Application of Neural Networks to Signal Prediction in Nuclear Power Plant

Wan Joo Kim, Soon Heung Chang, and Byung Ho Lee

Abstract—This paper describes the feasibility study of an artificial neural network for signal prediction. The purpose of signal prediction is to estimate the value of undetected next time step signal. As the prediction method, based on the idea of auto regression, a few previous signals are inputs to the artificial neural network and the signal value of next time step is estimated with the outputs of the network. The artificial neural network can be applied to the nonlinear system and answers in short time. The training algorithm is a modified backpropagation model, which can effectively reduce the training time. The target signal of the simulation is the steam generator water level. The basic idea is a variation of the auto regression (AR) [9], which is a powerful representation technique for the linear systems. In the typical AR method, the time series system is represented as an n-dimensional polynomial. Two well-known methods to compute the polynomial are the finite impulse response (FIR) method and the infinite impulse response (IIR) method. The computing time is so long for the FIR method that it cannot be applied to severe dynamic systems. The IIR method has the covariance windup problem when the perturbations are not persistently exciting. In addition, those two methods cannot be applied to the nonlinear system, which can be well manipulated by the artificial neural network.

In nuclear power plants, a differential method is usually used as in the PID (proportional-integral-differential) controller. The PID controller which is widely used in various areas in nuclear power plants analyzes the signal trend using the time-series differential. However, it does not inform the operator of the next time step value. In addition, the differential method is sensitive to the noise.

The developed ANN is a three layer feedforward network. The number of neurons for the input, output and hidden layer are nine (9), eight (8) and ninety (90), respectively. The large number of neurons for the hidden layer is due to the fact that the weights must store much information because of uncertainty and complexity in the trend of the steam generator water level. A modified backpropagation algorithm is used for training to reduce the training time.

II. SIGNAL PREDICTION AND ARTIFICIAL NEURAL NETWORK

A. Basic Concept

There are three types of signal estimation as compared in Fig. 1, they are smoothing, filtering, and prediction. Signal prediction is the estimation of the signal value to be detected in the near future by using the last available measurement data.

Signal prediction can be accomplished through the measured data and the physical model of a system.
ever, it is difficult to set up all required physical models for a nuclear power plant which has a large number of signals. Therefore, in this study, the ANN is applied without the physical model based on the idea of auto regression. (1) is auto regression and moving average (ARM) [8]:

$$s_n = - \sum_{k=1}^{p} a_k s_{n-k} + G \sum_{l=0}^{q} b_l u_{n-l}.$$  (1)

A signal $s_n$ is the output of a system with input $u_n$. The $a_k$, $b_l$ and gain $G$ are the parameters of the hypothesized system. In (1), the $u_{n-l}$ except $l=0$ can be neglected because their effects have been already reflected in the previous output values $s_{n-k}$ ($k=1,p$). This concept is called all pole model which reduces (1) as follows:

$$s_n = - \sum_{k=1}^{p} a_k s_{n-k} + Gb_0 u_n.$$  (2)

If the effect of $u_n$ on $s_n$ is assumed to be not significant, (2) becomes:

$$s_n = - \sum_{k=1}^{p} a_k s_{n-k}.$$  (3)

Fig. 2 shows the concept of the ANN structure based on (3).

B. Problems related to the ANN Application

For the training of the ANN, the previous data at the time of various trend are needed because the prediction system must be used at various circumstances. When the binary number is manipulated, the training patterns can be confirmed, e.g., the exclusive OR problem has four training patterns. When the analog number is manipulated, however, the training patterns are numerous. Therefore, the first problem is to determine the training patterns because they must include all the necessary data without repetition.

Another problem is the prediction error related to the modeling and the noise from the detector. As shown in (2) and (3), the predicted value is different from the real next time value because the input factors are not considered between the present time and the next prediction time. For this reason, the made training patterns include errors. In addition, the noise makes the situation worse since noise is included in the accumulated data of nuclear power plants.

Because of the above problems, the signal prediction by the ANN is not so simple that a large network is required and the local minima become a significant issue. Therefore, the modified algorithm should be developed to resolve these problems.

III. DEVELOPMENT OF THE ADVANCED TRAINING ALGORITHM

C. Backpropagation Algorithm

The basic backpropagation training algorithm is the generalized delta rule (GDR) suggested by Rumelhart and McClelland [6] in which the sum of squared errors of outputs is minimized by changing the weights. In the output layer, the error can be easily computed by using the difference between the desired and the actual outputs. For the hidden layer, however, the error cannot be computed directly because the desired outputs are not known. The GDR enables to compute the error at any neurons in any layers using back error propagation from the output layer. To do this, we compute the derivative of the error function with respect to each weight in the network and then change the weight as follows:

$$\Delta_j w_{ij} = - \eta \frac{\partial E_p}{\partial w_{ij}}.$$  (4)

In (4), $\Delta_j w_{ij}$ is the change of weight from the $i$th neuron to the $j$th neuron, where $E_p$ is the sum of squared errors.
for the training pattern $p$ and $\eta$ is a constant.

$$E_p = \frac{1}{2} \sum_j (t_{pj} - o_{pj})^2,$$

(5)

$t_{pj}$ is the desired output at the $j$th neuron of the output layer for the training pattern $p$, and $o_{pj}$ is the actual output produced through the network for the pattern $p$. $o_{pj}$ is represented as

$$o_{pj} = f(\text{net}_{pj}),$$

(6)

where $f(.)$ is an activation function and $w_{ji}$ is the weight from the $i$th neuron to the $j$th neuron. In this study, a sigmoid function, $f(x) = \frac{1}{1 + e^{-x}}$, is used for $f(.)$.

To obtain a rule for adjusting weights, the gradient of $E_p$ with respect to $w_{ji}$ is replaced as follows:

$$\Delta w_{ji} = \eta \delta_{pj} o_{pj},$$

(8)

where $\delta_{pj}$ is a random weight change in a certain range. The statistical method gives the weights the ability to go over the threshold between point A and B. However, it is a terribly time-consuming method for a complex network because the number of repeated trials required by (12) increases sharply with the number of weights.

In this study, a method to vary $\eta$ and $\alpha$ in (11) is proposed to avoid the local minima problem within a reasonable computation time by adopting some idea of the statistical method. Basic equations are as follows:

$$\eta(n+1) = \eta(n) + \Delta \eta(n+1)$$

(13)

$$\alpha(n+1) = \alpha(n) + \Delta \alpha(n+1)$$

(14)

where $\Delta \eta$ and $\Delta \alpha$ are random values in a certain range. The modified algorithm gives some random variation to $\alpha$ and $\eta$ instead of the weight itself in the statistical method. The result compared with ordinary backpropagation algorithm is represented in Fig. 4. The comparison test is executed with reduced sample training patterns for the prediction and then the convergence criteria is 2.3% RMS (root-mean-square) error. The modified backpropagation algorithm reaches the convergence criteria within 900 iterations and the ordinary backpropagation algorithm cannot reach the target within 5000 iterations.
detectors. The S/G water level is an important parameter for plant operation but is difficult to detect accurately. Actually, a large portion of the nuclear power plant trip during start up is related to the S/G water level. Thus the estimation of S/G water level has been studied frequently and there are many data for the ANN training.

Outputs from a computer code [12] are used as the training patterns since it is difficult to obtain actual plant data with severe changes of parameters. In that code, the S/G water level is analyzed based on the one-dimensional two-fluid model by dividing the S/G into three region: a steam dome, a downcomer and a boiling region. The code can simulate water level in a reasonable accuracy for wide variation of feedwater flow. The program outputs are transformed into adequate training patterns with incorporating noise which is not included in the program outputs. When the outputs are transformed into the training patterns, it is important to distribute the training patterns well because the training opportunity must be uniformly given to every case.

Fig. 5 shows the result of simulation by the developed ANN. It seems likely that the predicted value follows the real value trend well. If we look into the result, we can discover more important things. When the five signals make cup-shape in the Fig. 6, the abnormal high value appears like those of time 15 and 20. From the time 20 to 30, the inactive trend is shown and the noises are comparatively uniform. The prediction result is relatively good at that area.

The Fig. 7 shows slow decreasing and the noises are not dynamic. However, the detected signals show slow increasing trend because of the noises. As shown at the time 126, the prediction value reflects the cap-shape trend and at the time 127, the sudden increase of the value is less reflected.

In the Fig 6 and 7, the severe fluctuation is shown from the time 15 to 22 and from the time 118 to 122. At that time, the network think that the trend is sudden increasing or decreasing because the detected signals make cup-shape or cap-shape and then the noise is amplified in that status. However, it is not frequent happening in the real system.

As shown in the simulation result, the prediction trend does not run counter to the real value trend without the fluctuation. The fluctuation is the cause of that the signals appear to get out severely in the Fig. 5. At the section from the time 22 to 27 or from the time 123 to 135, the prediction trend is similar to detected value trend and the noise does not affect the result significantly.

B. Further Discussion

A modified algorithm is proposed for the ANN training in this paper. It can find better weights than the case of using constant \( \eta \) and \( \alpha \) as shown in Fig. 4. For the stopping rule, the limitation of error and iteration is used. However, more adequate stopping rule is needed to make the algorithm better [11].

Another issue is the change of \( \eta \). After training, the \( \alpha \) is almost unchanged and the \( \eta \) continues decreasing. If the initial \( \eta \) is 0.5, the final value is 0.2 - 0.3. If the initial value is 0.4, the \( \eta \) increases to about 0.45 and then begins...
to decrease $0.2 \sim 0.3$. It is the contrary behavior to the general $\eta$-acceleration algorithm. In order to explain the generalized $\eta$ acceleration algorithm, the (9) is reminded. As the weights are trained, the difference of the desired and the actual output decreases. And then, $\eta$ must increase relatively for the acceleration. However, this study shows the contrary trend. It is thought that the cause results from the error and noise included in the training patterns. As shown in (11), if the $\eta$ is small, the weight change is less sensitive to the difference of the desired and the actual output. Because the training pattern has error and noise, it is better that the weight is less sensitive to the $\delta$ after the network is trained to some degree. It may be an interesting problem whether adequate $\eta$ is a certain constant or not.

V. CONCLUSIONS

An artificial neural network for signal prediction is developed and applied to the S/G water level. The modified backpropagation algorithm adopted in this study shows better ability to escape from the local minima compared with the ordinary backpropagation algorithm. As the result, the training time is reduced effectively and the network performance is expected better, too. The target signal to simulate is S/G water level and the noise is considered. The simulation results show that the prediction system's performance is not significantly degraded by noise and the predicted value does not severely deviate from real value. By the prediction like this, the plant operator can predict the next state of the plant and it can help the determination of the operation strategy and the prohibition of an undesirable situation.

REFERENCES