Evolutionary Design Methodology of Fuzzy Set-based Polynomial Neural Networks with the Information Granule

Seok-Beom Roh[1], Tae-Chon Ahn[1], and Sung-Kwun Oh[2]

[1]Department of Electrical Electronic and Information Engineering, Wonkwang University, 344-2, Shinyong-Dong, Iksan, Chon-Buk, 570-749, South Korea

[2]Department of Electrical Engineering, The University of Suwon, South Korea

ohsk@suwon.ac.kr

Abstract

In this paper, we propose a new fuzzy set-based polynomial neuron (FSPN) involving the information granule, and new fuzzy-neural networks - Fuzzy Set based Polynomial Neural Networks (FSPNN). We have developed a design methodology (genetic optimization using Genetic Algorithms) to find the optimal structure for fuzzy-neural networks that expanded from Group Method of Data Handling (GMDH). It is the number of input variables, the order of the polynomial, the number of membership functions, and a collection of the specific subset of input variables that are the parameters of FSPNN fixed by aid of genetic optimization that has search capability to find the optimal solution on the solution space. We have been interested in the architecture of fuzzy rules that mimic the real world, namely sub-model (node) composing the fuzzy-neural networks. We adopt fuzzy set-based fuzzy rules as substitute for fuzzy relation-based fuzzy rules and apply the concept of Information Granulation to the proposed fuzzy set-based rules.

1. Introduction

A lot of researchers on systemmodeling have been interested in the multitude of challenging and conflicting objectives such as compactness, approximation ability, generalization capability and so on which they wish to satisfy. Fuzzy sets emphasize the aspect of linguistic transparency of models and a role of a model designer whose prior knowledge about the system may be very helpful in facilitating all identification pursuits. In addition, to build models with substantial approximation capabilities, there should be a need for advanced tools. As one of the representative advanced design approaches comes a family of self-organizingnetworks with fuzzy polynomial neuron (FPN) (called "FPNN" as a new category of neuro-fuzzy networks) [6]. The design procedure of the FPNNs exhibits some tendency to produce overly complex networks as well as comes with a repetitive computation load caused by the trial and error method being a part of the development process. In this paper, in considering the above problems coming with the conventional FPNN [6], we introduce a new structure of fuzzy rules as well as a new genetic design approach. The new structure of fuzzy rules based on the fuzzy set-based approach changes the viewpoint of input space division. In other hand, from a point of view of a new understandingof fuzzy rules, information granules seem to melt into the fuzzy rules respectively. The determination of the optimal values of the parameters available within an individual FSPN leads to a structurally and parametrically optimized network through the genetic approach.

2. The architecture of FSPNN

The FSPN encapsulates a family of nonlinear "if-then" rules. When put together, FSPNs results in a self-organizing Fuzzy Set-based Polynomial Neural Networks (FSPNN). As visualizedin Fig. 1, the FSPN consists of two basic functional modules. The first one, labeled by F, is a collection of fuzzy sets (here denoted by $\{A_k\}$ and $\{B_k\}$) that form an interface between the input numeric variables and the processing part realized by the neuron. The second module (denoted here by **P**) refers to the function based nonlinear (polynomial) processing that involves some input variables This nonlinear processing involves some input variables (x_i and x_j), which are capable of being the input variables (Here, x_p and x_q), or entire system input variables. Each rule reads in the form

$$\text{if } x_p \text{ is } \boldsymbol{A}_k \text{ then } z \text{ is } P_{pk}(x_i,\ x_j,\ \boldsymbol{a}_{pk})$$
$$\text{if } x_q \text{ is } \boldsymbol{A}_k \text{ then } z \text{ is } P_{qk}(x_i,\ x_j,\ \boldsymbol{a}_{qk}) \qquad (1)$$

where $\boldsymbol{a}_{qk}$ is a vector of the parameters of the conclusion part of the rule while $P(x_i,\ x_j,\ \boldsymbol{a})$denoted the regression polynomial forming the consequence part of the fuzzy rule. The activation levels of the rules contribute to the output of the FSPN being computed as a weighted average of the individual condition parts (functional transformations) P_K (note that the index of the rule, namely "K" is a shorthand notation for the two indexes of fuzzy sets used in the rule (1), that is $K = (l,\ k)$).

$$z=\sum_{l=1}^{m}\sum_{k=1}^{l}\mu_{(l,k)}P_{(l,k)}(x_i,x_j,\boldsymbol{a}_{(i,k)})/\sum_{k=1}^{l}\mu_{(l,k)}$$
$$=\sum_{l=1}^{m}\sum_{k=1}^{l}\bar{\mu}_{(l,k)}P_{(l,k)}(x_i,x_j,\boldsymbol{a}_{(i,k)}) \qquad (2)$$

When developing an FSPN, we use genetic algorithms to produce the optimized network. This is realizedby selecting such parameters as the number of input variables, the order of polynomial, and choosing a specific subset of input variables. Based on the genetically optimized number of the nodes (input variables) and the polynomial order, refer to Table 1, we construct the optimized self-organizing network architectures of the FSPNNs.

Table 1. Different forms of the regression polynomials forming the consequence part of the fuzzy rules.

No. of inputs / Order of the polynomial	1	2	3
0 (Type 1)	Constant	Constant	Constant
1 (Type 2)	linear	Bilinear	Trilinear
2 (Type 3)	Quadratic	Biquadratic-1	Triquadratic-1
2 (Type 4)		Biquadratic-2	Triquadratic-2

1: Basic type, 2: Modified type

3. Information Granulation through Hard C-Means clustering algorithm

Information granules are defined informally as linked collections of objects (data points, in particular) drawn together by the criteria of indistinguishability, similarity or functionality. Granulation of information is a procedure to extract meaningful concepts from numeric data and an inherent activity of human being carried out with intend of better understanding of the problem. We granulate information into some classes with the aid of Hard C-means clustering algorithm, which deals with the conventional crisp sets.

We assume that given a set of data $X=\{\mathbf{x}_1,\mathbf{x}_2,\cdot,\mathbf{x}_n\}$ related to a certain application, there are some clusters revealed by the HCM. Each cluster is represented by its center and all elements, which belong to it. In order to construct fuzzy sets on the basis of such clusters, we follow a construct shown in Fig. 1. Here the center point stand for the apex of the membership function of the fuzzy set.

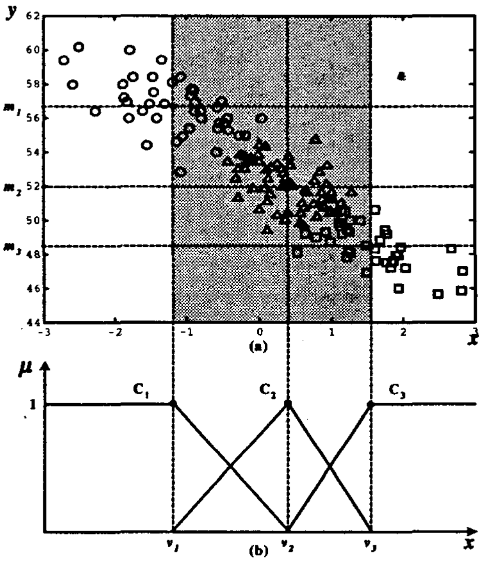


Fig. 1 Forming membership functions with the use of information granules

Where, C is the cluster, v means the center point of input variable x and m denotes the center point of output variable y. In Fig. 2, denotes the element belonging to 1st cluster, is the element allocated to the 2nd cluster, and to 3rd cluster. Each membership function in the premise part of the rule is assigned to be complementary with neighboring ones in the form being shown in the above figure. Let us consider building the consequent part of fuzzy rule. We can think of each cluster as a sub-model composing the overall system. The fuzzy rules of Information Granulation-based FSPN are as followings.

if x_p is $\boldsymbol{A}^*_k$ then $z-m_{pk} = P_{pk}((x_i-v^i_{pk}),(x_j-v^j_{pk}),a_{pk})$

if x_q is $\boldsymbol{B}^*_k$ then $z-m_{qk} = P_{qk}((x_i-v^i_{qk}),(x_j-v^j_{qk}),a_{qk})$ (3)

Where, $\boldsymbol{A}^*_k$ and $\boldsymbol{B}^*_k$ mean the fuzzy set, the apex of which is defined as the center point of information granule (cluster) and m_{pk} is the center point related to the output variable on cluster$_{pk}$, v^{i}_{pk} is the center point related to the i-th input variable on cluster$_{pk}$ and $\boldsymbol{a}_{qk}$ is a vector of the parameters of the conclusion part of the rule while $P((x_i-v_i),(x_j-v_j),a)$ denoted the regression polynomial forming the consequence part of the fuzzy rule which uses several types of high-order polynomials (linear, quadratic, and modified quadratic) besides the constant function forming the simplest version of the consequence; refer to Table 1. If we are given m inputs and one output system and the consequent part of fuzzy rules is linear, the overall procedure of modification of the generic fuzzy rules is as followings.

The given inputs are $\mathbf{X}=[\mathbf{x}_1\ \mathbf{x}_2\cdots\mathbf{x}_m]$ related to a certain application, where $\mathbf{x}_k=[x_{k1}\cdots x_{kn}]^T$, n is the number of data and m is the number of variables and the output is $\mathbf{Y}=[y_1\ y_2\ \cdots\ y_n]^T$.

Step 1) build the universe set

Universe set U={$\{x_{11}, x_{12},\cdots, x_{1m}, y_1\}$, $\{x_{21}, x_{22},\cdots, x_{2m}, y_2\}$, , $\{x_{n1}, x_{n2},\cdots, x_{nm}, y_n\}$}

Step 2) build m reference data pairs composed of $[\mathbf{x}_1;Y]$, $[\mathbf{x}_2;Y]$, and $[\mathbf{x}_m;Y]$.

Step 3) classify the universe set U into l clusters such as c_{i1}, c_{i2}, , c_{il} (subsets) by using HCM according to the reference data pair $[\mathbf{x}_i;Y]$. Where c_{ij} means the j-th cluster (subset) according to the reference data pair $[\mathbf{x}_i;Y]$.

Step 4) construct the premise part of the fuzzy rules related to the i-th input variable (x_i) using the directly obtained center points from HCM.

Step 5) construct the consequent part of the fuzzy rules related to the i-th input variable (x_i). On this step, we need the center points related to all input variables. We should obtain the other center points through the indirect method as followings.

Sub-step1) make a matrix as equation (4) according to the clustered subsets

$$A^i_j=\begin{vmatrix} x_{21} & x_{22} & \cdots & x_{2m} & y_2 \\ x_{51} & x_{52} & \cdots & x_{5m} & y_5 \\ \vdots & \vdots & & \vdots & \vdots \\ x_{k1} & x_{k2} & \cdots & x_{km} & y_k \\ \vdots & \vdots & & \vdots & \vdots \end{vmatrix} \tag{4}$$

Where, $\{x_{k1}, x_{k2}, , x_{km}, y_k\}\in c_{ij}$ and A_{ij} means the membership matrix of $\boldsymbol{j}$-$\boldsymbol{th}$ subset related to the $\boldsymbol{i}$-$\boldsymbol{th}$ input variable.

Sub-step2) take an arithmetic mean of each column on A_{ij}. The mean of each column is the additional center point of subset cij. The arithmetic means of column is equation (5)

$$centerpoints=[\ v^1_{ij}\ v^2_{ij}\cdots\ v^m_{ij}\ m_{ij}] \tag{5}$$

4. Genetic Optimization of FPNN

GAs are aimed at the global exploration of a solution space. The main features of genetic algorithms concern individuals viewed as strings, population-based optimization and stochastic search mechanism (selection and crossover). GAs use serial method of binary type, roulette-wheel as the selection operator, one-point crossover, and an invert operation in the mutation operator [2]. In this study, for the optimization of the FPNN model, GA uses the serial method of binary type, roulette-wheel used in the selection process, one-point crossover in the crossover operation, and a binary inversion operation in the mutation operator. To retain the best individual and carry it over to the next generation, we use elitist strategy. The overall genetically-driven structural optimization process of FPNN is shown in Fig. 2.

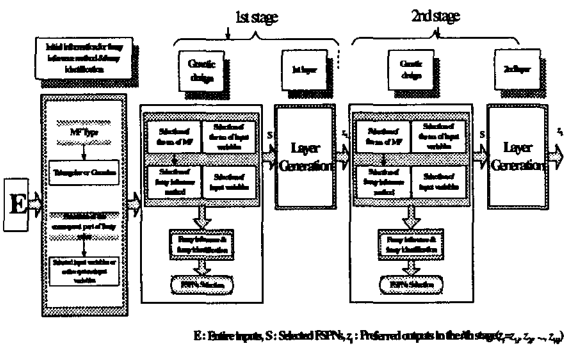


Fig. 2. Overall genetically-driven structural optimization process of FPNN

The framework of the design procedure of the genetically optimized FSPNN comprises the following steps

[Step 1] Determine systems input variables

[Step 2] Form training and testing data

[Step 3] specify initial design parameters

[Step 4] Decide FSPN structure using genetic design

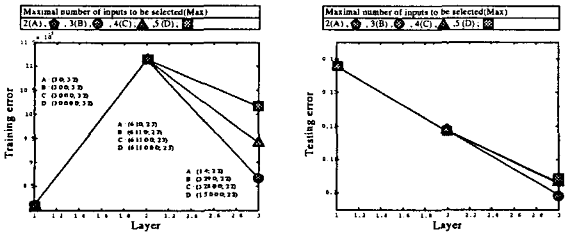
[Step 5] Carry out fuzzy-set based fuzzy inference and coefficient parameters estimation for fuzzy identification in the selected node (FSPN)

[Step 7] Check the termination criterion

[Step 8] Determine new input variables for the next layer

5. Experimental studies

We illustrate the performance of the network and elaborate on its development by experimenting with data coming from the gas furnace process. The time series data (296 input-output pairs) resulting from the gas furnace process has been intensively studied in the previous literature. Fig. 3 depicts the performance index of each layer of IG-gFSPNN with Type T* according to the increase of maximal number of inputs to be selected.



(a-1) Training error (a-2) Testing error

(a) Gaussian-like MF

Fig. 3. Performance index of IG_gFSPNN (with Type T*) with respect to the increase of number of layers

Table 2. Comparative analysis of the performance of the network; considered are models reported in the literature

Model				Performance Index		
				PI	PI	EPI
sugeno and yasukawa'smodel				0.190		
Oh and Pedrycz's model				0.123	0.020	0.271
Kim et al.' model					0.034	0.244
Proposed IG_gFSPNN	TypeII (I=6)	Triangular	3rd layer(Max=3)		0.008	0.110
		Gaussian	3rd layer(Max=3)		0.008	0.099

6. Conclusion

In this study, we have developed the new structure and formed the semantics of fuzzy rules and investigated the GA-based design procedure of Fuzzy Set-based Polynomial Neural Networks (FSPNN) driven to information granulation along with its architectural considerations. The entire system is divided into some sub-systems that are classified according to the data characteristics named information granules. Each information granule is a sound representative of the related subsystems. A new fuzzy rule with information granule describes a sub-system as a stand-alone system. A fuzzy system with some new fuzzy rules depicts the whole system as a combination of some stand-alone components.

The comprehensive experimental studies involving well-known datasets quantify a superb performance of the network in comparison to the existing fuzzy and neuro-fuzzy models. More importantly, through the proposed framework of genetic optimization we can efficiently search for the optimal network architecture (being both structurally and parametrically optimized) and this design facet becomes crucial in improving the performance of the resulting model.

감사의 글

본 연구는 산업자원부의 지원에 의하여 기초전력연구원(R-2004-B-133-01) 주관으로 수행된 과제임.

7. Reference

[1] V. Cherkassky, D. Gehring, and F. Mulier, Comparison of adaptive methods for function estimation from samples , *IEEE Trans. Neural Networks*, Vol. 7, pp. 969 984, July 1996.Electrical Engineering, Vol. 29, Issue 6, pp. 703-725, 2003.

[2] J. A. Dickerson and B. Kosko, Fuzzy function approximation with ellipsoidal rules , *IEEE Trans. Syst., Man, Cybernetics*, Part B, Vol. 26, pp.542 560, Aug. 1996.

[3] V. Sommer, P. Tobias, D. Kohl, H. Sundgren, and L. Lundstrom, Neural networks and abductive networks for chemical sensor signals: A case comparison , *Sensors and Actuators*, B. 28, pp. 217-222, 1995.

[4] S. Kleinsteuber and N. Sepehri, A polynomial network modeling approach to a class of large-scale hydraulic systems , *Computers Elect. Eng.*, 22, pp. 151-168, 1996.

[5] S.-K. Oh and W. Pedrycz, Fuzzy Polynomial Neuron-Based Self-Organizing Neural Networks , *Int. J. of General Systems*, Vol. 32, No. 3, pp. 237-250, May, 2003.

[6] S.-K. Oh, W. Pedrycz and T.-C. Ahn, Self-organizing neural networks with fuzzy polynomial neurons , *Applied Soft Computing*, Vol. 2, Issue 1F, pp. 1-10, Aug. 2002.

[7] M. Sugeno and T. Yasukawa, A Fuzzy-Logic-Based Approach to Qualitative Modeling , *IEEE Trans. Fuzzy Systems*, Vol. 1, No. 1, pp. 7-31, 1993.

[8] E.-T. Kim, et al, A new approach to fuzzy modeling , *IEEE Trans. Fuzzy Systems*, Vol. 5, No. 3, pp. 328-337, 1997.