Soft-Sensing for Feedwater Flowrate Using a GMDH Algorithm

Dong Hyuk Lim, In Ho Bae, and Man Gyun Na
Department of Nuclear Engineering, Chosun University, 375 Seosuk-dong, Dong-gu, Gwangju 501-759
Corresponding author: magyna@chosun.ac.kr

1. Introduction

Now that thermal reactor power is typically assessed by secondary system calorimetric calculations based on the accurate measurement of feedwater flow rate, it is very necessary to exactly measure the feedwater flowrate. In most pressurized water reactors, Venturi meters are used to take the measurement of the feedwater flowrate. But the fouling phenomena of the Venturi meter make the accuracy of the existing hardware sensors worse. Consequently, it is essential to resolve the inaccurate measurement issue of the feedwater flowrate. In this study, In order to precisely estimate online the feedwater flowrate, a soft sensing model for feedwater flowrate monitoring is developed by using the Group Method of Data Handling (GMDH) Algorithm [1].

2. Inferential Sensing for Feedwater Flowrate

In order to solve engineering system problems with control, monitoring, diagnostics and so on, a lot of exact mathematical models have been studied. Among them, the GMDH method which is one of the data driven models such as ANN (Artificial Neural Networks) is used for accurate estimation in this paper. It is because data driven models are famous for superior capability in modeling complex systems and have the two strong points of easy implementation and accuracy.

2.1 Basic GMDH Algorithm

The GMDH algorithm is the way to find a function that can well express a dependent variable from independent variables. This method uses a data structure similar to that of multiple regression models. The data set can be divided into the training data and test data. The reason of dividing the data set is to prevent over-fitting and maintain model parsimony. Figure 1 shows the data structure used in the GMDH method with n being the number of observations and m the number of prediction inputs. Data sets from 1 to t can be denoted as the training data set and are used for model fitting. And sets from t+1 to n can be denoted as the test data set and check the error.

The GMDH uses a self-organizing modeling algorithm with the flexibility of deciding nonlinear forms of the basic inputs $\{x_1, x_2, \ldots, x_m\}$. Figure 2 shows the branch architecture of the GMDH module. It starts with the basic inputs at the first layer and becomes more complex according to the increasing number of layers.

$$\begin{array}{lcccc}
y_1 & x_{11} & x_{12} & \cdots & x_{1m} \\
y_2 & x_{21} & x_{22} & \cdots & x_{2m} \\
. & . & . & . & . \\
. & . & . & . & . \\
. & . & . & . & . \\
y_t & x_{t1} & x_{t2} & \cdots & x_{tm} \\
. & . & . & . & . \\
. & . & . & . & . \\
. & . & . & . & . \\
y_n & x_{n1} & x_{n2} & \cdots & x_{nm}
\end{array}$$

Fig.1 GMDH data structure

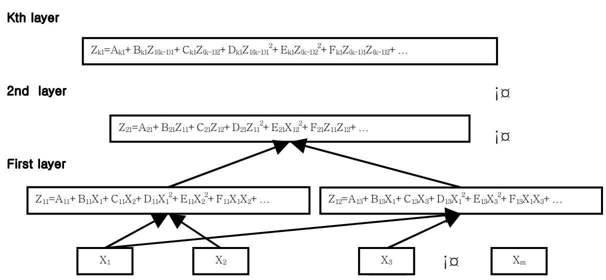

Fig.2 GMDH structure

The original GMDH method includes the following Eq. (1) at each level of the successive approximation:

$$Z = A + Bx + Cy + Dx^2 + Ey^2 + Fxy + \ldots \quad (1)$$

The reference function above can be solved by using the method of least squares using the data of training set about arbitrary pairs of the independent variables $\{x_1, x_2, \ldots, x_m\}$. Although only quadratic terms are shown in Eq. (1), more complex functional forms such as ratio terms, trigonometric, logic and exponential terms, could also be incorporated as input terms according to the complexity of system. The GMDH algorithm constructs a high-order polynomial of Kolmogorov-Gabor form:

$$y = a_0 + \sum_{i=1}^{m} a_i x_i + \sum_{i=1}^{m}\sum_{j=1}^{m} a_{ij} x_i x_j + \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{m} a_{ijk} x_i x_j x_k \ldots \quad (2)$$

The configuration of lower order polynomials shown above is used. That is, the GMDH algorithm amalgamates lower order regression type polynomials at each generation to reach the next generation. This uses the composition of lower order polynomials mentioned above. That means the GMDH algorithm combines lower order regression type polynomials at each generation to arrive at the next generation. This process continues until the GMDH model starts to simulate the noise in training or it exceeds maximum calculation time.

2.2 Implementation

The GMDH algorithm has been developed and improved in many applications. The main steps in its implementation are given below.

First step is to classify the data. That is, after constructing the input and corresponding output data for GMDH model, it is divided into training and test sets.

Second step is to establish a new variable. The external inputs have to be chosen to the GMDH network. And then calculate regression polynomial parameters for each pair of input variables x and combined output y in the training sets. Thus Least-squares error (LSE) linear regression parameters are calculated.

Next step is to choose the variables which have little contribution. That is, the output calculated in the second step from the test data set checks and then replaces the variable as variables of next generation. A measure used to evaluate the new variables at each generation is the fractional error defined as:

$$r_j^2 = \frac{\sum_{i=1}^{n}\left(y_i - z_{ij}\right)^2}{\sum_{i=1}^{n} y_i^2} \quad \text{for } j=1,2,\cdots \tag{3}$$

where j is the number of layer. Arranging the columns of Z by increasing order of r_j. An arbitrary cut-off value 'R' needs to be selected by the analyst. All the columns of Z satisfying $r_j < R$ are picked to replace the input terms in the previous layer, and all the variables with $r_j > R$ are screened out and are not passed onto the next generation of the algorithm.

Last step is to take the optimality test. The process above is performed repeatedly until over-fitting is found through cross checking; that is, when the root mean square error (RMSE) of current layer is larger than the last layer. In detail the procedure is: the minimum value of those rj's for generation k is denoted as $RMIN_k$, if $RMIN_k > RMIN$, then the training and testing processes of the algorithm stop and the polynomial with the minimum value of the error criterion in layer k-1 is selected to be the final approximate model. Otherwise, the algorithm moves to the next layer and repeats the above steps.

It has been shown that RMIN curve has the quadratic shape and the GMDH converges to a global minimum value.

3. Application to the Feedwater Flow Measurement

The proposed method was verified by applying to the real values measured from the primary and secondary systems of Yonggwang nuclear power plant unit 3. The data includes sixteen types of measured signals as follow: SG pressure, SG steam flow rate, SG feedwater flow rate and so on. The optimum fit can be obtained by applying entire steps to the acquired data, focused on the steam generator (SG).

Table 1 summarizes the performance results of the soft sensing model for feedwater flowrate. If the GMDH model is optimally identified at first by using the training data and the verification data, the GMDH model can be appropriately used to estimate the feedwater flowrate. In the simulations, the RMS error and the relative maximum error are 0.1532% and 0.9906% for the test data, respectively.

Table 1. Performance results of inferential sensing for feedwater flowrate.

Data type	Root mean square error (%)	Relative maximum error (%)	Number of data
Training Data	0.1694	1.5113	1198
Verification Data	0.1694	1.4841	300
Validation Data	0.1532	0.9906	300

4. Conclusion

In order to prevent the fouling phenomena caused in using the Venturi meter, a soft sensing model which estimates the feedwater flowrate signal has been developed to validate and monitor the existing feedwater flowrate. In this paper, the GMDH method which features easy implementation and accuracy is applied to the soft sensing model for feedwater flowrate monitoring. By using sixteen measured signals it is possible to calculate the solution of the reference function. And then we can get the optimum fit by repeating the entire steps. It is expected that this model can be applied to validate and monitor the existing feedwater flow meters successfully.

References

[1] Baofu Lu, "Development of an Incipient Fault Detection and Isolation Method for the Steam Generator System of a Nuclear Power Plant," Ph.D. Dissertation, The University of Tennessee, Knoxville, 2002.