to detect and isolate faulty sensors. Many neural networks and fuzzy inference methods have recently been presented to diagnose sensors. Through training, fuzzy-neural networks are known to have the capability for performing complex mappings between input and output data, especially when expert diagnostic knowledge and the prior relation of fault symptom model are not clear. The direct use of transient signals in the time domain to the input of a neural network can be difficult since subtle differences may occur between different transients.

In this work, the focus is on identifying important input features to fuzzy neural networks. There are several feature selection methods including principal component analysis (PCA), genetic algorithm (GA), and others [1-5]. PCA, GA and probability theory are combined to select important input features and this combined feature selection method will be called a PGP method hereafter. A newly developed PGP feature selection method will be applied to and verified by the signal estimation of the steam generator water level, the hot-leg flowrate, the pressurizer water level and the pressurizer pressure sensors in pressurized water reactors and compared with other algorithms.

2. Input Feature Selection Methods

The number of variables has to be reduced for several reasons. However, this seems to be paradoxical at first since a dimension reduction decreases the information content. A reduction of the number of variables can lead to an improved performance due to at least three reasons. First, including features that contain irrelevant information about measurements can cause problems. Thus, it becomes important to use only high information descriptors. Secondly, studies have shown that if colinearity is present among the variables, the prediction results can get worse. Hence, it is necessary to remove highly correlated variables. Finally, when making a model containing many input parameters, a large number of observations are required to span the complete input space. The number of required observations grows exponentially with the number of input variables, which makes a dimensional reduction necessary to get a good model [1]. In this section, two methods using PCA and genetic algorithm are described and a combined method will be proposed.

2.1. Principal Component Analysis

The principal component analysis approach attempts to reduce the dimensionality of the feature space by creating new features that are linear combinations of the original features. Thus the PCA method involves linearly transforming the input space into an orthogonal space that can be chosen to be of lower dimension with minimal loss of information. The standardized input data are projected onto the eigenvectors called the principal components of the covariance matrix of the original data. The method also makes the transformed vectors (principal components) orthogonal and uncorrelated [6-7].

Given a signal vector $\mathbf{x}$ of p dimensions, $\mathbf{x}=[x_1\ x_2 \cdots x_p]^T$, its mean vector μ and covariance matrix Σ are described by

$$\mu = E\{\mathbf{x}\} = [m_1\ m_2 \cdots m_p]^T, \tag{1}$$

$$\Sigma = E\ \{(\mathbf{x} - \mu)(\mathbf{x} - \mu)^T\} \tag{2}$$

Since the true mean and the true covariance matrix are seldom known, the mean and covariance matrix are replaced with the sample mean $\mathbf{m}$ and the sample covariance matrix $\mathbf{S}$. The eigenvalues $\lambda_1, \lambda_2, \cdots \lambda_p$, and the corresponding orthonormal eigenvectors $\mathbf{p}_1, \mathbf{p}_2, \cdots \mathbf{p}_p$ of the sample covariance matrix $\mathbf{S}$ are calculated, and then ranked according to the magnitude of their associated eigenvalues.

$$\lambda_1 \geq \lambda_2 \geq \cdots \lambda_p \tag{3}$$

The eigenvectors $\mathbf{p}_1, \mathbf{p}_2, \cdots \mathbf{p}_p$ are called the principal components. The eigenvalues are proportional to the amount of variance (information) represented by the corresponding principal component. The transformation to the principal component space can be written as:

$$\mathbf{z} = \mathbf{x}^T\mathbf{P}, \tag{4}$$

where $\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2, \cdots \mathbf{p}_p]$.

The feature vector $\mathbf{z}$ can be transformed back into the original data vector $\mathbf{x}$ without a loss of information as long as the number of features, m, is equal to the dimension of the original space, p. For $m < p$, some information is usually lost. The objective is to choose a small m that does not lose much information.

However, this method has a disadvantage of not reducing the number of the input signals to a PCA block. As it were, PCA does not necessarily reduce the number of features that must be measured since each new feature of lower dimensionality may be a linear combination of all of the features in the original data vector. If we use many input signals to the PCA block, a possibility that we use unreliable and faulty sensor signals increases. Although in this method the number of inputs to the fuzzy neural network decreases, the number of the actually used inputs does not decrease.

2.2. Genetic Algorithm

In optimization problems using genetic algorithms, the term *chromosome* refers to a candidate solution which minimizes a cost function, generally encoded as a bit string. Each chromosome can be thought of as a point in the search space of candidate solutions. Genetic algorithm is an optimization technique which imitates the evolutionary process of a living organism. An initial population of chromosomes is iteratively altered by mechanisms inspired by natural evolution such as selection, crossover and mutation. Thus genetic algorithms process populations of chromosomes, successively replacing one such population with another. The genetic algorithms require a fitness function that assigns a score to each chromosome in the current population. The fitness of a chromosome (individual) depends on how well that chromosome solves the problem at hand [8-9].

In this work, a fitness function that evaluates the extent to which each individual is suitable for the given objectives such as small maximum error together with small total squared error and the small number of input variables, is suggested as follows:

$$F = \exp(-\mu_1 E_1 - \mu_2 E_2 - \mu_3 E_3), \tag{5}$$

where μ_1, μ_2 and μ_3 are the weighting coefficients, and E_1, E_2 and E_3 are the overall sum of squared prediction errors, the maximum absolute prediction error and the number of used input variables, respectively, defined as

$$E_1 = \sum_{k=1}^{N} (y_d(k) - y(k))^2 , \tag{6}$$

$$E_2 = \max_k \{ |y_d(k) - y(k)| \} , \tag{7}$$

$$E_3 = N_{input} . \tag{8}$$

$y_d(k)$ and $y(k)$ denote the measured signal and the estimated signal, respectively.

Each chromosome is encoded as a bit string which is composed of the same bit number as the number of input variables, and one '1' in each bit string represents that the corresponding input is selected and zero '0' represents that the corresponding input is not selected.

Genetic algorithms start from many points simultaneously climbing many peaks in parallel, and hence the probability of finding a false peak is reduced compared to the methods that move from one point to another. Accordingly, genetic algorithms are less susceptible to being stuck at local optima than conventional search methods. On the other hand, genetic algorithms have a disadvantage that it requires too much computational time.

2.3. PGP Method

In this proposed method called PGP method, PCA, GA and a correlation between input variables and an output variable are combined to select an optimal input set for signal estimation. PCA and GA are already explained in above subsections. Note that the correlation coefficient matrix of the original data set is equal to the covariance matrix of the data after the data have been standardized. This correlation matrix indicates a close or distant relationship among

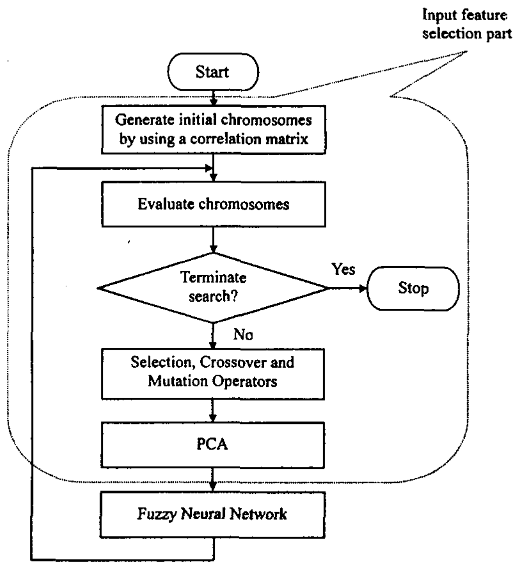

Fig. 1. Schematic Diagram of the Input Features Selection Algorithm (PGP method)

variables (signals). The high specific (i, j) component of the correlation matrix means that the related two variables are closely related to each other. These values between the input variables and the output variable are used to initialize the chromosomes of the genetic algorithm. To run a usual genetic algorithm, each bit of the chromosomes is usually randomly assigned one or zero. However, in this work, there is a high probability that the corresponding bit is assigned one in case that a correlation between a specific input and an output is high by using the correlation matrix. On the contrary, there is a high probability that the corresponding bit is assigned zero in case a correlation between the specific input and the output is low. This helps reduce the computational time by not considering from the first time the inputs which are not almost related with an output. But this does not mean that this PGP method removes completely the inputs with low correlation values from the first time. Each input is selected with a probability of which its selection is proportional to the square of the correlation value.

The algorithm for input feature selection is described in Figure. 1. First, the initial chromosomes of genetic algorithm are selected by a probabilistic weighting technique using the correlation coefficient matrix, which reduces the computational burden of the genetic algorithm which requires much computational time. Also, the input signals realistically used for signal estimation decrease in number through the genetic algorithm. A set of input signals to the PCA block which are important or closely related with estimating the output sensor signal are selected. The input signals coming into the PCA block are converted into new variables, called score vectors, which are orthogonal and span the multidimensional space of original signals. For the application of the PCA method, in this work, the number of input features from the PCA block is selected so that above 5 percent of original information into the PCA block will not be lost. As it were, the selected features include at least 95 percent of the initial information. Also, the time-delayed signal of the first score vector which usually includes most of the information of input signals is used to describe the sequential characteristics of a signal well.

3. Sensor Signal Estimation Using Fuzzy Neural Networks

3.1. Fuzzy Inference System

Neuronal improvements of fuzzy inference systems which mean that fuzzy systems is trained aim at exploiting the complementary nature of the two approaches; the fuzzy and neural network systems. Their composite is usually called as a fuzzy neural network or a neuro-fuzzy inference system. Each fuzzy rule of a fuzzy inference system is expressed as an *if/then* conditional rule. Thus, the arbitrary i-th rule can be described using the first-order Sugeno-Takagi type [10] as follows:

$$\begin{aligned}&\text{If } x_1 \text{ is } A_{i1} \; AND \cdots AND \; x_m \text{ is } A_{im},\\ &\quad \text{then } y_i \text{ is } f_i(x_1, \cdots, x_m),\end{aligned} \tag{9}$$

where

$x_1, \cdots, x_m$ = input variables to the neuro-fuzzy inference system (m = number of input variables),

$A_{i1}, \cdots A_{im}$ = antecedent membership function of each input variable for the i-th rule (i = 1, 2, ..., n),

y_i = output of the i-th rule,

$$f_i(x_1,\cdots,x_m)=\sum_{j=1}^{m} q_{ij}x_j + r_i, \tag{10}$$

q_{ij} = weighting value of the j-th input onto the i-th rule output,

r_i = bias of the i-th output,

n = number of rules.

In this work, the following Gaussian and sigmoid membership functions are used for each input variable:

$$A_{ij}(x_j)=\exp\left(-\frac{(x_j-c_{ij})^2}{2s_{ij}^2}\right), \tag{11}$$

$$A_{ij}(x_j)=\frac{1}{\exp\left(-\dfrac{x_j-c_{ij}}{s_{ij}}\right)+1}, \tag{12}$$

where

c_{ij} = center position of a membership function for the i-th rule and the j-th input,

s_{ij} = sharpness of a membership function for the i-th rule and the j-th input.

The sigmoid membership function is used for the maximum and minimum center values in each input variable and the Gaussian membership function is used for other center values. The output of an arbitrary i-th rule, f_i, consists of the first-order polynomial of inputs as given in Eq. (10).

The output of a fuzzy inference system with n rules is obtained by weighting the real values of consequent part for all rules with the corresponding membership grade. The output is obtained as follows:

$$y = \sum_{i=1}^{n} \overline{w}_i f_i, \tag{13}$$

where

$$\overline{w}_i = \frac{w_i}{\sum_{i=1}^{n} w_i}, \tag{14}$$

$$w_i = \prod_{j=1}^{m} A_{ij}(x_j). \tag{15}$$

3.2 Training of the Fuzzy Inference System

The fuzzy inference system is optimized by adapting the antecedent parameters (membership function parameters) and consequent parameters (the polynomial coefficients of the consequent part) so that a specified objective function is minimized. The adaptation methods of most fuzzy inference systems rely on the back-propagation algorithm [11]. The back-propagation algorithm is a general method for recursively solving for parameter optimization. Since this conventional optimization algorithm is susceptible to getting stuck at local optima, the genetic algorithm is used in this work. However, the genetic algorithm requires much computational time if there are many parameters to be optimized. Therefore, the least-squares method that is a one-pass optimization method is combined for a part of the parameters. The genetic algorithm is used to optimize the antecedent parameters c_{ij} and s_{ij}, and the least-squares algorithm is used to solve the consequent parameters q_{ij} and r_{ij}. A simple explanation on the genetic algorithm was given in the above subsection.

In this work, to increase the efficiency of the conventional genetic algorithm, three schemes are applied to accomplish the following good performance of the genetic algorithm: (a) initial coarse tuning and final fine tuning by changing the bit number of chromosomes versus generation; (b) prevention of an initial premature convergence without reaching optimal solutions and the acceleration of a final convergence by using two different selection methods of the crossover site that is randomly selected anywhere in a chromosome or randomly selected between only parameters in a chromosome; (c) prevention of final drifting without convergence by maintaining a certain part of chromosomes with higher fitness (refer to [11] for details).

If we fix some parameters of the fuzzy inference system by the genetic algorithm, the resulting fuzzy inference system is equivalent to a series expansion of some basis functions. This basis function expansion is linear in its adjustable parameters. Therefore, we can use the least-squares method to determine the remaining parameters. If a total of input-output pattern data for training are given, from Eq. (13) the consequent parameters can be chosen such that the pattern data satisfy the following equation:

$$\mathbf{y} = \mathbf{W}\mathbf{q}, \tag{16}$$

where $\mathbf{y}$ is the output data, $\mathbf{q}$ is the parameter vector, and the matrix $\mathbf{W}$ includes the input data defined as follows:

$$\mathbf{y} = \left[y^1 \; y^2 \cdots \; y^N\right]^T,$$

$$\mathbf{q} = \left[q_{11} \cdots q_{n1} \cdots\cdots q_{1m} \cdots q_{nm} \; r_1 \cdots r_n\right]^T,$$

$$\mathbf{W} = \left[\mathbf{w}^1 \; \mathbf{w}^2 \cdots \; \mathbf{w}^N\right]^T,$$

$$\mathbf{w}^k = \left[\overline{w}_1 x_1^k \cdots \overline{w}_n x_1^k \cdots\cdots \overline{w}_1 x_m^k \cdots \overline{w}_n x_m^k \; \overline{w}_1 \cdots \overline{w}_n\right]^T,$$
$$k = 1, 2, \cdots, N.$$

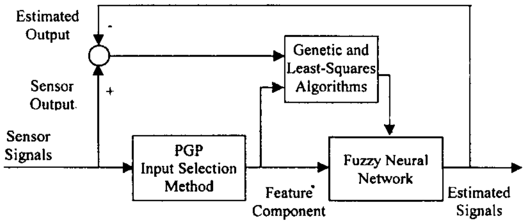

Fig. 2. Schematic Diagram of the Proposed Estimation Algorithm for Sensor Signals

The superscripts in the above notations indicate one of N input-output pattern data. The fuzzy neural network output is represented by the $N\times(m+1)n$ dimensional matrix $\mathbf{W}$ and the $(m+1)n$-dimensional parameter vector $\mathbf{q}$. In order to solve the parameter vector $\mathbf{q}$ in Eq. (16), the matrix $\mathbf{W}$ should be invertible but is not usually a square matrix. Therefore, we solve the vector by using the pseudo-inverse as follows:

$$\mathbf{q} = (\mathbf{W}^T\mathbf{W})^{-1}\mathbf{W}^T\mathbf{y}. \tag{17}$$

4. Applications

The signal estimation algorithm that combines the above-mentioned methods is described in Figure. 2. The input-output data were obtained for the load-decrease transients from the simulation of the MARS code [12] which is a unified version of the COBRA/TF and RELAP5/MOD3 codes, which consist of a total of 14 different signals. These data were standardized for being input to the fuzzy neural network. The proposed algorithm was applied to the steam generator water level, the hot-leg flowrate, the pressurizer water level and the pressurizer pressure sensors. Noise is added to model the real data of the nuclear power plant. The noise is proportional to the maximum variation σ_{max} of each signal and is chosen from a uniform distribution on the interval $(-0.02\sigma_{max}, 0.02\sigma_{max})$. In all computer simulations, the wavelet denoising technique was applied to all measurement signals and the Daubechies wavelet function was used [13]. Each signal consists of a total of 700 discrete time points where the sampling period is 1sec.

Table 1 shows the correlation coefficient matrix which indicates the relationships among all gathered signals. The fuzzy neural network was trained using one fifth of the given data in the training stage and was verified using the remaining data in the verification stage. Table 2 shows the results of all application cases and the proposed PGP method is compared with other two methods; genetic method and PCA method.

In the application to steam generator narrow-range water level estimation, the maximum error and the total squared error have similar magnitude for the three methods but PCA method is worse in the aspects of using many inputs, which is indicated by a fitness value. The steam generator narrow-range water level is almost closely related with the wide-range water level as shown in Table 1 and also, as we expect. The PGP method and the genetic method use the wide-range water level signal including other one or two signals, respectively, when their methods are applied to estimating of the narrow-range water level.

In the application to hot-leg flowrate estimation, there is no input signal closely related with the hot-leg flowrate as shown in Table 1. Therefore, the maximum error and the total squared error are relatively larger than those of other application cases. The PCA method is very bad and the genetic method is best except for using many signals. In the application to the pressurizer water level estimation, there are many signals closely related with the pressurizer water level even though they are correlated with each other. The PGP and genetic methods show similar performance. In the application to pressurizer

Table 1. Correlation Coefficient Matrix for Gathered Signals

	SF	FF	SP	ST	NL	WL	HT	CT	HF	AT	PP	PL	RP	PT
steam flowrate(SF)	1.0000	0.9897	-0.9566	-0.9573	-0.0031	-0.0039	0.9851	-0.4858	0.3202	0.9541	-0.1417	0.9544	0.9946	-0.1441
feed flowrate(FF)	0.9897	1.0000	-0.9353	-0.9362	-0.0919	-0.0916	0.9843	-0.4388	-0.3475	0.9605	-0.0896	0.9599	0.9877	-0.0920
steam pres.(SP)	-0.9566	-0.9353	1.0000	0.9999	-0.0260	-0.0252	-0.9078	0.7168	0.0436	-0.8368	0.2918	-0.8317	-0.9564	0.2942
steam temp.(ST)	-0.9573	-0.9362	0.9999	1.0000	-0.0263	-0.0254	-0.9085	0.7152	0.0476	-0.8377	0.2932	-0.8328	-0.9568	0.2956
S/G water level(NL)	-0.0031	-0.0919	-0.0260	-0.0263	1.0000	0.9985	-0.0368	-0.1170	0.0570	-0.0560	-0.2076	-0.0495	-0.0207	-0.2076
S/G wide-range level(WL)	-0.0039	-0.0916	-0.0252	-0.0254	0.9985	1.0000	-0.0372	0.1153	0.0565	-0.0562	-0.2065	-0.0496	-0.0212	-0.2065
hot-leg temp.(HT)	0.9851	0.9843	-0.9078	-0.9085	-0.0368	-0.0372	1.0000	-0.3655	0.4000	0.9891	0.0180	0.9852	0.9904	0.0156
cold-leg temp.(CT)	-0.4858	-0.4388	0.7168	0.7152	-0.1170	-0.1153	-0.3655	1.0000	0.6178	-0.2247	0.5781	-0.2110	-0.4899	0.5796
hot-leg flowrate(HF)	-0.3202	-0.3475	0.0436	0.0476	0.0570	0.0565	-0.4000	-0.6178	1.0000	-0.5152	-0.2655	-0.5446	-0.2837	-0.2644
RCS average temp.(AT)	0.9541	0.9605	-0.8368	-0.8377	-0.0560	-0.0562	0.9891	0.2247	-0.5152	1.0000	0.1131	0.9978	0.9593	0.1109
PZR pressure(PP)	-0.1417	-0.0896	0.2918	0.2932	-0.2076	-0.2065	0.0180	0.5781	0.2655	0.1131	1.0000	0.0727	-0.0727	0.9999
PZR water level(PL)	0.9544	0.9599	-0.8317	-0.8328	-0.0495	-0.0496	0.9852	-0.2110	0.5446	0.9978	0.0727	1.0000	0.9540	0.0705
reactor power (RP)	0.9946	0.9877	-0.9564	-0.9568	-0.0207	-0.0212	0.9904	-0.4899	-0.2837	0.9593	-0.0727	0.9540	1.0000	-0.0751
PZR temp. (PT)	-0.1441	-0.0920	0.2942	0.2956	-0.2076	-0.2065	0.0156	0.5796	-0.2644	0.1109	0.9999	0.0705	-0.0751	1.0000

pressure estimation, the pressurizer pressure has a close relationship with the pressurizer temperature because the pressurizer is in a saturated state, which is also shown in Table 1. In the genetic and PGP methods this pressurizer temperature signal is used and in the genetic method additional two signals are used.

In the summary, from Table 2, it is shown that although PCA method uses the largest number of input signals, PCA method is the worst of the three methods. But the PCA method is the fastest. It is determined that the genetic and PGP methods show similar performance as the input selection methods of fuzzy neural networks with application

Table 2. Final Results for Four Application Cases [after 30 generations training for input selection (genetic and PGP methods only) and After 50 Generations Training for the Fuzzy Neural network]

Sensors		S/G water level			Hot-leg flowrate			Pressurizer water level			Pressurizer pressure		
Methods		PCA	Genetic	PGP	PCA	Genetic	PGP	PCA	Genetic	PGP	PCA	Genetic	PGP
Training data	Total squared error	0.0045	0.0096	0.0098	4.2039	0.0310	0.2843	0264	0.0084	0.0073	0.0321	0.0228	0.0214
	Maximum error	0.0127	0.0184	0.0175	0.4041	0.0399	0.1056	0.0177	0.0161	0.0149	0.0355	0.0294	0.0306
	Fitness[1)]	0.7419	0.8903	0.9108	0.1507	0.7689	0.6669	0.7110	0.8789	0.8822	0.6910	0.8602	0.8923
Verification data	Total squared errors	0.0176	0.0375	0.0389	16.8232	0.1220	1.1524	0.0610	0.0338	0.0290	0.1325	0.0982	0.0906
	Maximum error	0.0128	0.0185	0.0180	0.4104	0.0437	0.1331	0.0310	0.0164	0.0152	0.0557	0.0590	0.0381
Input number to PCA		13	N/A	2	13	N/A	3	13	N/A	4	13	N/A	1
Number of FNN inputs		5	3	3	4	7	4	5	4	4	5	3	2
Input signals used		all signals	WL, HF, AT	WL, AT	all signals	SF,FF,SP NL,CT AT, RP	FF, CT PT	all signals	NL, CT AT, PT	ST, HT CT, PP	all signals	ST, NL PT	PT

1) This fitness was calculated in combination with the genetic optimization of the fuzzy neural network parameters.

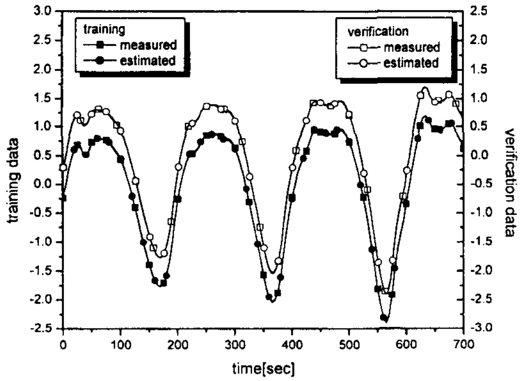

Fig. 3. Estimation of a (standardized) Steam Generator Water Level Signal Using the PGP Method

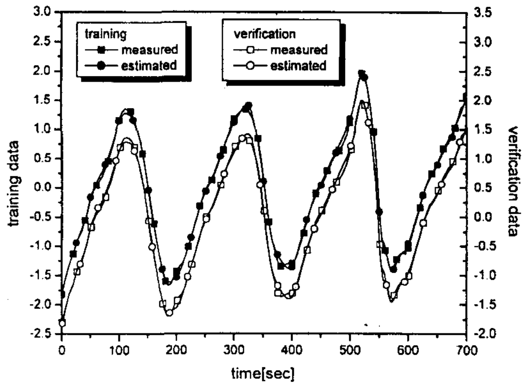

Fig. 4. Estimation of a (standardized) Hot-leg Flowrate Signal Using the PGP Method

to sensor signal estimation. However, the genetic method is about three times slower than the PGP method even though calculation time depends on the test cases and the relationship between input signals and an output signal. Figures 3 through 6 show the measured and estimated signals of each

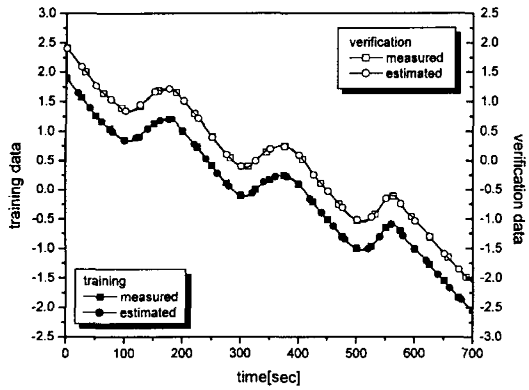

Fig. 5. Estimation of a (standardized) Pressurizer Water Level Signal Using the PGP Method

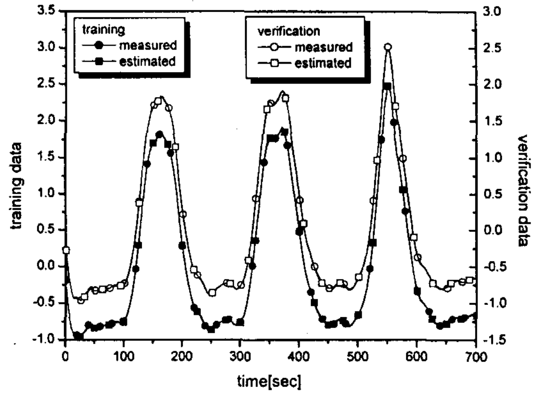

Fig. 6. Estimation of a (standardized) Pressurizer Pressure Signal Using the PGP Method

test case. From these figures, it is shown that a fuzzy neural network with the proposed input feature selection method actually estimates the relevant sensor signal using other sensor signals.

5. Conclusions

In this work, an input features selection method was proposed for the application to signal estimation using a fuzzy neural network. This proposed method combines the PCA, the genetic algorithm and the probabilistic concept. The reduction of number of input signals is actually accomplished by the genetic algorithm which requires the substantial computational burden. Thus, the heavy computational burden is reduced by using the correlation coefficient matrix which provides information on a close or distant relationship between input signals and an output signal. Each element of the correlation coefficient matrix connotes a relationship between the corresponding signals. Each input is selected with a probability of which its selection is proportional to the square of the correlation value. The usage of the correlation coefficient matrix takes effect on excluding from the first time the input signals coming into a fuzzy neural network which are almost irrelevant with the output signal. The proposed PGP method is about three times faster than a conventional genetic method even though calculation time depends on the test cases and the relationship between input signals and an output signal. Also, by combining the PCA method which can reduce the number of signals to the fuzzy neural network, the computation time for optimization of a fuzzy neural network can be shortened.

The proposed input feature selection method was applied to the estimation of the steam generator water level, the hot-leg flowrate, the pressurizer water level and the pressurizer pressure signals in pressurized water reactors and compared with other algorithms (PCA method, genetic method) and showed better performance compared to the other two methods.

Acknowledgment

This work has been conducted under an advanced nuclear technology acquisition project supported by the Korea Institute of Science & Technology Evaluation and Planning (KISTEP) which is funded by Korea Ministry of Science and Technology (Korea MOST).

References

1. Tomas Eklov, Per Martensson, and Ingemar Lundstorm, " Selection of variables for interpreting multivariate gas sensor data," *Analytica Chimica Acta*, vol. 381, pp. 221-232, (1999).
2. B. K. Lavine, A. Moores, and L. K. Helfend, "A genetic algorithm for pattern recognition analysis of pyrolysis gas chromatographic data," *J. Anal. Appl. Pyrolysis*, vol. 50, pp. 47-62, (1999).
3. K. W. Bauer Jr., S. G. Alsing, and K. A. Greene, "Feature screening using signal-to-noise ratios," *Neurocomputing*, vol. 31, pp. 29-44, (2000).
4. P. van de Laar and T. Heskes, "Input selection based on an ensemble," *Neurocomputing*, vol. 34, pp. 227-238, (2000).
5. N. R. Pal, "Soft computing for feature analysis," *Fuzzy Sets and Systems*, vol. 103, pp. 201-221, (1999).
6. X. Z. Wang and R. F. Li, "Combining conceptual clustering and principal component analysis for state space based process montoring," *Ind. Eng. Chem. Res.*, vol. 38, no. 11, pp. 4345-4358, (1999).
7. Junghui Chen and Jialin Liu, "Mixture principal component analysis models for process monitoring," *Ind. Eng. Chem. Res.*, vol. 38, no. 4, pp. 1478-1488, (1999).
8. D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison Wesley, Reading, Massachusetts, (1989).
9. M. Mitchell, *An Introduction to Genetic Algorithms*, The MIT Press, Cambridge, Massachusetts, (1996).
10. T. Takagi and M. Sugeno, " Fuzzy identification of systems and its applications to modeling and control," *IEEE Trans. System, Man, Cybern.*, vol. 1, pp. 116-132, (1985).
11. Man Gyun Na, Neuro-fuzzy control applications in pressurized water reactors: in Da Ruan (ed) *Fuzzy Systems and Soft Computing in Nuclear Engineering*, Springer-Verlag, Berlin Heidelberg New York, pp.172-207, (1999).
12. Won-Jae Lee, Bub-Dong Chung, Jae-Jun Jeong, Kwi-Seok Ha, and Moon-Kyu Hwang, Improved Features of MARS 1.4 and Verification, Korea Atomic Energy Research Institute, KAERI/TR-1386-99, (1999).
13. Michel Misiti, Yves Misiti, Georges Oppenheim, and Jean-Michel Poggi, *Wavelet Toolbox User's Guide*, MathWorks, Natick, MA, (1996).