A study of selection operator using distance information between individuals in genetic algorithm

Minoru Ito * and Masanori SUGISAKA **

* Graduate School of Engineering, Oita University, 700 Oaza Dannoharu, Oita, 870-1192, JAPAN
(Tel: +81-97-554-7831; Fax: +81-97-554-7841; E-mail: minoru@cc.oita-u.ac.jp)
** Faculty of Engineering, Oita University, 700 Oaza Dannoharu, Oita, 870-1192, JAPAN
(Tel: +81-97-554-7831; Fax: +81-97-554-7841; E-mail: msugi@cc.oita-u.ac.jp)

Abstract: In this paper, we propose a "Distance Correlation Selection operator (DCS)" as a new selection operator. For Genetic Algorithm (GA), many improvements have been proposed. The MGG (Minimal Generation Gap) model proposed by Satoh et.al. shows good performance. The MGG model has all advantages of conventional models and the ability of avoiding the premature convergence and suppressing the evolutionary stagnation. The proposed method is an extension of selection operator in the original MGG model. Generally, GA has two types of selection operators, one is "selection for reproduction", and the other is "selection for survival"; the former is for crossover and the latter is the individuals which survive to the next generation. The proposed method is an extension of the former. The proposed method utilizes distance information between individuals. From this extension, the proposed method aims to expand a search area and improve ability to search solution. The performance of the proposed method is examined with several standard test functions. The experimental results show good performance better than the original MGG model.

Keywords: genetic algorithm, real-coded GA, selection operator, function optimization

1. Introduction

Genetic Algorithm (GA) is a stochastic search and optimization algorithm. The GA is based on the principles of biological evolution[1], [2]. The GA is applied to many optimization problems and the effectiveness are reported. Recently, in this GA, real coded GA attract the attention as optimization methods for nonlinear function and the effectiveness are reported. For this real-coded GA, many crossover operator and generation alternation model have been proposed. As a major crossover operator, the Unimodal Normal Distribution Crossover (UNDX) proposed by Ono et al. shows good performance in optimization of various functions includings multimodal functions and benchmark functions with epistasis among parameters[4], [5], [6]. Also, as for the generation alternation models, Satoh et al. proposed a generation alternation model called the minimal generation gap (MGG) model which is more effective than conventional models[7], [8]. The MGG model has all advantages of conventional models and the ability of avoiding the premature convergence and suppressing the evolutionary stagnation.

In this paper, we propose the a distance correlation selection (DCS) operator as a new selection operator. The proposed method is an extension of selection for reproduction in the original MGG model. In the proposed method, we select parents utilizing distance information between individuals. From this extension, the proposed method aims to expand a search area and improve ability to search solution. In this paper, we evaluate the proposed method using standard test functions[9].

2. Real-coded GA

In this section, we review the UNDX and the MGG model as the basic crossover operator and generation alternation model for real-coded GA.

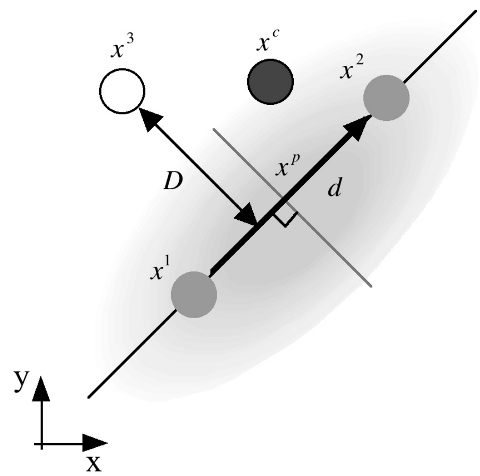

Fig. 1. Unimodal Normal Distribution Crossover, UNDX

2.1. UNDX

The UNDX proposed by Ono et al. shows good performance in optimization problems[4], [5], [6]. The UNDX generates children obeying a normal distribution around the parents as shown in Fig1. The normal distribution is centered at the middle point of parents $\boldsymbol{x}^1$ and $\boldsymbol{x}^2$. The standard deviation is proportion to the distance between parents $\boldsymbol{x}^1$ and $\boldsymbol{x}^2$ in the direction of the line connecting the parents. Moreover, the standard deviation in the verticality directions are proportion to the distance of the parents $\boldsymbol{x}^3$ and the line connecting the parents $\boldsymbol{x}^1$ and $\boldsymbol{x}^2$.

The mathematical descriptions of the UNDX show the following : where $\boldsymbol{x}^c$ is children, and $\boldsymbol{x}^p$ is the center of

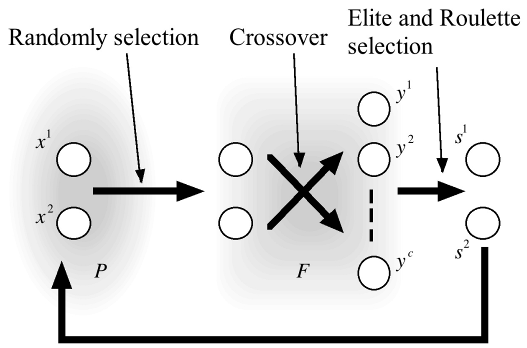

Fig. 2. Minimal Generation Gap, MGG

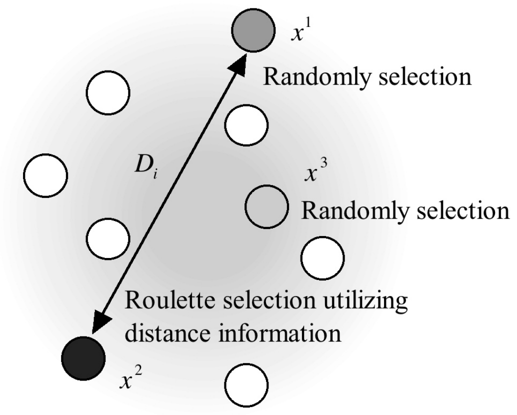

Fig. 3. Distance Correlation Selection, DCS

parents $\boldsymbol{x^1}$ and $\boldsymbol{x^2}$, and D is distance from parent $\boldsymbol{x^3}$ to the line connecting the parents $\boldsymbol{x^1}$ and $\boldsymbol{x^2}$, and ξ and η_i are normal distribution random number, and $\boldsymbol{e}_i$ is unit vector.

$$\boldsymbol{x}^c = \boldsymbol{x}^p + \xi \boldsymbol{d} + D \sum_{i=1}^{n-1} \eta_i \boldsymbol{e}_i \tag{1}$$

$$\boldsymbol{x^p} = (\boldsymbol{x^1} + \boldsymbol{x^2})/2 \tag{2}$$

$$\boldsymbol{d} = \boldsymbol{x^1} - \boldsymbol{x^2} \tag{3}$$

$$\xi \sim N(0, \sigma_\xi^2), \eta_i \sim N(0, \sigma_\eta^2)$$

2.2. MGG

The MGG model proposed by Satoh et al. is an effective one of generation alternation models[7], [8]. In the MGG model, we randomly select parents from population, and we generate children by crossover. From the parents and their children, we select an elite individual and a random one using the roulette selection based on fitness value. The original parents are replaced by the selected children. In this model, the original parents are two individuals, and replacing individuals are also two. Fig2 shows the outline of the MGG model.

3. Proposed method

In this section, we propose a Distance Correlation Selection (DCS) as a new selection operator. The proposed method is an extension of the original MGG model. Generally, the GA has two types of the selection operator. One is the selection for reproduction, the other is the selection for survival; the former is for crossover and the latter is the individuals which survive to the next generation. The proposed method is an extension of the former. The proposed method selects parents utilizing distance information between individuals. From this extension, the proposed method aims to expand a search area and improve ability to search solution. Fig3 shows the outline of the proposed method.

The procedure of the proposed method is following:

1. We randomly select parent $\boldsymbol{x^1}$.
2. We calculate a distance D_i between the parent $\boldsymbol{x^1}$ and other individuals.

$$D_i = |\boldsymbol{x}^1 - \boldsymbol{x}^i|, (i = 1, \ldots, m-1) \tag{4}$$

3. We select other one parent $\boldsymbol{x^2}$ using roulette selection utilizing distance information D_i.
4. We randomly select parent $\boldsymbol{x^3}$.

From this selection operator, the proposed method is more effective than the original MGG model in avoiding premature convergence.

Table 1. Test functions and its domains

$$F_1(x_i) = \sum_{i=1}^{n} [100(x_1^2 - x_i)^2 + (x_i - 1)^2]$$
$$-2.048 \le x_i < 2.048$$

$$F_2(x_i) = 10n + \sum_{i=1}^{n} \left(x_i^2 - 10\cos(2\pi x_i)\right)$$
$$-5.12 \le x_i < 5.12$$

$$F_3(x_i) = 1 + \sum_{i=1}^{n} \frac{x_i^2}{4000} - \prod_{i=1}^{n} \left(\cos\left(\frac{x_i}{\sqrt{i}}\right)\right)$$
$$-512 \le x_i < 512$$

$$F_4(x_i) = \sum_{i=1}^{n} \left(\sum_{j=1}^{i} x_j\right)^2$$
$$-64 \le x_i < 64$$

4. Simulation and Results

In order to confirm the effectiveness of the proposed method, we compare the proposed method with the original MGG model. In the first subsection, we show the test functions. In the second subsection, we show the experimental setting. In the last subsection, we show the simulation results and discuss the simulation results.

4.1. Test functions

In this paper, the performance of the proposed method is examined with standard test functions. The optimization problems used here are the minimization of the Rosenbrock

Table 2. Experimental setting

Number of individuals	(Unimodal)50 (Multimodal)300
Number of variables	n=20
Crossover	UNDX
σ_ξ	1/2
σ_η	$0.35/\sqrt{n}$
Crossover times	200
Number of max generations	30000

function (F1), the Rastrigin function (F2), the Griewank function (F3), the Ridge function (F4). Table1 shows the test functions and its domains.

The Rosenbrock function is the unimodal function and it has dependence on its variables. The Rastrigin function is the multimodal function and it has independence on its variables. The Griewank function is the multimodal function and it has dependence on its variables. The Ridge function is unimodal function and it has dependence on its variables.

The unimodal function has only one global optima. On the other hand, the multimodal function has many local optima, but only one global optima.

4.2. Experimental setting

In this paper, we compare the performance of the proposed method with the original MGG model. An experimental setting is shown in Table2. On the unimodal functions (F1,F4), number of individual is 50. On the multimodal functions (F2,F3), number of individual is 300. On all functions, number of variable is 20, number of applying crossover (UNDX) is 200, number of generation is 30000, parameters of UNDX σ_ξ is 1/2 and σ_η is $0.35/\sqrt{n}$. The initial population is generated randomly following the uniform distribution inside the region described in Table3. All experiments were carried out by changing the seed for random number.

4.3. Simulation results

Table3 shows the average evaluations where the global optima is discovered and the number of convergence times. These results are the average of 20 trials. From these results, the proposed method shows the best performance on the whole test functions. Especially, the proposed method can discover the optimal solution with the number of evaluations less than the MGG model in F1. Additionally, the proposed method can discover the optimal solution in all trials.

Fig4,6 show the histories of the best function value of the Rosenbrock function (F1), and Fig5,7 show the histories of the best function value of the Rastrigin function (F2). In these figures, the y axis shows the function value, and the x axis shows the generation. We see from these figures, we compare the proposed method with the original MGG model, the proposed method can stably search solution.

Generally, we are unable to know a characteristic of a optimization problem in advance. Therefore, if we use the method with dependence on the characteristic, we have possibilities that we don't discover a optimal solution. However, the proposed method improve a stability of search solution for the characteristic of problems. From this reason, the proposed method is effective method better than the original MGG model.

Table 3. The average evaluations where the global optima is discovered and the number of convergence times

	$F1$	$F2$	$F3$	$F4$
Proposed method	935000 20/20	2302200 20/20	1806400 20/20	670200 20/20
MGG model	1501000 20/20	2489800 18/20	1806800 20/20	720000 20/20

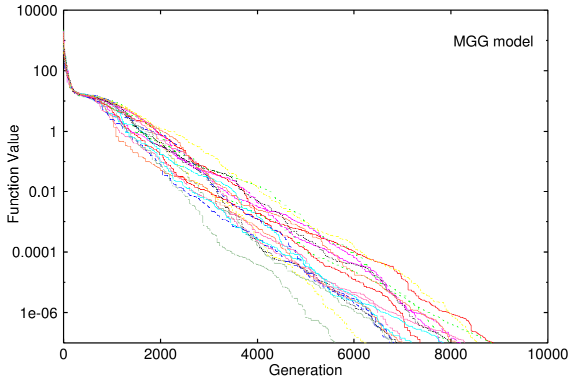

Fig. 4. History of best function value in F1

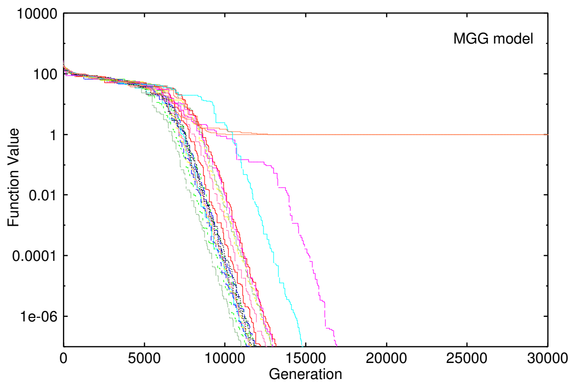

Fig. 5. History of best function value in F2

5. Conclusion

In this paper, we proposed a distance correlation selection (DCS) operator as a new selection operator for reproduction. We compared the performance of the proposed method with the original MGG model. The proposed method showed higher performance than the original MGG model.

References

[1] H.Kitano : Genetic Algorithm, Sangyou-tosyo, 1993, in Japanese

[2] D.E.Goldberg : Genetic Algorithm in Search, Optimization and Machine Learning, Addison-Wesley, 1989

[3] L.Davis : The Handbook of Genetic Algorithm, Van Nostrand, 1990

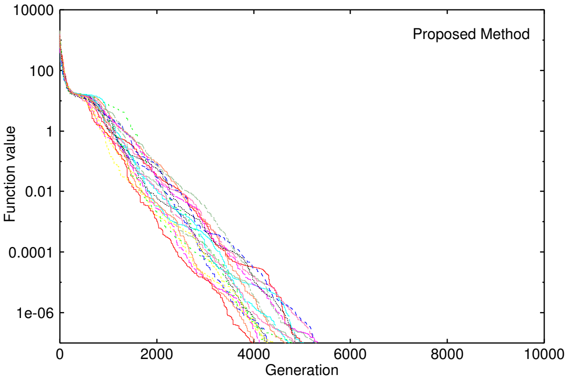

Fig. 6. History of best function value in F1

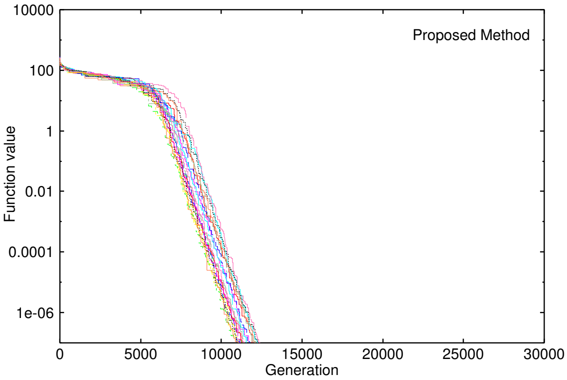

Fig. 7. History of best function value in F2

[4] I.Ono, S.Kobayashi : A Real-coded Genetic Algorithm for Function Optimization Using Unimodal Normal Distribution Crossover, Proceedings of 7th International Conference on Genetic Algorithms, pp246–253, 1997

[5] I.Ono, H.Satoh and S.Kobayashi : A Real-Coded Genetic Algorithm for Function Optimization Using the Unimodal Normal Crossover, Journal of Japanese Society for Artificial Intelligence, 14, pp1146–1155, 1999, in Japanese

[6] H.Kita, I.Ono and S.Kobayashi : Theoretical Analysis of the Unimodal Normal Distribution Crossover for Real-coded Genetic algorithm, Transactions of the SICE, 34-8, pp1333–1339, 1999, in Japanese

[7] H.Satoh, M.Yamamura and S.Kobayashi : Minimal Generation Gap Model for GAs considering Both Exploration and Exploitation, Proceedings of IIZUKA '96, pp494–497, 1996

[8] H.Satoh, I.Ono and S.Kobayashi : A New Generation Alternation Model for Genetic Algorithms and Its Assessment, Journal of Japanese Society for Artificial Intelligence, 12-5, pp734–744, 1997, in Japanese

[9] D.Whitley, K.Mathias, S.Rana, and J.Dzubera : Building better test functions, Proceedings of the 6th International Conference on Genetic Algorithms, pp239–246, 1995