Transactions of the Korean Nuclear Society Spring Meeting
Jeju, Korea, May 22, 2009

Statistical Learning Framework with Adaptive Retraining for Condition-Based Maintenance

Sang Ha An[a], Gyunyoung Heo[b*], Ho Joon Seo[c], Su Young Kim[c], Soon Heung Chang[a]
[a] *Korea Advanced Institute of Science and Technology, Yuseong-gu, Daejeon, 305-701, Korea*
[b] *Kyung Hee University, Yongin-si, Gyeonggi-do, 446-701, Korea*
[c] *BNF Technology Inc., Yuseong-gu, Daejeon, 305-500, Korea*
[*] *Corresponding author: gheo@khu.ac.kr*

1. Introduction

As systems become more complex and more critical in our daily lives, the need for the maintenance based on the reliable monitoring and diagnosis has become more apparent. However, in reality, the general opinion has been that 'maintenance is a necessary evil' or 'nothing can be done to improve maintenance costs'. Perhaps these were true statements twenty years ago when many of the diagnostic technologies were not fully developed. The developments of microprocessor or computer based instrumentation that can be used to monitor the operating condition of plant equipment, machinery and systems have provided the means to manage the maintenance operation. They have provided the means to reduce or eliminate unnecessary repairs, prevent catastrophic machine failures and reduce the negative impact of the maintenance operation on the profitability of manufacturing and production plants. Condition-based maintenance (CBM) techniques help determine the condition of in-service equipment in order to predict when maintenance should be performed.

Most of the statistical learning techniques are only valid as long as the physics of a system does not change. If any significant change such as the replacement of a component or equipment occurs in the system, the statistical learning model should be re-trained or re-developed to adapt the new system.

In this research, authors will propose a statistical learning framework which can be applicable for various CBMs, and the concept of the adaptive retraining technique will be described to support the execution of the framework so that the monitoring system does not need to be re-developed or re-trained even though there are any significant changes in the system or component.

2. Methods and Results

Figure 1 shows the framework of statistical learning technique composed of two streamlines: The prediction model generates the model estimates through the collected process variables by using a regression algorithm. And the detection model detects system anomaly by comparison with a measured signal and the model estimates.

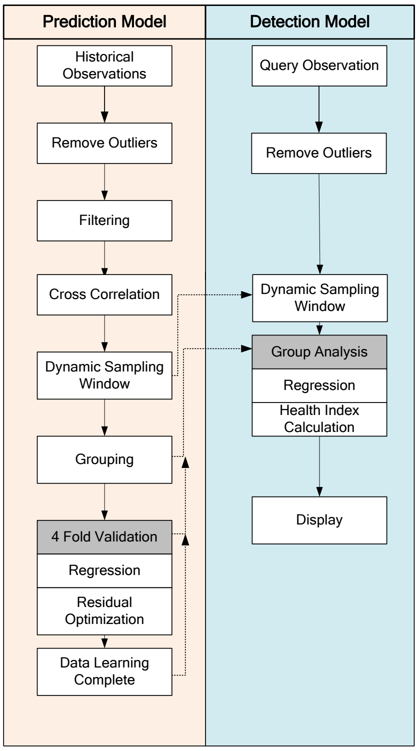

Fig. 1. Framework of a monitoring system

2.1 Prediction Model

In the prediction model, the collected data throughout data qualification processes are stored to generate the model estimates. There are three parts: 1) to collect historical observations for statistical learning, 2) to go through the multi-steps of pre-processing, the removal of outliers and the filtering of signals to guarantee the better quality of observations, the cross correlation analysis between process variables and the dynamic sampling window to eliminate time lag between process variables, and the grouping process to divide the process variables into several groups based on the cross correlation analysis, and n-fold validation to find and optimize the regression parameters.

2.2 Detection Model

Query observation signals throughout the data qualification are sent to the detection model and compared with model estimates generated by prediction model. Sequential probability ratio tests (SPRT) are performed to detect system anomaly by using the residuals between measured signal and model estimates.

2.3 Sustainable Monitoring System

Figure 2 shows framework of typical monitoring system. Prediction model process is performed as offline process before detection model monitors system.

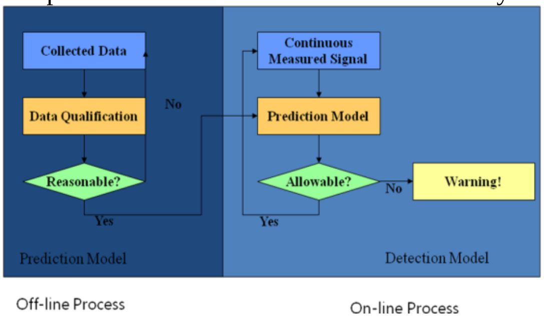

Fig. 2. Framework of a typical monitoring system

New system has been proposed to support sustainable monitoring system. Prediction model is not performed as pre-process before detection model but as online process. Continuous measured signal is sent to detection model and also stored in the historical exemplar observation so it will be used for next detection. Figure 3 shows the concept of sustainable monitoring system.

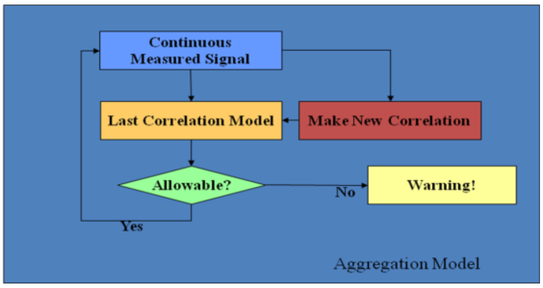

Fig. 3. Concept of the sustainable monitoring system

2.4 Time Adaptive Kernel Regression

In this research, kernel regression algorithm has been introduced for the prediction model. Kernel regression (KR) is a nonparametric, empirical modeling technique that uses historical, exemplar observations to make predictions.

In statistics and empirical modeling, the process of estimating a parameter's value by calculating a weighted average of historical, exemplar observations is known as kernel regression. Generally, KR may be most compactly represented by the so-called Nardaraya [1964]-Watson [1964] estimator [1]. For a simple single-input, single-output (SISO) regression model, where the input x is used to estimate the output y.

Step 1 – Distance Calculation

$$d(X_i, x) = \sqrt{(X_i - x)^2} \qquad (1)$$

Step 2 – Similarity Quantification

Step 3 – Output Estimation

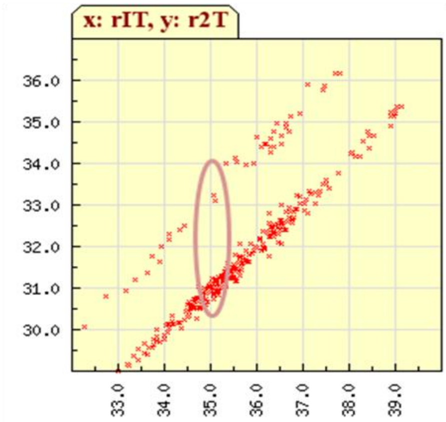

Fig. 4. Remaining information after system change

Collected information before system change still remains. So prediction algorithm should adapt to new system. Figure 4 shows the effect of remaining information before system change.

In equation (2), time gap or difference of sequential order between query vector and memory vector should be considered as distance of kernel regression.

$$d(X_i, x) = \sqrt{(X_{i,1} - x_1)^2 + c_t (X_{i,2} - x_2)^2} \quad (2)$$

where c_t is time weighting coefficient, X is query observation and x is historical exemplar observation.

From this modification of distance calculation in the kernel regression, the recent historical exemplar observation data has more weight than past one. And as time weighting coefficient become larger, this effect will become more powerful.

3. Conclusions

Time adaptive kernel regression developed in this research enables a monitoring system to learn by itself, so models do not need to be re-trained or re-developed periodically after any significant change to the system.

ACKNOWLEDGEMENT

This work is the outcome of a Manpower Development Program for Energy & Resources supported by the Ministry of Knowledge and Economy (MKE). This work was also supported and funded by BNF Technology Inc.

REFERENCES

[1] Nadaraya, E. A. (1964). "On Estimating Regression". *Theory of Probability and its Applications* **9** (1): 141–142(1964).