MODELLING HIGH-DIMENSIONAL SYSTEMS WITH FUZZY RULES

Cherif Remache[1], *Ramdane Maamri*[2], *Zaidi Sahnoun*[3]
[1]Computer science department, Oum-El-Bouaghi University. Algeria
[2,3]LIRE Laboratory- Constantine University- Algeria
[1]remachec@yahoo.fr [2]rmaamri@yahoo.fr [3]sahnounz@yahoo.fr

ABSTRACT

The use of large number of variables to describe complex systems generates redundant elements and an excessive complexity with a poor transparency of the model describing the system. A particular interest is necessary to select relevant input variables to identify the input structure. We propose a method to select this structure in order to provide a compromise between the complexity, the transparency and the numerical accuracy of the studied model.

Index Terms— fuzzy rules, fuzzy inference systems, interpretative rules, complex systems, High-dimensional Systems, complexity reduction.

1. INTRODUCTION

To model high-dimensional systems, the use of fuzzy logic provides many facilities, the behavior of the system is described as a conditional statement using fuzzy rules on the form: If X is C Then Y is D, where X and Y are linguistic variables; the condition C and the conclusion D are linguistic values determined by fuzzy sets on the universe of discourse of X and Y, respectively. In real worlds, the systems to be modeled are complex, we use a large number of variables for their description and the expertise is often absent; data representing the behavior of the system is the only information available. Methods like fuzzy rules induction and learning process are used to formalize the model describing this system. Different class of methods to generate rules formalizing the relationship between the inputs and the outputs of the system are related in the literature:

Partitioning methods: We implement all the rules corresponding to each possible combination of input variables. The number of rules required to cover the antecedent domain in the rule base grows exponentially. Despite the simplicity of the model obtained, this technique is computationally demanding for complex and high dimensional systems and limits its use as a rapid development tool.

Clustering methods: The identification data are partitioned in clusters; the premise of a fuzzy rule describes the characteristics of the cluster, the consequent part describes the output of the model. Different criteria are used to evaluate the quality of the partition obtained.

The drawback of the clustering methods is that the obtained fuzzy sets are not shared by all the rules and the transparency of the obtained model is greatly reduced.

Hybrid methods: This third Class includes different approaches: Genetic algorithms: they provide a global optimum and give good performances: they achieve a high numerical accuracy with a low complexity and a compact fuzzy rule base, their drawback is that they neglect the semantic aspect.

Neuro-fuzzy models are widely used, [1], [2], they combine the formalism of fuzzy inference systems and the ability of neural networks giving them a power to acquire knowledge from data. The accuracy is favored over semantic aspects [3]. The interpretability of the knowledge acquired during the learning phase may be compromised if no special attention is paid during this process.

The rest of the paper presents in section 2 some structure selection for complex systems. In section3, we describe the proposed method for structure selection providing a compromise between the complexity and the numerical accuracy of the model. The performance of the proposed method is presented in section 4 and some concluding remarks are presented in section 5.

2. STRUCTURE SELECTION FOR COMPLEX SYSTEMS

The variables used to describe high-dimensional and complex systems can contain redundancies and may be correlated each other [4], [5]. To obtain models with a low complexity we must eliminate the redundant ones and reduce their number at the minimum. [6].

To select the relevant variables in the initial structure, different methods are proposed. These methods develop different models and compare them according to heuristic criteria in order to select the best combination of variables and retain the most relevant ones. For these methods, the numerical accuracy is favored and the semantic aspect is neglected. [3]. A growing interest to develop methods of structure selection providing a good compromise between the numerical accuracy and the complexity of the model. Different learning techniques using neuro-fuzzy systems are developed [7], [8].

In [9] a method using relational fuzzy rules is proposed. The interactions between the input variables are expressed through binary relations and are incorporated during the inference process.

In [10], the relevant input variables are identified by using information encoded in the parameters of the fuzzy model. Iteratively, the least significant variable is removed according to a criterion of dimensionality.

In [3], the method performs an exhaustive search in the space of all possible model structures and the combination with the relevant variables satisfying the criterion of the total fuzzy variance is selected.

To produce interpretable rules and to assure an acceptable compromise between the complexity and the accuracy of the model, different approaches integrating a set of constraints in the learning process have been proposed. [11], [12], [13]

3. PROPOSED METHOD

According to the constraints of dimensionality (the number of variables to be selected in the initial structure); the method proposed selects the most significant variables directly from data by choosing in the initial combination the variables the most correlated with the output but not correlated with each other. The constraints of the dimensionality are fixed by the user. The method generates interpretable fuzzy rules from the available data on the form of a set of N input-output data pairs (xk;yk); describing the behaviour of the system to be studied. According to the some observations and remarks:

Using many variables affects negatively the interpretability of the model [10]

• The perspicacity of knowledge acquired from data is greatly reduced when we use a large number of rules and it results in a low generalization.

• The complexity of the model grows greatly with the number of linguistic terms and it results in a low transparency.

Therefore, some constraints will be considered:

• Selecting only the most relevant variables fixing their number to the minimum possible in the input structure.

• minimize as possible as the number of linguistic terms per variables.

• The transparency of the partitioning of data space must be saved in the learning phase.

These constraints are fixed by the user.

3.1. Selection of the relevant variables

The most relevant variables are selected in the initial structure of the model according to the algorithm of figure 1.

• Z = [X | Y] is the initial vector data where:
N: number of samples; n: number of variables
X is an (N × n) matrix; and Y is an (N × 1) vector corresponding to the output of the system.
• C (n +1) × (n+1) represents the correlation matrix; the coefficients
C (i, n+1): i=1; … n are sorted in descending order
• The first variable corresponding to the maximum C (i, n+1) is Select in the initial combination.
• Iteratively and in descending order, we select another variable if it presents a low correlation with the variables already included in the combination until satisfying the constraints initially fixed by the user.

Figure 1: Selection of the Initial Structure of the model.

3.2. structure dentification.

We adopt the Takagi-Sugeno fuzzy model [7]; the quality of this model is its simplicity and its good numerical accuracy. It is very appropriate for the approximation of complex and dynamic systems. The interpretability of the model is assured by choosing a transparent and readable partitioning of the input space. Gaussian membership functions are used to describe the fuzzy sets. These membership functions satisfy all the properties required to ensure the interpretability of the model [12]. It is useful to work with a reduced dimensionality parameter space, the constraints on the number of input variables were be a priori fixed by the user.

In the Takagi-Sugeno fuzzy model, the rules are of the form:

$$\mathbf{R_k}: \textit{If } \mathbf{x_1} \text{ is } \mathbf{A}_1^k \textit{ and } \mathbf{x_2} \text{ is } \mathbf{A}_2^k \textit{ and} \ldots \textit{and } \mathbf{x_n} \text{ is } \mathbf{A}_n^k$$
$$\textit{Then } \mathbf{y_k} \text{ is } \mathbf{b_k}, \quad \mathbf{k} = 1 \ldots \mathbf{K} \qquad (1)$$

$\mathbf{X} = [\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n}]^t$: is the input vector

$\mathbf{y_k}$ **:** is the output of the k^{th} rule.

$\mathbf{A_i^k}$ **:** are fuzzy sets describing variables $\mathbf{x_i}$ in the k^{th} rule, they are defined in the antecedent space by their membership function: $\mu_{A_i^k}(\mathbf{x_i})$.

$K = \prod_{i=1}^{n} ki$ is the number of rules : n is the number of variables and $\mathbf{k_i}$ is the number of linguistic terms of the variable **xi.**

$\mathbf{y_c}$ is the value inferred by the model:

$$\mathbf{YC} = \sum_{k=1}^{K} \beta_k y_k \Big/ \sum_{k=1}^{K} \beta_k \qquad \mathbf{(2)}$$

β_k is the degree of activation of the kth rule, and yk is the contribution of this rule.

$$\beta_k = \Pi_{i=1}^{n} \mu_{A_i^k}(X_I), \quad K = 1,2, \ldots, K \quad (3)$$

$\mu_{A_i^k}$ is the membership function of the input **xi** to fuzzy subset $\mathbf{A}_i^k$ in the premise of the $\mathbf{k}^{th}$ rule.

The variable $\mathbf{x_i}$ is defined in the interval $[\boldsymbol{a_i}, \boldsymbol{b_i}]$:
$\boldsymbol{a_i} = \inf \mathbf{x_i}$ and $\boldsymbol{b_i} = \sup \mathbf{x_i}$.
The variable $\mathbf{x_i}$ is defined by $\mathbf{A_{ij}}$; j=1,2,...,ki linguistic terms defined by the Gaussian membership functions $\mu_{\mathbf{Aij}}$ with center ω_{ij} and standard deviation σ_{ij} .

$$\mu_{\mathbf{Aij}}(x_i) = \exp\left[-\frac{(x_i - \omega_{ij})^2}{2\sigma_{ij}^2} \right];$$

$$\mathbf{i} = 1,2,\ldots,\mathbf{n} \quad \text{and} \quad \mathbf{j} = 1,2,\ldots,\mathbf{k_i} \quad (4)$$

The averages are uniformly distributed in the interval $[ai, bi]$ and they are calculated by **(5):**

$$\omega_{ij} = (j-1)\frac{b_i - a_i}{k_i - 1}, \quad \mathbf{j} = 1,2, \ldots, \mathbf{K_i} \quad (5)$$

The standard deviations σ_{ij} are calculated by **(6):**

$$\sigma_{ij} = \frac{b_i - a_i}{2(k_i - 1)\sqrt{-2\ln\varepsilon}}, \quad \mathbf{j} = 1, 2, \ldots, \mathbf{K_i} \quad (6)$$

$\varepsilon \in [0, 1]$ to assure a good coverage of the input space.
The number of variables in input and the number of linguistic terms for each variable xi are a priori fixed by the user. The performance and the accuracy of the model are measured in terms of the *Mean Square Error:*

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - yc_i)^2 \quad (7)$$

1. Construct Z = [X | Y] the initial data set

2. Compute correlation matrix and sort variables $\mathbf{x_i}$ according to Cov (x_i , y); i =1 to n.

3. Select the initial combination of the model according to the criteria of selection.

4. Initial averages and standard deviations: ω_{ij} and σ_{ij} are computed conform to (5) and (6).

5. Compute $\mu_{\mathbf{Aij}}$ conform to (4) and yc conform to (2).

6. Compute MSE conform to (7).

7. While *MSE* greater than an *Epsi*; repeat from step 5.

Figure 2: The global Fuzzy Modelling Algorithm.

4. EVALUATION OF THE PERFORMANCE

We use the following nonlinear function [9] [10] to examine the performance of the proposed method.

$$y = 10\sin\left(\frac{\pi}{10}x_1\right)x_2 + 2(x_3 - 0.5)^2 + 3x_4 + 5x_5 \quad \textbf{(8)}$$

We construct a set of 1200 samples as follows:
five variables: x1 ; x2 ; x3 ; x4 and x5 are randomly generated according to the uniform distribution $\boldsymbol{U}(0,1)$ and the corresponding value of y is calculated according to (8); five other variables: x6 = x1 + 2 x2 ; x7 = 2x1 and x8 ; x9 ; x10 are randomly generated according to the uniform distribution U(0, 1). So then we have 10 variables candidates to construct the input structure of the model. The constraints on the number of inputs and linguistic terms per variable will be set by the user. The structure (with the relevant variables) of the model is obtained by performing algorithm Figure 1.

TABLE 1
CORRELATION COEFFICIENT OF X_i WITH Y

x_i	Cov($\mathbf{x_i}$,**y**)	x_i	Cov(**xj**,**y**)
x_1	0.0602	x_6	0.0705
x_2	0.0576	x_7	0.0529
x_3	0.9466	x_8	0.0430
x_4	0.0786	x_9	0.0257

When we retain three inputs; the variables $\mathbf{x_3}$, $\mathbf{x_4}$ and $\mathbf{x_2}$ are selected in the initial structure of the model. In 9 iterations, a mean square error less than 3% is obtained.

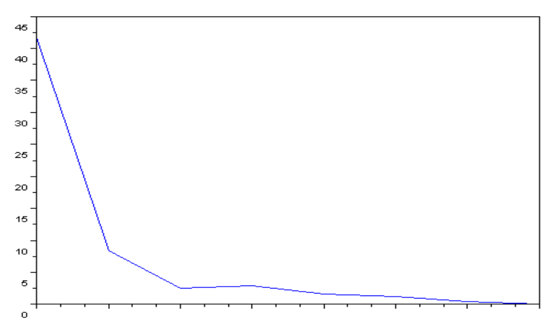

Figure. 3. number of iterations according to the MSE (3 variables and 3 linguistic termes by variable)

Figure 4 shows the first 60 values of the computed y according to (8) and the inferred value yc by the model obtained by our method.

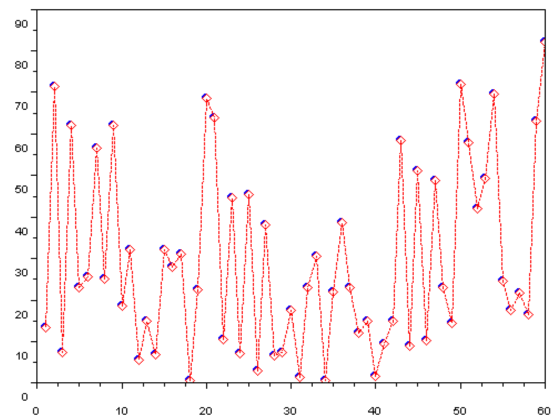

Figure 4: the first 60 values of y and yc

.

5. CONCLUSION AND PERSPECTIVE

The aim of the method presented is to simplify a complex system and to obtain a better compromise between the accuracy, the transparency and the complexity of the model. The interpretability of the system is provided by the transparency of the partitioning input space. The advantage of the proposed method is that the initial structure is identified in an incremental manner and there is no need to perform time demanding heuristics or clustering algorithms.

6. REFERENCES

[1] D. Nauk and R. Kruse. *A Neuro-Fuzzy Network to Learn Fuzzy Classification Rules from Data.* Fuzzy Sets and Systeems. 1997.

[2] J. Zurada and All. *Generating Linguistic Rules from Data using Neuro- Fuzzy Bramework*. 4th. Int. Conf. on Soft Computing. Fukuoda, Japan. 1996.

[3] D. Girimonte , R. Babuska and J. Abonyi. *Fuzzy Clustering for Selecting Stucture of Nonlinear Models with Mixed Discrete and Continuous Inputs*. IEEE Int. Conf. on Fuzzy Systems. Budapest, Hungary, 2004.

[4]M. Setnes,R. Babuska and H.B. Verbruggen. *Rule-Based Modeling: Precision and Transparency.* IEEE Transactions on Systems, Man and Cybernetics, 1998.

[5] A.E. Gaweda and J. Zurada. *Data-Driven Linguistic Modeling Using Relational Fuzzy Rules.* IEEE Transactions on Fuzzy Systems, 2003.

[6] R. Sindelar and R. Babuska. *Input Selection for Nonlinear Regression Models*. Transactions on Fuzzy Systems, Vol. 12, No 5, 2004.

[7] T. Takagi and M. Sugeno. *Fuzzy identification systems and its application to modelling and control.* IEEE,Transaction on Systems, Man and Cybernetics 1985.

[8]. J.J.Abonyi and R. Babuska. *Local and Global Identification and Interpretation of Parameters in Takagi-Sugeno Fuzzy Models.* IEEE Int. Conf. on Fuzzy Systems, San Antonio,USA 2000.

[9]. A.E. Gaweda and J. Zurada. *Fuzzy Neural Network with Relational Fuzzy Rules.* Int. Conf. on Neural Network. IJCNN' 2000.

[10] A.E. Gaweda, J. Zurada and R. Setiono. *Input Selection in Data-Driven Fuzzy Modeling.* Int. Conf. on Fuzzy Systems. IEEE 2001

[11] G. Ccastellano, A.M. Fanelli and C. Mencar. *Discovring Interpretable Classification Rules from Neural Processed Data.* Cognitive Systems, 2002

[12] G. Ccastellano, A.M. Fanelli and C. *A* Mencar. *Neuro-Fuzzy Network to Generate Human Understandable Knowledge from Data*", Cognitive Systems, 2002.

[13] J.V. De Olivera. *Semantic contraints for membership function optimization*. IEEE Trans on systems, Man and cybernetics, Part A, vol 29, 1999.