새로운 동적 멀티레이어 퍼셉트론 구조

김 동 원*, 오 성 권*
*원광대학교 공과대학 제어계측공학과

A new Dynamic Multilayer Perceptron Structure

Dong-Won Kim*, Sung-Kwun Oh*
*School of Electrical and Electronic Engineering, Wonkwang Univ.

Abstract - We propose a new Dynamic Multilayer Perceptorn(DMP) architecture for optimal model identification of complex and nonlinear system in this paper. The proposed DMP scheme is presented as the generic and advanced type based on the GMDH(Group Method of Data Handling) method for the limitation of GMDH under only two system input variables. It is worth stressing that the number of the layers and the nodes in each layer of the DMP are not predetermined, unlike in the case of the popular multilayer perceptron structue, but these are generated in a dynamic manner. The experimental part of the study comes with representative nonlinear static system.
Comparative analysis is included and shows that a new DMP can produce the model with higher accuracy than previous other works.

1. Introduction

The challenging quest for constructing models of systems that come with significant approximation and generalization abilities as well as are easy to comprehend has been within the community for decades. When dealing with high-order nonlinear and multivariable equations of the model, we require a vast amount of data for estimating all its parameters. The Group Method of Data Handling (GMDH)[1] introduced by A.G. Ivakhnenko is one of the approaches that help alleviate the problem. The GMDH algorithm is carried out to generate an optimal architecture through a successive generation of both the nodes at each layer and the layers themselves by using the partial descriptions of the data. These come in a form of some quadratic regression polynomials of the two input variables. But, if only two system input variables are considered, it cannot generate the polynomial neural networks. For dealing with the fatal weak point, we introduce a new dynamic Multilayer Perceptron Structure.

2. Dynamic Multilayer Perceptron(DMP) algorithm and structure.

2.1 DMP algorithm.

The DMP algorithm is based on the GMDH method and utilizes a class of polynomials such as linear, modified quadratic, cubic, etc. By choosing the most significant input variables and polynomial order among these various kinds of forms, we can obtain the best ones from the extracted partial descriptions according to both selecting nodes of each layer and generating additional layers until the best performance is taken. A new methodology which includes these design procedure leads to the optimal DMP structure. The input-output data are given as follows.

$$(\mathbf{X}_i, y_i) = (x_{1i}, x_{2i}, \mathrm{K}, x_{Ni}, y_i),\quad i = 1,2,3,\mathrm{K}, n \qquad (1)$$

The input-output relationship of the above data by DMP algorithm can be described in the following fashion

$$y = f(x_1, x_2, \mathrm{K}, x_N) \qquad (2)$$

The estimated output $\hat{y}$ of the above output y is as follows

$$\hat{y} = \hat{f}(x_1, x_2 \Lambda, x_N) = c_0 + \sum_{k1} c_{k1} x_{k1} + \sum_{k1k2} c_{k1k2} x_{k1} x_{k2} + \sum_{k1k2k3} c_{k1k2k3} x_{k1} x_{k2} x_{k3} + \mathrm{K} \qquad (3)$$

wher c_k denotes the coefficients of the model.

The framework of the dynamic multilayer perceptron(DMP) design procedure can be summarized as follows.

[Step 1] Determine system input variables.
[Step 2] Split input and output data set.
[Step 3] Decide upon the DMP structure.
[Step 4] Determine the number of input variables and the order of the polynomial forming a partial description (PD) of data.
[Step 5] Estimate the coefficients of the PD.
[Step 6] Select PDs with the best predictive capability.
[Step 7] Check the stopping condition.
[Step 8] Determine new input variables for the next layer.

2.2 DMP structure.

The DMP structure for getting the best estimate model continues to generate additional layers and consists of PDs that the number of input variables is same in every layer. Some

PD in the form of regression polynomial is shown in Table 1.
Two cases for the regression polynomial in each layer are considered as the following.

Case 1. The polynomial order of PDs is same in every layer.

Case 2. The polynomial order of PDs in the 2nd layer or more has a different or modified type in comparison with that one of PDs in the 1st layer.

Two types, (a) the generic and (b) the advanced type, of the basic and modified DMP architectures are shown in Figs. 1-2, where z_i(Case 2) of the 2nd layer denote that polynomial order of the PD of each node has a different or modified type each other in comparison with z_i(Case 1) of the 1st layer.
In (b) the advanced type of Figs. 1-2, the node of dotted line mean the nodes of the previous layer. The superscript A of PDBA denotes the layer number and the subscript B of PDBA denotes the node sequence number of new nodes generated by the combination of the node outputs(outputs of PDs) of the preceding layer at the Ath layer.

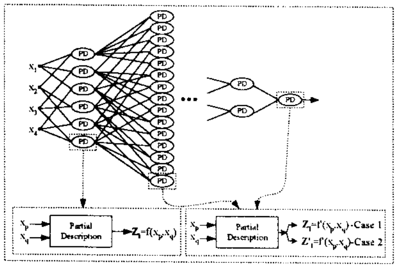

Fig. 1. Generic Type

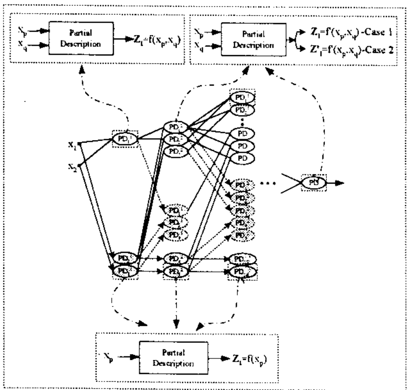

Fig. 2. Advanced Type

3. Experimental studies

This section is devoted . The DMP algorithm and its structure are illustrated with the aid of nonlinear static system.

3.1. Nonlinear static system

In this section, we perform a simulation to illustrate the validity of the proposed algorithm.

The training data in this example are obtained from a two-input nonlinear equation defined by

$$y = (1 + x_1^{-2} + x_2^{-1.5})^2 \text{ , } 1 \le x_1 \cdot x_2 \le 5 \text{ .} \qquad (4)$$

This nonlinear static equation is widely used to evaluate modeling performance. This equation was also used by Sugeno and Yasukawa [3], Nakanishi [4], and Kim[9-10] to test their modeling approaches.
This system represents the nonlinear characteristic as shown in Fig. 3, which shows a three-dimensional input-output graph of this system. From this system equation, 50 input-output data are obtained.

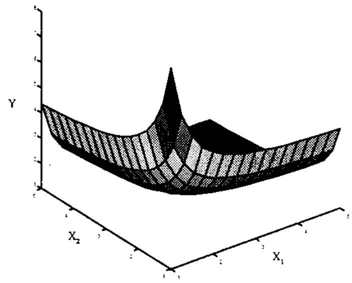

Fig. 3. Input-output relation of nonlinear system

Case 1 - This case is that the number of input variables of PDs is same in every layer and the polynomial order of PDs is also same in every layer.
The training result plotted in Fig. 4 depend on both the number of node input variables and three different types of polynomial order such as linear, quadratic, or modified quadratic, that is, Type 1, Type 2, or Type 3. Here Type 1, Type 2, and Type 3 are listed in Table 1. Fig. 4 show the performance index (identification error) of DMP architecture with 2 node inputs and Type 1, Type 2, and Type 3. In that case the best result of 2 node inputs, PI=0.0212 is obtained by using Type 2.

Table 1. Form of regression polynomials

inputs	order	Node equation considered
2	Type 1	$c_0+c_1x_1+c_2x_2$
	Type 2	$c_0+c_1x_1+c_2x_2+c_3x_1^2+c_4x_2^2+c_5x_1x_2$
	Type 3	$c_0+c_1x_1+c_2x_2+c_3x_1x_2$
3	Type 1	$c_0+c_1x_1+c_2x_2+c_3x_3$
	Type 2	$c_0+c_1x_1+c_2x_2+c_3x_3+c_4x_1^2+c_5x_2^2+c_6x_3^2+c_7x_1x_2+c_8x_1x_3+c_9x_2x_3$
	Type 3	$c_0+c_1x_1+c_2x_2+c_3x_3+c_4x_1x_2+c_5x_1x_3+c_6x_2x_3$

Case 2 - This case is that the number of input variables of each PD is same in every layer but the polynomial order of PDs in the 2nd layer or more has a different or modified type in comparison with that one of PDs in the 1st layer.
The training result is plotted in Fig. 5. In Fig. Type a b(a, b=1, 2, 3) means that the

polynomial order of PDs changes from Type a in the 1st layer to Type b in the 2nd layer or more.
Fig. 5 shows the performance index of DMP architecture with 2 node inputs and Type 1→2, Type 1→3, Type 2→1, Type 2→3, and Type 3→2. In that case the best result of 2 node inputs, PI=0.0212 is obtained by using Type 1→2.(Type a→b: 1st layer→2nd layer)

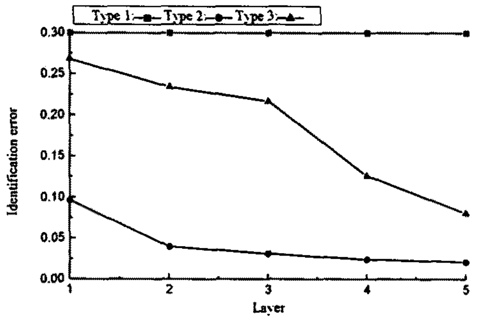

Fig. 4. Identification error (Every layer: 2 Input)

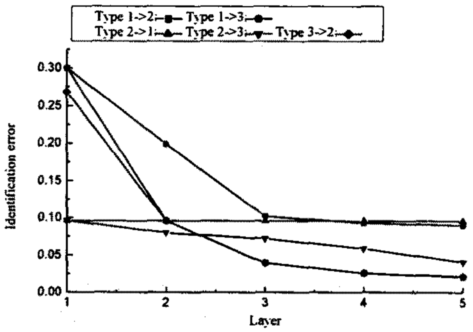

Fig.5. Identification error (Every layer: 2 Input)

Table 2 shows a comparison of identification errors with previous modeling methods. The experiment results show output performances of the proposed DMP model according to two kinds of DMP architectures.
As we know from Table 2 the performance results of DMP architecture are quite satisfactory. Compared with approaches presented in previous literatures [3-4,9-10], our modeling method has much more accuracy.

Table 2. Comparison of identification error with previous modeling methods

Model		Performance Index
		Mean Squared Error
Sugeno and Yasukawa[3]		0.079
Kim et al.[9]		0.019
Gomez-Skarmeta et al. [11]		0.070
Our model	Case 1	0.0212
	Case 2	0.0212

Case 1: 2 node inputs, Type 2.
Case 2: 2 node inputs, Type 1→2.

4. Conclusion

In this paper, The design procedure of Dynamic Multilayer Peceptron(DMP) is proposed to build a optimal model architecture for nonlinear and complex system modeling. Nonlinear static system are used for the purpose of evaluating the performance of the proposed DMP modeling method. Experimental results show that the proposed method is superior to other previous works from the viewpoint of the identification errors.

Acknowledgement

This work was supported by grant No. 2000-1-30300-009-3 from the Basic Research Program of the Korea Science&Engineering Foundation

(References)

[1] A.G. Ivahnenko, "Polynomial theory of complex systems", *IEEE Trans. On Systems, Man and Cybernetics*, Vol. SMC-1, pp. 364-378, 1971.
[2] S.J. Farlow, The GMDH algorithm, In: S.J. Farlow (Ed), *Self-organizing Methods in Modeling: GMDH Type Algorithms*, Marcel Dekker, New York, pp. 1-24, 1984.
[3] M. Sugeno and T. Yasukawa, "A fuzzy-logic-based approach to qualitative modeling", *IEEE Trans. Fuzzy Systems*, Vol. 1, No. 1, 7-31, 1993.
[4] H. Nakanishi, I.B. Turksen, M. sugeno, A review and comparison of six reasoning methods , *Fuzzy sets and Systems*, Vol. 57, pp. 257-294, 1992.
[5] S.K. Oh, and W.Pedrycz, "Identification of Fuzzy Systems by means of an Auto-Tuning Algorithm and Its Application to Nonlinear Systems", *Fuzzy Sets and Systems*, Vol.115, No. 2, pp.205-230, 2000.
[6] C.S. Park, S.K. Oh, and W. Pedrycz, "Fuzzy Identification by means of Auto - Tuning Algorithm and Weighting Factor", *The Third Asian Fuzzy Systems Symposium*(AFSS), pp.701-706, 1998.
[7] S.K. Oh, T.C. Ahn and W. Pedrycz, "FUZZY POLYNOMIAL NEURAL NETWORKS-BASED ARCHITECTURE AND ITS APPLICATION TO NONLINEAR PROCESS SYSTEMS", *Seventh IFSA World Congress, Prague*, pp.495-499, 1997.
[8] S.K. Oh, K.C. Yoon, and H.K. Kim, The Design of Optimal Fuzzy-Neural Networks Architecture by Means of GA and an Aggregate Weighted Performance Index , *Journal of Control, Automation and Systems Engineering*, Vol. 6, No. 3, March, 2000
[9] E.T. Kim, M.K. Park, S.H. Ji, and M. Park, "A new approach to fuzzy modeling", *IEEE Trans. Fuzzy Systems*, Vol. 5, No. 3, pp. 328-337, 1997.
[10] E. Kim, H. Lee, M. Park, M. Park, "A simple identified Sugeno-type fuzzy model via double clustering", *Information Science* 110, 25-39, 1998.
[11] A.F. Gomez-Skarmeta, M. Delgado, and M.A. Vila, "About the use of fuzzy clustering techniques for fuzzy model identification", *Fuzzy Sets and Systems*, Vol. 106, 179-188, 1999.