

A Probabilistic Modeling Clonal Selection Algorithm and Its Application to Traveling Salesman Problems

Huiqin Chen[†], Shangce Gao, Sheng Li and Zheng Tang^{††},

Graduate School of Innovative Life Science, University of Toyama, Toyama-shi, Japan

Summary

In this paper, we propose a probabilistic modeling clonal selection algorithm which combines the clonal selection algorithm (CSA) and the probabilistic modeling (PM) for traveling salesman problem (TSP). The clonal selection algorithm is employed by the natural immune system to define the basic features of an immune response to an antigenic stimulus, can initialize antibodies and maintain the population diversity. Furthermore, the PM phase attempts to reduce the computational complexity, generate new solution offspring for the CSA phase. Simulations on traveling salesman problems show that the proposed algorithm has better performance when compared with other traditional algorithms.

Key words:

Clonal selection algorithm, probabilistic modeling, receptor gene, traveling salesman problems

1. Introduction

Artificial immune system (AIS) has been drawn significant attention in the last few years. Based upon theory of the artificial immune systems, some AIS modes have been created, such as clonal selection algorithms [1], negative selection algorithms [2], and immune network algorithms [3] and so on. These models have developed adaptive systems capable of performing a wide range of tasks in various engineering applications, such as anomaly detection [4], pattern recognition [5], data mining [6], computer security [7-9], adaptive control [10-11] and fault detection [12].

In particular, the clonal selection algorithms (CSA) proposed by Burent in 1959, which tries to imitate the mechanisms in the clonal selection principle to better understand its natural presence and simulate its dynamical behavior in the presence of antigens, has received a rapid increasing interest. Recently, a lot of works have been done on the development of a variety of domains, such as pattern recognition and optimization problems. An adaptive polyclonal programming algorithm proposed by Du et al. [13], which added a clonal recombination operator in order to realize the cooperation and communication among different antibodies and thus obtain the diversity and high convergence speed. Furthermore, an adaptive dynamic clone selection algorithm which integrated the local search with global and the probability

evolution searching with the stochastic searching was presented in [14]. Besides, a new classification model which can carry out the global search and the local search in many directions rather than one direction around the same antibody simultaneously was presented in [15].

Although the clonal selection algorithm can exploit and explore the solution space parallelly and effectively, there are still a few disadvantages such as the poor convergence properties and difficulties in reaching high-quality solutions in reasonable time [16]. In CSA, an antibody repertoire is randomly generated firstly, and then the low affinity antibodies have to be replaced by new random antibodies during the mutation process. However, the genetic material between members of the repertoire has no crossover in the immune response process. As a result, the diversity of repertoire will be influenced, and may fail to get the better solutions. In order to solve this problem, a crossover process which can successfully achieve the communication during different antibodies should be constructed.

In this paper, a probabilistic modeling clonal selection algorithm (PMCSA) is proposed. The probabilistic modeling [17] as a crossover mechanism can make the optimization results better. To prove the effectiveness of the proposed algorithm, the simulation based on the traveling salesman problems is implemented.

The paper is organized as follows: in the next section 2, a general architecture of the clonal selection algorithms are explained. Section 3 introduces the proposed probabilistic modeling clonal selection algorithm in detail. Moreover, section 4 presents the PMCSA model by applying it to TSP. In additions, Section 5 gives the conclusions of the paper.

2. The clonal selection

The feature of the clonal selection theory includes sufficient diversity, discrimination of self and non-self and long-lasting immunologic memory.

2.1 Immune Cells

Lymphocytes are the most important cells in the immune system. B lymphocytes (B cells) and T lymphocytes (T cells) are the major groups of immune cells which are responsible for adaptive immune response and the immunologic attributes. B cells play a large role in the humoral immune response, and apply to make antibodies against antigens, perform the role of Antigen Presenting Cells (APCs) and eventually develop into memory B cells after activation by antigen interaction.

T cells play a central role in cell-mediated immunity. They can activate and direct B cells and recognize only antigen that is bound to cell-membrane proteins called major histocompatibility complex (MHC) molecules through a unique antigen-binding molecule.

2.2 Antibodies (Ab) and Antigens (Ag)

Antibodies which are produced by a kind of white blood cell are used by the immune system to identify and neutralize foreign objects, such as bacteria and viruses. An antigen originally defined as any molecule that binds specifically to an antibody, the term now also refers to any molecule or molecular fragment that can be bound by an MHC molecule and presented to a T-cell receptor [18]. "Self" antigens are usually tolerated by the immune system; whereas "Non-self" antigens are identified as intruders and attacked by the immune system.

2.3 The clonal selection principle

The clonal selection theory interprets the response of lymphocytes in the face of an antigenic stimulus and is shown in Fig. 1.

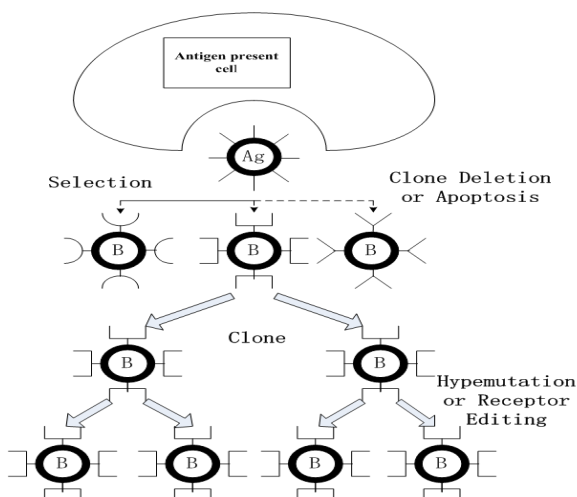


Fig. 1 The clonal selection theory.

When an antigen invades the human body, only one type of B lymphocytes can match it with a certain affinity. Once combined, the B lymphocytes will be stimulated to produce a large number of lymphocytes with the same group of antibodies, and then terminal differentiation into plasma cells. The B lymphocytes complete the proliferation by dividing themselves, and then bear receptors of identical specificity as the parental cell. The selected cells are subject to the affinity level. The individuals with lower affinity in clonal selection mechanisms experience the proliferation and mutation processes, its affinity will be increased and mature.

2.4 The clonal selection algorithm

The clonal selection algorithm inspired by the clonal selection theory of acquired immunity is one of the common algorithms of the artificial immune system.

In the clonal selection algorithm, the problem generally defined as antigen, the solution is antibody, through clonal selection, cloning, somatic hypermutation and receptor editing and other major operation to improve the antibody population. Known by the mechanism of biological immune system, the antibody variable regions rearrange the various genetic segments by somatic hypermutation and receptor editing to achieve affinity maturation. The function of somatic hypermutation mechanism is to accelerate the maturation of the immune response. The greater part of the cloned population usually becomes dysfunctional or develops into harmful anti-self cells after the mutation. However, occasionally an effective change enables the offspring cell to bind better with the antigen, and then a mutated cloned cell with higher affinity is found, it in turn will be activated to undergo proliferation. The rate of the mutation of the cloned cells is inversely proportional to the antigen affinity. In other words, those cells with low affinity receptors may be further mutated and led to eventual elimination through the apoptosis process of the immune system. According to receptor editing, the recombinant of both B and T lymphocytes that carry antigen receptors will be occurred. In this mechanism, the antibody genes and the genetic DNA gene fragments will be reorganized to form a new specific antibody.

As a result, the affinity of the offspring cell may be higher than the pro-B cells. Besides these two mechanisms (somatic hypermutation and receptor), a fraction of the least stimulated lymphocytes is replaced per generation by the newcomer cells which can recognize the pool of available antigen and thus obtain the diversity of the population.

3. A Probabilistic Modeling Clonal Selection Algorithm

In this section, we present and analyze the proposed probabilistic modeling clonal selection algorithm by the clonal operator.

3.1 A Brief Introduction to PM

Probability and statistics are the fundamental mathematical tools that allow us to model, reason and proceed with inference in uncertain environments. The Probabilistic Modeling, which has been recently drawn significant attention, is considered as a new area of evolutionary computation, signaling a paradigm shift in genetic and evolutionary computation research [17] [19]. Incorporating (automated) linkage learning techniques into a graphical probabilistic model, The PM is an adaptive model which built around superior solutions found thus far while efficiently traversing the search space [17].

3.2 The PMCS model

As shown in Fig. 2, the general approaches of the PMCS model involve initialization, affinity evaluation, clonal operator, hypermutation or receptor editing, clonal selection, Estimation, and Sampling. These above steps are iterated until a pre-specified termination criterion is satisfied.

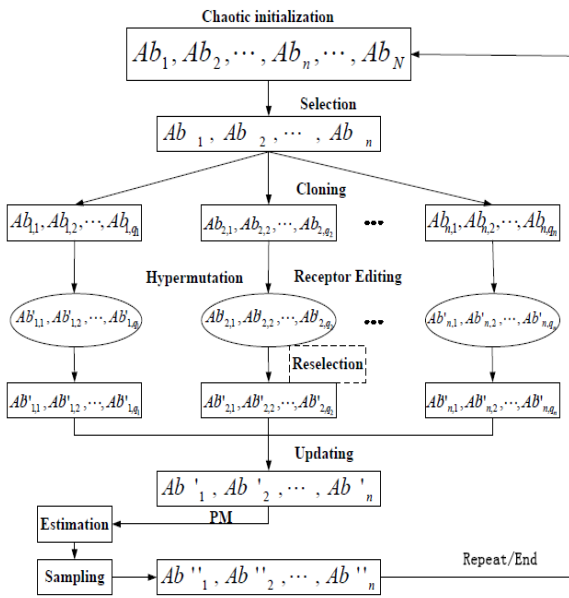


Fig. 2 The process of the proposed PMCSA.

Step 1. Initialize the population of the antibodies, then the collection $(Ab_1, Ab_2, \dots, Ab_N)$ of N antibodies is generated. For order-based combinational optimization problems such as TSP, VRP (Vehicle Routing Problem), JSP (Job-shop Scheduling Problem) and so on, $Ab_i (i=1, 2, \dots, N)$ can be expressed as a permutation of the set $[M] = \{1, 2, \dots, m\}$, i.e.:

$$Ab_i = (x_1, x_2, \dots, x_m) \quad \forall x_i \in [M] \quad (1)$$

where x_i is a receptor gene.

Step 2. According to the clonal selection principle, select a portion of the antibodies $(Ab_1, Ab_2, \dots, Ab_n)$ ($n < N$) with higher affinity to enter into the elite pools. The affinity of the antibodies is supposed as follows:

$$\alpha(Ab_1) > \alpha(Ab_2) > \dots > \alpha(Ab_N) \quad (2)$$

$$\alpha(Ab_i) = \sum_{i=1}^{m-1} d(x_i, x_{i+1}) + d(x_1, x_m) \quad (3)$$

Step 3. Clone the elites in each elite pool, where the clones are exactly identical with their parent antibody. All the antibodies in the elite pool are deleted with the clone operator Ω :

$$\Omega(Ab_i) = (Ab_{i,1}, Ab_{i,2}, \dots, Ab_{i,q_i}) \quad (i=1, 2, \dots, n) \quad (4)$$

where

$$q_i = \text{round}((n-i) \cdot Q/n) \quad (5)$$

and q_i is the clonal size of the antibody Ab_i , Q is the total clonal range, $\text{round}(\cdot)$ is the operator that rounds its argument towards the closest integer.

After the clone process, the antibody which has higher affinity will have more clones.

Step 4. Subject each clone of the antibody into the hypermutation (HM) or receptor editing (RE) process. The mutation number (C_{hm} and C_{re} for hypermutation and receptor editing, respectively) are defined as follows:

$$C_{hm} = \beta \cdot q_i \quad (6)$$

$$C_{re} = (1 - \beta) \cdot q_i \quad (7)$$

where β is a user-defined parameter which determines the complementary intensity between the hypermutation and receptor editing.

By the previous work [20], $\beta = 0.5$ will lead the CSA algorithm to a better performance. After this step, we obtain the mutated antibodies just as $(Ab'_{1,1}, Ab'_{1,2}, \dots, Ab'_{1,q_1}; \dots; Ab'_{n,1}, Ab'_{n,2}, \dots, Ab'_{n,q_n})$.

Step 5. All of the mutated antibodies enter into a reselect process where the mutated ones $Ab'_{i,j} (i=1, 2, \dots, n, j=1, 2, \dots, q_i)$ are judged to compare with their parent antibody $Ab_{i,j}$ according to the following updating rule:

$$Ab''_{i,j} = \begin{cases} Ab_{i,j} & \text{if } \alpha(Ab_{i,j}) > \alpha(Ab'_{i,j}) \\ Ab'_{i,j} & \text{if } \alpha(Ab_{i,j}) \leq \alpha(Ab'_{i,j}) \end{cases} \quad (8)$$

Then we can obtain the updated antibodies just as $(Ab''_{1,1}, Ab''_{1,2}, \dots, Ab''_{1,q_1}, \dots, Ab''_{n,1}, Ab''_{n,2}, \dots, Ab''_{n,q_n})$.

Step 6. Update the parent antibody in each elite pool with the fittest individual of clones according to Eq. (9), Eq. (10):

$$\alpha(D_i) = \max \{ \alpha(Ab''_{i,1}), \dots, \alpha(Ab''_{i,q_i}) \} \quad (9)$$

$$T = \begin{cases} 1 & \alpha(Ab_i) < \alpha(D_i) \\ 0 & \alpha(Ab_i) \geq \alpha(D_i) \\ \exp\left(\frac{\alpha(D_i) - \alpha(Ab_i)}{\mu}\right) & \text{otherwise} \end{cases} \quad (10)$$

where D_n the fittest individual of clones, D_1 is the best one during D_n , Ab_n is the parent antibody, Ab_1 is the best one, μ is a user-defined parameter which related to the diversity; the better the diversity is, the bigger μ is. Otherwise, μ is smaller.

Step 7. Estimate the probability matrix P which featuring the distribution of promising solution in the search space, given by:

$$P = \begin{pmatrix} p_{1,1} & p_{1,2} & \dots & p_{1,m} \\ p_{2,1} & p_{2,2} & \dots & p_{2,m} \\ & & \dots & \\ p_{m,1} & p_{m,2} & \dots & p_{m,m} \end{pmatrix} \quad (11)$$

Where m is the number of the receptor gene in the gene sequence $[M]$, $p_{i,j}$ ($i, j \in \{1, 2, \dots, m\}$) denotes the probability that the edge (i, j) is selected in the solution $Ab_k, \forall k \in \{1, 2, \dots, n\}$.

Step 8. Sample the updated solution distribution model to generate new solution offspring. Since some elements of the offspring are sampled from the probability vector P , it can be expected that should fall in or close to a promising area.

For the traveling salesman problem, the sample algorithm is a recursive procedure. Firstly, the departure city is selected randomly, and then the next city will be selected based on the well-known Roulette rule [21]. The rule of the sample is implemented according to the probability vector P .

Step 9. Terminate and return the best antibody which reach the maximum number of generation, otherwise, go back to Step 2.

4. Simulation results

In this section, the proposed PMCS algorithm is tested on the traveling salesman problem which is a classic combinatorial optimization problem. Given a list of cities and their pairwise distances, the task is to find a shortest possible tour that visits each city exactly once. We analyzed the characteristics of the algorithm through constructing several traditional algorithms and comparing the performances during them. The problem instances of TSP that we adopted in our test are taken from TSPLIB [22]. Then we listed these instances in Table 1 and indicated their corresponding maximum numbers of generation (K_{\max}) used in our implementation.

We compare the experimental results of the four algorithms involving KL (KNIES-TSP), KG (KNIES-TSP-Global) [23], GA (Genetic Algorithm) [21], RECS (Receptor Editing Clonal Selection) [20], and PMCS based on 9 instances of TSP which is outlined in Table 2. All the results are over 20 replications, and then make statistic comparisons to prevent the stochastic effect.

Table 1: Problem instances used in the paper and the maximum number

Instance Name	Cities' Number	Optimum	K_{\max}
eil51	51	426	1000
st70	70	675	1000
eil76	76	538	1000
rd100	100	7910	1000
eil101	101	629	1000
lin105	105	14379	1000
pr107	107	44303	1000
pr124	124	59030	2000
bier127	127	118282	2000

In Table 2, the symbols P_{Dm} and P_{Db} were defined as follows:

$$P_{Dm} = \frac{D_m - D_{op}}{D_{op}} \times 100 \quad (10)$$

$$P_{Db} = \frac{D_b - D_{op}}{D_{op}} \times 100 \quad (11)$$

where D_m is the mean distance, D_b is the best distance, D_{op} is the optimum tour length that listed in Table 1. And $T(s)$ denotes the average computation time over 20 runs.

From Table 2, the average best solutions found by PMCS are at least three times better than those of the other four algorithms, the average P_{Dm} of PMCS are at least half less than that of the others.

Furthermore, the average quality of the solutions for all tested instances during the three algorithms is outlined in Fig. 3. It is obvious that PMCS can produce much better

solutions and construct a much smoother curve than the others. As a result, the efficiency of PMCSA had been demonstrated better than the compared algorithms.

5. Conclusion

In this paper, we proposed a probabilistic modeling clonal selection algorithm for traveling salesman problems. The proposed algorithm combines the clonal selection

algorithm and the probabilistic modeling to create promising solutions. The simulation results on the traveling salesman problem show that the PMCSA can significantly improve the quality of the final solutions. CSAPM has superior performance to other traditional algorithms. Furthermore, we plan to hybridize the proposed algorithm with suitable applications to solve more optimization problems.

Table 2: Simulation results during KL, KG, GA, RECS and PMCS.

Instance	KL	KG	GA			RECS			PMCS		
	P_{Db}	P_{Db}	P_{Db}	P_{Dm}	T(s)	P_{Db}	P_{Dm}	T(s)	P_{Db}	P_{Dm}	T(s)
eil51	2.86	2.86	1.26	2.58	45.26	0.31	1.78	5.81	0.00	0.26	9.91
st70	1.51	2.33	1.54	2.35	65.23	0.86	2.03	6.92	0.30	0.97	13.62
eil76	4.98	5.48	2.15	3.21	68.95	3.26	4.58	7.21	0.56	1.99	15.29
rd100	2.09	3.38	3.14	5.27	100.52	2.53	4.25	8.56	0.71	2.13	21.59
eil101	4.65	5.63	4.24	6.21	122.84	3.39	5.08	8.25	1.27	2.85	21.89
lin105	1.98	1.29	1.06	1.98	98.48	1.25	2.01	9.38	0.62	1.63	23.58
pr107	0.73	0.42	2.38	3.59	119.57	0.98	2.36	9.45	0.48	1.84	22.42
pr124	0.07	0.49	1.96	2.74	152.65	0.67	1.86	19.58	0.08	0.74	53.83
bier127	2.76	3.08	3.51	6.51	157.29	2.59	3.74	21.35	0.56	2.18	56.72
Average	2.40	2.77	2.36	3.83	103.42	1.76	3.08	10.72	0.51	1.62	26.54

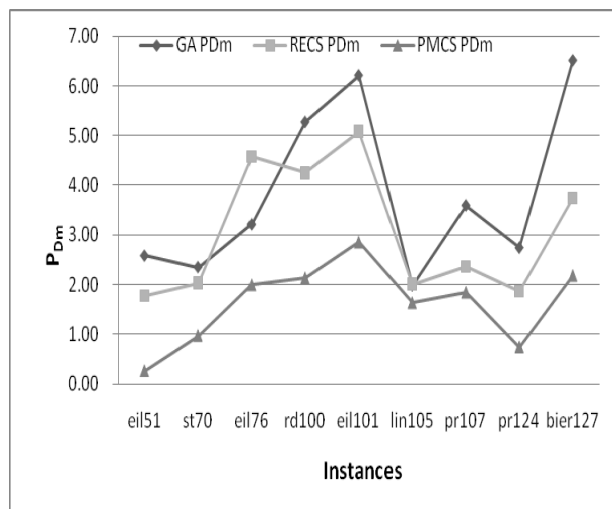


Fig. 3 Comparison of P_{Dm} during GA, RECS and PMCS

References

- [1] L. de Castro, and F. Von Zuben, "Learning and optimization using the clonal selection principle," IEEE Trans. Evol. Comput., vol.6, no.3, pp.239-251, 2002.
- [2] S. Forrest, A. Perelson, L. Allen, and R. Cherukuri, "Self-nonspecific discrimination in a computer," IEEE Symposium on Research in Security and Privacy, pp. 202-212, 1994.
- [3] J. Timmis, M. Neal, and J. Hunt, "An artificial immune system for data analysis," BioSystems 55 (1): 143-150, 2000.

- [4] D. Dasgupta and S. Forrest, "An anomaly detection algorithm inspired by the immune system," in Artificial Immune System and Their Applications, pp.262-277, Springer-Verlag, 1999.
- [5] Y. Cao and D. Dasgupta, "An immunogenetic approach in chemical spectrum recognition," in Advances in Evolutionary Computing, Springer-Verlag, 2003.
- [6] T. Knight and J. Timmis, "A multi-layered immune inspired approach to data mining," Proc. 4th International Conference on Recent Advances in Soft Computing, pp.266-271, 2002.
- [7] S. Hofmeyr and S. Forrest, "Architecture for an artificial immune system," Evolutionary Computation, vol.8, no.4, pp.443-473, 2000.
- [8] D. Dasgupta, "Immune-based intrusion detection system: A general framework," Proc.22nd National Information Systems Security Conference (NISSC), pp.147-160, 1999.
- [9] J. Kim and P.J. Bentley, "Towards an artificial immune system for network intrusion detection: An investigation of dynamic clonal selection," Proc. Congress on Evolutionary Computation, pp.1015-1020, 2002.
- [10] K. KrishnaKumar and J. Neidhoefer, "Immunized adaptive critic for an autonomous aircraft control application," in Artificial Immune System and Their Applications, pp.221-240, Springer-Verlag, 1999.
- [11] K. KrishnaKumar, "Artificial immune system approaches for aerospace application," 41st Aerospace Sciences Meeting & Exhibit, AIAA-2003-0457, 2003.
- [12] D. Bradley and A. Tyrell, "Immunotronics: Hardware fault tolerance inspired by the immune system," Proc. 3rd International Conference on Evolvable System (ICES2000), vol.1801, pp.11-20, Springer-Verlag, 2000.
- [13] H.F. Du, L.C. Jiao, and R.C. Liu, "Adaptive polyclonal programming algorithm with applications," ICCIMA, pp.350-355, 2003.

- [14] H.F. Du, L.C. Jiao, M.G. Gong, and R.C. Lin, "Adaptive Dynamic Clone Selection Algorithms," Lecture Notes in Computer Science, vol.3066, pp.768-773, 2004.
- [15] W. Ma, and R. Shang, "SAR image classification based on clonal selection algorithm," ICNC (1), pp.927-934, 2006.
- [16] Y.F. Zhong, L.P. Zhang, and P.X. Li, "Multispectral remote sensing image classification based on simulated annealing clonal selection algorithm," IGARSS (05), vol.6, pp.3745-3748, 2005.
- [17] M. Pelikan, D.E. Goldberg, and F.G. Lobo, "A survey of optimization by building and using probabilistic models," Computational Optimization and Applications, vol. 21, pp.5-20, 2002.
- [18] P. Parham, "The Immune System," 3rd Edition, Garland Science, Taylor and Francis Group, LLC, pg. G: 2, 2009.
- [19] P. Larranaga and J.A. Lozano, Estimation of Distribution Algorithms: A New Tool for Evolutionary Computation, Kluwer Academic Press, 2001.
- [20] S. Gao, H. Dai, G. Yang, and Z.Tang, "A novel clonal selection algorithm and its application to traveling salesman problems," IEICE Trans. Fundamentals, vol.E90-A, no.10, pp.2318-2325, Oct. 2007.
- [21] D.E. Goldberg, "Genetic Algorithms in Search, Optimization, and Machine Learning," Addison-Wesley, Reading, MA, 1989.
- [22] <http://www.iwr.uni-heidelberg.de/groups/comopt/software/TSPLIB95/>.
- [23] N. Aras, B.J. Oommen, and I.K. Altinel, "The Kohonen network incorporating explicit statistics and its applications to the traveling salesman problem," Neural Networks, vol.12, pp.1273-1284, 1999.



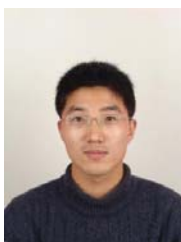
Sheng Li received the B.S. degree from Nanjing University of Science and Technology, Nanjing, China in 2006. Now, he is working toward the Ph.D. degree at Toyama University, Toyama, Japan. His main research interests are intelligent algorithms and artificial immune system.



Zheng Tang received the B.S. degree from Zhejiang University, Zhejiang, China in 1982 and an M.S. degree and a D.E. degree from Tsinghua University, Beijing, China in 1984 and 1988, respectively. From 1988 to 1989, he was an instructor in the institute of microelectronics at Tsinghua University. From 1990 to 1999, he was an associate professor in the department of electrical and electronic engineering, Miyazaki University, Miyazaki, Japan. In 2000, he joined Toyama University, Toyama, Japan, where he is currently a professor in the department of intellectual information systems



Huiqin Chen received the B.S. degree from HoHai University, Nanjing, China in 2006. Now, she is working toward the Ph.D. degree at Toyama University, Toyama, Japan. Her main research interests are intelligent algorithms and artificial immune system.



Shangce Gao received the B.S. degree from Southeast University, Nanjing, China in 2005. Now, he is working toward the Ph.D. degree at Toyama University, Toyama, Japan. His main research interests are multiple-valued logic, artificial immune system and artificial neural system.