A Multi-objective Genetic-based Method for Design Fuzzy Classification Systems

***Chen Ji-lin*[†]*, Hou Yuan-long*[†]*, Xing Zong-yi*[†]*, Jia Li-min*[††]*, and Tong Zhong-zhi*[†]**

*[†] **School of Mechanical Engineering, Nanjing University of Scientific and Technology, Jiangsu,210094 China***
*[††] **School of Traffic and Transportation, Beijing Jiaotong University, 100044 China***

Summary
An approach based on multi-objective genetic algorithms is proposed to construct interpretable and precision fuzzy classification system from data. First, a multi-objective genetic algorithm is used to accomplish feature selection and dynamic grid partition with three objectives: maximization of precision, minimization of the number of features, and minimization of the number of fuzzy rules. The parameters of the membership functions are determined by neighboring overlap method, so the obtained initial fuzzy classification system is highly interpretable. Second, a compact fuzzy classification system is obtained using a genetic algorithm with three objectives: maximization of accuracy, minimization of the number of features, and minimization of the average length of fuzzy rules. Third, a constrained genetic algorithm is used to optimize the compact fuzzy classification system to improve its precision, while preserves its interpretability. The proposed approach is applied to the Iris and Wine benchmark classification problems, and the results show its validity.
Key words:
Fuzzy classification system, Genetic algorithm, Interpretability, Fuzzy partition, Grid Partition, Multi-objective

1. Introduction

Fuzzy systems have been successfully applied to various areas such as classification, simulation, data mining, pattern recognition, prediction and control, etc. Traditionally, for a simple system, we can obtain a fuzzy system from human experts; however it is difficult to construct a fuzzy system for a complex system, where the expert experience is incomplete. So how to build a fuzzy system from data has become research focus in recent years [1-4]. Several fuzzy modeling methods have been proposed, including fuzzy clustering-based algorithms [1], neuro-fuzzy systems [2, 3] and genetic fuzzy systems [4]. However all these methods only focus on precision that 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 [5-13].

Papadakis [5] proposed a genetic algorithm based modeling method for building fuzzy system with scatter-type partitions. The method manages all attributes concerning the structure of fuzzy system simultaneously, including the number of fuzzy rules, the input partition, the participating inputs of each fuzzy rule and the consequent parameters. The structure learning task is formulated as a multi-objective optimization problem which is resolved using a novel genetic-based structure learning scheme; and a genetic-based parameter learning scheme is performed for fine-tuning of the initial fuzzy system. Xing [6] gave an approach to identify interpretable fuzzy systems. The number of fuzzy rules is determined by validity indices, and the initial fuzzy system is identified by a modified fuzzy clustering algorithm and the lease square method. An orthogonal least square method and a method of merging similar fuzzy sets are then used to remove the redundancy of the fuzzy systems. Finally, a constraint Levenberg-Marquardt method is used to optimize the precision of the fuzzy system. Chang [7] addressed an automatic method to design fuzzy systems for classification via evolutionary optimization. At the beginning of the algorithm, the fuzzy system is empty with no rules in the rule base and no membership functions assigned to fuzzy variables. Then, different rules and membership functions are automatically created via VISIT algorithm by randomly assigning different initial parameters. At last, the evolutionary algorithm is used to find the optimal fuzzy system through simultaneously optimizing all the parameters of the system. Castro [8] presented an approach to construct genetic fuzzy system considering both accuracy and interpretability. First, a data pre-processing step for feature selection is performed and the membership functions are generated using FCM algorithm. After that, a genetic algorithm is used for fuzzy rule base generation considering both comprehensibility and simplicity criteria. Finally, another genetic algorithm is executed for optimization of the previous rule base obtained, excluding unnecessary and redundant fuzzy rules. Wang [9] proposed a new scheme based on multi-objective hierarchical genetic algorithm to extract interpretable

Manuscript received
Manuscript revised

fuzzy systems from data. Some important concepts about the interpretability are introduced firstly. Then fuzzy clustering is applied to an initial fuzzy system. Finally, the multi-objective hierarchical genetic algorithm and an interpretability-driven simplification method are used to obtain the optimized fuzzy systems.

The paper presents a method based on multi-objective genetic algorithms to build fuzzy classification system. First, in order to relieve the "curse of dimensionality", a multi-objective genetic algorithm is used to accomplish feature selection and fuzzy 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

The paper is organized as follows. In section 2, we review the fuzzy classification system. Construction of initial fuzzy system, including feature selection and fuzzy partition, is introduced in section 3. Section 4 shows how to select significant fuzzy rules and optimize the parameters of the fuzzy system by genetic algorithms. Section 5 provides experiments and results before concluding in section 6.

2. Fuzzy Classification System

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

$$R_i: \text{If } x_1 \text{ is } A_{(1,k)},\ \ x_2 \text{ is } A_{(2,k)},\cdots,\ x_n \text{ is } A_{(n,k)} \\ \text{Then the pattern x belongs to class } C_l \text{ with } CF = CF_i \tag{1}$$

where $A_{(1,k)},\cdots,A_{(n,k)}$ are membership functions defined on the domain 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:

$$A_{(j,k)}=\begin{cases}\exp\left(-\left(\dfrac{x_j-c_{jk}}{w_{jk}^l}\right)^2\right) & x_j \le c_{jk}\\ \exp\left(-\left(\dfrac{x_j-c_{jk}}{w_{jk}^r}\right)^2\right) & x_j > c_{jk}\end{cases} \tag{2}$$

where $c_{(j,k)}$, $w_{(j,k)}^l$ and $w_{(j,k)}^r$ 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:

$$\mathrm{x_k} \in C_l,\quad l=\arg(\max(\beta_i(\mathrm{x_k})))\quad 1\le l\le M \tag{3}$$

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

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

3. Building Initial Fuzzy Classification 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.

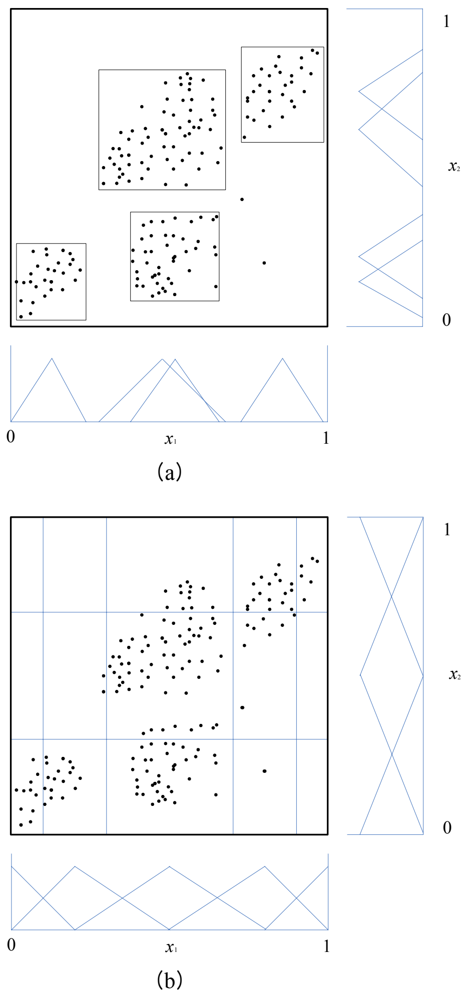


Fig. 1 Fuzzy partition: (a) scatter partition (b) grid partition

3.1 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 [14]. 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 equaling 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

3.2 Fuzzy Partition

Fuzzy clustering algorithm is a well-recognized technique to identify fuzzy models. As illustrated in Fig.1, the projection of multi-dimensional fuzzy clustering results the possible heavy overlap between membership functions and the obtained fuzzy system is less interpretable for it is difficult to assign linguistic terms to the fuzzy sets. However, a fuzzy system constructed by human experts is more interpretable, because the grid partition determined based on experts' knowledge is easy to interpret. In this paper, a modified grid partition is proposed.

Without experts' knowledge, the simple grid partition is often used for fuzzy partition of the feature spaces. There are two drawbacks to this method. First, the measure data are not always uniformly distributed. A fine fuzzy partition is required in one region, while perhaps a coarse fuzzy partition is needed in another region. So the simple grid partition cannot describe the actual distribution characteristic of the data. Second, the number of fuzzy rules increases exponentially with the number of features, i.e., it suffers from the curse-of-dimensionality problem.

Ishibuchi [15] proposed a multiple grid partition approach, and selected the most suitable fuzzy partitions to fit with the data. However, this method aggravates the "curse-of-dimensionality" problem. For example, for a three-dimensional feature space with five grid partitions, there are 125 and 228 fuzzy rules for simple grid partition and multiple grid partition, respectively. The number can increase enormously, especially for high-dimensional feature spaces. Wong [17] introduced a dynamic grid partition. First, a simple grid partition is used to obtain the centers of membership functions. All these centers are candidate centers of the dynamic grid partition. Second, a genetic algorithm is used to select a subset of these centers, and the selected centers are adopted as the final centers of the dynamic grid partition. For the example mentioned above, if the centers of the dynamic grid partition are 3-3-5 respectively, the number of fuzzy rules is only 45, which is far less than that of the simple grid partition and the multiple grid partition. However, the leftmost and the rightmost membership values of this method do not equal to one. This will hold back interpretability of the obtained fuzzy system. In addition, the interval-overlap method used to determine the parameters of the membership functions lacks visual interpretability.

A modified dynamic grid partition is proposed in this study. The neighboring overlap method is used to calculate the parameters of the membership functions. The leftmost and the rightmost membership values are equal to one.

Given K_j is the number of simple grid partition of feature x_j; we denote the new grid partition of x_j as the string $\{I_1, I_2, \cdots, I_{K_j}\}$ 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:

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

where $\{x_j^{\min}, x_j^{\max}\}$ are the minimum and the maximum of x_j, and I_{jk} is the sequence number of the k-th "1" in the string.

For example, if the domain of the j-th feature (i.e. x_j) is the unit interval [0,1], the centers of the five simple grid partitions are {0, 0.25, 0.5, 0.75, 1}. If the fuzzy partition string is {10110}, then the centers of the modified grid partition are {0, 0.5, 0.75}, as illustrated in Fig. 2.

In order to guarantee the interpretability of the obtained fuzzy system, the neighboring overlap method is used to calculate the parameters of the membership functions. The neighboring memberships functions $A_{(j,k)}(\frac{c_{(j,k)}+c_{(j,k+1)}}{2})$ and $A_{(j,k)}(\frac{c_{(j,k)}+c_{(j,k-1)}}{2})$ are overlapped at membership value of μ :

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

$$A_{(j,k)}\left(\frac{c_{(j,k)}+c_{(j,k-1)}}{2}\right)=\exp\left(\left(-\frac{\frac{c_{(j,k)}+c_{(j,k-1)}}{2}-c_{(j,k)}}{w_{(j,k)}^{l}}\right)^{2}\right)=\mu \quad (7)$$

The left width value and the right width value of the membership function are obtained as follows:

$$w_{(j,k)}^{r}=\sqrt{\frac{(c_{(j,k+1)}-c_{(j,k)})^{2}}{-2\ln(\mu)}} \quad (8)$$

$$w_{(j,k)}^{l}=\sqrt{\frac{(c_{(j,k-1)}-c_{(j,k)})^{2}}{-2\ln(\mu)}} \quad (9)$$

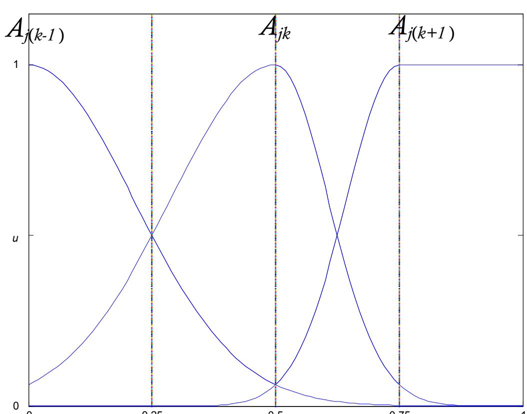


Fig.2. Simple grid partition and modified dynamic grid partition

If the first number of the string $\{I_1, I_2, \cdots, I_{K_j}\}$ does not equal to one, in order to satisfy interpretability, the leftmost membership function should be modified as:

$$A_{(j,1)}=\begin{cases}1 & x_j \le c_{(j,k)}\\ \exp\left(\left(-\frac{x_j-c_{(j,1)}}{w_{(j,1)}^{r}}\right)^{2}\right) & x_j > c_{(j,k)}\end{cases} \quad (10)$$

In like manner, if the last number of the string $\{I_1, I_2, \cdots, I_{K_j}\}$ does not equal to one, the rightmost membership function should be modified as:

$$A_{(j,K_j)}=\begin{cases}\exp\left(\left(-\frac{x_j-c_{(j,K_j)}}{w_{(j,K_j)}^{r}}\right)^{2}\right) & x_j \le c_{(j,k)}\\ 1 & x_j > c_{(j,k)}\end{cases} \quad (11)$$

In order to obtain incomplete fuzzy rules with high interpretability, a special membership function whose membership value is always unit in the domain of the feature value, named "don't care", is introduced [15]:

$$\mu(x_j)=\begin{cases}1 & x_j^{\min} \le x_j \le x_j^{\max}\\ 0 & \text{others}\end{cases} \quad (12)$$

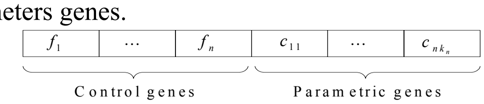
where $x_j^{\min}$ and $x_j^{\max}$ are the minimum and maximum of x_j, respectively.

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].

3.3 The Multi-objective Genetic Algorithm

The multi-objective genetic algorithm is utilized to accomplish feature selection and dynamic grid partition simultaneously.

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



The control genes implement feature selection with the length equaling 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 equaling to the total number of simple 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 modeling 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

$$\min F_1 = w_1 * J_{ERR} + w_2 * J_f + w_3 * J_r \tag{13}$$

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

$$J_r = \prod_{j=1}^{n} k_j \tag{14}$$

where k_j is the number of the dynamic 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 in Fig.3.

4. Optimize 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 centers of the dynamic grid partition are select from the centers of the simple grid partition, and the neighboring overlap value is predetermined experientially. So it is also essential to optimize the obtained fuzzy system to improve its classification performance.

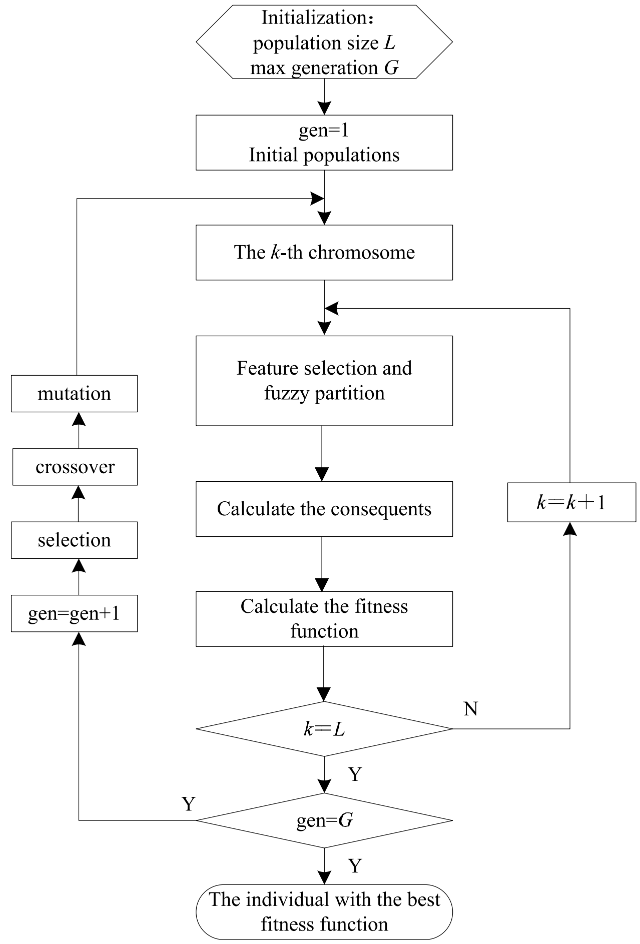


Fig.3. Construction of initial fuzzy classification system

4.1 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 [18] 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 equaling 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 patterns and average length of fuzzy rules are combined into a single scalar fitness function as

$$\min F_2 = w_4 * J_{ERR} + w_5 * J_r + w_6 * J_m \tag{10}$$

where J_m is average length of fuzzy rules, w_4, w_5 and w_6 are positive weights respectively.

4.2 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 centers of the membership functions, the neighboring overlap values, 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 centers are limited to change in a range of $\pm\alpha\%$ around their initial values. The search spaces of the neighboring overlap values are constrained in [0.02 0.45], and the certainty degrees can only vary form 0 to 1.

5. 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.

5.1 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*, *pental length*, *pental 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. 4 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.

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 50; the crossover probability is 0.8; the mutation probability is 0.1; the

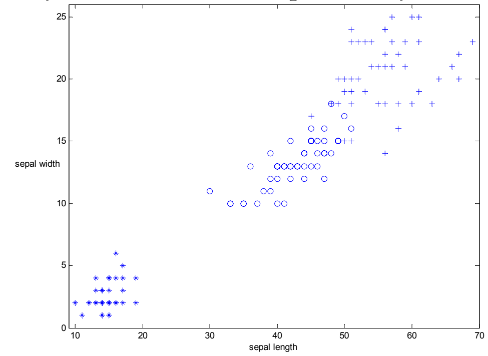


Fig.4. Iris data: *setosa* (*), *versicolor* (o), *virinica* (+)

neighboring membership values equal to 0.5, and the number of simple grid partition is 5. Three weights of fitness function are. $\omega_1 = 1, \omega_2 = 0.5, \omega_3 = 0.1$.

The multi-objective genetic algorithm runs 10 times, and the results are concluded in Table 1, where the *sepal length* and the *sepal width* are selected.

Table 1. Initial fuzzy classification system（Iris）

	Number of correctly classified patterns	Number of features	Number of rules
Average	145.8	2	16
Best	146	2	11
Worst	145	2	24

Table 2 details a obtained initial fuzzy system with 146 correctly classified patterns and 2 features and 15 rules. Fig. 5 gives the membership functions of the fuzzy classification system.

Table 2 Initial fuzzy classification system (Iris)

	Sepal length			Sepal Width			#	CF
	C	L	R	C	L	R		
1	4	0	1.089	0	0	0.7507	1	0.9117
2	4	0	1.089	1.25	0.7507	0.7507	2	0.6366
3	4	0	1.089	2.5	0.37535	0	2	0.2654
4	4	0	1.089	DC			1	0.2779
5	5.5	1.089	1.089	0	0	0.7507	1	0.7818
6	5.5	1.089	1.089	1.25	0.7507	0.7507	3	0.3429
7	5.5	1.089	1.089	2.5	0.37535	0	3	0.8921
8	5.5	1.089	1.089	DC			3	0.5729
9	7	1.089	0	0	0	0.7507	3	0.7835
10	7	1.089	0	1.25	0.7507	0.7507	3	0.9779
11	7	1.089	0	2.5	0.37535	0	3	0.9977
12	7	1.089	0	DC			3	0.9918
13	DC			0	0	0.7507	1	0.9005
14	DC			1.25	0.7507	0.7507	2	0.7849
15	DC			2.5	0.37535	0	3	0.2907

The genetic algorithm is used to select fuzzy rules to compose compact fuzzy system from the initial fuzzy system detailed in Table2. The parameters of the genetic algorithm are: the population size is 20, the maximum number of generation is 50, the crossover probability is 0.8, the mutation probability is 0.1, three weights of fitness function are $\omega_4 = 1, \omega_5 = 1, \omega_6 = 0.5$

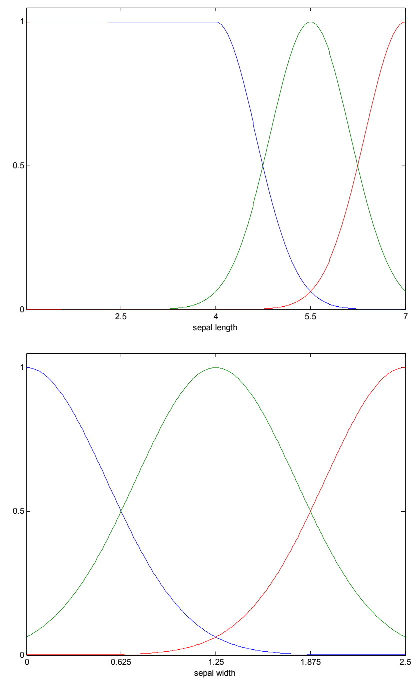


Fig.5. Membership functions of the initial fuzzy system (Iris)

The genetic algorithm runs 10 times, and the results are concluded in Table 3.

Table3. Compact fuzzy classification system (Iris)

	Number of correctly classified patterns	Number of rules	Average rule length
Average	144.7	3.4	1.5455
Best	146	3	1
Worst	144	4	2

The constrained genetic algorithm is then used to optimize the compact fuzzy systems to improve their classification performance. After optimization, the final fuzzy classification systems can all classify 147 patterns correctly.

In order to illustrate the performance of the proposed method, Table 4 compares the obtained 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.

Table 4 Performance comparison of different methods (Iris)

	Number of features	Number of rules	Correctly classification (%)
Wang[19]	11	3	97.5
Wu[20]	9	3	96.2
Shi[21]	12	4	98
Ishibuchi[15]	7	5	98
Tong[22]	12	3	98
Russo[23]	18	5	100
This paper	2	3	98

5.2 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_2, x_{13} to represent these attributes: *alcohol*, *malic acid*, *ash*, *alcalin*ity *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 multi-objective genetic algorithm to construct initial fuzzy classification system are described following: the population size is 30; the maximum number of generation is 100; the crossover probability is 0.8; the mutation probability is 0.1; the neighboring membership values equal to 0.5, and the number of simple grid partition is 5. Three weights of fitness function are. $\omega_1 = 1, \omega_2 = 0.5, \omega_3 = 0.1$.

The multi-objective algorithm runs 10 times, and the results are concluded in Table 5. If the algorithm selects four features, then the features are x_1, x_7, x_2, and x_{13}. If the algorithm selects five features, then the x_{11} should be added. If the algorithm selects six features, then the x_3 should also be added.

Table 5. Initial fuzzy classification system (Wine)

	Number of correctly classified patterns	Number of features	Number of rules
Average	163.6	5.2	653.4
Best	165	4	191
Worst	160	6	1296

The fuzzy system with 164 correctly classified patterns and 5 features and 242 rules is optimized by rule-selection genetic algorithm. The algorithm runs 10 times, and the results are identical. The compact fuzzy classification system classifies 165 patterns correctly with 4 rules, and the average length of fuzzy rules is 5.

The constrained genetic algorithm is then used to optimize the compact fuzzy systems to improve their classification performance. After optimization, the final fuzzy classification systems can all classify 175 patterns correctly. Table 6 compares the obtained results and the result of other studies.

Table 6 Performance comparison of different methods (Wine)

	Num. of features	Num. of fuzzy sets	Num. of rules	Correctly classification (%)
Setnes[24]	9	21	3	98.3
Wang[19]	13	34	3	99.4
Ishibuchi [15]	-	9	6	100
Robos[25]	5	15,11, 10	3	98.9, 98.3, 99.4
Chang[7]	6	13	5	98.9
This paper	5	10	4	98.3

6. Conclusion

The paper presents an approach to build fuzzy system from data. 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 three 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 centers of fuzzy partition are selected from the centers of the simple grid partition which are predefined, further improvement should be focus on how to determine the centers 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)

References

[1] A. F. Gomez-Skarmeta, M. Delgado, M. A. Vila, "About the use of fuzzy clustering techniques for fuzzy model identification," Fuzzy Sets and Systems, vol. 106, pp. 179-188, 1999.
[2] H. T. Lefteri, E. U. Robert, Fuzzy and Neural Approaches in Engineering, New York, Wiley, 1997.
[3] J. S. R. Jang, C. T. Sun, E. Mizutani, Neuro-Fuzzy and Soft Computing, New Jersey, Prentice Hall, 1996.
[4] O. Cordon, F. Herrera, F. Hoffmann, L. Magdalena, Genetic Fuzzy Systems: Evolutionary Tuning and Learning of Fuzzy Rule Bases, Singapore, World Scientific, 2000.
[5] S.E. Papadakis, J. B. Theocharis , "A GA-based fuzzy modeling approach for generating TSK models," Fuzzy Sets and Systems, vol. 131 pp. 121-152, 2002.
[6] Z-Y. Xing, Y. Zhang, L-M. Jia, W-L. Hu, "A case study of data-driven interpretable fuzzy modeling," ACTA AUTOMATICA SINICA, vol. 31: pp. 815-824, 2005.
[7] X. Chang, J. H. Lilly, "Evolutionary design of a fuzzy classifier from data," IEEE Trans System Man Cybernetic Part B, vol. 34, pp. 1894-906, 2004.
[8] P. A. D. Castro, H. A. Camargo, "Focusing on Interpretability and Accuracy of a Genetic Fuzzy system," Proc. IEEE Int. Conf. on fuzzy systems, RenoNevada, pp. 696-701, 2005.
[9] H L. Wang, S. Kwong, Y. C. Jin, et al, "Multi-objective Hierarchical Genetic Algorithm for Interpretable Fuzzy Rule-based Knowledge Extraction," Fuzzy Sets and Systems, vol. 149, pp. 149-186, 2005.
[10] J. Abonyi, J. A. Roubos, F. Szeifert, "Data-driven generation of compact, accurate, and linguistically sound fuzzy classifiers based on a decision-tree initialization," International journal of approximate reasoning, vol. 32, pp. 1-21, 2003.
[11] R. P. Paiva, A. Dourado, "Interpretability and learning in neuro-fuzzy system," Fuzzy sets and systems. 2004, vol. 147, pp. 17-38.
[12] D. D. Nauck, "Fuzzy data analysis with NEFCLASS," International journal of approximate reasoning, 2003, vol. 32, pp. 103-130.
[13] Y. C. Jin, Advanced fuzzy systems design and applications, New York, Physical-Verl, 2003.
[14] R. Kohavi, G. John, "Wrapper for feature subset selection," Artificial intelligence, vol. 97: pp. 273-324, 1997.
[15] H. Ishibuchi, T. Nakashima, T Murata, "Three-objective genetics-based machine learning for linguistic rule extraction," Information Scienc, vol. 136, pp. 109-133 e, 2001.
[16] S-Y. Ho, H-M. Chen, "Design of accurate classifiers with a compact fuzzy-rule base using an evolutionary scatter partition of feature space," IEEE Trans. on Systems, Man and Cybernetics, Part B, vol. 34, pp. 1031-1044, 2004.
[17] C-C. Wong, C-C. Chen, "A GA-based method for constructing fuzzy systems directly from numerical data," IEEE Trans. on Systems, Man and Cybernetics, Part B, vol. 30, pp. 904-911, 2000.
[18] J. Casillas, O. Cordón, F. Herrera, L. Magdalena, "Interpretability improvements to find the balance interpretability-accuracy in fuzzy modeling: an overview. Chapter of Interpretability Issues in Fuzzy Modeling. Springer, pp. 3-22, 2003.
[19] J. S. Wang, G. C. S. Lee, "Self-adaptive neuro-fuzzy inference system for classification application," IEEE Trans. Fuzzy System, vol. 10, pp. 790–802, 2002.
[20] T. P. Wu, S. M. Chen, "A new method for constructing membership functions and fuzzy rules from training examples," IEEE Trans System Man Cybernetic Part B, vol. 29, pp. 25–40,1999.
[21] Y. Shi, R. Eberhart, Y. Chen, "Implementation of evolutionary fuzzy system," IEEE Trans. Fuzzy Systems, vol. 7, pp. 109–119, 1999.
[22] S. Tong, Y. Shen, Z. Liu, "Approach to construct fuzzy classification system with clustering," Control and Decision, vol. 16, pp. 737-740, 2001 (in Chinese).

[23] M. Russo, “Genetic fuzzy learning,” IEEE Trans. Evolutionary Computation, vol. 4(3): pp. 259–273, 2000.
[24] M. Setnes, H. Roubos, “GA-fuzzy modeling and classification: complexity and performance,” IEEE Trans. Fuzzy systems, vol. 8(5): pp. 509–522, 2000.
[25] J. A. Roubos, M. Setnes, “Learning fuzzy classification rules from labeled data. Information Science, vol. 150, pp. 77–93, 2003.

Chen Ji-lin received the M.E. degrees from Nanjing University of Science and Technology. He is an assistant professor in School of Mechanical Engineering, Nanjing University of Science and Technology. His research interest includes intelligent system, servo control system..

Xing Zong-yi received the Dr. Degree from China Academy of Railway Science in 2003. He is an assistant professor in School of Mechanical Engineering, Nanjing University of Science and Technology. His research interest includes intelligent system, electro-hydraulic control system, industry process control, and fuzzy modeling.