Jurnal Teknologi, 43(D) Dis. 2005: 143–153
© Universiti Teknologi Malaysia

GENERATION OF FUZZY RULES WITH SUBTRACTIVE CLUSTERING

AGUS PRIYONO[1], MUHAMMAD RIDWAN[2], AHMAD JAIS ALIAS[1], RIZA ATIQ O. K. RAHMAT[3], AZMI HASSAN[2] & MOHD. ALAUDDIN MOHD. ALI[1]

Abstrak. Learning fuzzy rule-based systems with genetic algorithms can lead to very useful descriptions of several problems. Fuzzy logic (FL) provides a simple way to arrive at a definite conclusion based upon vague, ambiguous, imprecise, noisy or missing input information. The FL model is empirically based, relying on an operator's experience rather than their technical understanding of the system. In the FL method, any reasonable number of inputs can be processed and numerous outputs will be generated, although defining the rule-base quickly becomes complex if too many inputs and outputs are chosen for a single implementation since rules defining their interrelations must also be defined. This will increase the number of fuzzy rules and complexity but may also increase the quality of the control. Many methods were proposed to generate fuzzy rules-base. The basic idea is to study and generate the optimum rules needed to control the input without compromising the quality of control. The paper proposed the generation of fuzzy rule base by subtractive clustering technique in Takagi-Sugeno-Kang (TSK) fuzzy method for traffic signal control system.

Keywords: TSK fuzzy logic, fuzzy rule base system, subtractive clustering technique

Abstract. Pembelajaran sistem pangkalan peraturan kabur menggunakan algoritma genetik mempunyai masa depan yang cerah bagi menyelesaikan beberapa masalah. Lojik kabur menawarkan cara sederhana bagi menyimpulkan maklumat input yang kasar, kabur, cacat atau tidak jelas. Model lojik kabur adalah berasaskan kaedah-kaedah empirik bergantung kepada pengalaman operator berbanding dengan pengetahuan teknikal daripada sistem. Dalam metod lojik kabur, sebarang input yang munasabah dapat diproses dan sebilangan output dapat dijana meskipun penakrifan pangkalan peraturan secara cepat dapat menjadi rumit sekiranya terlalu banyak input dan output yang dipilih untuk sebuah penggunaan. Bergantung kepada sistem, semakin rumit input dan output yang ingin diselesaikan oleh sistem, maka akan semakin banyak jumlah bilangan peraturan dan kerumitan tetapi juga akan menambah mutu kawalan dari sistem. Banyak kaedah telah dicadangkan bagi menjana peraturan kabur. Idea asa daripada penyelidikan ini adalah untuk mempelajari serta menjana peraturan paling optimum yang diperlukan bagi mengawal input tanpa mengurangi mutu kawalan. Kertas kerja ini yang mencadangkan penjanaan peraturan kabur menggunakan penggugusan subtraktif pada lojik kabur Takasi-Sugeno-Kang (TSK) bagi kegunaan kawalan lampu isyarat lalu lintas.

Kata kunci:: Lojik kabur TSK, sistem pangkalan peraturan kabur, teknik penggugusan subtraktif

[1] Department of Electrical, Electronic and System Engineering.
[2] Department of Mechanical and Material Engineering, Faculty of Engineering.
[3] Department of Civil and Structural Engineering.
Universiti Kebangsaan Malaysia. 43600, Bangi Selangor Darul Ehsan, Malaysia

1.0 INTRODUCTION

In recent years, fuzzy modeling technique have become an active research area due to its successful application to complex system model, where classical methods such as mathematical and model-free methods are difficult to apply because of lack of sufficient knowledge. The FL model is empirically-based, relying on an operator's experience rather than their technical understanding of the system. For a little more complex system, but for which significant data exist, model free method such as neural networks provide a powerful and robust means to reduce some uncertainty through learning, based on patterns in the available data. For the most complex system where few numerical data exist and only ambiguous or imprecise information may be available, fuzzy reasoning provides a way to understand system behavior by allowing us to interpolate approximately between observed input and output situation. The imprecision in fuzzy models is therefore, generally quite high. Fuzzy systems can implement crisp input and output, and produce a non-linear functional mapping just as do algorithms. Depending on the system, it may not be necessary to evaluate every possible input combination since some may rarely or never occur. This will increase the number of fuzzy rules and complexity but may also increase the quality of the control.

1.1 Generation of Fuzzy Rule Base

Learning algorithms for the automatic generation of fuzzy rules are well known [1, 2]. These learning algorithms generate fuzzy rules based on learning vectors and learn/optimize the I/O behavior of the system. A very simple rule-base structure and they restrict the membership functions which leads to solutions far away from an optimum used by Cheong *et al*. [3]. Fuzzy rules have a simple geometry in the input-output space, normally having the shape of ellipsoids [4]. Fuzzy rules induce in this way a fuzzy partition of the product space of the input-output variables. In addition, fuzzy clustering algorithms are very suitable techniques to detect this fuzzy partition. Different authors have proposed the use of the fuzzy clustering techniques in this process [5-7]. Some authors have applied this technique separately to each domain of the input and output variables of the system, and then the fuzzy model is formulated in terms of fuzzy relational equations [8]. Other authors use this only on the input variables and combine the results with a Takagi-Sugeno-Kang (TSK)-like consequent using the fuzzy sets of the antecedent [7].

The fuzzy model suggested by Takagi *et al*. [9, 10], also known as the TSK model, has gained increasing interest in theoretical analysis and applications of fuzzy modeling and control. The TSK model is associated with fuzzy rules that have a special format with a functional-type consequent instead of the fuzzy consequent that normally appears in the Mamdani model [11]. In this way the TSK approach tries to decompose the input space into subspaces and then approximate the system in each subspace by a simple linear regression model. This characteristic provides efficient models to deal with a complex system although the generation of the corresponding fuzzy rules, specially the premise structure is technically difficult and may lead to a nonlinear programming problem.

Since the TSK-Fuzzy system is used to predict and control the complexity problems [12, 13], fuzzy system needs the antecedent and consequence to express the logical connection between the input and output which is used as a basis to produce the desired output [14]. Hence, to produce the desired output, the consequence is a logical connection from antecedent that should be a variable of the antecedent. The TSK fuzzy system has a high output sensitivity to input data, because the consequence of fuzzy is variables-function system at the antecedent.

Some authors work with fuzzy clustering methods in the product space of the input-output space in order to detect the interaction between the input and output variables. Others have extended the use of fuzzy clustering to detect multidimensional fuzzy sets in the product space of the input variables to identify the premise of the fuzzy rules and then assigning a linear consequent to each rule. The identification of fuzzy models can be improved using these multi-dimensional reference fuzzy sets [15]. Hence, fuzzy clusters give rise to "local" regression models. The model is then structured into a set of IF-THEN statements.

The TSK model is composed of IF-THEN rules of the following form:

$$R_{(r)}: \textbf{if } x_1 \text{ is } A_r^1 \textbf{ and } x_2 \text{ is } A_r^2 \textbf{ and...and } x_m \text{ is } A_r^m \textbf{ then } y_r \text{ is } fr\,(x) \tag{1}$$

$$\text{where : } f_r\,(x) = \alpha_r^0 + \alpha_r^1\ x_1 + \ldots\ldots + \alpha_r^m\ x_m, \tag{2}$$

in which (r = 1,....., n) and x_j ($1 \le j \le m$) are the input variables, y_r is the output variable, A_r^m are fuzzy sets (usually corresponding to linguistic labels), and $f_r(x)$ is a linear function. By this equation, the TSK model of each fuzzy rule describes a local linear behavior associated to the fuzzy input region characterized by the antecedent of the fuzzy rule.

For any input, say $\hat{x} = \left(x_r^1, x_r^2 \ldots\ldots x_r^m\right)$, the inferred value of the TSK fuzzy model, is calculated as:

$$\gamma = \frac{\sum_{r=1}^{m} A_r\left(\hat{x}\right)^*}{\sum_{r=1}^{m} A_r\left(\hat{x}\right)} = \frac{\sum_{r=1}^{m} \tau_r * f_r\left(\hat{x}\right)}{\sum_{r=1}^{m} \tau_r} \tag{3}$$

where:

$$A_r\left(\hat{x}\right) = \tau_r = A_r^1\left(x_r^1\right) * A_r^2\left(x_r^2\right) * \ldots\ldots * A_r^m\left(x_r^m\right) \tag{4}$$

where τ_r being the level of firing of the rth rule for the current input $\hat{x}$. This algorithm is very appealing as the inference and defuzzification process are integrated into a single-step procedure [16].

The rule interpretation of TSK fuzzy model in [10] depends on the choice of the center and standard deviation. Only if the premise input space partitioning is performed in an axis-orthogonal manner, the multivariable membership function can be projected to a one-dimensional fuzzy sets A. Note that the model output is linear in the weight w but is non-linear in the centers c and standard deviation, σ.

1.2 Subtractive Clustering

The idea of fuzzy clustering is to divide the data space into fuzzy clusters, each representing one specific part of the system behavior. After projecting the clusters onto the input space, the antecedent parts of the fuzzy rules can be found. The consequent parts of the rules can then be simple functions. In this way, one cluster corresponds to one rule of the TSK model.

Using a fuzzy clustering algorithm, membership functions can be determined according to two possible methods. In the first method, the clusters are projected orthogonally onto the axes of the antecedent variables, and the membership functions are fitted to these projections. The second method uses multi-dimensional antecedent membership functions, i.e. the fuzzy clusters are projected onto the input space. Figure 1 illustrates a schematic overview of the latter method. Using this method, the membership degree of a data point is directly computed in this projected cluster according to its distance from the projected cluster center. From Figure 1, d_i denotes the projected cluster center and c_i denotes the data point.

Several clustering methods are well known [17, 18]. The first method is *k-means*, second is fuzzy *c-means* method, third is mountain and then fourth is subtractive clustering (SC) method, which is a noniterative algorithm [18].

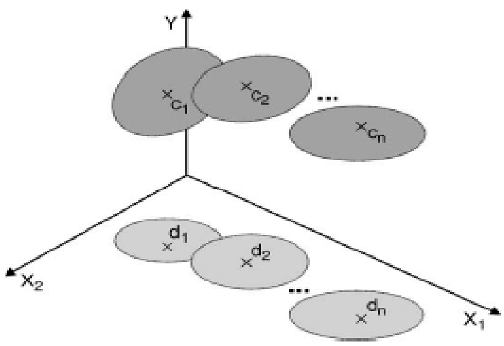

Figure 1 Projection of the fuzzy clusters onto the antecedent space in the case of a three-dimensional input-output space

Subtractive clustering was introduced by [18]. For this method, data points have to be rescaled to [0,1] in each dimension. Each data point $z_j = (x_j, y_j)$ is assigned a potential P_j, according to its location to all other data points:

$$P_i^* = \sum_{j=1}^{n} e^{-\alpha \left\| x^i - x^j \right\|^2} \tag{5}$$

where

$$\alpha = \frac{\gamma}{r_a} \tag{6}$$

P_i^* is the potential-value i-data as a cluster centre.
α is the weight between i-data to j-data
x is the data point
γ is variables (commonly set 4)
r_a is a positive constant called cluster radius

The potential of a data point to be a cluster center is higher when more data points are closer. The data point with the highest potential, denoted by P_i^* is considered as the first cluster center $\boldsymbol{c}_1 = (\boldsymbol{d}_1, e_1)$. The potential is then recalculated for all other points excluding the influence of the first cluster center according to:

$$P_i^* = P_i^* - P_k^* \zeta \tag{7}$$

where:

$$\zeta = e^{-\beta \left\| x^i - c^k \right\|^2} \tag{8}$$

$$\beta = \frac{4}{r_b^2} \tag{9}$$

$$r_b = r_a * \eta \tag{10}$$

P_i^* is the new potential-value i-data.
P_k^* is the potential-value data as cluster centre
c is the cluster center of data
β is the weight of i-data to cluster centre
r_i is the distance between cluster centre
η is the quash factor

Again, the data point with the highest potential P_k^* is considered to be the next cluster center $\boldsymbol{c}_k$, if

$$\frac{d_{\min}}{r_a} + \frac{P_k^*}{P_1^*} \geq 1 \tag{11}$$

with d-min is the minimal distance between c_1 and all previously found cluster centers, the data point is still accepted as the next cluster center c_1. Further iterations can then be performed to obtain new cluster centers c_2. If a possible cluster center does not fulfill the above described conditions, it is rejected as a cluster center and its potential is set to 0. The data point with the next highest potential P_k^* is selected as the new possible cluster center and re-tested. The clustering ends if the following condition is fulfilled:

$$P_k^* < \varepsilon P_i^* \tag{12}$$

where : ε is the reject ratio.

Indicative parameters values for r_a, η, ε and ε^* have been suggested by [18]. Each cluster center is considered as a fuzzy rule that describes the system behavior of the distance to the defined cluster centers:

$$\mu_j^{ik} = e^{-\alpha \left\| x_j^i - c_j^k \right\|^2} \tag{13}$$

Equation (13) is a common form of subtractive clustering, hence it needs to create an algorithm to process data clustering. Thus, this paper proposed an algorithm to cluster the data to train a traffic control system:

Step 1: Calculate the input data to be clusters.

$$X_{ij},\ i = 1,2,\ldots,n;\ j = 1,2,\ldots,m$$

with: n is the number of data
m is the type of data

Step 2: Set the variables value:

- i- r_j, $j = 1, 2, \ldots, m$
- ii- η *quash factor*
- iii- ε^* *accept ratio*
- iv- ε *reject ratio*
- v- $X_{j\text{-min}}$
- vi- $X_{j\text{-max}}$

Step 3: Set the normal data value based on $X_{j\text{-min}}$ dan $X_{j\text{-max}}$ use with the following model:

$$X_{ij}^{norm} = \frac{X_{ij} - X_{j-\min}}{X_{j-\min} - X_{j-\min}}, i = 1,2,\ldots,n; j = 1.2,\ldots,m \tag{14}$$

Step 4: Set the potential of each data point by the formula:

a = 1, revise to a = n

if m = 1, set

$$P_i^{'} = \sum_{i=1}^{n} e^{-4\left\|\frac{x_i - x_k}{r_a}\right\|^2}, i = 1,2,\ldots,n; k = 1,2,\ldots,n; i \neq k \tag{15}$$

if m > 1, set

$$P_i^{'} = \sum_{i=1}^{n} e^{-4\left\|\sum_{j=1}^{m}\frac{x_i^j - x_k^j}{r_a}\right\|^2}, i = 1,2,\ldots,n; j = 1,2,\ldots,n; i \neq j \tag{16}$$

Step 5: Set the highest potential value of data:

$$\mathrm{M} =_{\max} \left\lfloor P_i^{'} | i = 1,2,\ldots,n \right\rfloor \tag{17}$$

$$h = i, \text{ so that } D_i = M$$

Step 6: Set cluster centre and update the potential value that correspond to another data:

i- Cnt = [ ]
ii- $V_j = X_{hj}$; j = 1,2,...m
iii- C = 0 (number of clusters)
iv- Cnd = 1
v- $z = m$
vi- Do $Cnd \neq 0$ and $Z \neq 0$

Step 7: Put the real data:

$$Cnt_{ij} = Cnt_{ij} * (X_{j\text{-max}}\text{-}X_{j\text{-min}}) + X_{j\text{-min}} \tag{18}$$

Step 8: Set the cluster sigma:

$$\sigma_j = \frac{r_j^* \left(X_{j-\max} - X_{j-\min}\right)}{\sqrt{8}} \tag{19}$$

The algorithm produced vector-centre and sigma value of each cluster based on the model developed by Chiu [18], as shown in Figure 2. The proposed algorithm used to decide the membership of learning data of each cluster is as follows:

$$\mu_{ik} = e^{-\sum_{j=1}^{m}\frac{\left(x_{ij}-Cnt_{kj}\right)^2}{2\sigma_j^2}} \tag{20}$$

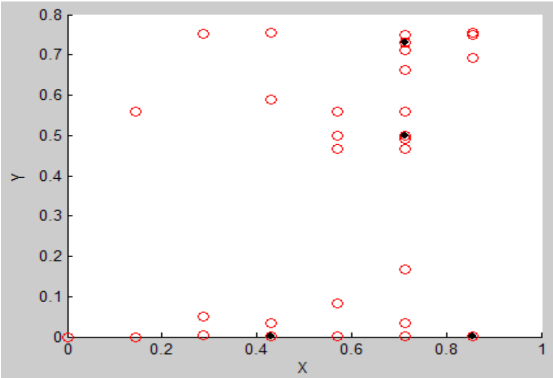

Figure 2 Fuzzy clustering data with subtractive method. The black points are data centers

1.3 Fuzzy Logic Inference

FL inference system is used to collect all of fuzzy rules base to set up the crisp output. Therefore, developing fuzzy inference system must be based on fuzzy rules base.

In case of optimization inference system TSK model, the number of desired rules is equal to the number of clustering center. Clustering process produce cluster center value and cluster sigma which will be used to perform the fuzzy logic rules. From this fuzzy rules, the membership of each data on each cluster can also be performed, and the antecedent of each rule can be quantified with union theory [19] as follows:

(a) AND relation use min operation
(b) OR relation use max operation

This quantification process for each rules produces fitness-limit value of each rules. These value is a weight for each fuzzy rule base to set the fuzzy output. In this paper, fuzzy output will be calculated using weight-average method [20] as follows:

$$O = \frac{\sum_{r=1}^{R}(\alpha_r * y_r)}{\sum_{r=1}^{R}\alpha_r} \tag{22}$$

where:

O is the fuzzy output
a is the fitness-limit value
R is the number of rules (equal to number of data cluster)

Generally, the performance of supervised learning process is determined by mean square error (MSE), calculated based on the output split green time and the split green time calculated by the human expert. The MSE is determined by the equation as follows:

$$MSE = \frac{\sum_{i=1}^{n} (t_i - y_r)^2}{n}; \forall r \tag{23}$$
$$y^r \Rightarrow f(x)$$

where:

MSE is mean square error
n is the number of data

Equation (23) needs $y^r \Rightarrow f(x)$ function with fitness-limit value variable, a, which is calculated using genetic algorithm.

1.4 Fuzzy Rules Result

As shown in Figure 2, we started with 30 (3 x 30) data and set the cluster radius at 0.5. After clustering the data, the system produced 4 cluster centers. This means if each cluster center is equal with one rule, we have only four rules to represent 30 data as shown in Figure 3.

On the subtractive clustering process, we set the data cluster radius to 0.3 and 0.5. Using 0.3 as cluster radius, the process result is 8 rules, but using 0.5 as cluster

IF Kendr Utama is Sedikit AND Jarak Kepala is Dekat THEN Masa Hijau is 1.523 (X1) + -0.035 (X2) + -1.349
IF Kendr Utama is Agk Banyak AND Jarak Kepala is Agk Jauh THEN Masa Hijau is 0.818 (X1) + -0.769 (X2) + 2.048
IF Kendr Utama is Banyak AND Jarak Kepala is Jauh THEN Masa Hijau is -0.436 (X1) + -0.521 (X2) + 5.357
IF Kendr Utama is Sgt Banyak AND Jarak Kepala is Sgt Jauh THEN Masa Hijau is -4.105 (X1) + 1.283 (X2) + 6.359

Figure 3 Result of fuzzy rule base

radius, the result is 4 rules. Comparing with other methods in the literature, the fuzzy rule base that is proposed in this paper is relatively better, as shown in Table 1. This result will be optimized later by using the parallel hybrid genetic algorithm to optimize the weight of fuzzy inference.

Table 1 Result for generation of fuzzy rule base data compared with other work

Author	Initial rules	Rules after clustering	MSE
Tong [20]	7×6	19	0.469
Sugeno *et al.* [21]	5×5	6	0.355
Pedryz [22]	9×9	81	0.320
Xu - Lu [23]	5×5	25	0.328
Abreu *et al.* [24]	7×7	38	0.172
Surmann [7]	3×5	15	0.138
Agus *et al.*	**3×30**	**4**	**0.005**

2.0 CONCLUSION

The above technique is proposed as an expert system part of the urban traffic control system (UTCS) that is developed and implemented to control multi-junctions in Bangi, Selangor Darul Ehsan, Malaysia. The expert system parts of the traffic control such as the two-stage neural network system to recognize the traffic pattern and the fuzzy-genetic system to optimize the split green time and offset were simulated in the laboratory. The part developed here is included in the parallel hybrid genetic algorithm to optimize the phase timing.

Two-stage neural network model is used to recognize the traffic pattern and then decide the traffic control strategies. A fuzzy-genetic model is used to estimate the objective values in the optimization process with iterative adjustment of signal timings and offset. The method proposed here would be applicable to an on-line system because it is trained for extensive traffic condition.

The proposed system can be integrated in a control room. Due to its flexible and modular structure, the system is the core for the local traffic control system at each junction and can contribute to an improvement in traffic performance, reliability, and human expert satisfaction.

ACKNOWLEDGEMENTS

The system described in this paper has been developed at Traffic Laboratory in the Department of Civil and Structural Engineering, Faculty of Engineering, Universiti Kebangsaan Malaysia. The authors acknowledged the Ministry of Science, Technology and Innovation of Malaysia No: 04-02-02-0047-PR0059/09-01 for having provided the financial support for this work.

REFERENCES

[1] Herrera, F., and J. L. Verdegay. 1996. Studies in Fuzziness and Soft Computing. *Genetic Algorithm and Soft Computing Journal*. 8.

[2] Salehfar, H., N. Bengiamin, and J. Huang. 2000. A Systematic Approach To Linguistic Fuzzy Modeling Based on Input-Output Data. Proceedings of the 2000 Winter Simulation Conference.

[3] Cheong, F., and R. Lai. 1996. Constraining the Optimization of a Fuzzy Logic Controller Using Enhanced Genetic Algorithm. *IEEE Trans. on Systems Man and Cybernetics, part B Cybernetics*. 30.

[4] Dickerson, J. A., and B. Kosko. 1993. Fuzzy Function Approximation with Ellipsoidal Rules. *IEEE Trans. System Man and Cybernetics*. 26(4): 542-560.

[5] Sugeno, M., and T. Yasukawa. 1993. A Fuzzy Logic-Based Approach to Qualitative Modeling. *IEEE Trans. Fuzzy Systems*. 1(1): 7-31.

[6] Yager, R. T., and D. P. Filev. 1993. Unified Structure and Parameter Identification of Fuzzy Models. *IEEE Trans. System Man and Cybernetics*. 23: 157-165.

[7] Wang, L., and R. Langari. 1996. Complex System Modeling via Fuzzy Logic. *IEEE Trans. System Man and Cybernetics*. 26: 100-106.

[8] Surmann, H., and A. Selenschtschikow. 2002. Automatic Generation of Fuzzy Rule Bases: Example I. Proc. of the NF2002 First International ICSC on Neuro-Fuzzy Technologies. January 2002.

[9] Takagi, T., and M. Sugeno. 1985. Fuzzy Identification of Systems and its Application to Modeling and Control. *IEEE Trans. Systems Man and Cybernetics*. 15: 116-132.

[10] Sugeno, M., and G. T. Kang. 1988. Structure Identification of Fuzzy Model. *Fuzzy Sets and Systems*. 28: 15-33.

[11] Mamdani, E. 1976. Advances in the Linguistic Synthesis of Fuzzy Controllers. *Int. Journal in Man-Machine Studies*. 8: 669-678.

[12] Papadakis, S. E., and J. B. Theocharis. 2002. A GA-based Fuzzy Modeling Approach for Generating TSK Models. *Fuzzy Sets and Systems*. 131: 121-152.

[13] Cox, E. 1994. *The Fuzzy Systems Handbook: A Practitioner's Guide to Building, Using, and Maintaining Fuzzy Systems*. Academic Press.

[14] Sin, S. K., and De Figueiredo. 1993. Fuzzy System Designing Through Fuzzy Clustering and Optimal preDefuzzification. Proc. IEEE International Conference on Fuzzy Systems. 190-195.

[15] Gomez, A. F., M. Delgado, and M. A. Vila. 1999. About the Use of Fuzzy Clustering Techniques for Fuzzy Model Identification. *Fuzzy Set and Systems*. 106: 179-188.

[16] Demirli, K., S. X. Cheng, and P. Muthukumaran. 2003. Subtractive Clustering Based Modeling of Job Sequencing with Parametric Search. *Fuzzy Sets and Systems*. 137: 235-270.

[17] Chiu, S. L. 1994. A Cluster Estimation Method with Extension to Fuzzy Model Identification. Proc. IEEE Internat. Conf. on Fuzzy Systems. 1240-1245.

[18] Hongxing, L. *et al*. 2001. *Fuzzy Neural Intelligent System, Mathematical Foundation and the Application in Engineering.* CRC Press LLC.

[19] Jyh-Sing Roger Jang *et al*. 1997. Neuro-Fuzzy and Soft Computing. *Prentice-Hall Press*.

[20] Tong, R. M. 1980. The Evaluation of Fuzzy Model Derived from Experimental Data. *Fuzzy Sets and System*. 4: 1-12.

[21] Sugeno, M., and T. Yasukawa. 1991. Linguistic Modelling based on Numerical Data. *Proc. of IFSA*. Brussel.

[22] Pedryz, W. 1984. An Identification Algorithm in Fuzzy Relational System. *Fuzzy Sets and Systems*. 13: 153-167.

[23] Zu, C. W., and Y. Z. Lu. 1987. Fuzzy Model Identification and Self-learning for Dynamic System. *IEEE Trans. on Systems Man and Cybernetics*. 683-689.

[24] Abreu, A., L. Custodio, and C. P. Frreria. 1996. A Fuzzy Modelling: A Rule Based Approach. Fifth IEEE International Conf. in Fuzzy Systems, Fuzzy-IEEE. 162-168.