Transactions of the Korean Nuclear Society Autumn Meeting
PyeongChang, Korea, October 30-31, 2008

Principal Component based Auto-Associative Support Vector Regression for Signal Validation in Nuclear Power Plants

In-Yong Seo[a*], Seong-Jun Kim[b]
[a]*Korea Electric Power Research Inc. , 65 Moonji-Ro, Yuseong, Daejeon, 305-380, Korea*
[b]*Dept. of Industrial Engr. Kangnung National Uni., 120 Daehangno Gangneung Gangwondo 210-702*
[*]*Corresponding author : iyseo@kepri.re.kr*

1. Introduction

For the past two decades, the nuclear industry has attempted to move towards condition-based maintenance philosophies using new technologies developed to ascertain the condition of plant equipment during operation. From the early 1980's the application of artificial intelligence techniques to nuclear power plants were investigated for instrument condition monitoring [1]. The Multivariate State Estimation System (MSET) was developed in the late 1980s [2]. And the Plant Evaluation and Analysis by Neural Operators (PEANO) was developed [3]; it uses auto-associative neural networks (AANN) and applies them to the monitoring of nuclear power plant sensors.

In this paper, a method that utilizes the attractive merits of principal component analysis (PCA) for extracting predominant feature vectors and Auto-Associative support vector regression (AASVR) for databased statistical learning is proposed for the on-line monitoring and signal validation with the use of real plant data.

2. PC based AASVR

An auto-associative model is a model whose outputs are trained to emulate its inputs over an appropriate dynamic range. An auto-associative model will estimate the correct input values using the correlations embedded in the model during its training. The estimated correct value from the auto-associative model can then be compared to the actual process parameter to determine if a sensor has drifted or has been degraded by another fault type. Fig. 1 shows the schematic diagram of the proposed PCA-AASVR (PCSVR) method for modeling measurements in Nuclear Power plant.

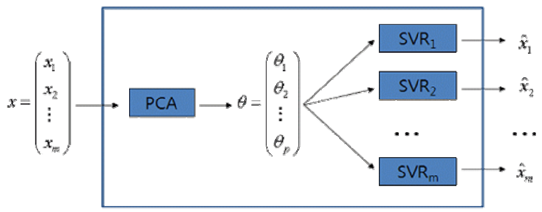

Fig. 1. The schematic diagram of PC-AASVR

In this paper, an SVM regression method is used for signal validation of the measurements in NPPs. The SVM regression is to map nonlinearly the original data into a higher dimensional feature space. Hence, given a set of data $\{(x_i, y_i)\}_{i=1}^{n} \in R^m \times R^m$ where $\mathbf{x}_i$ is the input vector to support vector machines, $\mathbf{y}_i$ is the actual output vector and n is the total number of data patterns. The multivariate regression function for each output signal is approximated by the following function,

$$y_k = f_k(\boldsymbol{x}) = \boldsymbol{w}_k{}^T \phi(\boldsymbol{x}) + b_k \quad (1)$$

where $\boldsymbol{w}_k = [w_1, w_2, \cdots, w_n]^T$, $\phi = [\phi_1, \phi_2, \cdots, \phi_n]^T$, $k = 1,2,\cdots,m$ and m is the number of sensor measurements. Also, the function $\phi_i(\boldsymbol{x})$ is called the feature. Equation (1) is a nonlinear regression model because the resulting hyper-surface is a nonlinear surface hanging over the m-dimensional input space. The parameters $\mathbf{w}$ and b are a support vector weight and a bias that are calculated by minimizing the following regularized risk function:

$$R(\boldsymbol{w}_k) = \frac{1}{2}\boldsymbol{w}_k{}^T \boldsymbol{w}_k + C_k \sum_{i=1}^{n} L_k(y_{k,i}) \quad (2)$$

where

$$L_k(y_{k,i}) = \begin{cases} 0, & |y_{k,i} - f_k(\boldsymbol{x})| < \varepsilon_k \\ |y_{k,i} - f_k(\boldsymbol{x})| - \varepsilon_k, & otherwise \end{cases} \quad (3)$$

Finally, the regression function of (1) becomes

$$y_k = \sum_{i=1}^{n} (\lambda_{k,i} - \lambda_{k,i}^*) \boldsymbol{K}(\boldsymbol{x}_i, \boldsymbol{x}) + b_k^* \quad (4)$$

where $\boldsymbol{K}(\boldsymbol{x}_i, \boldsymbol{x}) = \phi^T(\boldsymbol{x}_i)\phi(\boldsymbol{x})$ is called the kernel function.

$$b_k^* = -\frac{1}{2}\sum_{i=1}^{n} (\lambda_{k,i} - \lambda_{k,i}^*)[K(\mathbf{x}_r, x_i) + K(\mathbf{x}_s, x_i)] \quad (5)$$

where $\mathbf{x}_r$ and $\mathbf{x}_s$ are called support vectors (SVs)

The three most relevant design parameters for the AASVR model are the insensitivity zone, ε, the regularization parameter, C, the kernel function parameter, σ . In this paper those parameters were optimized by response surface methodology (RSM). In this study, they are assumed common in every model of SVR. The optimal point was searched on the response surface which minimizes mean squared error (MSE). The optimum point of the response surface is obtained as $(\sigma, \varepsilon, C) = (1.4, 0.0005, 6.8)$.

In this study. 7 PCs were selected that explains more than 99.98% of total variation.

3. Application Results

The real plant startup data of the Kori Nuclear Power Plant Unit 3 were applied to the PCSVR. The data is derived from the following 11 types of measured signals: the reactor power (the ex-core neutron detector signal), the pressurizer water level, the SG steam flow rate, the SG narrow range level, the SG pressure, the SG wide-range level, the SG main feedwater flow rate, the turbine power, the charging flow rate, residual heat removal flow rate and the reactor head coolant temperature. The data was divided into 5 subsets of equal size, named KR1 through KR5. Each subset is composed of 458 patterns which sampled every 5 minutes. KR3 is used for training and the rest 4 subsets are for test. Fig.1 shows normalized 11 input signals.

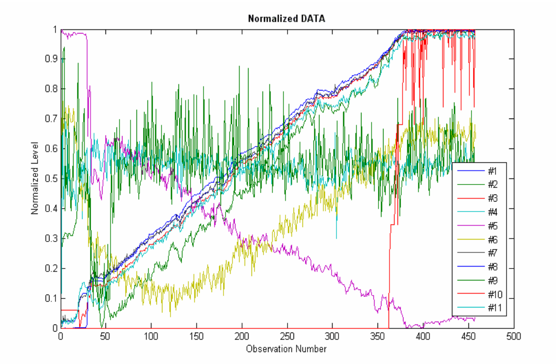

Fig. 1. Normalized input signals

Fig. 2 represents the RMS error and relative RMS errors compared with the ranges of sensors. They are [120, 100, 2.2, 100, 100, 100, 2.2, 1500, 45.4, 1140, 450]. The RMS error for the turbine power is higher than those of other signals. This is caused by the big standard deviation of the signal.

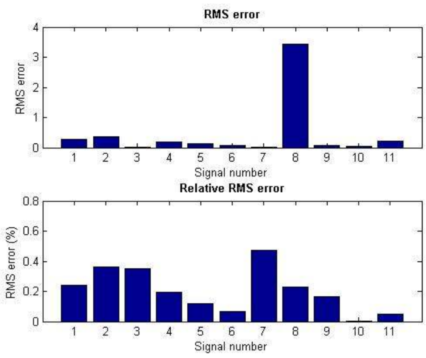

Fig. 2. RMS error

Fig. 3. represents hitogram of training and test data of the SG main feedwater flow rate. The standard deviations of the estimated error for training and test data are 0.3492% and 0.4711%, respectively.

In order to investigate the system's drift detection ability, we artificially degraded the SG main feed water flow rate signal (unit : Mkg/hr) in test subset KR1. The amount of drift is linearly increasing with time and its total amount of drift at the end point is assumed 5%.

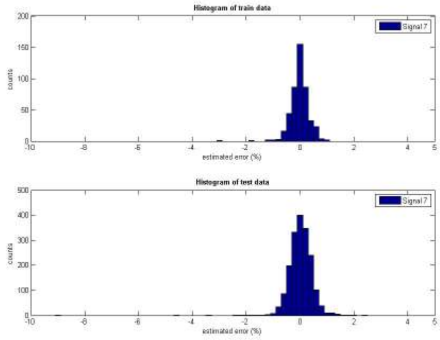

Fig. 3. Histogram of train and test data

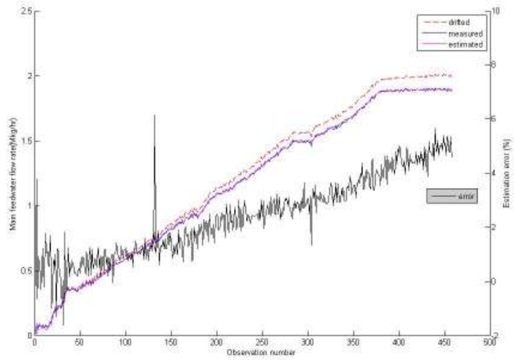

Fig. 4. Drifted signal and residual of estimate

Drifted, its original and estimated signals are depicted on left Y-axis while estimation error is on right Y-axis in Fig. 4. The predicted feedwater flow rate is almost the same as the actual feedwater flow rate even though the measured feedwater flow rate had degraded. Also, the estimation error is increasing as the drift progresses so that we could detect the sensor drift.

4. Conclusion

We developed PC based AASVR for signal validation and applied Kori Nuclear Power Plant unit 3 startup data to the developed algorithm. It shows very accurate estimate for each signal and even for the drifted signal. This proposed algorithm can be used for the calibration monitoring system in nuclear power plants. For further study grouping or clustering technique could be adopted.

REFERENCES

[1] Upadhyaya, B. R., and E. Eryurek (1992), "Application of Neural Networks for Sensor Validation and Plant Monitoring," *Nuclear Technology*, 97, 170–176 (February 1992).

[2] Mott, Y., and R. W. King (1987), *Pattern Recognition Software for Plant Surveillance*, U.S. DOE Report.

[3] Fantoni, P., S. Figedy, and A. Racz (1998), "A Neuro-Fuzzy Model Applied to Full Range Signal Validation of PWR Nuclear Power Plant Data", FLINS-98, Antwerpen, Belgium.