

# Rule-based fuzzy classifier based on quantum ant optimization algorithm

Jue Wu<sup>a,b</sup>, Lei Yang<sup>b,\*</sup>, Tianrui Li<sup>a</sup>, Changjiang Zhang<sup>c</sup> and Zhihui Li<sup>d</sup>

<sup>a</sup>*School of Information Science and Technology, Southwest Jiaotong University, Chengdu, China*

<sup>b</sup>*School of Computer Science and Technology, Southwest University of Science and Technology, Mianyang, China*

<sup>c</sup>*Department of Computer Science, Wenzhou-Kean University, Wenzhou, China*

<sup>d</sup>*Hypervelocity Aerodynamics Institute of China, Aerodynamics Research and Development Center Mianyang, China*

**Abstract.** Fuzzy rule-based classification systems have been used extensively in data mining. This paper proposes a fuzzy rule-based classification algorithm based on a quantum ant optimization algorithm. A method of generating the hierarchical rules with different granularity hybridization is used to generate the initial rule set. This method can obtain an original rule set with a smaller number of rules. The modified quantum ant optimization algorithm is used to generate the optimal individual. Compared to other similar algorithms, the algorithm proposed in this paper demonstrates higher classification accuracy and a higher convergence rate. The algorithm is proved to be convergent on theory. Some experiments have been conducted on the algorithm, and the results proved that the algorithm is feasible.

**Keywords:** Rule-based, fuzzy classification, quantum ant optimization

## 1. Introduction

During the past several decades, there has been much reported research on fuzzy control systems to control widely diverse systems, such as car automation systems, classification systems of data mining, cart centering problems, truck backing problems, etc. Previous research particularly focuses on the application of fuzzy systems to these various control problems; fuzzy if-then rules are typically employed, and are usually derived from human experts. Recent research has addressed the automatic generation of fuzzy if-then rules from measured data in real-world situations, without human experts. These self-learning methods have been proposed for adjusting membership functions of fuzzy sets in fuzzy rules, using methods such as gradient descent, neural networks, and genetic algorithms.

The fuzzy rule-based classification system (FRBCS) is considered to be a very useful tool in computational intelligence, since it provides a very interpretable model for end users. The FRBCSs consist of two main components: the Inference System and the Knowledge Base (KB). The KB is composed of the Rule Base (RB) constituted by the collection of fuzzy rules, and the Data Base (DB), which contains the membership functions of the fuzzy partitions associated with the linguistic variables. The composition of the KB depends directly on the problem to be solved. If there is no expert information regarding the problem to be solved, an automatic learning process must be employed to derive the KB from examples.

Genetic algorithms have been widely proposed to generate fuzzy if-then rules and tune the membership function of fuzzy sets in fuzzy rules [1–10]. J.R. Kim and D.U. Jeong proposed a genetic algorithm-based approach to a fuzzy rule-based system for fuzzy classification with a minimum of fuzzy rules [3]. Z.Y. Xing, Y.L. Hou, et al. proposed a novel method based on a

---

\*Corresponding author. Lei Yang, School of Computer Science and Technology, Southwest University of Science and Technology, Mianyang, China. Tel./Fax: +86 13808113631; E-mail: fieldyang@163.com.

multi-objective genetic algorithm to construct a fuzzy classification system [10].

This paper proposes a fuzzy rule-based classification algorithm based on quantum ant optimization algorithm is proposed (FRBCA-QANO). The method of generating the hierarchical rules with various granularity hybridizations is used to generate the initial rule set. The modified quantum ant optimization is used to solve the problem of rule selection. The algorithm is proved to be convergent on theory, and experimental result confirms that the algorithm is convergent and feasible.

## 2. Algorithm model

### 2.1. Fussy classification

Considering an object which can be described as a point in a characteristic space ( $x \in S^N$ ), the classification assigns the point a class out of  $M$  possible classes  $\{C_1, C_2, \dots, C_M\}$ . The key problem of the classification is to find an optimal map  $D: S^N \rightarrow C$ .

The fuzzy rules are used to describe the knowledge in the fuzzy classification system, and fuzzy inference methods are used to construct a knowledge base. The fuzzy rule set is included in the rule base. The antecedents of the fuzzy rules define local fuzzy regions, while the consequents describe the classification labels within those regions. Supposed that there is a data set with  $k$  samples,  $T = \{(x^1, c^1), \dots, (x^k, c^k)\}$ , where  $x^k = \{x_1^k, \dots, x_N^k\}$  represents a sample of an attribute space  $\{X_1, \dots, X_N\}$ ,  $C_i \in \{C_1, \dots, C_M\}$  represents the class of the sample, and  $x_N^k$  represents the  $N$ th attribute of the training samples. The consequent of the rule is the class and the accuracy of the classification. Therefore, rule  $i$  can be described as follows.

$R_i: \text{IF } X_1 \text{ is } A_{i1} \text{ AND } \dots \text{ } X_N \text{ is } A_{iN} \text{ then } Y \text{ is } C_i$  with  $r_i$ , where  $r_i$  represents the accuracy of an example which belongs to to class  $c_i$ ; and  $A_{ji}$  is the membership function defined on the domain of the attribute, which can be triangular, Gaussian-shaped, a trapezoidal function, etc. The support and confidence of the association rules are used to choose the rules. The fuzzy rule is considered to be the associate rule, which is described as follows.

$$A_q \Rightarrow C_q \quad (1)$$

where  $A_q$  is the antecedent of the rule, and  $C_q$  is the consequent of the rule. These two conceptions are extended to the fuzzy association rule. The extended forms proposed by Isibuchi and Yamamoto are used in this paper.

The confidence of  $A_q \Rightarrow C_q$  can be defined as follows.

$$c(A_q \Rightarrow C_q) = \frac{|D(A_q) \cap D(C_q)|}{|D(A_q)|} \quad (2)$$

where  $D$  is a training set with  $m$  samples;  $x^p = (x_1^p, x_2^p, \dots, x_N^p)$ ,  $p = 1, \dots, m$ ;  $|D(A_q)|$  represents the number of the training samples which match  $A_q$ ; and  $|D(A_q) \cap D(C_q)|$  represents the number of the training samples which match both  $A_q$  and  $C_q$ . In the standard associate rule, the antecedent and the consequence of the rule are not fuzzy, and  $|D(A_q)|$  and  $|D(A_q) \cap D(C_q)|$  only calculate the number of the matched samples. The training model of the antecedent  $A_q$  has a different compatibility degree  $\mu_{A_q}(x^p)$ , so the rule  $A_q \Rightarrow C_q$  is a fuzzy associate rule. The rule confidence can be redefined as follows.

$$c(A_q \Rightarrow C_q) = \frac{|D(A_q) \cap D(C_q)|}{|D(A_q)|} = \frac{\sum_{p \in \text{class } C_q} \mu_{A_q}(x^p)}{\sum_{p=1}^m \mu_{A_q}(x^p)} \quad (3)$$

where  $\mu_{A_q}(x^p)$  can be a product or minimum; the product is used in this paper.

$$\mu_{A_q}(x^p) = \mu_{A_{1q}}(x_1^p) \times \mu_{A_{12q}}(x_2^p) \times \dots \times \mu_{A_{Nq}}(x_N^p) \quad (4)$$

where  $\mu_{A_{iq}}(x_i^p)$  is the membership degree of the antecedent of rule  $A_{iq}$ .

The support degree of the rule  $A_q \Rightarrow C_q$  can be described as follows.

$$s(A_q \Rightarrow C_q) = \frac{|D(A_q) \cap D(C_q)|}{|D|} \quad (5)$$

where  $|D|$  represents the number of samples in the training set. In the same way, the support degree can be redefined as follows.

$$s(A_q \Rightarrow C_q) = \frac{|D(A_q) \cap D(C_q)|}{|D|} = \frac{\sum_{p \in \text{class } C_q} \mu_{A_q}(x^p)}{m} \quad (6)$$

The antecedent is the fuzzy rule of  $A_q$ , therefore the consequence  $C_q$  can be calculated as follows.

$$c(A_q \Rightarrow C_q) = \max\{c(A \Rightarrow \text{Class1}), \dots, c(A \Rightarrow \text{ClassM})\} \quad (7)$$

The consequence of the rule is the class which has the maximum confidence degree out of  $M$  classes. When the support degree is used instead of the confidence

degree, the result identical, because they demonstrate the following relationship.

$$s(A_q \Rightarrow Classh) = c(A_q \Rightarrow Classh) \times \frac{|D(A_q)|}{|D|} \quad (8)$$

$\frac{|D(A_q)|}{|D|}$  is independent from the class of the consequence. The class  $C_q$  of every  $A_q$  can be determined in a common situation, but when more than one classe has the same confidence degree, the consequence of the rule (class  $C_q$ ) cannot be determined, and thus rule  $R_q$  can be generated.

### 2.2. Encoding method

A quantum bit as an information storage unit represents a two-state quantum system. It is a unit vector defined in a two-dimensional complex vector space. The space is composed of standard orthogonal basis  $\{|0\rangle, |1\rangle\}$ . Therefore, it can be in the quantum superposition simultaneously. The state can be represented as follows [13, 14].

$$|\varphi\rangle = \alpha |0\rangle + \beta |1\rangle \quad (9)$$

where  $\alpha$  and  $\beta$  are complex numbers that specify the probability amplitudes of the corresponding states.  $|\alpha|^2$  is the probability that the Q-bit will be found in the “0” state and  $|\beta|^2$  is the probability that the Q-bit will be found in the “1” state. Normalization of the state to unity guarantees Equation (10).

$$|\alpha|^2 + |\beta|^2 = 1 \quad (10)$$

If there is a system of  $m$  Q-bits, the system can simultaneously represent  $2^m$  states. However, in the act of observing a quantum state, it collapses into a single state.

In this paper, there are two parts in the chromosome ( $C=C_1+C_2$ ). There are  $m$  bits in the  $C_1$ , and if the bit is 1, the corresponding rule is selected; otherwise is the rule is not used. There are  $m$  real numbers in the  $C_2$  portion; the real number represents the weight of the related rule. The real number in the interval  $[0, 1]$  can be directly encoded with the quantum bit of probability amplitude. The encoding can be described as follows.

$$q_i = \left[ \begin{array}{c|c|c|c|c|c} \alpha_{i1} & \alpha_{i2} & \dots & \alpha_{im} & \alpha_{im+1} & \alpha_{i2m} \\ \beta_{i1} & \beta_{i2} & \dots & \beta_{im} & \beta_{im+1} & \beta_{i2m} \end{array} \right] \quad (11)$$

The position of an ant represents a solution of the problem in the continuous ant colony optimization

algorithm. Suppose that there are  $n$  ants, which are distributed randomly in the  $2m$ -dimensional space  $[-1, 1]^T$ . Each ant has  $2m$  quantum bits, and the probability amplitude represents the current position of the ant.

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 training patterns, which are classified correctly and denoted as  $f_1$ . Interpretability is the number of fuzzy rules, denoted as  $f_2$ . The object function is described as follows.

$$F_{fitness} = f_1\omega_1 - f_2\omega_2 \quad (12)$$

where  $\omega_1$  and  $\omega_2$  are the weighs of  $f_1$  and  $f_2$ , respectively.

### 2.3. Select ants moving to target location

Set  $\tau(x_r)$  as the pheromone intensity of the  $k$  ant, which is at the position of  $x_k$ . It is initialized as a constant. Set  $\eta(x_r)$  as the visibility of the position of  $x_k$ . The probability of the ant  $k$  moving from  $x_r$  to  $x_s$  is described as follows.

$$p_{xs} = \frac{[\tau(x_s)]^\alpha [\eta(x_s)]^\beta}{\sum_{x_s, x_u \in X} [\tau(x_u)]^\alpha [\eta(x_u)]^\beta} \quad (13)$$

### 2.4. Quantum rotation gate

The rotation operation updates the quantum bit (Q-bit) by using rotational gate. This operation makes the ant population develop into the best individual. The Q-gate is defined by Equation (14).

$$U(\Delta\theta) = \begin{bmatrix} \cos(\Delta\theta) & -\sin(\Delta\theta) \\ \sin(\Delta\theta) & \cos(\Delta\theta) \end{bmatrix} \quad (14)$$

The Q-bit is updated as follows.

$$\begin{bmatrix} \cos(\Delta\theta) & -\sin(\Delta\theta) \\ \sin(\Delta\theta) & \cos(\Delta\theta) \end{bmatrix} \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix} = \begin{bmatrix} \cos(t + \Delta\theta) \\ \sin(t + \Delta\theta) \end{bmatrix} \quad (15)$$

Equation (15) indicates that this updated operation only changes the phase of the Q-bit, but does not alter the length of the Q-bit.

where  $\Delta\theta$  is the rotation angle of each Q-bit. The magnitude of  $\Delta\theta$  has an effect on the speed of convergence; if it is too large, the solution may diverge or converge prematurely to a local optimum. The sign of  $\Delta\theta$  determines the direction of convergence, and  $\Delta\theta$  can be determined

according to the following method.  $\alpha_0\beta_0$  is the probability amplitude of the global optimal solution in the current search, and  $\alpha_1\beta_1$  is the probability amplitude of the Q-bit in the current solution. Define  $A$  as follows.

$$A = \begin{vmatrix} \alpha_0 & \alpha_1 \\ \beta_0 & \beta_1 \end{vmatrix} \quad (16)$$

The direction of  $\Delta\theta$  can be determined as follows. If  $A \neq 0$ , the direction is  $-\text{sgn}(A)$ ; if  $A = 0$ , the direction can be selected randomly.

In order to avoid premature convergence, the size of  $\Delta\theta$  can be determined according to Equation (10), as described. This method is a dynamic adjustment strategy and is unrelated to the problem;  $pi$  is the circumferential ratio, and  $\max\text{Gen}$  represents the maximal number of iterations.

$$\Delta\theta = 0.5 * pi * \exp(-gen / \max Gen) \quad (17)$$

### 2.5. Update rules of pheromone intensity and visibility

The function of the optimal problem is responsible for evaluating the ant position, and the primary goal of pheromone intensity updating is to integrate function value with the pheromone. Better function values result in higher pheromone intensities. The gradient information of the function is also integrated with the visibility, indicating that a position with greater gradient also has greater visibility. The point determined according to this method not only has higher fitness, but also a higher rate of change.

Each ant computes the function value after one step research, and updates the local pheromone intensity and the visibility according to the rules given as follows. Suppose that the previous position of the ant is  $x_q$ , the current position is  $x_r$ , and the next position is  $x_s$ . The updating rule is described as follows.

$$\tau(x_s) = \tau(x_r) + \text{sgn}(f) * |f|^\alpha \quad (18)$$

$$f = f(x_s) - f(x_r) \quad (19)$$

$$\eta(x_s) = \eta(x_r) + \text{sgn}(\partial f) * |\partial f|^\beta \quad (20)$$

$$\partial f = \max_{1 \leq i \leq n} \left( \frac{\partial f}{\partial x_{si}} \right) - \max_{1 \leq i \leq n} \left( \frac{\partial f}{\partial x_{ri}} \right) \quad (21)$$

If the function  $f$  is a non-differentiable function, the first difference can be used. It is described as follow.

$$\begin{aligned} \partial f = & \max_{1 \leq i \leq n} \left( \frac{f(x_s) - f(x_r)}{x_{si} - x_{ri}} \right) \\ & - \max_{1 \leq i \leq n} \left( \frac{f(x_r) - f(x_q)}{x_{ri} - x_{qi}} \right) \end{aligned} \quad (22)$$

where  $0 < \alpha < 1$ ,  $0 < \beta < 1$ . After all ants have completed a cycle, the global pheromone intensity is updated according to Equation (23). The  $(1 - \rho)$  is the volatility coefficient of the pheromone where  $0 < \rho < 1$ , and  $\tilde{x}$  is the current optimal solution.

$$\tau(x_u) = \begin{cases} (1 - \rho)\tau(x_u) + \rho \text{fit}(x_u) & x = \tilde{x} \\ (1 - \rho)\tau(x_u) & x \neq \tilde{x} \end{cases} \quad (23)$$

### 3. Convergence analysis

**Theorem 1:** The FRBCA-QANO algorithm ant population sequence  $\{Q(t), t > 0\}$  is the finite homogeneous Markov chain.

**Proof:** The Q-bits are used in the algorithm. In the ant colony evolution algorithm, the value of the ants is discrete. Assume that the length of the ant is 2 m and the ant population size is  $n$ , therefore the state space size is  $2^{\text{mn}}$ . Due to the continuity of the variable value, the size of the state space size is infinite in theory, but the accuracy is finite during calculation. Assuming that the dimension is  $V$ , then the state size of the population is  $V^{2\text{mn}}$ , and so the population is finite. In the algorithm, the rotational operation is entirely independent of the generation number, the pheromone intensity is updated according to Equation (18), and the visibility is updated according to Equation (20). All of these are independent of the generation number. Therefore the population sequence is a finite homogeneous Markov chain.

**Definition 2.**  $f_k = \max\{f(x_i), i = 1, 2, \dots, N\}$  is a random variant sequence. The variable represents the best solution in the  $k$  generation. If the condition satisfies Equation (24), then the algorithm is convergent to the global optimal solution with a probability of 1.

$$\lim_{k \rightarrow \infty} P\{f_k = f^*\} = 1 (f^* = \max\{f(b) | b \in \Omega\}) \quad (24)$$

**Theorem 2.** The FRBCA-QANO algorithm is convergent to the global optimal solution with a probability of 1.

**Proof.** According to Theorem 1, the ant is a finite homogeneous Markov chain. Suppose that the ant population is  $2m$ , and the ants are points in the research space  $\Omega$ .  $X_i \in \Omega$  indicates that  $X_i$  is a point in  $\Omega$ .  $X_i = \{x_1, x_2, \dots, x_n\}$ ,  $X_k^i$  indicates that in the  $k$  iteration, the ant  $X$  is at point  $X_i$ . The ant moves to point  $j$  from point  $i$  in the search space after a step of iteration in the randomizing procedure. The transfer possibility is  $P \{X_{k+1}^j | X_k^i\}$ .

During the transfer procedure, there are two special situations.

- The first situation is that  $X_{k+1}^j \notin f^*$ ,  $X_k^i \in f^*$ . The optimal solution is preserved in the algorithm. This operation ensures that each iteration will not degenerate, thus the transfer possibility  $P \{X_{k+1}^j | X_k^i\} = 0$ .
- The second special situation is when  $X_{k+1}^j \in f^*$ ,  $X_k^i \notin f^*$ . According to the algorithm operation, a conclusion can be reached that the transfer possibility  $P \{X_{k+1}^j | X_k^i\} > 0$ .

The proof of convergence is described below, noting that it does not apply to the two above-mentioned special situations. Suppose that  $X_k$  is at position  $s_i$ , and the possibility is  $p_k^i$ .

$$p_k = \sum_{X_i \notin f^*} p_k^i \quad (25)$$

$$p_{k+1} = \sum_{X_i \in f^*, X_j \notin f^*} P \{X_{k+1}^j | X_k^i\} + \sum_{X_i \notin f^*, X_j \in f^*} P \{X_{k+1}^j | X_k^i\} \quad (26)$$

$$\begin{aligned} & \sum_{X_i \notin f^*, X_j \in f^*} P \{X_{k+1}^j | X_k^i\} \\ & + \sum_{X_i \notin f^*, X_j \notin f^*} P \{X_{k+1}^j | X_k^i\} \quad (27) \\ & = \sum_{X_i \notin f^*} P \{X_k^i\} = p_k \end{aligned}$$

According to Equations (27, 28) is obtained.

$$\begin{aligned} & \sum_{X_i \notin f^*, X_j \notin f^*} P \{X_{k+1}^j | X_k^i\} \\ & = p_k - \sum_{X_i \notin f^*, X_j \in f^*} P \{X_{k+1}^j | X_k^i\} \quad (28) \end{aligned}$$

Equation (28) is substituted into Equations (21), and (29) is obtained.

$$\begin{aligned} 0 \leq p_{k+1} & < p_k - \sum_{X_i \notin f^*, X_j \in f^*} P \{X_{k+1}^j | X_k^i\} \\ & + \sum_{X_i \in f^*, X_j \in f^*} P \{X_{k+1}^j | X_k^i\} \\ & < p_k + \sum_{X_i \in f^*, X_j \in f^*} P \{X_{k+1}^j | X_k^i\} = p_k \quad (29) \end{aligned}$$

According to the two special situation described above,  $\lim_{k \rightarrow \infty} p_k = 0$  can be obtained, and  $\lim_{k \rightarrow \infty} p(f_k = f^*) = 1 - \lim_{k \rightarrow \infty} p_k = 1$ . This indicates that the algorithm represents global convergence.

#### 4. Simulation experiment

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, Virginica). Therefore, there are four independent variables that form the antecedent of the rule, and three dependent variables that form the consequence of the rule. From each species, there are 50 observations regarding sepal length, sepal width, petal length and petal width (in cm).

In the simulation experiment, the variables are set as follows. The ant populate size  $m$  is 50, the size of quantum  $n=8$ , volatility coefficient  $(1 - \rho)=0.3$ , the pheromone update exponent  $\alpha=0.5$ , the visibility exponent  $\beta=0.5$ , the corner step  $\theta_0=0.05*pi$ , and  $\omega_1 = 1$ ,  $\omega_2 = 0.1$ . In order to avoid influence of change in the experiment, all programs were run 20 times. The final fuzzy classification system is described as follows.

R1: if  $x_3$  is A31,  $x_4$  is A41, then  $x$  belongs to Class 1, CF=0.21.

R2: if  $x_3$  is A32,  $x_4$  is A42, then  $x$  belongs to Class 2, CF=0.42.

Table 1  
Comparison of three algorithms

	Rule-based classifier	FCMF-LSGA	FRBCA-QANO
Number of rules	3	3	3
Correct classification	144	144	146
Number of Variable	2	1	1
Accuracy	96%	96%	96.7%

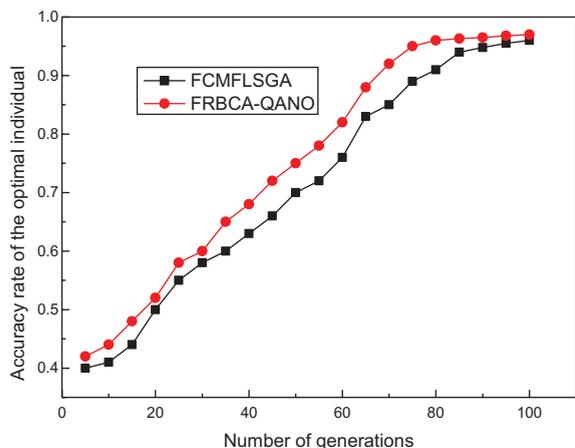


Fig. 1. The accuracy rate of the optimal rule set.

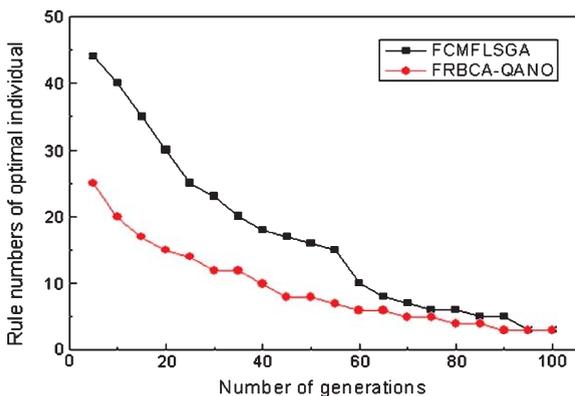


Fig. 2. Rule number of the optimal individual with different iterations.

R3: if  $x_3$  is A33,  $x_4$  is A43, then  $x$  belongs to Class 3, CF = 0.38.

To prove the validity of QACO, it is compared to the conventional classification algorithm; results are shown in Table 1. As shown in Table 1, the three algorithms can obtain the same rule number, but the FRBCA-QANO demonstrates higher accuracy and less variability.

In order to prove the efficiency of the algorithm, it is compared to a fuzzy classifier based on the Mamdani fuzzy logic system and the genetic algorithm (FCMFLSGA) [9]. The standard genetic algorithm is used in the algorithm FCMFLSGA. In the FCMFLSGA experiment, the population size is 100, and cross-probability  $P_c$  is 0.9. The mutation probability  $P_m$  is 0.1. Experimental results are described in Figs. 1 and 2. Figure 1 indicates that the FRBCA-QANO demonstrates a higher convergence rate and accuracy than FCMFLSGA. According to Fig. 2, a conclusion can be drawn that the algorithm FRBCA-QANO also demonstrates a higher convergence rate than algorithm FCMFLSGA.

### 5. Conclusion

A fuzzy rule-based classification algorithm based on the quantum ant optimization algorithm is proposed. The method of generating the hierarchical rules with various granularity hybridizations is used to generate the initial rule set. This method can obtain an original rule set with fewer rules, and enhance the convergence rate of the algorithm. The modified quantum ant colony optimization (MQACO) is used to solve the problem of rule selection from the original rule set. The MQACO ensure that the algorithm has a higher classification accuracy and higher convergence rate than similar algorithms. The algorithm is proved to be convergent on theory, and experimental results also confirm that the algorithm is convergent and feasible.

### Acknowledgments

This research was supported by the National Key Basic Research and Development Program (2014CB744100), the foundation of continuing education of Southwest University of Science (13JYF01), the doctoral fund of the Southwest University of Science (13ZX7102) and the research project of Sichuan Education Bureau (12ZB328, 12ZB327).

### References

- [1] C.G. Lim, Y.M. Jung and E.K. Kim, A genetic algorithm for generating optimal fuzzy rules, *International Journal of KIMICS* 7 (2003), 765–778.
- [2] H. Pooya and A.S.M. Taghi, Online adaptive fuzzy logic controller using genetic algorithm and neural network for networked control systems, *The Proceedings of 15th International Conference on Advanced Communications Technology (ICTACT)*, PyeongChang, Korea, 2013, pp. 88–98.

- [3] J.R. Kim and D.U. Jeong, Optimizing the fuzzy classification system through genetic algorithm, *The Proceedings of the Third International Conference on Convergence and Hybrid Information Technology*, Busan, South Korea, 2008, pp. 903–908.
- [4] M.I. Ray, M. Galende, G.I. Sainz and M.J. Fuente, Selection of rules by orthogonal transformation and genetic algorithm to improve the interpretability in fuzzy rule based systems, *The Proceedings of IEEE International Conference on Fuzzy Systems (FUZZ)*, Hyderabad, India, 2013, pp. 1–8.
- [5] P. Villar, A. Fernandez and F. Herrera, Studying the behavior of a multiobjective genetic algorithm to design fuzzy rule-based classification systems for imbalanced data-sets, *The Proceedings of the IEEE International Conference on Fuzzy Systems*, Taipei, Taiwan, 2011, pp. 1239–1246.
- [6] S.M. Chen, Y.C. Chang and J.S. Pan, Fuzzy rules interpolation for sparse fuzzy rule-based systems based on interval type-2 Gaussian fuzzy sets and genetic algorithms, *IEEE Transactions on Fuzzy Systems* **21** (2013), 412–425.
- [7] S. Morteza, A. Ali, S. Zahra and S. Pazhoeshfar, A knowledge management system based on artificial intelligence (AI) methods: A flexible fuzzy regression-analysis of variance algorithm for natural gas consumption estimation, *The Proceedings of the International Conference on Information Retrieval & Knowledge Management (CAMP)*, Malaysia, 2012, pp. 143–147.
- [8] T. Murada, *Genetic algorithm for multi-objective optimization*, Doctoral thesis, Osaka Prefecture University, 1997.
- [9] W.H. Zhou, S.Q. Xiong and T. Ma, A fuzzy classification based on mamdini logic system and genetic algorithm, *The Proceedings of the IEEE Youth Conference on Information Computing and Telecommunications (YC-ICT)*, Beijing, China, 2010, pp. 198–201.
- [10] Z.Y. Xing, Y.L. Hou, Z.Z. Tong and L.M. Jia, Intelligent systems design and applications, *The Proceeding of the Sixth International Conference on Intelligent Systems Design and Applications* **2** (2006), 1029–1034.