Intelligent System Predictor using Virtual Neural Predictive Model

박상민, 배재호, 왕지남
Mechanical and Industrial Engineering Division
Ajou University, Suwon Korea 442-749
(Fax) 82-331-219-1610 (Tel) 82-331-219-2425
(Email Add: gnwang@madang.ajou.ac.kr)

Abstract

A large system predictor, which can perform prediction of sales trend in a huge number of distribution centers, is presented using neural predictive model. There are 20,000 number of distribution centers, and each distribution center need to forecast future demand in order to establish a reasonable inventory policy. Therefore, the number of forecasting models corresponds to the number of distribution centers, which is not possible to estimate that kind of huge number of accurate models in ERP (Enterprise Resource Planning) module. Multilayer neural net as universal approximation is employed for fitting the prediction model. In order to improve prediction accuracy, a sequential simulation procedure is performed to get appropriate network structure and also to improve forecasting accuracy. The proposed simulation procedure includes neural structure identification and virtual predictive model generation. The predictive model generation consists of generating virtual signals and estimating predictive model. The virtual predictive model plays a key role in tuning the real model by absorbing the real model errors. The complement approach, based on real and virtual model, could forecast the future demands of various distribution centers.

I. Introduction

Modeling and identification method, which is to find the characterization of the structure of a mathematical input-output relationship from observed data, are needed for interpretation of observation and measurements. The interpretation could be utilized directly for monitoring, diagnosing, and analyzing process. However, there are no general methods that always can be used to get a complete model. A process cannot not be characterized by one mathematical model [Astrom and Wittenmark 1990].

In modeling a time-vary system like a non-stationary process, multiple models are frequently used. Since the mapping relation of a non-stationary process could be time dependent, time dependent multiple models are utilized for describing dynamic processes. This concept is interpreted as task decomposition in that a non-stationary process is decomposed into a set of stationary process, and each sub-stationary process is characterized by one model [Wang 1996].

The situation is occurred that there is no prior information on a process and even multiple models are not appropriate for characterizing a process. It might be difficult to establish a mathematical relationship of a process from the collected observations and measurements. The difficulties in modeling can be caused by the complexities of process nature. The difficulties could be also caused by that observations and measurements are not suitable for describing a process.

This paper is motivated by modeling a process when a process is not easily characterized either by one model nor by collected variables. The collected variables could be decomposed into other sub-signals which are not available for measurements and correlated each others. Real and virtual modeling approach are complementarily developed for increasing modeling accuracy. Real modeling, which is interpreted as a coarse modeling, uses real measurements for modeling an input-output system relation and the proposing virtual modeling is also performed continuously using the real modeling errors. In virtual modeling, newly generated input signals, which are obtained summing the real modeling error and artificially generated uniform series, are utilized for characterizing mapping relation.

The paper consists of five sections. Some background and a problem definition are described in section 2. In section III, the multi-phase system identification method is presented. The proposed on-line system identification using neural network and virtual system generation is suggested. In section IV, computer simulation is given to verify the proposed method. In final section, conclusion and further research issues are discussed.

II. Problem Statements and Background

Problem Statements

Consider the discrete-time, non-linear stochastic model.

$$y_t = g(X^1{}_{t-1},..X^1{}_{t-n_1},X^2{}_{t-1},..,X^2{}_{t-n_2},..,X^k{}_{t-1}.. \\ ..,X^k{}_{t-n_k},u_{t-1},..,u_{t-n_u},..,e_{t-1},..,e_{t-n_e},\Theta_g)+e_t \quad (1)$$

where y_t is the output measurement, u_t is the control input, e_t is the Gaussian white-noise process with known variance σ_e^2, Θ is the parameter vector, and t is the discrete time index. The model orders n_k, n_u, n_e, and the function g are unknown. $X^i{}_t$ is sub-system measurement which is not available for measurement. In this case, a process model should be characterized by the measurable existing observations. The main issue of this paper is to establish model structure using multi-phase system identification as follows.

$$y_t = R_1(y_{t-1},...,y_{t-n_y},u_{t-1},..,u_{t-n_u},..,e^1{}_{t-1},..,e^1{}_{t-n_e},\Theta_{R_1}) \\ + V_1(A^1{}_{t-1},...,A^1{}_{t-n_A},e^1_{t-1},...,e^1_{t-n_{e1}},\Theta_{V_1}) \\ + V_2(A^2{}_{t-1},...,A^2{}_{t-n_A},e^2_{t-1},...,e^2_{t-n_{e2}},\Theta_{V_2})+..... \\ + V_k(A^k{}_{t-1},...,A^k{}_{t-n_A},e^{k-1}_{t-1},...,e^{k-1}_{t-n_{e-1}},\Theta_{V_k})+w_t$$

(2)

$\bar{A}^i = (A^i_{t-1}, \ldots, A^i_{t-n_A})$ is a collection of time series which is followed by the given compact uniform distribution from $U[0, 1]$ for all $i = 1, 2, \ldots k$. R_1 and V_i are an unknown mapping function. W_t is the present modeling error at time t. The focus is given to verify that the process model in equation (1) could be approximated by the multi-phase model described in equation (2). The multi-phase model approximation is more practical and applicable in that it is based on all measurable and obtainable variables. Even though the multi-phase model relies on virtual series of a given compact distribution, all elements are computed or observed while the model in equation (1) is based on some no-measurable elements.

$$e^1_{t-1} = y_{t-1} - R_1(y_{t-2}, \ldots, y_{t-n_y-1}, u_{t-2}, \ldots, u_{t-n_u-1}, \ldots, e^1_{t-2}, \ldots, e^1_{t-n_e-1}, \Theta_{R_1}) \quad (3)$$

$$e^2_{t-1} = y_{t-1} - R_1(y_{t-2}, \ldots, y_{t-n_y-1}, u_{t-2}, \ldots, u_{t-n_u-1}, \ldots, e^1_{t-2}, \ldots, e^1_{t-n_e-1}, \Theta_{R_1}) - V_1(A^1_{t-1}, \ldots, A^1_{t-n_A}, e^1_{t-2}, \ldots, e^1_{t-n_e-1}, \Theta_{V_1})$$
$$= e^1_{t-1} - V_1(A^1_{t-1}, \ldots, A^1_{t-n_A}, e^1_{t-2}, \ldots, e^1_{t-n_e-1}, \Theta_{V_1})$$

and

$$e^{k-1}_{t-1} = y_{t-1} - R_1(y_{t-2}, \ldots, y_{t-n_y-1}, u_{t-2}, \ldots, u_{t-n_u-1}, \ldots, e^1_{t-2}, \ldots, e^1_{t-n_e-1}, \Theta_{R_1}) - V_1(A^1_{t-1}, \ldots, A^1_{t-n_A}, e^1_{t-2}, \ldots, e^1_{t-n_e-1}, \Theta_{V_1})$$
$$- V_2(A^2_{t-1}, \ldots, A^2_{t-n_A}, e^2_{t-2}, \ldots, e^2_{t-n_e-1}, \Theta_{V_2}) - \ldots$$
$$- V_{k-2}(A^{k-2}_{t-1}, \ldots, A^{k-2}_{t-n_A}, e^{k-3}_{t-2}, \ldots, e^{k-3}_{t-n_e-1}, \Theta_{V_{k-2}})$$
$$= e^{k-2}_{t-1} - V_{k-2}(A^{k-2}_{t-1}, \ldots, A^{k-2}_{t-n_A}, e^{k-2}_{t-2}, \ldots, e^{k-2}_{t-n_e-1}, \Theta_{V_{k-2}})$$

Now, a recursive relation in modeling error is obtained as follows

$$e^{k-1}_{t-1} = e^{k-2}_{t-1} - V_{k-2}(A^{k-2}_{t-1}, \ldots, A^{k-2}_{t-n_A}, e^{k-2}_{t-2}, \ldots, e^{k-2}_{t-n_e-1}, \Theta_{V_{k-2}}) \quad (4)$$

From the equation (2), the present multi-phase modeling error, W_t, is obtained as

$$W_t = y_t - R_1(y_{t-1}, \ldots, y_{t-n_y}, u_{t-1}, \ldots, u_{t-n_u}, \ldots, e^1_{t-1}, \ldots, e^1_{t-n_e}, \Theta_{R_1})$$
$$- V_1(A^1_{t-1}, \ldots, A^1_{t-n_A}, e^1_{t-1}, \ldots, e^1_{t-n_{e1}}, \Theta_{V_1})$$
$$- V_2(A^2_{t-1}, \ldots, A^2_{t-n_A}, e^2_{t-1}, \ldots, e^2_{t-n_{e2}}, \Theta_{V_2}) + \ldots$$
$$- V_k(A^k_{t-1}, \ldots, A^k_{t-n_A}, e^{k-1}_{t-1}, \ldots, e^{k-1}_{t-n_{e-1}}, \Theta_{V_k})$$
$$= e^1_t - V_1(A^1_{t-1}, \ldots, A^1_{t-n_A}, e^1_{t-1}, \ldots, e^1_{t-n_{e1}}, \Theta_{V_1})$$
$$- V_2(A^2_{t-1}, \ldots, A^2_{t-n_A}, e^2_{t-1}, \ldots, e^2_{t-n_{e2}}, \Theta_{V_2}) + \ldots$$
$$- V_k(A^k_{t-1}, \ldots, A^k_{t-n_A}, e^{k-1}_{t-1}, \ldots, e^{k-1}_{t-n_{e-1}}, \Theta_{V_k})$$
$$= e^{k-1}_{t-1} - V_k(A^k_{t-1}, \ldots, A^k_{t-n_A}, e^{k-1}_{t-1}, \ldots, e^{k-1}_{t-n_{e-1}}, \Theta_{V_k}).$$

Related works

Lapedes and Farber [1987] presented , who used a 1-15-1(i.e. one input node, 15 hidden nodes and one output node) feed-forward back-propagation network to map the deterministic input-output functions,

$$y_t = 4x_{t-1}(1 - x_{t-1})$$

which are quadratic and smooth in the interval $0 \le x \le 1$. It was a clear demonstration of the power of a multilayer network to approximate a non-linear function without a prior information about the system model. Similar results are obtained by Hecht-Nielsen [1989] and Funahashi [1989]. They show that any continuous function is approximately realizable by network with monotone increasing continuous or a sinusoidal activation function.

Narendra and Parthasarathy (1990) presented a novel and clear neural network application for both nonlinear-dynamical system identification and control. They considered SISO four models which can also be generalized to multivariable case. They gave a clear demonstration how neural network could be used for estimating mapping function $f()$ and $g()$, which could be used effectively for identification and control of nonlinear dynamical systems.

Recently, Wang(1996) presented multi-phase modeling approach to prediction of a non-stationary process. His approach was based on decomposing a non-stationary series into a set of stationary sub-series. After a non-stationary process is decomposed into appropriate numbers of stationary processes, an multi-phase RBF neural network is designed for modeling the decomposed stationary process.

III. Multi-phase Neuro-identification by Virtual System Generation

1) Multi-phase System Identification

Since a process might not be easily characterized by one mathematical model, one could be represented effectively by multiple models ranging from detailed and complex simulation models to very simple models [Astrom and Wittenmark 1990]. The existing works showed that neural network could be utilizes for approximating a mapping function. Multiple models, which is aimed at increasing modeling accuracy, could be associated with multiple neural networks. Following equations show how multi-phase system identification is designed using multiple neural networks. Suppose the

$$y_t = f(y_{t-1}, \ldots y_{t-n_y}, u_{t-1}, \ldots, u_{t-n_u}, e_{t-1}, \ldots, e_{t-n_e}, \Theta_f) + e_t$$

$$\hat{y}_t = N_1(y_{t-1}, \ldots y_{t-n_y}, u_{t-1}, \ldots, u_{t-n_u}, e_{t-1}, \ldots, e_{t-n_e}, W_{N_1}) = N_1(I_1^t, W_{N_1}) \quad (5)$$

$$e_t^{N_1} = y_t - \hat{y}_t^{N_1} \quad (6)$$

In equation (5), a single neural network, which is associated with single phase modeling, is employed for stochastic system identification. In this case the mapping structure has $(n_y + n_u + n_e)$ to one mapping structure where the corresponding input vector is equal to $I_1^t = (y_{t-1}, \ldots y_{t-n_y}, u_{t-1}, \ldots, u_{t-n_u}, e_{t-1}, \ldots, e_{t-n_e})$ and W_{N_1} is the weight vector. Since $\hat{y}_t^N$ is the estimated system output, the modeling error between the output generated the first phase neural network and the observed output y_t is $e_t^{N_1}$ which is expressed in equation (6).

Since a complex process is not easily characterized by one mapping model, the first phase modeling errors might be required to perform re-modeling procedure in equation (7). The re-modeling procedure is based on the fact that the series of modeling errors have another mapping characterization, which means they are not described by identically independent Normal distribution.

$$e_t^{N_1} = f_{N_1}(e_{t-1}^{N_1}, e_{t-2}^{N_1}, \ldots, e_{t-N_2^1}^{N_1}) + w_t \quad (7)$$

where $f_{N_1}(\,.\,)$ is an unknown function and w_t is white noise.

$$\hat{e}_t^{N_1} = N_2(e_{t-1}^{N_1}, e_{t-2}^{N_1}, \ldots, e_{t-N_2^1}^{N_1}, W_{N_2^1}) \quad (8)$$

$$e_t^{N_2} = e_t^{N_1} - \hat{e}_t^{N_1} \quad (9)$$

The two-phase model approximation could be expressed as

$$\hat{y}_t^{N_1 . N_2} = N_1(I_1^t, W_{N_1}) + N_2(e_{t-1}^{N_1}, e_{t-2}^{N_1}, \ldots, e_{t-N}^{N_1}, W_{N_2^1}) = \hat{y}_t^{N_1} + N_2(I_2^t, W_{N_2})$$

and it's overall modeling error is

$$y_t - \hat{y}_t^{N_1 . N_2} = y_t - [\hat{y}_t^{N1} + N_2(I_2^t, w_{N_2})] = e_t^{N_1} - \hat{e}_t^{N_1} = e_t^{N_2} \quad (10)$$

Without loss of generality, the k-phase system modeling could be extended as followings.

$$\hat{y}_t^{N_1 . N_2 \ldots N_i} = N_1(I_1^t, W_{N_1}) + N_2(I_2^t, W_{N_2}) + \ldots + N_k(I_k^t, W_{N_k})$$

(11)

$$e_t^{N_k} = y_t - \hat{y}_t^{N_1 . N_2 \ldots . N_k}$$

where

$$I_k^t = (e_{t-1}^{N_{k-1}}, e_{t-2}^{N_{k-1}}, e_{t-3}^{N_{k-1}}, \ldots, e_{t-N_{ik}^{k-1}}^{N_{k-1}}) \quad (12)$$

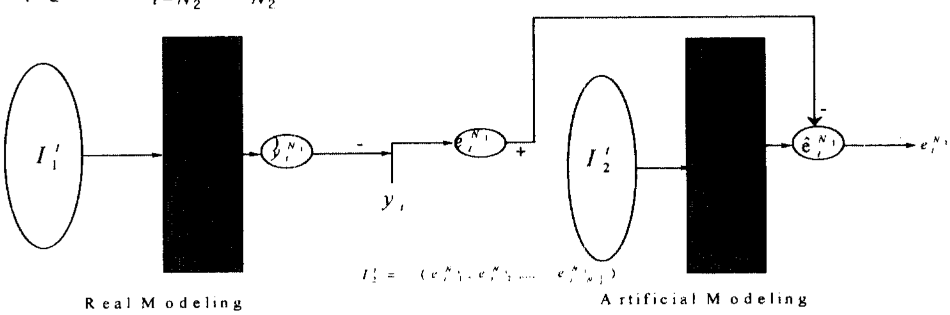
where $N_k(I_k^t, W_{N_k})$ denotes the mapping structure by the phase-*k* neural network. I_k^t is the corresponding network's input vector and W_{N_k} is the weight vector.

Equation (11) and (12) describe multi-phase model identification using multiple neural networks. The input-ouput mapping elements are obtained by computing previous modeling error recursively as follows.

$$e_t^{N_k} = e_t^{N_{k-1}} - \hat{e}_t^{N_{k-1}} = e_t^{N_{k-1}} - N_k(I_k^t, W_{N_k})$$

$$I_k^t = (e_{t-1}^{N_{k-1}}, e_{t-2}^{N_{k-1}}, e_{t-3}^{N_{k-1}}, \ldots, e_{t-N_{ik}^{k-1}}^{N_{k-1}})$$

Figure 1 represents the block diagram of multi-phase modeling system. The phase-1 neural network is the real model of the only existing real system, while the next-phase neural network (i.e., the phase-2, the phase-3, .., and the *k*-phase neural network) are not existing systems, which are synthetically generated.



(Figurel)

2) Virtual System Generation

The problem of the multi-phase identification, considered in the previous section, consists of iterative setting up a suitable neural network model at each phase and also consist of a proper approximation of $e_t^{N_{i-1}}$ by the network output $\hat{e}_t^{N_{i-1}}$ at each phase $i \in [1, K]$. The approximation capability will be subject to an implicit relation between the input $I_i^t = (e_{t-1}^{N_{i-1}}, e_{t-2}^{N_{i-1}}, e_{t-3}^{N_{i-1}}, \ldots, e_{t-N_{ik}^{k-1}}^{N_{i-1}})$ and the desired output $e_t^{N_{i-1}}$. If any system identification model does not exist between the input I_i^t and output $e_t^{N_{i-1}}$, multi-phase modeling does not have any superior advantage to single-phase modeling. The multi-phase modeling errors does not either guarantee that they are smaller than the single-phase modeling error nor guarantee that they are asymptotically stable with the white noise property.

The main concern of this research is to improve the multi-phase modeling accuracy by artificial constructing the mapping structure using newly generated I-O variables. A compact deterministic structure is employed as a basic model structure, and this structure is modified by unifying artificial and real system variable. The virtual compact mapping is used for absorbing the multi-phase modeling errors. Consider following virtual mapping structure and I-O variable.

$$Z_t = H(x_t, x_{t-1}, \ldots x_{t-n_x})$$

where x_t is in the interval $0 \le x_t \le 1$ and is followed by the given compact uniform distribution. The previous feed-forward network, which is employed by Lapedes and Farber [1987], could map the deterministic input-output functions effectively a

$\hat{Z}_t = NH(x_t, x_{t-1}, \ldots x_{t-n_x-1}, W^{Nf})$ where $NH()$ is the output of neural network and W^{Nf} is the corresponding weight vector. The output errors, $Z_t - \hat{Z}_t$, tends to be very small since the mapping structure could be characterized by a compact deterministic function. Using this virtually well-determined system, new system identification model will be created for reducing the multi-phase modeling errors by unification of the identified system model with each-phase mapping system generated from multi-phase modeling procedure.

IV. Computer Simulation

In this section simulation results are given to verify the presented idea.

Example: The process models to be considered are auto-regressive models which are synthetically generated.

$$AR(2): X_t^1 = 1.49X_{t-1}^1 - 0.653X_{t-2}^1 + e_t^1$$
$$AR(3): X_t^2 = 2.146X_{t-1}^2 - 1.598X_{-2}^2 + 0.409X_{t-3}^2 + e_t^2$$
$$AR(4): X_t^3 = 1.876X_{t-1}^3 - 1.781X_{t-2}^3 - 1.201X_{t-3}^3 + 0.373X_{t-4}^3 + e_t^3$$
$$AR(5): X_t^4 = 1.840X_{t-1}^4 - 0.893X_{t-2}^4 - 0.613X_{t-3}^4 - 0.879X_{t-4}^4 + 0.350X_{t-5}^4 + e_t^4$$

Where, $e_t^i \sim N(0,1)$ for i=1,2,3,4.

Using the above stationary process model, a new process plant, which has a similar process model in equation (1), is generated for simulation.

$$y_t = X_t^1 + X_t^2 + X_t^3 + X_t^4$$
$$\hat{y}_t = R_1(y_{t-1}, \ldots, y_{t-5}, , e_{t-1}, \ldots, e_{t-5}, \Theta_{R_1}) + V_1(A_{t-1}^1, \ldots, A_{t-5}^1, e_{t-1}^{v_1}, \ldots, e_{t-5}^{v_1}, \Theta_{V_1})$$

$\bar{A}^1 = (A_{t-1}^1, \ldots, A_{t-5}^1)$ is a collection of time series which is followed by the given compact uniform distribution from *U[0, 1]* for all i=*1,2,..k*. R_1 and V_1 are an neural network mapping. The weight vectors in neural networks, Θ_{R_1} and Θ_{V_1} were adjusted. Table 1 represents modeling accuracy of the single-phase model, which is described as real model, and multi-phase model as described as artificial model.

V. Conclusion

In this paper multi-phase model identification is presented when a process is not easily characterized either by one model nor by collected variables. The collected variables could be decomposed into other sub-signals which are not available for measurements and correlated each others. Real and virtual modeling approach are complementarily developed for increasing modeling accuracy. Real modeling, which is interpreted as a coarse modeling, uses real measurements for modeling an input-output system relation and the proposing virtual modeling is also performed continuously using the real modeling errors. In virtual modeling, newly generated input signals, which are obtained summing the real modeling error and artificially generated uniform series, are utilized for characterizing mapping relation.

The complementarily relationship between real and virtual modeling is connected to coarse and fine modeling approach. The multi-phase modeling procedure offers a practical and effective alternative that can be applied to very complicate non-stationary and non-linear dynamic process specially when a process is not easily `described by the collected measurements.

References

[1] K. J. Astrom and B. Wittenmark, *Computer-Controlled Systems Theory and Design*, Englewood- Cliffs, Prentice Hall, New Jersey, 1984.

[2] S. A. Billings, H. B. Jamaluddin, and S. Chen, Properties of Neural Networks with Applications to Modeling Non-linear Dynamical Systems, *International Journal of Control* **55(1)** 193-224, 1992

[3] G. E. P. Box, and G. M. Jenkins, *Time Series Analysis Forecasting and Control*, Holden-Day Inc., San Francisco, 1976.

[4] R. Hecht-Neilsen, Theory of the Back-Propagation Neural Network. *In Proceedings of the international joint conference on neural networks*, Washington, D.C (pp 1:217-1:225). New York: IEEE Press, 1989.

[5] R. A. Jacobs, M. I. Jordan, and A. G. Barto, Task Decomposition through Competition in a Modular Connectionist Architecture: the What and Where Vision Tasks, *Cognitive Science* **15**, 219-250, 1991.

[6] A. Lapedes and R. Farber, How Neural Nets Work, *Technical Report Los Almos National Laboratory*, . 331-346, January 1988.

[7] L. Ljung and T. Soderstrom, *Theory and Practice of Recursive Identification*, MIT Press, Cambridge, MA., 1983.

[8] K. S. Narendra and K. Parthasarathy, Identification and Control of Dynamical Systems using Neural Networks, *IEEE. Transactions on Neural Networks* **1**, 4-27, 1990.

[9] T. Sastri, J.R. English and G.N. Wang, Neural Networks for Time Series Model Identification and Change Detection, to appear in *Advances in Artificial Intelligence on E.F.M.*, 1993.

[10] Gi-Nam Wang, "An Adaptive Hybrid Neural Network Approach to Prediction of Nonstationary Processes," *Hybrid Intelligent System Applications*, Edited by Dr. Jay Liebowitz, Cognizant Communication Corporation, 140-161, 1996.

[11] G.N. Wang, Gwang-Sup Kim, and Y.S. Jung, "Two-Phase Radial Basis Hybrid Neural Network for On-line Modeling of Machine Condition," *Journal of the Society of Plant Engineers Japan*. Vol. 6, No. 3, 20-28, 1994.

[12] G. N. Wang and Y.C. Go, "On-line Neuro-Tracking of Nonstationary Manufacturing Processes," *Computers & Industrial Engineering,* Vol. 30, No. 3, 449-461, 1996.