On System Identification via Fuzzy Clustering for Fuzzy Modeling *

Hsuan-Shih Lee
Department of Shipping and Transportation Management
National Taiwan Ocean University
Keelung, 202, Taiwan
Republic of China

Abstract

The area of fuzzy modeling encompasses several kinds of models that use various concepts from fuzzy sets theory to construct abstract models of systems. Typically, the fuzzy model of the process can be obtained directly from process measurements using an identification method based on fuzzy clustering. Domains of the problem, i.e., input, output and state variables, are partitioned into overlapping regions via fuzzy clustering. However, the number of clusters, which determines the degree of accuracy and complexity of the approximation, must be specified a priori before clustering. In this paper, an efficient method is proposed to determine the number of clusters for fuzzy clustering. A fuzzy equivalence relation is constructed for data points, and a hierarchy of partitions is obtained. To determine appropriate number of clusters, cohesion is defined to be the relatedness between data points in a cluster and coupling is defined to be the relatedness between data points in different clusters. An objective function is developed so as to maximize the cohesion between data points in the same cluster and minimize the coupling between data points in different clusters. The number of clusters when objective function is optimized is used in fuzzy clustering for input and output spaces. The method proposed is used as a preprocessing step for constructing fuzzy models of EDM (electrical discharge machining) process. Based on the results of simulation, we find that our hierarchical clustering approach serves as a good method for determining the number of clusters for fuzzy clustering.

Keywords: electrical discharge machining, fuzzy clustering, fuzzy control logic, hierarchical clustering

1 Introduction

The area of fuzzy modeling encompasses several kinds of models that use various concepts from fuzzy sets theory to construct abstract models of systems. Typically, the fuzzy model of the process can be obtained directly from process measurements using an identification method based on fuzzy clustering. Domains of the problem, i.e., input, output and state variables, are partitioned into overlapping regions via fuzzy clustering. However, the number of clusters, which determines the degree of accuracy and complexity of the approximation, must be specified a priori before clustering. In this paper, an efficient method is proposed to determine the number of clusters for fuzzy clustering. A fuzzy equivalence relation is constructed for data points, and a hierarchy of partitions is obtained. When a α-cut is taken, a partition is

*This research work was partially supported by the National Science Council of the Republic of China under grant No. NSC87-2416-h-019-001-

got. To determine appropriate number of clusters, cohesion is defined to be the relatedness between data points in a cluster and coupling is defined to be the relatedness between data points in different clusters. An objective function is developed so as to maximize the cohesion between data points in the same cluster and minimize the coupling between data points in different clusters. The number of clusters when objective function is optimized is used in fuzzy clustering for input and output spaces. It is shown that the objective function is optimized when $\alpha = 0.5$.

The method proposed is used as a preprocessing step for constructing fuzzy models of EDM (electrical discharge machining) process. Based on the results of simulation, we find that our hierarchical clustering approach serves as a good method for determining the number of clusters for fuzzy clustering.

In section 2, the structure of fuzzy models is discussed. The method proposed to determine the number of clusters is introduced in section 3. Finally, an application example of our method is shown in section 4.

2 The Structure of Fuzzy Models

Various classes of fuzzy models exist, depending on the form of the rules and the way the rules are organized in the rule base. For both Mamdani model and Sugeno-Takagi model [4], the fuzzy rule base corresponds to a static nonlinear map f that can be used to model a dynamic system, for instance in an input-output NARX (Nonlinear AutoRegressive with eXogenous input) form:

$$y(k+1) = f(y(k), y(k-1), \ldots, y(k-n+1), u(k), u(k-1), \ldots, u(k-m+1))$$

where $y(k), \ldots, y(k-n+1)$ and $u(k), \ldots, u(k-m+1)$ denote the lagged model outputs and inputs respectively and n, m are integers related to the model order. The lagged inputs and outputs are called the regressors. The NARX structure applies directly also to Multi-Input, Single-Output (MISO) systems and to systems with transport delays.

A Mamdani fuzzy model consists of rules of the following form:

$$\begin{aligned} R_i : &\text{ If } y(k) \text{ is } A_{i,1} \text{ and } y(k-1) \text{ is } A_{i,2} \text{ and}, \ldots \\ & y(k-n+1) \text{ is } A_{i,n} \text{ and } u(k) \text{ is } B_{i,1} \text{ and } u(k-1) \text{ is } \\ & B_{i,2} \text{ and}, \ldots, u(k-m+1) \text{ is } B_{i,m} \\ & \text{then } y(k+1) \text{ is } C_i \end{aligned} \tag{1}$$

where R_i denotes the ith rule, $A_{i,1}, \ldots, A_{i,n}$, $B_{i,1}, \ldots, B_{i,m}$, and C_i are fuzzy sets that usually have some linguistic meaning such as *Zero, Small, Big,* etc. In a variant of the Mamdani model, the consequent fuzzy sets C_i are reduced to singletons and can be represented as real numbers c_i.

A Sugeno-Takagi fuzzy model combines logical rule antecedents with numerical consequents that are know functions of the input variables:

$$\begin{aligned} R_i : &\text{ If } y(k) \text{ is } A_{i,1} \text{ and } y(k-1) \text{ is } A_{i,2} \text{ and}, \ldots \\ & y(k-n+1) \text{ is } A_{i,n} \text{ and } u(k) \text{ is } B_{i,1} \text{ and } u(k-1) \text{ is } \\ & B_{i,2} \text{ and}, \ldots, u(k-m+1) \text{ is } B_{i,m} \text{ then } \\ & y(k+1) = g_i(y(k), y(k-1), \ldots, u(k), u(k-1), \ldots) \end{aligned} \tag{2}$$

Both the Mamdani and Sugeno-Takagi fuzzy models can approximate any smooth function to any degree of accuracy [5], (1) and (2) can approximate any observable and controllable modes of a large class of discrete-time nonlinear systems [1].

For systems without lagged inputs and outputs, the model of simplified fuzzy reasoning [3] may suffice:

$$R_i : \text{If } x_1 \text{ is } A_{i1} \text{ and } \ldots \text{ and } x_m \text{ is } A_{im}$$

$$\text{then } y \text{ is } w_i \ (i = 1, \ldots, n) \tag{3}$$

where x_1, x_2, ..., x_m are input variables, A_{i1}, ..., A_{im} are the membership functions of the antecedent part, and w_i is a singleton for consequent part. The membership function A_{ij} of the antecedent part is assumed to a triangle. The parameters determining the triangle are the center a_{ij}, the left width b_{ij} and the right width c_{ij}. The output of fuzzy model y can be derived from the equations shown below.

$$A_{ij}(x_j) = \begin{cases} 1 - \frac{a_{ij} - x_j}{b_{ij}} & \text{if } a_{ij} - b_{ij} < x_j < a_{ij} \\ 1 - \frac{x_j - a_{ij}}{c_{ij}} & \text{if } a_{ij} < x_j < a_{ij} + c_{ij} \\ 1 & \text{if } x_j = a_{ij} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

$$u_i = A_{i1}(x_1) \cdot A_{i2}(x_2) \cdot \ldots \cdot A_{im}(x_m) \tag{5}$$

$$y = \frac{\sum_{i=1}^{n} u_i w_i}{\sum_{i=1}^{n} u_i} \tag{6}$$

where u_i is the membership value of antecedent part for rule i.

These models can be constructed by fuzzy clustering in the product space of the model input and output variables. The identification algorithm requires the structure of the model (inputs, outputs, order of the system) and the number of clusters to be fixed beforehand. The antecedent membership functions are extracted by projecting the clusters and the consequent parameters are estimated using least squares.

3 Determining Number of Clusters

Domains of problem consist of input, output and state variables. To partition the domains into overlapping regions, the number of clusters must be known beforehand. Here we propose a hierarchical clustering algorithm to determine the number of clusters.

Let the product space of model input and output be

$$R^{n'+m'+1} = Y \times \cdots \times Y \times U \times \cdots \times U \times Y$$

Let

$$X = \{x_1, x_2, \ldots, x_n\} \subset R^{n'+m'+1}$$

be the collection of process data. Each element x_i is itself a vector of lenght $m = n' + m' + 1$, that is,

$$x_i = \{x_{i_1}, x_{i_2}, \ldots, x_{i_m}\}.$$

Similarity metric can be applied to X to construct a tolerance relation R on X. Similarity between x_i and x_j can be measured with Correlation coefficient

$$r_{ij} = \frac{\sum_{k=1}^{m} | x_{i_k} - \overline{x_i} || x_{j_k} - \overline{x_j} |}{(\sum_{k=1}^{m}(x_{i_k} - \overline{x_i})^2)^{1/2} \cdot (\sum_{k=1}^{m}(x_{j_k} - \overline{x_j})^2)^{1/2}}$$

where

$$\overline{x_i} = \frac{1}{m}\sum_{k=1}^{m} x_{i_k} \text{ and } \overline{x_j} = \frac{1}{m}\sum_{k=1}^{m} x_{j_k}.$$

Then tolerance relation R can be defined as

$$\mu_R(x_i, x_j) = r_{ij}.$$

R is symmetric and reflexive, but not transitive. A fuzzy equivalence relation can be derived by computing transitive closure of R. Let R^* be the transitive closure of R. It can be computed as follows:

$$\mu_{R^2}(x_i, x_j) = \max_k \min[\mu_R(x_i, x_k), \mu_R(x_k, x_j)]$$

$$\mu_{R^n}(x_i, x_j) = \max_k \min[\mu_{R^{n-1}}(x_i, x_k), \mu_R(x_k, x_j)]$$

and

$$\mu_{R^*}(x_i, x_j) = \max_{k \geq 1} \mu_{R^k}(x_i, x_j).$$

Since R^* satisfies reflexive, symmetric and transitive law as stated in the following, R^* is a fuzzy equivalence relation:

1. reflexive: $\mu_{R^*}(x_i, x_i) = \max_{k \geq 1} \mu_{R^k}(x_i, x_i) \geq \mu_R(x_i, x_i) = 1$

2. symmetric: $\mu_{R^*}(x_i,x_j) = \max_{k\geq 1}\mu_{R^k}(x_i,x_j) = \max_{k\geq 1}\mu_{R^k}(x_j,x_i) = \mu_{R^*}(x_j,x_i)$

3. transitive: $\mu_{R^*}(x_i,x_j) \geq \max_k \min[\mu_{R^*}(x_i,x_k), \mu_{R^*}(x_k,x_j)]$

□

Thus $R^*{}_\alpha$ is an equivalence relation:

$$R^*{}_\alpha = \{(x_i,x_j) \mid \mu_{R^*}(x_i,x_j) \geq \alpha\}$$

,where $0 \leq \alpha \leq 1$. In other words, given a α, we can get a partition on X. So we can use α as a threshold for partitioning. However it is difficult to tell which value for α is suitable. Usually, you have to do the experiments for different α. There is no guideline for determining suitable α.

In structured programming, the term cohesion has been adopted for referring to the internal binding, or the degree of relatedness of a module's internal parts, and the connection between modules is referred to as coupling. To achieve a good modular design, coupling must be minimized and cohesion must be maximized. Likewise, cohesion within a cluster and coupling between clusters can be defined as follows.

Definition 3.1 *Let the cohesion of an equivalence class $Y \in E_{R^*{}_\alpha}$ is defined to be*

$$\sum_{x_i\in Y}\ \sum_{(x_j\in Y\, and\ x_j\neq x_i)} \mu_{R^*}(x_i,x_j).$$

And the coupling between equivalence classes of $E_{R^*{}_\alpha}$ is defined to be

$$\sum_{Y\in E_{R^*{}_\alpha}}\ \sum_{x_i\in Y}\ \sum_{(Z\in E_{R^*{}_\alpha} and\ Z\neq Y)}\ \sum_{x_j\in Z} \mu_{R^*}(x_i,x_j).$$

If we only consider the cohesion within clusters and coupling between clusters, it is obvious that all processes in one cluster would achieve maximum cohesion and minimum coupling. Therefore, the degree of unrelatedness within clusters which is in favor of dissolving the clusters, and the degree of unrelatedness between clusters which is in favor of aggregation of clusters should be taken into consideration. The degree of unrelatedness between processes i and j is denoted as $1-\mu_{R^*}(x_i,x_j)$. The net cohesion and net coupling of a partition can be defined as follows.

Definition 3.2 *The net cohesion of clustering $E_{R^*{}_\alpha}$ can be defined as*

$$\sum_{Y\in E_{R^*{}_\alpha}}\ \sum_{x_i\in Y}\ \sum_{(x_j\in Y\, and\ x_j\neq x_i)} \mu_{R^*}(x_i,x_j) + \sum_{Y\in E_{R^*{}_\alpha}}\ \sum_{x_i\in Y}\ \sum_{(Z\in E_{R^*{}_\alpha} and\ Z\neq Y)}\ \sum_{x_j\in Z}(1-\mu_{R^*}(x_i,x_j)). \quad (7)$$

And the net coupling of $E_{R^*{}_\alpha}$ can be defined as

$$\sum_{Y\in E_{R^*{}_\alpha}}\ \sum_{x_i\in Y}\ \sum_{(x_j\in Y\, and\ x_j\neq x_i)} (1-\mu_{R^*}(x_i,x_j)) + \sum_{Y\in E_{R^*{}_\alpha}}\ \sum_{x_i\in Y}\ \sum_{(Z\in E_{R^*{}_\alpha} and\ Z\neq Y)}\ \sum_{x_j\in Z}\mu_{R^*}(x_i,x_j). \quad (8)$$

Definition 3.3 *A partition of X is optimal or said to be in a state of equilibrium if the net cohesion is minimized and net coupling is maximized.*

Actually, according to the definition above, the optimal partition occurs at $\alpha = 0.5$.

Theorem 3.1 *Equation (7) is maximized and equation (8) is minimized if $\alpha = 0.5$.*

Proof: This can be shown by the following simple arguments: Each pair (x_i,x_j) contributes to the equation (8) either as the term $\mu_{R^*}(x_i,x_j)$ or $1-\mu_{R^*}(x_i,x_j)$. If $\mu_{R^*}(x_i,x_j) \geq \alpha$, (x_i,x_j) is counted as $(1-\mu_{R^*}(x_i,x_j))$. Otherwise, (x_i,x_j) is counted as $\mu_{R^*}(x_i,x_j)$. To minimize the contribution of (x_i,x_j) to equation (8), (x_i,x_j) should be counted as $(1-\mu_{R^*}(x_i,x_j))$ if $\mu_{R^*}(x_i,x_j) \geq 0.5$ and as $\mu_{R^*}(x_i,x_j)$ otherwise. In other words, α should be 0.5. In the same manner, we can show that equation (7) is maximized when $\alpha = 0.5$. □

4 Application

Electrical discharge machining (EDM) is the process of removing material in a closely controlled manner from an electrically conductive material immersed in a dielectric fluid with a series of transient electric sparkling discharge. EDM provides one of the best alternatives to machine the high strength and corrosion-and-wear materials into complex shapes.

There are three characteristics of EDM, i.e., machining speed, smoothness of workpiece, and consumption ratio of electrode to workpiece. Only two of them can be satisfied. If high machining speed and high quality of smoothness are required, the consumption ratio would be high. If high machining speed and low consumption ratio are required, then the quality of smoothness is low. Likewise, we can achieve high quality of smoothness and low consumption ratio at the expense of low machining speed. Since in the process of EDM, the machining speed is slower and longer than ordinary cutting speed, it is necessary to solve these difficulties by automatic control.

A fuzzy control model of simplified fuzzy reasoning is constructed for EDM. The consumption ratio, smoothness and speed are taken as control variables of the fuzzy model. Inference rules of the model generate output electrical current according to the values of these control variables.

Applying the method proposed in section 3 to an empirical data, we find that the number of the membership functions for control variables is three. Therefore, total number of inference rules is 27. These membership functions are determined by fuzzy c-mean clustering. Then a decent method [2] is applied to tune the parameters of membership functions. Simulation results are shown in Fig. 1.

5 Conclusion

In this paper, we proposed a hierarchical clustering method to determine the number of clusters for fuzzy clustering in constructing fuzzy models. An EDM application is applied to verify our method. The results of simulation show that our method serves as a good initial solution for fuzzy clustering.

References

[1] Leonaritis, I.J. and Billings, S.A. "Input-output parametric models for non-linear systems," International Journal of Control. Vol. 41, 1985, pp. 303-344.

[2] Nomura, H., Hayashi, I. and Wakami, N. "A Learning Method of Fuzzy Inference Rules by Descent Method," Proceedings of the First IEEE International Conference on Fuzzy Systems, pp. 203-210, 1992.

[3] Maeda, M. and Murakami, S. "Self-Tuning Fuzzy Logic Controlled," Transactions of the Society of Instrument and Control Engineers, Vol. 24, No. 2, pp. 191-197, 1988.

[4] Takagi, T. and Sugeno, M. "Fuzzy Identification of Systems and Its Application to Modeling and Control," IEEE Transactions on Systems, Man and Cybernetics. Vol. 15, No. 1, 1985, pp. 116-132.

[5] Wang,L.-X. "Fuzzy Systems Are Universal Approximators," Proc. IEEE Int. Conf. on Fuzzy Systems 1992, pp. 1163-1170, San Diego, CA, March 1992.

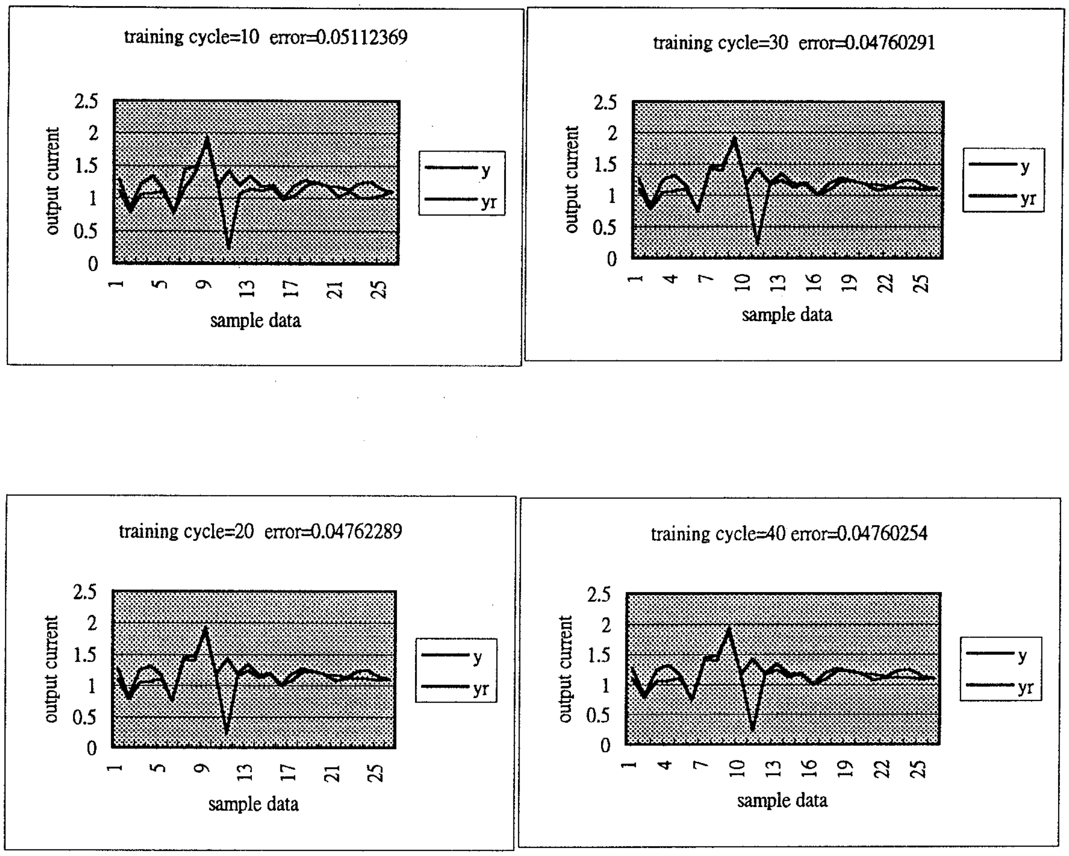

Fig. 1. Simulation Results of EDM fuzzy model.