

Construction of Fuzzy Classification System Based on Multi-objective Genetic Algorithm

Xing Zong-yi, Hou Yuan-long, Tong Zhong-zhi

School of Mechanical Engineering
University of Scientific and Technology
Nanjing, Jiangsu 210094, China
xingzongyi@gmail.com

Jia Li-min

School of Traffic and Transportation
Beijing Jiaotong University
Beijing 100044, China
jialm@vip.sina.com

Abstract

This paper present a novel method based on multi-objective genetic algorithm to construct fuzzy classification system. First, in order to relieve the problem of "curse of dimensionality", a multi-objective genetic algorithm is used to accomplish feature selection and fuzzy partition with maximum classification performance and minimum number of features and minimum number of fuzzy rules, thus an initial fuzzy system is obtained. Then, a genetic algorithm is employed to select significant fuzzy rules with two objectives to achieve a compact fuzzy system. In order to improve the classification performance of the compact fuzzy system, a constrained genetic algorithm is utilized to optimize the parameters of the compact fuzzy system. The proposed approach is applied to the Iris and Wine benchmark problems, and the results show its validity.

I. INTRODUCTION

Fuzzy classification system is used to describe pattern classification problem with IF-THEN fuzzy rules. The antecedents (IF- parts) of the fuzzy rules define local fuzzy regions, while the consequents (THEN- parts) describe the classification labels within these regions. It is desirable that the fuzzy classification system can correctly partition maximum number of patterns, while the number of features and the number of fuzzy rules can be kept low to increase the generalization of the system and improve its interpretability.

For a simple system, we can obtain a fuzzy classification system from human experts; however it is difficult to construct a fuzzy classification system for a complex system, where the expert experience is incomplete. So how to build a fuzzy classification system from data has become research focus in recent years^[1-4]. However all these technologies only focus on precision that simply fit data with highest possible accuracy, neglecting interpretability of the obtained fuzzy classification system. In order to improve interpretability of fuzzy classification system, some methods have been developed^[5-16].

This paper presents a novel method based on multi-objective genetic algorithm to construct fuzzy classification system. First, in order to relieve the "curse of dimensionality", a multi-objective genetic algorithm is used to accomplish feature selection and grid partition simultaneously. Then, rule selection is accomplished using a genetic algorithm, and the

obtained fuzzy classification system is optimized to improve its classification performance. The proposed approach is applied to two benchmark problems, and the results show its validity

This paper is organized as follows. In section II, we review the used fuzzy classification system. Construction of initial fuzzy system, including feature selection and fuzzy partition, is introduced in section III. Section IV shows how to select significant fuzzy rules and optimize the parameters of the fuzzy system by genetic algorithms. Section V provides experiments and results before concluding in section VI.

II. FUZZY CLASSIFICATION SYSTEM

Considering an n -dimensional classification problem for which N patterns, $x = (x_1, x_2, \dots, x_n)$ are given from M classes $\{C_1, C_2, \dots, C_M\}$. The classical fuzzy classification system is described as follows:

$$R_i : \text{If } x_1 \text{ is } A_{(1,k)}, x_2 \text{ is } A_{(2,k)}, \dots, x_n \text{ is } A_{(n,k)} \quad (1)$$

Then the pattern x belongs to class C_l with $CF = CF_l$ where $A_{(1,k)}, \dots, A_{(n,k)}$ are membership functions defined on the universe of features, M is the number of classes, n is the number of features, CF is certainty degree of the fuzzy rule..

In this paper, the k -th membership function (fuzzy set) of feature x_j is described as follows:

$$A_{(j,k)} = \begin{cases} \exp\left(-\frac{x_j - c_{jk}}{\omega_{jk}^l}\right)^2 & x_j \leq c_{jk} \\ \exp\left(-\frac{x_j - c_{jk}}{\omega_{jk}^r}\right)^2 & x_j > c_{jk} \end{cases} \quad (2)$$

where $c_{(j,k)}$, $\omega_{(j,k)}^l$ and $\omega_{(j,k)}^r$ are the centre, left width and right width of the membership function.

The output of the fuzzy classification system is determined by *winner takes all* strategy, i.e. the output is the class related to the consequent of the rule that reaches the highest degree of activation:

$$x_k \in C_l, \quad l = \arg(\max(\beta_i(x_k))) \quad 1 \leq l \leq M \quad (3)$$

where β_i is the degree of activation of the i -th rule

$$\beta_i(x_k) = CF_i \prod_{j=1}^n A_{ij}(x_k) \quad (4)$$

III. CONSTRUCTION OF THE INITIAL FUZZY SYSTEM

There are three types of fuzzy partition to form the antecedents of the fuzzy classification system: grid, tree, and scattering partition, where the grid partition has highest interpretability for it's easy to assign meaningful linguistic terms to membership functions. However the grid partition suffers from the "curse of dimensionality", *i.e.* the number of possible fuzzy rules tends to grow exponentially with the number of features. There are two ways to alleviate this problem. The first is to keep the dimensionality as low as possible by feature selection, and the second method is to employ lowest possible membership functions (fuzzy sets) for each feature. In this paper, we propose a multi-objective genetic algorithm to implement feature selection and fuzzy partition simultaneously.

A. Feature Selection

The main objective of feature selection is to reduce the dimensionality of the fuzzy classification system. There are two kinds of feature selection algorithm, including filter-model method and wrapper-model method^[17]. In this paper, we propose a feature selection method base on a genetic algorithm described as follows.

In the chromosome of the multi-objective genetic algorithm, all the candidate features are coded into the control genes with length equalling to the number of features. Each bit of control genes corresponds to a specific feature. If a bit is 0, the corresponding feature is excluded, otherwise it is included.

B. Fuzzy Partition and Membership Function

Without experts' knowledge, the homogenous grid partition is often used for fuzzy partition of feature space. There are two drawbacks of this method. First, the measured data are not always uniformly distributed, so the homogenous grid partition cannot express the actual characteristics of the data. Second, this method cannot alleviate the "curse of dimensionality". In this paper, we introduce a new type of grid partition technology described following^[18].

The homogenous grid partition is applied to the feature space, and the centres of the membership functions are then obtained. All these centres are used as candidate centres of the new grid partition. Subsequently, the multi-objective genetic algorithm is executed to select a subset of these centres, and these selected centres are adopted as the final centres of the new grid partition.

Given K_j is the number of homogenous grid partition of feature x_j , we denote the new grid partition of x_j as the string $\{I_1, I_2, \dots, I_{K_j}\}$ composed of 0 and 1. If a bit is 1, the corresponding centre is selected, otherwise it is disused. The total number of 1 in the string, denoted as k_j , is viewed as the number of membership functions of feature x_j . The centre of the k -th membership function is:

$$c_{(j,k)} = x_j^{\min} + (I_{jk} - 1) \frac{x_j^{\max} - x_j^{\min}}{k_j - 1} \quad k \in (1, 2, \dots, k_j) \quad (5)$$

where $\{x_j^{\min}, x_j^{\max}\}$ are the minimum and the maximum of x_j , I_{jk} is defined to represent the index of the string $\{I_1, I_2, \dots, I_{K_j}\}$ taking a value "1" with respect to x_j feature.

In order to guarantee the interpretability of the obtained fuzzy system, we define the overlapped factor, denoted as μ , where the membership functions are overlapped.

$$A_{(j,k)}(c_{(j,k-1)}) = \exp\left[-\frac{(c_{(j,k-1)} - c_{(j,k)})^2}{w_{(j,k)}^l}\right] = \mu \quad (6)$$

$$A_{(j,k)}(c_{(j,k+1)}) = \exp\left[-\frac{(c_{(j,k+1)} - c_{(j,k)})^2}{w_{(j,k)}^r}\right] = \mu$$

Then, the left and right width of the membership function can be calculated by-

$$w_{(j,k)}^l = \left[-\frac{(c_{(j,k)} - c_{(j,k-1)})^2}{\ln \mu}\right]^{1/2} \quad (7)$$

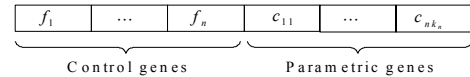
$$w_{(j,k)}^r = \left[-\frac{(c_{(j,k+1)} - c_{(j,k)})^2}{\ln \mu}\right]^{1/2}$$

After feature selection and fuzzy partition of feature spaces, the antecedents of the fuzzy system are determined. As for the consequents, please see the reference [19].

C. The Multi-objective Genetic Algorithm

The multi-objective genetic algorithm is used to accomplish feature selection and fuzzy partition simultaneously.

The chromosome is coded with sequence of binary digits including two parts: the control genes and the parametric genes.



The control genes implement feature selection with the length equalling to the number of candidate features. Each feature is coded to a bit in the control genes. If a bit is 1, then the corresponding feature is selected, otherwise it is disused.

The parametric genes implement fuzzy partition of feature spaces with the length equalling to the total number of homogenous grid partition of all candidate features. Each centre is coded to a bit in parametric genes. If a bit is 1, then the corresponding centre is selected as the centre of the new grid partition, otherwise, it is disused.

All genes of the first chromosome are assigned value 1, and the remaining chromosomes are generated randomly.

Fuzzy modelling requires the consideration of multiple objectives in the design process, including classification performance and interpretability. In this paper, classification performance is defined as the number of mistakenly classified training patterns (denote as J_{ERR}); interpretability is weighted by the number of used features (denote as J_f) and the number of fuzzy rules (denote as J_r).

These three objectives are combined into a single scalar fitness function as

$$\min F_1 = w_1 * J_{ERR} + w_2 * J_f + w_3 * J_r \quad (8)$$

where w_1 , w_2 and w_3 are positive weights which should be specified based on the users' preference. J_r is the number of rules:

$$J_r = \prod_{j=1}^n k_j \quad (9)$$

where k_j is the number of the new grid partition of feature x_j , i.e. the number of membership functions of feature x_j , n is the number of selected features.

Three genetic operators, i.e. selection, crossover and mutation are then employed to evolve the population.

The proposed approach is illustrated by Fig. 1.

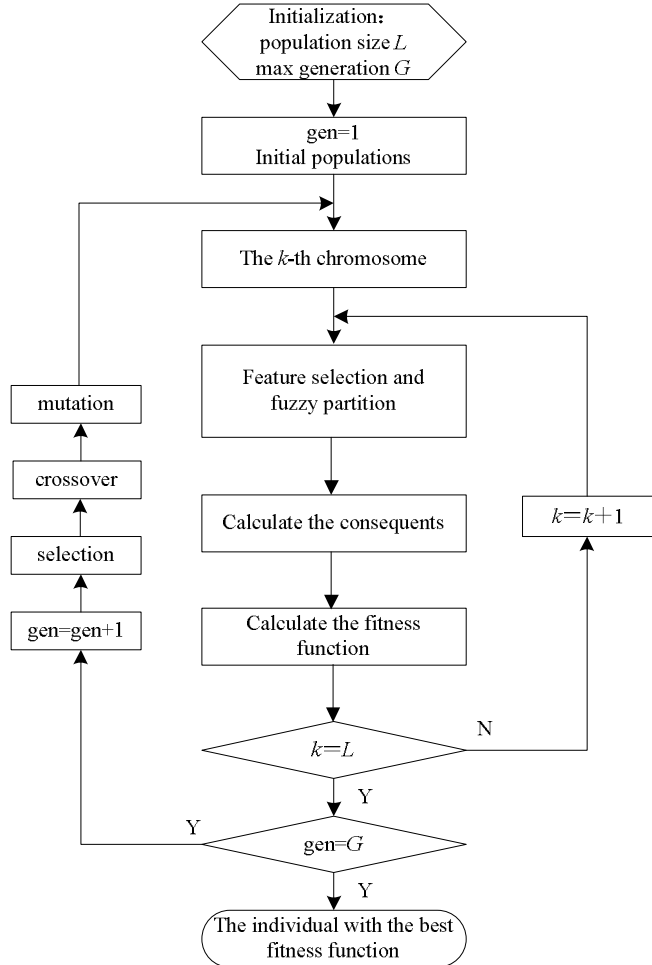


Fig. 1. Construction of initial fuzzy classification system

IV. OPTIMIZATION OF THE INITIAL FUZZY SYSTEM

The initial fuzzy classification system obtained in the preceding section owns a large number of fuzzy rules. Therefore, it is necessary to select a subset of significant fuzzy rules to construct compact fuzzy system with high classification performance. Additionally, the candidate centres of the new grid partition are select from the centres of the homogenous grid partition, and the overlapped factor is predetermined experientially. So it is also essential to optimize the obtained fuzzy system to improve its classification performance.

A. Rule Selection

Rule selection is to exclude the redundant fuzzy rules and extract the significant fuzzy rules to construct compact fuzzy

system with high classification performance. The orthogonal transformation methods^[7] are the most common technologies to accomplish this task, but it can only gives the importance factors about fuzzy rules, and the number of fuzzy rules is determined experientially. In this section, we use a genetic algorithm to selected significant fuzzy rules.

The chromosome is coded with sequence of binary digits with length equalling to the number of fuzzy rules. Each bit of the chromosome is associated with one fuzzy rule. If a bit is 1, then the corresponding rule is included in the rule base, otherwise will not.

All bits of the first chromosome are assigned value 1, which indicate that all the fuzzy rules are preserved in the rule base. The remaining chromosomes are generated randomly.

The objectives of the rule selection are selected a subset of fuzzy rules while keeping the classification performance. So the number of fuzzy rules and the number of mistakenly classified training patterns are combined into a single scalar fitness function as

$$\min F_2 = w_4 * J_{ERR} + w_5 * J_r \quad (10)$$

where w_4 and w_5 are positive weights respectively.

B. Parameters Optimization

In order to improve the classification performance of the compact fuzzy classifications system, it is essential to optimize the parameters of the obtained fuzzy system. The precision and the parameters of the fuzzy system are strongly nonlinear. So a genetic algorithm is used in this paper.

The parameters of the compact fuzzy systems, including the centres of the membership functions, the overlapped factors, and the certainty degrees of the consequents, are coded with a sequence of real number to form the chromosomes.

For the sake of maintaining the interpretability of the compact fuzzy classification system, the search spaces of the genetic algorithm are restricted. The centres are limited to change in a range of $\pm\alpha\%$ around their initial values. The search spaces of the overlapped factors are constrained in $[0.02 \ 0.45]$, and the certainty degrees can only vary form 0 to 1.

V. EXPERIMENTS AND RESULTS

In order to examine the performance of the proposed method, two benchmark problems, the Iris system and the Wine system are demonstrated in this section.

A. Iris System

The Iris system is a common benchmark problem in classification and pattern recognition studies. It contains 150 measurements of four features (*sepal length*, *sepal width*, *petal length*, *petal width*) from each of three species (*setosa*, *versicolor*, *virinica*). The first class is separate from others clearly, while the second and third class are overlap slightly. Fig. 2 shows the two-dimension (*sepal length*, *sepal width*) measurement, where “*” denotes data of *setosa* class, “o” denotes data of *versicolor* class, “+” denotes data of *virinica* class.

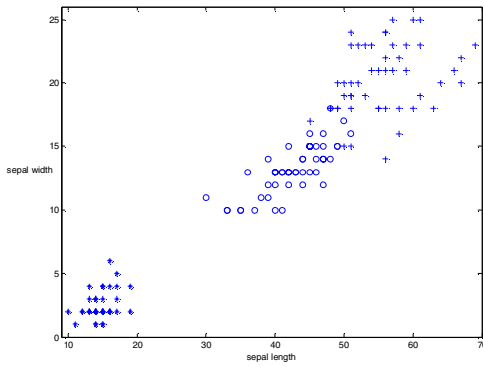


Fig.2. Iris data: *setosa* (*), *versicolor* (o), *virginica* (+)

The parameter setups of the multi-objective genetic algorithm to construct initial fuzzy classification system are described following: the population size is 20; the maximum number of generation is 100; the crossover probability is 0.5; the mutation probability is 0.1; the overlapped factor is 0.15; three weights of fitness function are $\omega_1 = 1, \omega_2 = 0.5, \omega_3 = 0.2$.

Table I details the obtained initial fuzzy system where x_3 and x_4 represent *sepal width* and *petal width* respectively, C , L and R represent centres, left width and right width of the membership functions respectively, # and CF represent the class labels and certainty degrees respectively. The obtained initial fuzzy classification system classifies 146 patterns correctly with 2 features and 12 fuzzy rules.

TABLE I
INITIAL FUZZY CLASSIFICATION SYSTEM OF IRIS

	x_3			x_4			#	CF
	C	L	R	C	L	R		
1	1	0	1.089	0	0	0.9075	1	0.9992
2	1	0	1.089	1.25	0.9075	0.9075	1	0.9903
3	1	0	1.089	2.5	0.9075	0	1	0.8948
4	2.5	1.089	1.089	0	0	0.9075	2	0.7808
5	2.5	1.089	1.089	1.25	0.9075	0.9075	2	0.5592
6	2.5	1.089	1.089	2.5	0.9075	0	2	0.8376
7	4	1.089	1.089	0	0	0.9075	1	0.8258
8	4	1.089	1.089	1.25	0.9075	0.9075	2	0.4786
9	4	1.089	1.089	2.5	0.9075	0	2	0.2892
10	5.5	1.089	1.089	0	0	0.9075	1	0.6324
11	5.5	1.089	1.089	1.25	0.9075	0.9075	3	0.2907
12	5.5	1.089	1.089	2.5	0.9075	0	3	0.8047

Feature selection is then executed by a genetic algorithm with the following parameters: the population size is 30; the maximum number of generation is 100; the crossover probability is 0.5; the mutation probability is 0.05; two weights of fitness function are $\omega_4 = 1, \omega_5 = 2$.

The genetic algorithm selects the 1st, 8th and the 12th fuzzy rules to build the compact fuzzy system:

$$\begin{aligned}
 R_1 &: \text{If } x_3 \text{ is } A_{(3,1)}, x_4 \text{ is } A_{(4,1)}, \text{ Then } x \text{ belongs to Class 1 with } CF = 0.9992 \\
 R_8 &: \text{If } x_3 \text{ is } A_{(3,2)}, x_4 \text{ is } A_{(4,2)}, \text{ Then } x \text{ belongs to Class 2 with } CF = 0.4786 \\
 R_{12} &: \text{If } x_3 \text{ is } A_{(3,3)}, x_4 \text{ is } A_{(4,3)}, \text{ Then } x \text{ belongs to Class 3 with } CF = 0.8047
 \end{aligned} \tag{11}$$

Fig.3 illustrates the membership functions of the compact fuzzy system.

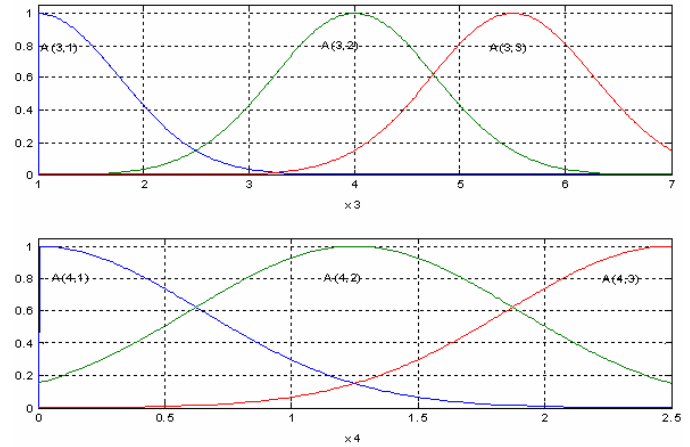


Fig.3. Membership functions of the compact fuzzy system (Iris)

The constrained genetic algorithm is then used to optimize the compact fuzzy system to improve its classification performance. After optimization, the final fuzzy classification system is described in (12), and the membership functions are displayed in Fig.4.

$$\begin{aligned}
 R_1 &: \text{If } x_3 \text{ is } A_{(3,1)}, x_4 \text{ is } A_{(4,1)}, \text{ Then } x \text{ belongs to Class 1 with } CF = 0.1992 \\
 R_2 &: \text{If } x_3 \text{ is } A_{(3,2)}, x_4 \text{ is } A_{(4,2)}, \text{ Then } x \text{ belongs to Class 2 with } CF = 0.4045 \\
 R_3 &: \text{If } x_3 \text{ is } A_{(3,3)}, x_4 \text{ is } A_{(4,3)}, \text{ Then } x \text{ belongs to Class 3 with } CF = 0.3816
 \end{aligned} \tag{12}$$

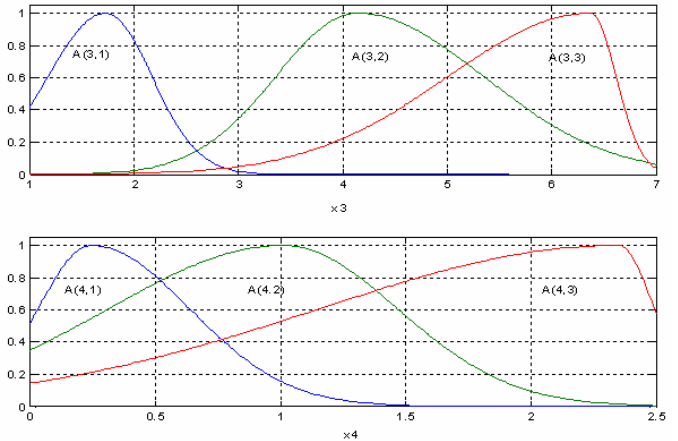


Fig.4 Membership functions of the final fuzzy system (Iris)

TABLE II
PERFORMANCE OF THE FUZZY SYSTEM BASED ON DIFFERENT METHOD (IRIS)

	Num. of features	Num. of rules	Correctly classification (%)
Wang[20]	11	3	97.5
Wu[21]	9	3	96.2
Shi[22]	12	4	98
Ishibuchi[14]	7	5	98
Tong[23]	12	3	98
Russo[24]	18	5	100
This paper	6	3	98

Table II compares our results and the result of other studies, which indicates that the proposed method can generate fuzzy classification system with higher classification performance and lesser number of features and less number of fuzzy rules.

B. Wine System

The wine data contains the chemical analysis of 178 wines grown in the same region in Italy but derived from three different cultivars. The 13 continuous attributes are available for classification. We use $x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}, x_{13}$ to represent these attributes: *alcohol, malic acid, ash, alkalinity of ash, magnesium, total phenols, Paranoids, nonflavanoids phenols, proanthocyaninsm, color intensity, hue, OD280/OD315 of dilluted wines and proline* respectively.

The parameter setups of the genetic algorithm to construct initial fuzzy system are identical to parameter setups aforementioned. The obtained initial fuzzy classification system classifies 164 patterns correctly with 5 features and 32 fuzzy rules.

After rule selection, the compact fuzzy classification system is obtained and is described in (13). Fig.5 shows the membership functions of the obtained compact fuzzy system.

$$\begin{aligned}
 R_1 : & \text{If } x_1 \text{ is } A_{(1,1)}, x_7 \text{ is } A_{(7,1)}, x_{11} \text{ is } A_{(11,2)}, x_{12} \text{ is } A_{(12,1)}, x_{13} \text{ is } A_{(13,1)} \\
 & \text{Then } x \text{ belongs to Class 2 with CF} = 0.4606 \\
 R_2 : & \text{If } x_1 \text{ is } A_{(1,1)}, x_7 \text{ is } A_{(7,2)}, x_{11} \text{ is } A_{(11,2)}, x_{12} \text{ is } A_{(12,2)}, x_{13} \text{ is } A_{(13,1)} \\
 & \text{Then } x \text{ belongs to Class 2 with CF} = 0.5584 \\
 R_3 : & \text{If } x_1 \text{ is } A_{(1,2)}, x_7 \text{ is } A_{(7,1)}, x_{11} \text{ is } A_{(11,1)}, x_{12} \text{ is } A_{(12,1)}, x_{13} \text{ is } A_{(13,1)} \\
 & \text{Then } x \text{ belongs to Class 3 with CF} = 0.5094 \\
 R_4 : & \text{If } x_1 \text{ is } A_{(1,2)}, x_7 \text{ is } A_{(7,2)}, x_{11} \text{ is } A_{(11,2)}, x_{12} \text{ is } A_{(12,2)}, x_{13} \text{ is } A_{(13,2)} \\
 & \text{Then } x \text{ belongs to Class 1 with CF} = 0.7945
 \end{aligned} \tag{13}$$

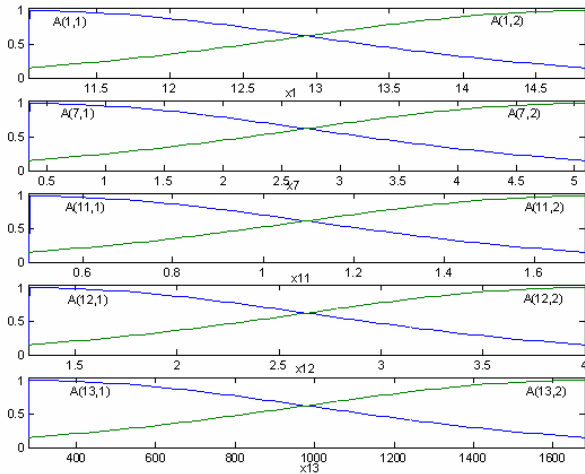


Fig.5 Membership functions of the compact fuzzy system (Wine)

The constrained genetic algorithm is then used to optimize the compact fuzzy system to improve its classification performance. After optimization, the final fuzzy classification system is described in (14) and the membership functions are displayed in Fig.6. Table III compares our results and the result of other studies

$$\begin{aligned}
 R_1 : & \text{If } x_1 \text{ is } A_{(1,1)}, x_7 \text{ is } A_{(7,1)}, x_{11} \text{ is } A_{(11,2)}, x_{12} \text{ is } A_{(12,1)}, x_{13} \text{ is } A_{(13,1)} \\
 & \text{Then } x \text{ belongs to Class 2 with CF} = 0.35086 \\
 R_2 : & \text{If } x_1 \text{ is } A_{(1,1)}, x_7 \text{ is } A_{(7,2)}, x_{11} \text{ is } A_{(11,2)}, x_{12} \text{ is } A_{(12,2)}, x_{13} \text{ is } A_{(13,1)} \\
 & \text{Then } x \text{ belongs to Class 2 with CF} = 0.5046 \\
 R_3 : & \text{If } x_1 \text{ is } A_{(1,2)}, x_7 \text{ is } A_{(7,1)}, x_{11} \text{ is } A_{(11,1)}, x_{12} \text{ is } A_{(12,1)}, x_{13} \text{ is } A_{(13,1)} \\
 & \text{Then } x \text{ belongs to Class 3 with CF} = 0.8915 \\
 R_4 : & \text{If } x_1 \text{ is } A_{(1,2)}, x_7 \text{ is } A_{(7,2)}, x_{11} \text{ is } A_{(11,2)}, x_{12} \text{ is } A_{(12,2)}, x_{13} \text{ is } A_{(13,2)} \\
 & \text{Then } x \text{ belongs to Class 1 with CF} = 0.9395
 \end{aligned} \tag{14}$$

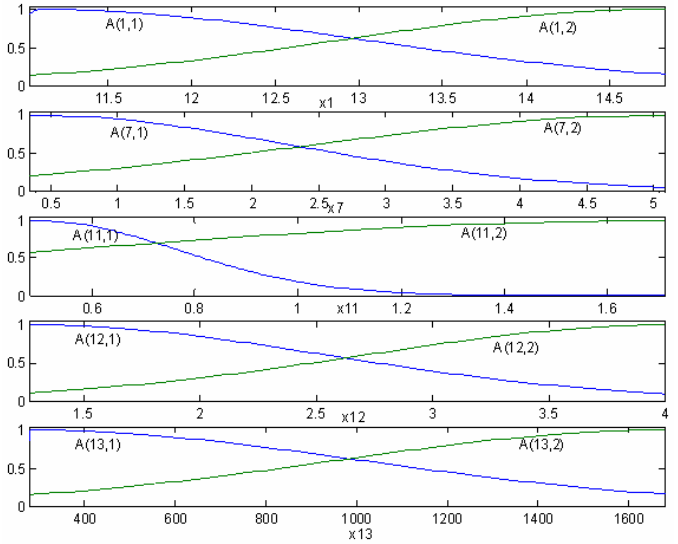


Fig.6 Membership functions of the final fuzzy system (Wine)

TABLE III
PERFORMANCE COMPARISONS OF DIFFERENT METHOD (WINE)

	Num. of features	Num. of fuzzy sets	Num. of rules	Correctly classification (%)
Setnes[25]	9	21	3	98.3
Wang[20]	13	34	3	99.4
Ishibuchi [14]	-	9	6	100
Robos[26]	5	15,11,10	3	98.9, 98.3, 99.4
Chang[27]	6	13	5	98.9
This paper	5	10	4	98.3

VI. CONCLUSIONS

The paper presents a novel approach to construct fuzzy classification system. First, a multi-objective genetic algorithm is used to accomplish feature selection and fuzzy partition with three objectives, thus an initial fuzzy system is obtained. Then, a genetic algorithm is employed to select significant fuzzy rules with two objectives to achieve a compact fuzzy system. In order to improve the classification performance of the compact fuzzy system, a constrained genetic algorithm is utilized to optimize the parameters of the compact fuzzy system. The proposed approach is applied to the Iris and Wine benchmark problems, and the results show its validity.

Since the centres of fuzzy partition are selected from the centres of the homogenous grid partition which are predefined, further improvement should be focus on how to determine the centres of fuzzy partition in the universe of features optionally.

ACKNOWLEDGMENT

This paper is partially supported by National Science Foundation of P. R. China (60332020), and Scientific Research Foundation of NUST (2005-AB96132).

IEEE Trans. Systems, Man, and Cybernetics, Part B, 2004, 34(4): 1894-1906.

REFERENCES

- [1] Lefteri H T, Robert E U. *Fuzzy and Neural Approaches in Engineering*. New York: Wiley, 1997.
- [2] Jang J S R, Sun C T, Mizutani E. *Neuro-Fuzzy and Soft Computing*. New Jersey: Prentice Hall, 1996.
- [3] Gomez-Skarmeta A F, Delgado M, Vila M A. About the use of fuzzy clustering techniques for fuzzy model identification. *Fuzzy Sets and Systems*, 1999, 106(2): 179-188.
- [4] Cordon O, Herrera F, Hoffmann F, Magdalena L. *Genetic Fuzzy Systems: Evolutionary Tuning and Learning of Fuzzy Rule Bases*. Singapore: World Scientific, 2000.
- [5] Jin Y. *Advanced fuzzy systems design and applications* [M]. New York: Physical-Verl, 2003.
- [6] Roubos H, Setnes M. Compact and transparent fuzzy models and classifiers through iterative complexity reduction [J]. *IEEE Trans on Fuzzy Systems*, 2001, 9(4): 516-524
- [7] Casillas J, Cordón O, Herrera F, Magdalena L. Interpretability improvements to find the balance interpretability-accuracy in fuzzy modeling: an overview. Chapter of Interpretability Issues in Fuzzy Modeling. Springer, 2003, pp. 3-22.
- [8] Guillaume S. Designing fuzzy inference systems from data: An interpretability-oriented review [J]. *IEEE Trans. Fuzzy Systems*, 2001, 9(3): 426-443.
- [9] Liu S-R, Yu J-S. A fuzzy inference system based on optimal fuzzy cluster and its application in product quality estimation [J]. *Information and Control*, 2000, 29(3): 272-279.
- [10] Abonyi J, Roubos J A, Szeifert F. Data-driven generation of compact, accurate, and linguistically sound fuzzy classifiers based on a decision-tree initialization [J]. *International journal of approximate reasoning*. 2003, (32): 1-21.
- [11] Wang H, Kwong S, Jin Y, et al. A multi-objective hierarchical genetic algorithm for interpretable rule-based knowledge extraction [J]. *Fuzzy Sets and Systems*, 2005, 149(1), 149-186.
- [12] Xing Z-Y, Jia L-M, Zhang Y, Hu W-L. A case study of data-driven interpretable fuzzy modeling [J]. *ACTA AUTOMATICA SINICA*, 2005, 31(6):815-824.
- [13] Ho S-Y, Chen H-M. Design of accurate classifiers with a compact fuzzy-rule base using an evolutionary scatter partition of feature space [J]. *IEEE Trans. on Systems, Man and Cybernetics, Part B*, 2004, 34(2): 1031-1044.
- [14] Ishibuchi H, Nakashima T, Murata T. Three-objective genetics-based machine learning for linguistic rule extraction [J]. *Information Science*, 2001, 136(1-4): 109-133.
- [15] Paiva R P, Dourado A. Interpretability and learning in neuro-fuzzy system [J]. *Fuzzy sets and systems*. 2004, 147: 17-38.
- [16] Nauck D D. Fuzzy data analysis with NEFCLASS [J]. *International journal of approximate reasoning*. 2003, 32:103-130.
- [17] Kohavi R, John, G. Wrapper for feature subset selection [J]. *Artificial intelligence*, 1997, 97(1-2):273-324.
- [18] Wong C-C, Chen C-C. A GA-based method for constructing fuzzy systems directly from numerical data [J]. *IEEE Trans. on Systems, Man and Cybernetics, Part B*, 2000, 30(6): 904-911.
- [19] Xing Z-Y, Zhang Y, Jia L-M, Hu W-L. Design of Interpretable and precise Fuzzy Classification System based on Fuzzy Clustering and Genetic Algorithm [J]. *Chinese Journal of Electronic*, 2006, 36(1): 97-102.
- [20] Wang J S, Lee G. C S. Self-adaptive neuro-fuzzy inference system for classification application [J]. *IEEE Trans. Fuzzy System*. 2002, 10: 790-802.
- [21] Wu T P, Chen S M. A new method for constructing membership functions and fuzzy rules from training examples [J]. *IEEE Trans System Man Cybernet Part B*.1999, 29: 25-40.
- [22] Shi Y, Eberhart R, Chen Y. Implementation of evolutionary fuzzy system [J]. *IEEE Trans. Fuzzy Systems*, 1999, 7:109-119.
- [23] Tong S, Shen Y, Liu Z. Approach to construct fuzzy classification system with clustering [J]. *Control and Decision*. 2001, 16(SUPP1):737-740. (in Chinese)
- [24] Russo M. Genetic fuzzy learning [J]. *IEEE Trans. Evolutionary Computation*. 2000, 4(3): 259-273.
- [25] Setnes M, Roubos H. GA-fuzzy modeling and classification: complexity and performance [J]. *IEEE Trans. Fuzzy systems*, 2000, 8(5): 509-522.
- [26] Roubos J A, Setnes M. Learning fuzzy classification rules from labeled data [J]. *Information Science*, 2003, 150(1-2): 77-93.
- [27] Chang X, Lilly JH. Evolutionary Design of a Fuzzy Classifier from Data [J].