ISBN 978-952-5726-00-8 (Print), 978-952-5726-01-5 (CD-ROM)
Proceedings of the 2009 International Symposium on Web Information Systems and Applications (WISA'09)
Nanchang, P. R. China, May 22-24, 2009, pp. 064-067

Using Niche Genetic Algorithm to Find Fuzzy Rules

Yan Li
College of Computer and Communication Engineering,Weifang University.Weifang, China
liyan@wfu.edu.cn

***Abstract*—The genetic algorithm has been widely used in many fields as an easy robust global search and optimization method. This paper introduces a new approach to find fuzzy rules using niche genetic algorithms with elitist migrating operator. It uses binary coding scheme. In order to make search results stable, elitist migrating operator reserves good individuals into an elitist sub-population. The simulation result shows that this approach can find fuzzy rules with less compute complexity.**

***Index Terms*—genetic algorithms; niche; fuzzy rule**

I. INTRODUCTION

Larger and larger amounts of data are collected and stored in databases, increasing the need for effective analysis methods to make use of the information contained implicitly in the data. Data mining is the exploration and analysisof data in order to discover meaningful patterns. Fuzzy rule based systems have been shown to be an important tool for modeling complex systems. Although simple atatistical techniques for data analysis have been used for many years, advanced techniques for intelligent data analysis are not yet mature. As a result, there is a growing gap between data generation and data understanding. Therefore it is very important to find a fast,effective and intelligent method for data analysis. In the last few years many different approaches have been presented taking the Genetic Algorithms (GAs) as a base of the learning process[1]. The aim of this paper is to propose an effective method that can find a compact set of fuzzy if–then rules for classification problems by using niche genetic algorithms (NGA).

Genetic algorithm (GA)is a stochastic and parallel search technique based on the mechanics of natural selection， genetics and evolution， which was first developed by Holland in 1970s[2]. GAs are known to be a powerful tool for performing search in complex spaces. In recent years，GA has been widely applied to different areas such as fuzzy systems ， neural networks,etc． Anyway, one drawback they present is that when dealing with multi-modal functions with peaks of unequal value, simple GAs are characterized by converging to the best peak of the space (or to a space zone containing several of the best peaks) and to lose an adequate individual sampling over other peaks in other space zones. This phenomenon is called genetic drift [3] and is not a correct behavior for several kinds of problems in which one may be interested in knowing the location of other function optima. The niche and species concepts were introduced in order to overcome this behavior [4,5]. In nature, a niche is viewed as an organism's task in the environment and a species is a collection of individuals with similar features. In this way, the formation of stable subpopulations of organisms surrounding separate niches by forcing similar individuals to share the available resources is induced. The niche techniques aim at gathering the individuals on several peaks of fitness function in the population according to genetic likeness and then permit GAs to investigate those peaks in parallel. The fittest individual in the niche is kept unchanged or high fitness value, while the others in the niche are changed to reduced their fitness values sharply. So the individuals in the population may be dispersed into the whole search space. Thus some diversity can be maintained effectively during the generations in the population. This paper presents an approach taking niche genetic algorithms (NGA) to find fuzzy rules. Each rule is coded as a chromosome. A rule is evaluated according to its own characteristic and its effect on the fuzzy rule set. Niche Genetic Algorithm with elitist migrating operator

In NGAs, the analogy with nature is straightforward, as in an ecosystem there are different subsystems (niches) that contain many diverse species. The number of individuals in a niche is determined by its resources and by the efficiency of each individual in taking profit of these resources[6]. Using this analogy, it is possible to maintain the population diversity in a GA. Each peak of the multi-modal function can be seen as a niche that supports a number of individuals directly proportional to its ''fertility'', which is measured by the fitness of this peak relatively to the fitness of the other peaks of the domain. The difficulty in implementing niching methods lies in the fact that the peaks are obviously not known beforehand. This complicates the process of populating each niche correctly according to its fitness[7,8]. Niche Genetic Algorithms as introduced in are often used to tackle static optimization problem of the type

$$\left\{ x \middle| f(x)\ \text{is local maximum}, x \in IX^l = \{0,1\}^l \right\} \qquad (1)$$

Where $f(x)$ is a multi-modal function, assume that $f(x) > 0$ and $f(x) \neq const$ .A population consisting of some n-tuple binary strings x_i of length l is established when resolving problem (1). Each individual x_i is a candidate solution of problem (1), the objective function value $f(x_i)$ is said to be fitness of individual x_i . A distance function d_{ij} which denoted the distance between

AP-PROC-CS-09CN001

phenotypes (decoded values) of x_i and of x_j. The distance function d_{ij} is defined as below:

$$d_{ij} = \sqrt{1 - \exp(-\beta \|x_i - x_j\|^2)} \qquad (2)$$

Where $\beta = 2/(\|x_i - \overline{x_i}\| + \|x_j - \overline{x_j}\|)$, $\overline{x_i}$ and $\overline{x_j}$ are median vectors of each gene's range in the individuals x_i and x_j. The operators of NGA are described as below:

- Selection (Proportional selection based on sharing scheme): The sharing scheme proposed by Goldberg [3]was based on an idea that the fitness of a design in a niche should be degraded due to the presence of other designs in the same niche. In this scheme, two designs are considered to be located inside the same niche if the distance between two designs is smaller than a predeter mined sharing radius. Sharing degree of individual x_i is defined as $m_i = \sum_{l=1}^{n} sh(x_i, x_j)$, where $sh(x_i, x_j)$ is sharing function of individual x_i, x_j, and sharing function is defined as below

$$sh(x_i, x_j) = \begin{cases} 1 - \left(\dfrac{d_{ij}}{\sigma_{share}}\right)^{\alpha}, & d_{ij} < \sigma_{share} \\ 0, & \text{otherwise} \end{cases} \qquad (3)$$

where σ_{share} and α are constants. σ_{share} is mating radius. α is a parameter to control the shape of the share function. d_{ij} is the distance between the x_i and the x_j. The sharing fitness function of individual x_i is defined as :

$$f'(x_i) = \frac{f(x_i)}{\sum_{j=1}^{N} sh_{ij}} \qquad (4)$$

the probability of individual x_i selected into next generation is

$$P(x_i \text{ is selected}) = \frac{f'(x_i)}{\sum_{j=1}^{N} f'(x_j)} \qquad (5)$$

- Crossover: crossover among individuals in remote niches often causes disruptions on the convergence to the optimum in each niches. Deb and Goldberg[9]devised the mating restriction scheme that aims to prevent individuals in different niche from crossing each other. In this paper, only two individuals that the distance between them is less than mating radius should be allowed to be mating parents. Then choose a crossover point randomly. Alternate the part behind crossover point according to the crossover probability p_c.

- Mutation (random mutation): Choose an individual randomly and rock over every binary bit (1 to 0, 0 to 1) according to a small probability p_m.

Elitist migrating operator is used to maintain the elitist individuals of population. In every generation, the population is checked if there are elitist individuals. The elitist individuals will be migrated into an elitist sub-population[10]. The running steps of Elitist migrating operator are summary as follows:

- Step1: Sort the individuals according to their sharing fitness;
- Step2: Find the individual which has the biggest sharing fitness, if it can migrate go on else go to step 5.
- Step3: Migrate this best individual into elitist sub-population, repeat gradient operation till the individual arrives at peak. Then remove the duplicate individuals out of sub-population.
- Step4: Remove the individual which has migrated into elitist sub-population and individuals near to it out of original population. Then go step 2；
- Step5: Fill the population with new individuals produced randomly.

II. CHROMOSOME CODE

A fuzzy rule consists of two parts: antecedent part and consequent part. Commonly, a fuzzy rule is showed as below:

$$if\ x_1\ is\ A_1^1 \text{ and } x_2\ is\ A_2^1 \ldots \text{ and } x_M\ is\ A_M^1 \quad \text{then y } is\ B^1$$

Let Φ denotes a sample data set. We use binary code scheme input variable x_i is divided into L_i fuzzy numbers within its domain. Output variable y is divided into L_y fuzzy numbers within its domain. Each rule is coded as a chromosome and the whole population is the rule set. Let k be a minimum integer which satisfies $\max(L_1, L_2, ..., L_M, L_y) + 1 < 2^k$. Then each variable can be denoted as a binary string of length k, where the string with 0 at every position denotes the variable does not affect in the rule. But the output can not be a string with 0 at every position, because the rule's output can not be null. So a rule can be coded as a binary string of length $(M+1)k$.

III. EVALUATION FUNCTION

We evaluate a rule through two aspects. One is the rule's own characteristic. The other is how the rule can affect the rule set. At first, we consider the rule's own characteristic. It contains two parts: one is adaptation degree between rule and sample datum, the other is confidence degree of rule.

Let $s = (s_1, s_2, ..., s_M, O(s))$ be a sample data. $O(s)$ denotes value of output variable y. e_i denotes value of input variable x_i. Let $\tilde{R}$ be a rule set and R be a rule in the rule set $\tilde{R}$. Let $A(R)$ denote antecedent of rule R and $O(R)$ be consequent of rule R. The adaptation degree between a rule R^i and sample datum is denoted as $U(R^i)$ and

$$U(R^i) = \sum_{e \in \Phi} u_{A_1^i}(s_1) \cdot u_{A_2^i}(s_2) \cdot ... \cdot u_{A_M^i}(s_M) \cdot u_{B^i}(O(s)) \quad (6)$$

We can say a sample data satisfies a rule and denoted as $e \in R$ when it satisfies : ① $(s_1, s_2, ..., s_M) \subseteq A(R)$, ② $O(s) \subseteq O(R)$. Let $\Phi_{A(R)}$ denote the sample data set which satisfies condition ①, and $\Phi_{A(R)O(R))}$ denote the sample data set which satisfies condition ①and condition ②. Let $\Phi_{A(R)\overline{O(R)}}$ denote the sample data set which satisfies condition ①but does not satisfy condition ② , and $|\Phi_x|$ denote the dimension of a sample data set Φ_x. Let $con(R)$ denote the confidence degree of a fuzzy rule and

$$con(R) = \begin{cases} \dfrac{|\Phi_{A(R)O(R)}|}{|\Phi_{A(R)}|}, & if\, |\Phi_{A(R)}| > 0 \\ 0, & else \end{cases} \quad (7)$$

The completeness degree of a rule set is denoted as $D(\tilde{R})$, and $D(\tilde{R}) = \dfrac{|\Phi_{\tilde{R}}|}{|\Phi|}$

The effect of a rule R on the completeness degree of rule set $\tilde{R}$ is denoted as $CD(R)$, and

$$CD(R) = \frac{D(\tilde{R}) - D(\tilde{R} - \{R\})}{D(\tilde{R})} \quad (8)$$

Where $D(\tilde{R} - \{R\})$ denotes the completeness degree of rule set $\tilde{R}$ without rule R.

According to (6)(7)(8), we can define the evaluation function as :

$$fitness(R_i) = \alpha_1 U(R_i) + \alpha_2 con(R_i) + \alpha_3 CD(R_i) + \alpha_0 \quad (9)$$

Where $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ are five positive real numbers.

IV. SIMULATION EXPERIMENT

We have studied the efficiency of the proposed algorithms. We use the meterology decision table (table I)[11]to verfify the validity of the algorithm. The NGA uses one-point crossover, 01 random mutation and proportion selection. The parameter of NGA is showed in table II.

TABLE I. DECISION TABLE OF WEATHER

U	Outlook	Temperature	Humidity	Windy	Class
1	Sunny	Hot	High	False	N
2	Sunny	Hot	High	True	N
3	Overcast	Hot	High	False	P
4	Rain	Mild	High	False	P
5	Rain	Cool	Normal	False	P
6	Rain	Cool	Normal	True	N
7	Overcast	Cool	Normal	True	P
8	Sunny	Mild	High	False	N
9	Sunny	Cool	Normal	False	P
10	Rail	Mild	Normal	False	P
11	Sunny	Mild	Normal	True	P
12	Overcast	Mild	High	True	P
13	Overcast	Hot	Normal	False	P
14	Rail	Mild	High	True	N

TABLE II. PARAMETER OF NGA

Parameter	Value
Population scale	150
Crossover proportion	0.8
Mutation proportion	0.02
Mate radius	3
Migrating threshold	0.8
Moving step	1
Max generation	300

There are 9 rules found after running over.

- Rule 1: IF Outlook=Overcast THEN P
- Rule 2: IF Outlook=Sunny AND Humidity=Normal THEN P
- Rule 3: IF Outlook=Sunny AND Humidity=High THEN N
- Rule 4: IF Outlook=Rain AND Windy=False THEN P
- Rule 5: IF Outlook=Rain AND Windy=False THEN N
- Rule 6: IF Outlook= Sunny AND Temperature=Hot THEN N
- Rule7: IF Outlook= Sunny AND Humidity=High THEN N
- Rule8: IF Outlook= Sunny AND Temperature=Mild AND Windy=False THEN N

- Rule9: IF Outlook= Rail AND Temperature=Mild AND Windy=False THEN P

V. CONCLUSION

As we known, many genetic operators relative to some special problems have been present. Choosing effective operators and deciding adequate population size for each operator plays an important role in avoiding the premature convergence of Gas[12]. In this paper , we used niche genetic algorithms with elitist migrating operator to find fuzzy rule.A chromosome consists only one rule. From the simulation result, we can conclude that this approach can find fuzzy rules with less compute complexity and more accuracy.

REFERENCES

[1] R. Chelouah and P. Siarry, Genetic and Nelder–Mead algorithms hybridized for a more accurate global optimization of continuous multi minima functions. European Journal of Operational Research 148 (2003), pp. 335–348

[2] Holland J H. Adaptation in natural and artificial system: An Introduction Analysis with Applications to Biology, Control , and Artificial Intelligence . USA; The University of Michigan Press , 1975.

[3] D.E.Goldberg, Genetic Algorithm in Search, Optimization and Machine Learning. , Addison-Wesley, New York (1989)

[4] G. Berthiau, P. Siarry, A genetic algorithm for globally minimizing functions of several continuous variables, in: Second International Conference on Meta-heuristics, Sophia-Antipolis, France, July 1997.

[5] Z. Michalewicz, Genetic algorithms + Data Structures=Evolution Programs. , Springer-Verlag, Heidelberg (1996).

[6] D. Beasly, D.R. Bull, and P.R. Martin. A sequential niche technique for multi-modal function optimization. Evolutionary Computation, vol 1, 1993, pp.101-125

[7] P. Siarry, A. Pétrowski and M. Bessaou, A multipopulation genetic algorithm aimed at multimodal optimization. Advances in Engineering Software 33 4 (2002), pp. 207–213..

[8] S.W. Mahfoud, Niching Methods for Genetic Algorithms, Ph.D. dissertation, Univ. of Illinois, Urbana-Champaign, 1995.

[9] Deb K, Goklberg DE. “An investigation of niche and species formation in genetic function optimization”. Proceedings of the Third International Conference on Genetic Algorithm;1989,pp.42-50.

[10] A. Pétrowski, A clearing procedure as a niching method for genetic algorithms, in: Proceedings of the 1996 IEEE International Conference on Evolutionary Computation, Nagoya, Japan, 1996, pp. 798–803.

[11] Liying Zheng, Youchang Li and Liyan Liu, “The Research of Data Mining Approach Based on Rough Set Theory” ,Ingormation, Management and Algorithms. Vol 2, 2007,pp.97-101.

[12] A. Pétrowski, A clearing procedure as a niching method for genetic algorithms, in: Proceedings of the 1996 IEEE International Conference on Evolutionary Computation, Nagoya, Japan, 1996, pp. 798–803.