

Empirical Evaluation of Changing Crossover Operators to Solve Function Optimization Problems

Ryouei Takahashi

Department of System and Information Engineering
Hachinohe Institute of Technology
Hachinohe, Japan
ryoei@hi-tech.ac.jp

Abstract— In this paper, the effectiveness of methodologies for changing crossover operators (CXOs) to solve function optimization problems (FOP) are empirically validated in order to solve the problems of premature convergence in genetic algorithms. CXOs are methods of finding solutions for combinatorial optimization problems through genetic algorithms (GAs) while maintaining the balance of satisfying the contrary requisites for GAs: to sustain the diversity of the population and to improve the efficiency of searching for solutions. CXOs measure the diversity of the population on each end of a generation and dynamically alternate global search methods and local search methods according to the degree of that diversity. With the above devices, CXOs can decrease the probability of falling into local optima compared to using only one kind of crossover operator in the crossover operation. In FOP, considering the continuity of functions, real-coded crossover operators such as BLX- α (Blend Crossover) and SPX (Simplex Crossover) are invented and added to general crossover operators such as Two-point Crossover operators. In our investigation, we studied CXOs that exchange Two-point crossover for BLX- α and SPX for BLX- α when the diversity of the generated children drops below the required threshold. We verified experimentally that CXOs can improve the accuracy of searched solutions beyond singly used crossover operators such as Two-point Crossover, BLX- α and SPX. Two well-known multimodal functions are also investigated in this report—Shubert’s test function and the Six-hump camel back function.

Keywords—Diversity Measurement; Changing Crossover Operators; TDGA; BLX- α ; SPX; Function Optimization Problem

I. INTRODUCTION

It is widely known that Genetic Algorithms (GAs) [1]—artificial adaptive systems which use a probabilistic method to search for optimal solutions—are effective for solving combinatorial optimization problems [2]. One of the issues faced by GA is overcoming the problem of premature convergence in which the result converges toward local optima at an early generational stage and becomes stuck there, meaning that the final global solution cannot be found [3]. This problem occurs because the diversity of the population cannot be conserved [4]. In other words, the population is not composed of chromosomes with a variety of genetic structures. If locally optimal solutions comprise a majority of the population, their chromosomal structures cannot be improved through mutation or crossover operations and the probability that they will be selected as offspring in subsequent generations increases with the roulette-wheel selection method. Hence, in

order to solve the premature convergence problem it is necessary to maintain the diversity of the population. But if we allow all individuals generated by GA operations to survive in the next generation without consideration even though they have lower fitness compared to other individuals in the current generation, convergence efficiency worsens and the optimal solutions cannot be found. We must therefore improve GA so that it satisfies the contrary conditions of improved search efficiency and sustained population diversity. TDGA (Thermodynamical Genetic Algorithms)[5] and immune-GA (immune Genetic Algorithms) [6],[7],[8] have been developed to satisfy the above conditions; individuals are selected from candidates generated by GA operations to be offspring in the next generation, maximizing population diversity and optimizing the value of the target function. TDGA is one selection strategy based on the principle of free energy minimization in thermodynamics. In TDGA, we select individuals for the next population from those candidates generated temporarily for the next generation to satisfy the above principle. In applying the principle to GAs, we select them so that the average energy $\langle E \rangle$ is at a minimum and the entropy H is at a maximum. Immune-GA is one optimization method based on Jerne’s idiotypic network, which is a hypothesis describing the homeostasis of organic immune systems. After calculating the similarity between individuals X produced by GA and suppressor cells, the production of X is suppressed in the immune system if similarity is high. Suppressor cells are individuals with various kinds of chromosomal structures and which have higher fitness values and different similarities among them. They are selected and updated successively by individuals generated on every crossover operation.

It has been experimentally shown [9] that TDGA is effective in solving the FOP (Function Optimization Problem)[10],[11]. TDGA and immune-GA have also been experimentally demonstrated [12] to be effective ways of solving the TSP (Traveling Salesman Problem). There, we measured the diversity of the population with the entropy defined in TDGA. We experimentally and quantitatively verified the hypothesis that the ratio of the number of optimum trials and the accuracy of the solution can be improved by maintaining diversity. In this study, we propose some methodologies for changing crossover operators (CXOs) and test their validity experimentally with regard to solving FOP. CXOs can strike the balance needed to satisfy the contrary requisites of GAs (to sustain population diversity and

improve search efficiency) on the crossover operation phase in GA. The effectiveness of CXOs called ECXO has been experimentally verified in TSP, where ACO (Ant Colony Optimization) [13] is used as a global search method and EAX (Edge Assembly Crossover) is used as a local search method [14]. In FOP, considering the continuity of functions, real-coded crossover operators [15] such as BLX- α (Blend Crossover) [16],[17] and SPX (Simplex Crossover)[18],[19] are invented and added to general crossover operators such as Two-point crossover operators. These represent β individuals selected from those in the current generation in n -dimensional real data space, where β is two for BLX- α and $n+1$ for SPX. They randomly reproduce several offspring in the rectangular area or $(n+1)$ polygonal area, including the above β real data. This reproduction enlarges the search space for the solution, ensuring a variety of chromosomal structures and potentially improving the efficiency of searching for more accurate solutions since they can search for solutions in neighbors of locally optimal solutions. They are effective in supplementing the global search through GAs such as Two-point Crossover. One such example is the Bee System [20], which uses Two-point Crossover for the global search and SPX for the local centralized search; it was proposed then experimentally verified as effective. In our investigation, we studied CXOs that exchange Two-point crossover for BLX- α and SPX for BLX- α when the diversity of the generated children drops below the required threshold. We verified experimentally that CXOs can improve the accuracy of searched solutions beyond singly used crossover operators such as Two-point crossover, BLX- α and SPX. We also quantitatively verified that the reason why CXOs can improve accuracy is that they maintain a higher degree of population diversity than single crossover operators. Two well-known multimodal functions are also investigated in this report—Shubert’s test function and the Six-hump camel back function. [21], [22]. This paper is organized as follows. In Section II, the presupposition of genotype and phenotype of individuals in GA to solve the function value minimization problem. In Section III, the metric derived from TDGA to measure the diversity of the population based on the structure of chromosomes of is defined. In Section IV, the concept of CXOs is described. In Section V, experimental results are reported. Section VI concludes the paper.

II. PHENOTYPE AND GENOTYPE OF FUNCTION OPTIMIZATION PROBLEM

A. Function optimization problem

We suppose that the function is denoted by $y=f(x_1, \dots, x_n)$, where y is a dependent variable in real data space R and x_i ($i=1, 2, \dots, n$) is an independent variable in R . The function optimization problem is to find the optimal real value $y_{opt}=f_{opt}$, which is the minimum value f_{min} or the maximum value f_{max} for the above function f , and to search for the value x_{i0} for each of the independent variables x_i with which the function f can take its optimal value f_{opt} in a certain region R^n_0 which is a subspace in R^n . In this paper, we investigate the function minimization problem.

B. Genotype and phenotype of individuals in GA that solves the function value minimization problem

When we solve the function value $y=f(x_1, \dots, x_n)$ minimization problem with GA, candidates for solutions to the problem are designed as individuals. Individuals are characterized by chromosomes and their fitness. In this paper, in order to measure the diversity of the population quantitatively based on TDGA, chromosomes are designed to have binary codes [23]. Chromosomes are gene arrays of binary data taking each of their values as 0 or 1, and they represent values of independent variables x_i of the function $y=f(x_1, \dots, x_n)$. They are arranged in order of the variables x_i ($i=1, 2, \dots, n$). These gene arrays are called the genotype of the individual. The numerical variable data represented by the above binary arrays are called the phenotype of the individual. The length L of the chromosome is determined by the accuracy of the independent variables x_i in the domain. If each variable x_i has its domain $-a \leq x_i < a$ ($a>0$) and if it is designed to have a length of l genes, the interval of the data observed is $a/2^{l-1}$, which is the accuracy of the solution-search space and means that the solution is found among discrete data represented as $x_i = -a + k \times a/2^{l-1}$ ($k=0, 1, \dots, 2^l-1$), and L is calculated as $L=n \times l$. For example, let Shubert’s test function F11 function’s domain be $-10 \leq x_1, x_2 < 10$. Here, if $l=30$, the accuracy of solution-search space is calculated as $10/2^{30-1} \doteq 1.8626 \times 10^{-8}$, and the solution searches for each x_i among $1073741824=2^{30}$ discrete values $x_i = -10 + k \times 10/2^{30-1}$ ($i=1, 2; k=0, 1, 2, \dots, 2^{30}-1$). For example, if x_i has its genotype of (0110011101 0110011101 0110011101), then k is calculated as $778803942 = (0 \times 2^0 + 1 \times 2^1 + 1 \times 2^2 + 0 \times 2^3 + 0 \times 2^4 + 1 \times 2^5 + 1 \times 2^6 + 1 \times 2^7 + 0 \times 2^8 + 1 \times 2^9 + \dots + 1 \times 2^{29})$ and its phenotype is calculated as $x_i = 4.50635384768247 = -10 + 778803942 \times 10/2^{30-1}$. Furthermore, although extended crossover operators such as BLX- α and SPX can enlarge the search space for each variable x_i from discrete binary data space to continuous real data space, the solution x_{i0} they search for is approximately represented again by the binary data with its accuracy of l . These approximated binary data are interpreted to be individuals generated by the above extended crossover operators.

III. THE METRIC H MEASURING THE DIVERSITY OF POPULATION

The metric H measuring the diversity of population P in this study is derived from TDGA. When we apply TDGA to solve any combinatorial optimization problem, we have to give a different interpretation of entropy H according to the problem. Examples of definitions of entropy H have been introduced in GA to solve the Traveling Salesman Problem (TSP)[5] and Knapsack problems[26]. Referring to this example, in this study the definition of entropy H for GA used to solve the function value minimization problem is defined as follows. This formula means that entropy H is defined as the summation of j -th gene’s entropy H_j , which is Shannon’s entropy calculated from the probability distribution function following each gene value on the j -th locus.

$$H = \sum_{j=1}^{chrom \ size} H_j, H_j = -\sum_{i=0}^1 p_{ij} \times \log_2 p_{ij}, p_{0j} = \frac{N_{0j}}{N}, p_{1j} = \frac{N_{1j}}{N} \quad (1)$$

In the above formula, N is the population size of population P . N_{0j} is the number of individuals in P whose gene value on the j -th locus is zero. N_{1j} is the number of individuals in P whose gene value on the j -th locus is 1. Hence, $N_{0j} + N_{1j} = N$. P_{0j} is the probability that the gene value on the j -th locus is 0. P_{1j} is the probability that the gene value on the j -th locus is 1. Hence, $P_{0j} + P_{1j} = 1$. Entropy H_j is the Shannon entropy determined by the probability distribution $n \{P_{0j}, P_{1j}\}$.

IV. CXO (CHANGING CROSSOVER OPERATOR)

A. The CXO Concept

CXO applies a hypothesis stemming from Darwin's well-known evolutionary theory on the origin of species by means of natural selection [24], [25]—that life evolves while maintaining a balance of selection and sustaining the diversity of the population—to GAs in order to solve complex combinatorial optimization problems. It prevents GAs from falling into locally optimal solutions. If we correlate GA procedures with the process of natural selection, then species selection during early generational stages coincides with global search in GAs while selection in later generations coincides with local search in GAs. In early generations, individuals which are represented by a kind of schema with various chromosomal structures with higher fitness are selected to be offspring for the next generation. A schema that has higher than average fitness with a short defining length and low degree is called a building block. In later generational stages, building blocks are combined to construct better and better individuals with higher fitness, whose number increases with exponential order of generations to occupy the majority of the population. These principles are known as the schema theorem and building block hypothesis. However, multimodal functions spread locally optimal solutions over a wide area and it is difficult to find the global optimum function value among those solutions, which causes GAs to fall into local optima in earlier generational stages. CXO solves this premature convergence problem by improving GAs. They capitalize on the fact that natural selection provides us with a mechanism through which life is permanently evolving, balancing both survival of the fittest and population diversity so that offspring can survive, even within changing environments. In it, there is room for individuals with various kinds of structures to be produced with a structural distribution that can survive environmental changes.

Fig. 1 shows CXO algorithm. CXO searches for solutions, while keeping the balance between maintaining the diversity of the population and efficiency to search for solutions. CXO measures the diversity of the population on each end of a generation then dynamically alternate the use of global and local search methods according to the degree of that diversity. In CXO, when the diversity of the population generated by global search operator gets below the required threshold β , the

- STEP1: Initialize population.
STEP2: Evaluate fitness of the individuals on the current generation.
Measure the diversity of the population on the current generation.
Alternate the use of global and local search crossover operators according to the degree of the diversity.
STEP3: Terminate condition is checked.
If target terminate criteria is satisfied, stop the job.
STEP4: Select individuals for reproduction with roulette wheel selection or HpS.
Perform crossover operation on parents' genotype to produce offspring.
Perform mutation operation on the offspring.
STEP5: Replace parents with the offspring, and alternate generation.
Go to STEP2.

Fig. 1 Algorithm of CXO

crossover operator is changed from global search operator to local search operator. If local search operator can recover the diversity of the population over the threshold β , the crossover operator is exchanged from local search operator to global search operator. Global search methods are crossover operators that take extensive views of the solution space to specify several subspaces in which optimal solutions exist. Their goal is to search for solutions efficiently, which means finding solutions with a minimum of computer time. Local search methods are crossover operators that focus their search on neighbors of locally optimal solutions. Their goal is to search for solutions that are more accurate. If the diversity of the population increases above the required threshold, CXOs determines that the search area is too wide and that it has to use the global search method to narrow down the search area in which local optima might exist. If the diversity of the population drops below the required threshold, CXOs judges the search area to be too specific with the global search method and attempts to concentrate the search area in neighbors of local optima using the local search method in order to improve the accuracy of solutions. Furthermore, the local search method is generally devised to maintain the diversity of the individuals generated by expanding the search space so that it includes locally optimal solutions and searches solutions randomly in the extended area. With the above devices, CXOs can decrease the likelihood of falling into local optima compared to using only one kind of crossover operator in the crossover operation. The effectiveness of CXOs called ECXO has been experimentally verified in TSP[14]. This experiment verifies the effectiveness of CXOs in FOP.

B. Crossover operators in FOP

This paper investigates the following three types of crossover operators for solving FOP.

1) Two-point crossover [11]: Two cross sites on chromosomes are randomly selected and genes are exchanged materially between the two sites.

2) BLX- α [9],[15][16]: In the function value $y=f(x_1, \dots, x_i, \dots, x_n)$ minimization problem, we assume that two individuals denoted by $X_1=(x_{11}, \dots, x_{1i}, \dots, x_{1n})$ and $X_2=(x_{21}, \dots, x_{2i}, \dots, x_{2n})$ are selected. BLX- α randomly searches for solutions $(x_1, \dots, x_i, \dots, x_n)$ in the intervals $(S_1, \dots, S_i, \dots, S_n)$ defined as follows, where $x_i \in S_i$ ($i=1, 2, \dots, n$), where α is an extended ratio in the search space is 0.5 in this paper.

$$S_i = [\min(x_{1i}, x_{2i}) - \alpha \times |x_{1i} - x_{2i}|, \max(x_{1i}, x_{2i}) + \alpha \times |x_{1i} - x_{2i}|]. \quad (2)$$

3) SPX[17],[18]: In the function value $y=f(x_1, \dots, x_i, \dots, x_n)$ minimization problem, SPX obtains $(n+1)$ n -dimensional real data from genotypes of a pair of parents selected via roulette-wheel selection or other methods then chooses random candidates for solutions in the extended simplex area which includes the above $(n+1)$ points whose extension ratio ε in the search space is $\sqrt{n+2}$. This random selection in the extended area can bring about diversity in the chromosomal structures of individuals generated as well as BLX- α . The details of SPX are as follows:

a) Let each $\vec{p}_i = (x_{i1}, \dots, x_{ij}, \dots, x_{in})$ denote a parent individual ($i=1, 2, \dots, n+1$). Let S_0 represent the simplex which is created from vertexes of those $(n+1)$ parent individuals.

b) Let $\vec{g} = (g_1, \dots, g_j, \dots, g_n)$ be the center of gravity of the simplex S_0 , calculated as follows.

$$\vec{g} = \frac{\sum_{i=1}^{n+1} \vec{p}_i}{n+1} \quad (3)$$

c) Calculate $(n+1)$ coordinates \vec{q}_i comprising the simplex S_1 which is an extension of S_0 that includes S_0 itself and has an extension ratio of ε as follows, where ($i=1, 2, \dots, n+1$):

$$\vec{q}_i = \vec{p}_i + \varepsilon \times (\vec{p}_i - \vec{g}) \quad (4)$$

d) An offspring $\vec{c} = (c_1, \dots, c_j, \dots, c_n)$ is randomly selected from the above simplex S_1 .

d1) Generate uniformly random number r in the interval $[0, 1]$. $r \in [0, 1]$

d2) Calculate reduction ratio r_i as follows for the relative vector \vec{v}_i that connects \vec{q}_{i+1} to \vec{q}_i in this direction, where ($i=1, 2, \dots, n+1$):

$$r_i = r^{(1/i)} \quad (5)$$

d3) Use the following to calculate harmonized vector \vec{h}_1 which reduces the relative vector \vec{v}_1 connecting \vec{q}_2 with \vec{q}_1 in this direction with a reduction ratio of r_1 .

$$\vec{h}_1 = (h_{11}, \dots, h_{1j}, \dots, h_{1n}) = r_1 \times \vec{v}_1 = r_1 \times (\vec{q}_2 - \vec{q}_1) \quad (6)$$

d4) Calculate a harmonized vector \vec{h}_i sequentially using the following recursion formula, where \vec{h}_i is a reduced harmonized vector of \vec{v}_i connecting \vec{q}_{i+1} with \vec{q}_i in this direction added to the previous harmonized vector \vec{h}_{i-1} with a reduction ratio of r_i .

$$\vec{h}_i = (h_{i1}, \dots, h_{ij}, \dots, h_{in}) = r_i \times ((\vec{q}_i - \vec{q}_{i+1}) + \vec{h}_{i-1}) \quad (7)$$

, where $i=2, 3, \dots, n$.

d5) Calculate an offspring \vec{c} as follows:

$$\vec{c} = (c_1, \dots, c_j, \dots, c_n) = \vec{q}_{n+1} + \vec{h}_n \quad (8)$$

C. CXOs in this study

In our investigation, we studied the following CXOs after considering the characteristics of each crossover operator. We experimentally verified that CXOs improve the accuracy of searched solutions compared to singly used crossover operators such as Two-point crossover, BLX- α and SPX. We also quantitatively verified that CXOs improve this accuracy because they maintain a higher degree of population diversity than the single crossover operators Two-point crossover + BLX- α .

1) CXO(Two-point crossover + BLX): This CXO alternates the use of Two-point crossover for global searches and BLX- α for local searches.

2) CXO(SPX+BLX): This CXO alternates the use of SPX for global searches and BLX- α for local searches.

D. Describing the CXO concept using the CXO(Two-point crossover + BLX) as an example

In this section the idea of CXOs is described using the example CXO(Two-point crossover + BLX) to solve the Shubert's test function F11 minimization problem. Here it is shown that the above CXO can improve accuracy compared to single crossover operators such as Two-point crossover or BLX.

The primary GA run parameters used in this example are:

- Individuals in subsequent generations consist of offspring generated with simple GAs. The fitness of an individual is calculated by 300 - function value. This GA model is called Non-SW (GA without Scaling Window) in Section V.

- The selection method is HpS [4]. In the HpS method, parent A is chosen using the roulette-wheel selection method, and parent B is selected randomly from among individuals whose similarities S_{AB}^s to A are lower than population P 's similarity S_A to A . Refer to the paper [9] for the definition of similarity.

- $Seed_id=1$, an initial value for generating uniformly random numbers which is the same for all three models.

- CXO (Two-point crossover + BLX- α) exchanges Two-point crossover for BLX- α when the diversity of generated children falls below the required threshold $30 = \text{chrom_size } L (=60) \times 0.5$.

- Population diversity is measured with the metric H defined in Section III.

Summary of results (Fig. 2, Fig. 3 and Fig. 4)

Fig. 2 shows how the best minimum value (represented by the Y-axis to a precision of seven decimal places) varies as the generation (represented by the X-axis) advances for CXO (Two-point crossover + BLX- α), BLX- α , and Two-point crossover respectively. From Fig. 2, we can see that both CXO and BLX- α were able to improve the accuracy of the solution found using Two-point crossover. However, it is not clear which of the two methods can find the more accurate solution. Hence, we calculated down to fourteen decimal places in order to improve the accuracy of solutions searched by CXO and BLX- α and compared their accuracy. The results are presented in Fig. 3, which shows that CXO can search for a smaller

function value than BLX- α and that the accuracy of the solution found through BLX- α is improved by CXO.

Fig. 4 shows how diversity (Y-axis) varies as the generation (X-axis) advances for CXO, BLX- α , and Two-point crossover respectively. Chromosomal length $L=60=30$ (bit) $\times 2$ (variables) can be obtained by rounding the initial diversity up to the nearest integer because each initial gene value is randomly determined. Fig. 4 illustrates that CXO maintains the highest degree of diversity until the optimum minimum value is found, while BLX- α maintains the second highest and Two-point crossover the third highest. From this figure, we see that the reason CXOs improve the accuracy of the solution is that they maintain a higher degree of population diversity than the single crossover operators Two-point crossover + BLX- α .

A detailed explanation of the three figures is provided below.

Two-point crossover: On the 36th generation, Two-point crossover finds the function value of -186.73090878512227 . It then continues to improve the function value gradually, finding its best minimum function value of -186.73090878708823 on the 49th generation after which it cannot improve the solution. A diversity of 10.34 is preserved on the 27th generation, decreases to 4.9 by the 36th generation when the best value is found, and finally declines to 0.18 by the 57th generation. This result shows that Two-point crossover has higher efficiency in searching for solutions but a low ability to maintain population diversity.

BLX- α : On the 22nd generation, BLX- α finds the function value of -186.73090882861166 while maintaining a diversity value of 28.69. It continues to improve the function value little by little while maintaining a diversity over 10. And on the 33rd generation it finally finds its best minimum function value of -186.73090883102358 with a diversity of 9.99. It is unable to improve the solution beyond that and becomes stuck in a locally optimal solution, with diversity declining to zero on the 41st generation.

CXO: CXO exchanges Two-point crossover for BLX- α on the 22nd generation, when the diversity of the population drops below the threshold of 30. In doing so it improves the accuracy of the solution from -186.73088015956179 to -186.73090879012304 . CXO then exchanges BLX- α for Two-point crossover on the 24th generation when the diversity of the population recovers to a value over the threshold of 30, which allows it to improve the accuracy of the solution from -186.73090879012304 to -186.73090880051976 . Similarly, exchanging crossover operations are done on the 25th, 26th and 27th generations and the minimum function value of -186.73090883102381 is finally arrived at on the 36th generation with a diversity of 12.37. This result demonstrates that CXOs can improve the accuracy of solutions found by single operators such as Two-point crossover and BLX- α , balancing the contrary requirements of ensuring both the survival of the fittest and the diversity of the population by dynamically changing crossover operators according to population diversity.

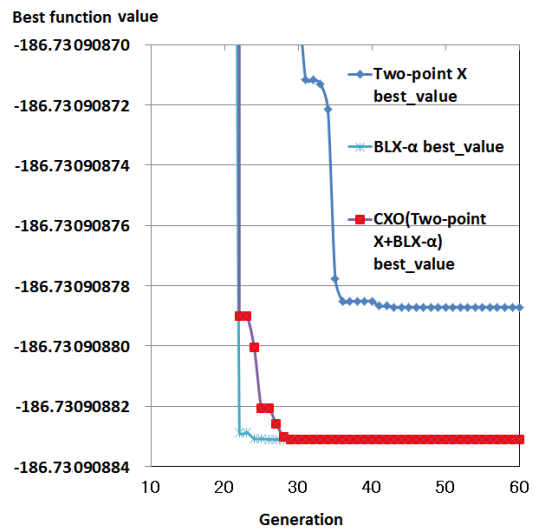


Fig. 2 Comparison of CXO to single operators such as BLX- α and Two-point crossover in terms of best minimum value versus generation

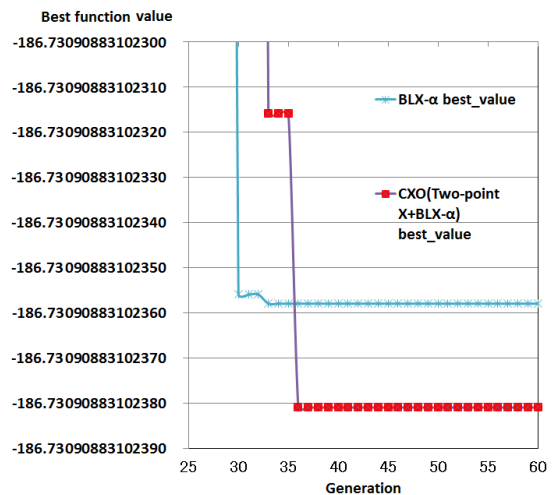


Fig. 3 Detailed comparison of CXO to single operator BLX- α in terms of best minimum value versus generation

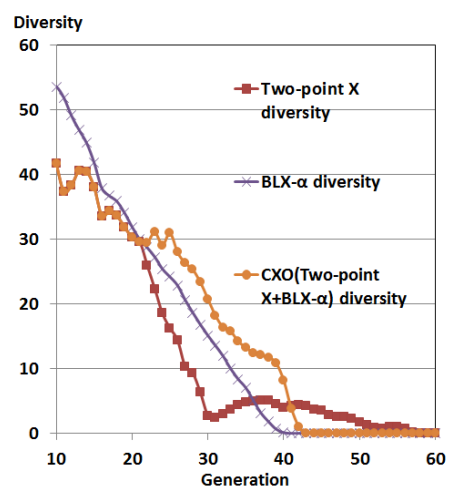


Fig. 4 Evaluation of CXO compared with single operator such as BLX- α and Two-point Crossover on Diversity maintained versus generation.

V. EXPERIMENTS

A. Test functions used for function minimization problems and precision of solutions for evaluating GA models

All of our 900 (15 independent trials \times 3 alternating generations \times 5 crossover operations \times 2 selections \times 2 functions) GA models were able to find the minimum function values with a precision of four decimal places for test functions such as Shubert's test function (F11) and the six-hump camel back function (F12). These can be found in papers [21] and [22]: -186.7309 for F11 and -1.0316 for F12. Under the above precision, the ratio of optimum trials in our experiments is 900/900 or 100%. In order to evaluate GA models more accurately with the accuracy and ratio of optimum trials, in this experiment we raised the precision to 14–16 decimal places. The details are as follows.

1) Shubert's test function F11 (multimodal):

a) Definition:

$$z = f(x_1, x_2) = \sum_{i=1}^5 i \cos(i + (i+1)x_1) \times \sum_{i=1}^5 i \cos(i + (i+1)x_2)$$

b) Domain: $-10 \leq x_j < 10$ ($j=1, 2$)

c) Chrom_size (L): 60 (bits)=30 (bits) \times 2 (variables)

d) The minimum value: -186.730908831023810 at $x_1 = -0.800321102142334$, $x_2 = -1.42528430128098$.

e) Remarks: We judged the model to be optimal if the function takes its value of -186.730908831023810 .

2) Six-hump camel back function F12 (multimodal):

a) Definition:

$$z = f(x_1, x_2) = (4 - 2.1x_1^2 + \frac{1}{3}x_1^4)x_1^2 + x_1x_2 + 4(x_2^2 - 1)x_2^2$$

b) Domain: $-3 \leq x_1 < 3$, $-2 \leq x_2 < 2$

c) Chrom_size (L): 60 (bits)=30 (bits) \times 2 (variables)

d) The minimum value: -1.0316284534898774 at $x_1 = -0.089842010289430618$, $x_2 = 0.712656401097774510$

e) Remarks: We judged the model to be optimal if the function takes its value of -1.0316284534898774 .

B. Machine and software environments for tests

Numerical experiments were performed on a Dell Precision T3500, Intel(R) Xeron (R) CPU W3680, 3.33 GHz, 6-GB RAM, running Windows 7. Our GAs were coded in C. The C language compiler was Microsoft Visual C++ 2010 development system. We should note that precision also depends on the compiler and excel used for calculations.

C. GA models evaluated

For two test functions F11 and F12, thirty GA models (combinations of five crossover operators, three alternating generation models and two reproduction methods) were evaluated statistically over fifteen independent runs.

a) Crossover operations:

(1) Two-point crossover, (2) BLX- α , (3) SPX, (4) CXO (Two-point crossover + BLX- α): If the diversity of the population falls below the threshold $\beta (=L \times 0.9$ or $L \times 0.5)$, CXO exchanges Two-point crossover for BLX- α . If the diversity recovers over the threshold β , CXO exchanges BLX- α for Two-point crossover. This alternation is done at each end of the generation. (5) CXO (SPX + BLX- α): If the diversity of the population falls below the threshold $\beta (=L \times 0.9$ or $L \times 0.5)$, CXO exchanges SPX for BLX- α . If the diversity recovers over the threshold β , CXO exchanges BLX- α for SPX. This alternation is done at each end of the generation.

b) Models of altering generations:

(1) Non-Scaling GA: This model is a simple Genetic Algorithm without a scaling window. All the individuals generated by GA are replaced with old ones. The fitness is calculated as (CONSTANT - function value) > 0 in the function minimization problem.

(2) Scaling window GA [27]: This model calculates the value of the individual to be its function value minus the worst (maximum) function value among individuals from several prior generations. It can update the fitness values of individuals at the end of all generations.

(3) TDGA: Our TDGA works on the scaling window GA.

c) Selection method: (1) Roulette-wheel selection [3], (2) HpS (Heterogeneous Paring Selection) [4]

D. Evaluated items

All of the evaluated items for each combination of run parameters of each GA model are as follows.

(1) Ratio of optimum trials (R_{opt}): ratio of times we found the optimum solution to all fifteen test cases

(2) Average minimum value \bar{f} : defined as $(f_1 + f_2 + \dots + f_m)/m$, where f_i is the minimum function value searched with GA for each trial and m is the number of test trials ($m = 15$ for each GA model's tests.)

(3) Best minimum value: the minimum function value of the fifteen test cases

(4) Sample Standard deviation s : defined as $((f_1 - \bar{f})^2 + (f_2 - \bar{f})^2 + \dots + (f_m - \bar{f})^2)/m)^{1/2}$

(5) Average gen.: average number of generations required to find each minimum function value

(6) Average initial diversity: average initial diversity of all fifteen trials

(7) Average best diversity (DIV_{best}): average diversity in the best generation when each minimum function value is found

(8) T-test value (T): T-test is performed for inferring the alternative hypothesis that any GA model does not have the same average minimum value as the best GA model which has

Table 1 Evaluation of Selection Methods

Test function	Selection method	Optimal minimum value	Ratio of optimum trials	Avg. minimum value	Avg. standard deviation	Avg. gen.	Avg. initial div.	Avg. best div.
Shubert's function (F11)	HpS	-186.73090883102381	0.19(=43/225)	-186.73090343521530	1.61×10^{-5}	58	59.77	20.79
	Roulette-Wheel		0.16(=35/225)	-186.73089017419291	3.75×10^{-5}	57.27	59.77	19.31
Six-hump camel back function (F12)	HpS	-1.0316284534898774	0.0(=0/225)	-1.0316283147116928	4.57×10^{-7}	40.40	59.77	31.96
	Roulette-Wheel		0.004(=1/225)	-1.0316284263015368	7.86×10^{-8}	40.87	59.77	31.53

Table 2. Detailed evaluation of changing crossover operators on Two-point-X, BLX- α , and SPX using Shubert's function (a) Selection method: Heterogeneous Paring Selection (in a case of F11)

Crossover operator	GA model	Ratio of optimum trials	Avg. min. value	Best min. value	Standard deviation	Avg. gen	Avg. initial div.	Avg. best div.	T-test value
Two-point-X	Non-SW*	0 (=0/15)	-186.73087732795594	-186.73090883102293	0.000117	48	59.86	3.70	-1.035×10^9
	SW-GA**	0 (=0/15)	-186.73085939696566	-186.73090883102375	0.000125	48	59.86	4.97	-1.624×10^9
	TDGA	0.27 (=4/15)	-186.73090883102370	-186.73090883102381	7.265×10^{-14}	86	59.59	41.10	0
BLX- α	Non-SW	0.13 (=2/15)	-186.73090883102361	-186.73090883102381	9.927×10^{-14}	35	59.86	14.77	-2.954
	SW-GA	0.33 (=5/15)	-186.73090883102364	-186.73090883102381	1.055×10^{-13}	36	59.86	15.47	-1.969
	TDGA	0.33 (=5/15)	-186.73090883102378	-186.73090883102381	1.048×10^{-13}	96	59.59	36.84	2.629
SPX	Non-SW	0.2 (=3/15)	-186.73090883102361	-186.73090883102381	1.32×10^{-13}	40	59.86	13.25	-2.954
	SW-GA	0.27 (=4/15)	-186.73090883102370	-186.73090883102381	8.46×10^{-14}	38	59.86	15.03	0
	TDGA	0.13 (=2/15)	-186.73090883102370	-186.73090883102381	7.448×10^{-14}	109	59.59	27.93	0
Two-point-X + BLX- α	Non-SW	0.13 (=2/15)	-186.73090883102367	-186.73090883102381	1.076×10^{-13}	35	59.86	16.18	-0.985
	SW-GA	0.13 (=2/15)	-186.73090883102367	-186.73090883102381	8.495×10^{-14}	37	59.86	16.00	-0.985
	TDGA	0.13 (=2/15)	-186.73090883102378	-186.73090883102381	5.732×10^{-14}	87	59.59	39.99	2.629
SPX + BLX- α	Non-SW	0.27 (=4/15)	-186.73090883102367	-186.73090883102381	1.206×10^{-13}	36	59.86	15.40	-0.985
	SW-GA	0.2 (=3/15)	-186.73090883102370	-186.73090883102381	9.791×10^{-14}	37	59.86	14.382	0
	TDGA	0.33 (=5/15)	-186.73090883102370	-186.73090883102381	1.139×10^{-13}	102	59.59	36.866	0

Table 3 Detailed evaluation of changing crossover operators on Two-point-X, BLX- α , and SPX using Six-hump camel back function (b) Selection method: Roulette-wheel Selection (in a case of F12)

Crossover operator	GA model	Ratio of optimum trials	Avg. min. value	Best min. value	Standard deviation	Avg. gen	Avg. initial div.	Avg. best div.	T-test value
Two-point-X	Non-SW*	0.0 (=0/15)	-1.0316281710517807	-1.0316284534898705	9.205×10^{-7}	50	59.86	6.89	-106423939
	SW-GA**	0.0 (=0/15)	-1.0316283281030889	-1.0316284534897917	2.591×10^{-7}	48	59.86	5.75	-47246299
	TDGA	0.0 (=0/15)	-1.0316284534898577	-1.0316284534898770	1.899×10^{-14}	53	59.59	55.98	-0.641
BLX- α	Non-SW	0.0 (=0/15)	-1.0316284534898579	-1.0316284534898739	9.558×10^{-15}	31	59.86	21.36	-0.565
	SW-GA	0.0 (=0/15)	-1.0316284534898628	-1.0316284534898750	9.736×10^{-15}	31	59.86	21.46	1.281
	TDGA	0.0 (=0/15)	-1.0316284534898601	-1.0316284534898714	8.867×10^{-15}	37	59.59	59.99	0.264
SPX	Non-SW	0.0 (=0/15)	-1.0316284534898637	-1.0316284534898765	9.635×10^{-15}	33	59.86	21.51	1.62
	SW-GA	0.0 (=0/15)	-1.0316284534898574	-1.0316284534898750	8.887×10^{-15}	34	59.86	20.00	-0.754
	TDGA	0.0 (=0/15)	-1.0316284534898605	-1.0316284534898750	8.737×10^{-15}	55	59.59	59.85	0.414
Two-point-X + BLX- α	Non-SW	0.0 (=0/15)	-1.0316284534898605	-1.0316284534898750	8.272×10^{-15}	33	59.86	20.57	0.414
	SW-GA	0.07 (=1/15)	-1.0316284534898594	-1.0316284534898774	9.930×10^{-15}	34	59.86	20.78	0
	TDGA	0.0 (=0/15)	-1.0316284534898579	-1.0316284534898770	1.904×10^{-14}	53	59.59	57.52	-0.565
SPX + BLX- α	Non-SW	0.0 (=0/15)	-1.0316284534898637	-1.0316284534898765	9.304×10^{-15}	34	59.86	20.36	1.62
	SW-GA	0.0 (=0/15)	-1.0316284534898610	-1.0316284534898765	1.084×10^{-14}	32	59.86	21.03	0.603
	TDGA	0.0 (=0/15)	-1.0316284534898605	-1.0316284534898750	8.737×10^{-15}	55	59.59	59.85	0.414

<Remarks> Non-SW*: GA without scaling window; SW-GA**: GA with scaling window.
 : the best CXO model of the highest R_{opt} selected for the T-test.

Table 4 Evaluation of changing crossover operators compared to single operators such as Two-point-X, BLX, and SPX using Shubert's function
(a) Selection method: Heterogeneous Paring Selection

Crossover operator	Ratio of optimum trials	Avg. min. value	Avg. best div.	Avg. [T-test value]
Two-point crossover	0.09 (=4/45)	-186.730881851981767	16.59	8.86×10^8
BLX- α	0.27 (=12/45)	-186.730908831023677	22.36	2.517
SPX	0.20 (=9/45)	-186.730908831023670	18.74	0.985
(Two-point crossover +BLX)/ (SPX+BLX)	0.20 (=18/90)	-186.730908831023698	23.14	1.533

the highest ratio of optimum trials and is determined experimentally. T is calculated as $(\bar{f}_{best} - \bar{f})/(s/(m-1)^{1/2})$,

where \bar{f}_{best} is the average minimum value of the above GA model and \bar{f} is the average minimum value of each GA model.

In the above items, the diversity of the population is measured by the TDGA's entropy H .

E. GA run parameters

The primary GA run parameters were: (1) Population size P : 300 (2) Probability of crossover operation: 0.95 (3) Probability of mutation: 0.05 (4) Scaling window size W : 7 (5) The job was stopped if the current minimum value was not improved for 200 generations or if the diversity of the population did not change for 5 generations. (6) TDGA run parameters for controlling temperature T [11]: $r=0.999$, $r_l=1.001$ and $r_u=0.99$, $r_l=1.01$, where r is the parameter used for decreasing T when system equilibrium conditions are satisfied, and r_l is the parameter used for increasing T when the system loses population diversity. With the former settings we can find solutions with higher precision, while the latter is superior in terms of search efficiency.

F. Comparison between HpS and roulette-wheel selection method

Table 1 illustrates a summary of Table 2(a) for Shubert's function and Table 3(b) for the six-hump camel back function from the viewpoints of selection methods (1) HpS and (2) Roulette wheel selection and with regard to the following evaluated items. In the table, the optimum minimum value is the minimum function value that is determined theoretically or searched out experimentally in the tests.

a) Shubert's function

- (1) Both selection methods were able to find the optimum minimum value of -186.73090883102381 .
- (2) HpS had a higher ratio of optimum trials ($R_{opt}=0.19$ (=43/225) than roulette-wheel selection ($R_{opt}=0.16$ (=35/225)).
- (3) HpS had less average minimum value and less average standard deviation than roulette-wheel selection as shown in the Table 1. This result shows that the accuracy of solutions

Table 5 Evaluation of changing crossover operators compared to single operators such as Two-point-X, BLX, and SPX using Six-hump camel back function.

(b) Selection method: Roulette-Wheel Selection				
Crossover operator	Ratio of optimum trials	Avg. min. value	Avg. best div.	Avg. [T-test value]
Two-point crossover	0.0 (=0/45)	-1.0316283175482424	22.87	5.12×10^7
BLX- α	0.0 (=0/45)	-1.0316284534898602	34.27	0.703
SPX	0.0 (=0/45)	-1.0316284534898605	33.79	0.929
(Two-point crossover +BLX)/ (SPX+BLX)	0.01 (=1/90)	-1.0316284534898605	33.35	0.603

found through HpS is better than those found through roulette-wheel selection.

- (4) HpS has higher R_{opt} and accuracy than roulette-wheel selection because it maintains higher population diversity while searching for solutions. Specifically, average best diversity (DIV_{best}), which is the diversity when the best minimum value is found, was 20.79 for HpS and 19.31 for roulette-wheel selection.

b) Six-hump camel back function

- (1) Roulette-Wheel selection method was able to find the optimum minimum value of -1.0316284534898774 .
- (2) HpS had a higher minimum function value and higher average standard deviation than roulette-wheel selection as shown in the table 1. This result seems to show that the accuracy of solutions found through HpS is worse than those found through roulette-wheel selection. But T-test with the level of significance of 0.05 showed that the null hypothesis that they have the same average minimum value was not rejected.
- (3) In fact, average best diversity DIV_{best} was 31.96 for HpS, and it had nearly the same DIV_{best} of 31.53 for roulette-wheel selection.

What we can learn from the above experiments?

The GA model that maintains higher population diversity when searching for solutions has a higher ratio of optimum trials and higher search accuracy. Whether the HpS method or roulette-wheel selection method will have a higher ratio of optimum trials and higher search accuracy depends on the particularity of test functions.

G. Experimental results on Shubert's test function F11

1) Evaluation of HpS (Table 2(a), Table 4(a)): Totally 225 samples were used for HpS evaluation. Results are provided in Table 2(a) and Table 4(a) and are summarized below.

a) Evaluation of fifteen GA models (Table 2(a)): The fifteen GA models are combinations of five crossover operators (Two-point crossover/BLX- α /SPX/CXO(Two-point crossover+BLX- α)/CXO(SPX+BLX- α)) and three alternating generation models (Non-SW/SW-GA/TDGA). Fifteen independent tests

were conducted for each GA model. The GA models were then evaluated in terms of R_{opt} , the average minimum value, and standard deviation. The results of these experiments are provided in Table 2(a) and summarized below.

(1) Among 15 GA models, (Two-point Crossover + BLX- α) \times TDGA model had the least minimum value (-186.73090883102378) with a less standard variance meaning that it had the highest accuracy for solutions searched.

(2) Among 15 GA models, (SPX+BLX- α) \times TDGA model as well as (BLX- α) \times (TDGA or SW-GA) model had the highest ratio of optimum trials ($R_{opt}=0.33$ (=5/15)).

(3) The (Two-point Crossover + BLX- α) \times TDGA model had the higher accuracy and the highest R_{opt} because it was able to strike a balance between maintaining population diversity and searching for solutions efficiently. Specifically, this model's DIV_{best} , which is the diversity maintained on the generation when the optimum minimum value is found, was 39.99, ranking second among the 15 GA models.

b) Evaluation of Changing Crossover Operators (CXO(Two-point crossover+BLX- α) and CXO(SPX+BLX- α)) (Table 2(a), Table 4(a)): Each of the two CXOs had 45 samples, so a total of 90 samples were analyzed for CXOs. The results of their statistical analysis are illustrated in Table 2(a) and Table 4(a) and are summarized below.

(1) We select (SPX+BLX- α) \times TDGA model as the best CXO model for the T-test, because it had the highest ratio of optimum trials ($R_{opt}=0.33$ (=5/15)) and the second minimum function value among all the GA models as shown in the Table 2(a).

(2) Each GA's T-test value calculated from above the best CXO model is shown in the Table 2(a). The null hypothesis that any GA model has the same minimum value with the best CXO model is rejected, if it has T-test value over 2.145 or below -2.145 which are points for the inference with the level of significance of 0.05 under the number of degrees of freedom of (15-1). Table 2(a) shows five GA models in Two-point Crossover, SPX and BLX- α reject the null hypothesis.

(3) CXOs had a relatively high R_{opt} : CXOs' R_{opt} of 0.20 (=18/90) is inferior to BLX- α 's R_{opt} of 0.27 (=12/45), the same as SPX's R_{opt} of 0.11(=5/45), and is superior to Two-point crossover's R_{opt} of 0.09 (=4/45) as shown in Table 4(a).

(4) CXOs had the highest average minimum value compared with other single operators such as Two-point crossover or BLX- α or SPX as shown in Table 4(a).

(5) CXOs had the highest solution accuracy, because they had the highest average DIV_{best} of 23.14 among the four crossover operators.

(6) T-test results using the absolute value of the T-test value denoted by |T-test value| in the Table 4(a) show that the significant difference of the average minimum value was found between CXOs and other operators such as Two-point crossover and BLX- α . Regards as SPX, it had less significant difference between CXOs.

2) Evaluation of roulette-wheel selection method: omitted.

H. Results on Six-hump camel back function F12

1) Evaluation of HpS: omitted.

2) Evaluation of roulette-wheel selection (Table 3(b), Table 5(b)): A total of 225 samples were used for evaluating roulette-wheel selection. The results of evaluating of roulette-wheel selection are shown in Table 3(b) and Table 5(b) and are summarized below:

a) Evaluation of fifteen GA models (Table 3(b)): Fifteen independent tests were conducted for each GA model. The models were evaluated in terms of RE and R_{opt} ; the results of those experiments are provided in Table 3(b) and are summarized as follows:

(1) Among the 15 GA models, the (SPX+BLX- α) \times Non-SW model had the minimum function value (-1.0316284534898637) with the less standard variance which means it had the highest accuracy of searched solutions.

(2) Among the 15 GA models, the (Two-point Crossover +BLX- α) \times SW-GA model had the highest ratio of optimum trials R_{opt} (=0.07(= 1/15)).

(3) The (SPX+BLX- α) \times Non-SW model had the highest accuracy because we suppose that it was able to strike a balance between maintaining population diversity and searching for solutions efficiently. Specifically, this model's DIV_{best} was 20.36 and ranked lower among ten GA models. Hence other models with higher DIV_{best} can have possibility to find more accurate solutions.

b) Evaluation of Changing Crossover Operators (CXO(Two-point crossover + BLX- α) and CXO(SPX+BLX- α)) (Table 3(b), Table 5(b)): Each of the two CXOs had 45 samples, so a total of 90 samples were analyzed for CXOs. The results of their statistical analysis are illustrated in Table 3(b) and Table 5(b) and are summarized below.

(1) We select (Two-point Crossover+BLX- α) \times SW-GA as the best CXO model for the T-test, because it had the highest ratio of optimum trials ($R_{opt}=0.07$ (=1/15)) shown in the Table 3(b).

(2) Each GA's T-test value calculated from above the best CXO model is shown in the Table 3(a). Table 3(b) shows that all of the GA models except for Two-point Crossover (Non-SW, SW-GA) cannot reject the null hypothesis that any GA model has the same average minimum value as the above selected best GA model. This result might show a kind of situations to which the NFL theorem [28] could be applied.

(3) CXOs had the highest R_{opt} : CXOs' R_{opt} of 0.01 (=1/90) as shown in Table 5(b).

(4) CXOs had the highest average minimum value compared with other single operators such as Two-point crossover or BLX- α or SPX as shown in Table 5(b). It is the same as SPX.

(5) CXOs had the highest solution accuracy, because they as well as SPX and BLX- α were able to maintain higher population diversity better than Two-point crossover: Specifically, CXOs had the higher average DIV_{best} of 33.35.

(6) From above higher maintainability of population diversity, T-tests in the Table 5(b) show that the significant difference of the average minimum value was not found between CXOs and other operators such as SPX and BLX- α except for Two-point crossover.

VI. CONCLUSION AND FUTURE WORK

In this paper, we used two well-known multimodal test functions, which are Shubert's test function F11 and the six-hump camel back function F12, to experimentally verify that when using genetic algorithms to solve function optimization problems, CXOs that exchange Two-point crossover for BLX- α and exchange SPX for BLX- α when the diversity of generated children falls below the required threshold can improve the accuracy of searched solutions compared to singly used crossover operators such as Two-point crossover, BLX- α , and SPX. All of our 900 (15 independent trials \times 3 alternating generations \times 5 crossover operations \times 2 selections \times 2 functions) GA models were able to find the minimum function values with a precision of four decimal places (-186.7309 for F11 and -1.0316 for F12). Under the above precision, the ratio of optimum trials in our experiments is 100% (=900/900). In order to evaluate GA models more accurately with the accuracy and ratio of optimum trials, in this experiment we raised the precision to 14–16 decimal places. Experiments show that CXOs can improve the accuracy of solutions found through BLX- α , which is the single crossover operator with the highest solution accuracy among single crossover operators, by 9% (= $[7.14/5.98 + 7.50/7.14 + 1.72/1.57 + 1.66/1.63]/4 - 1.0$) which is evaluated with the well-known absolute relative error defined as $|\bar{f} - f^*| / |f^*|$, where \bar{f} is the average minimum value found in the experiments and f^* is the optimum minimum value. This is because CXOs exchange crossover operators dynamically from global search to local search in order to maintain higher diversity when searching for solutions than single crossover operators. Our future work will involve extending sample spaces to verify the CXOs and improving CXOs to get more accurate solutions.

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number JP15K00347. We would like to thank them for supporting our work.

REFERENCES

- [1] J. H. Holland, "Adaptation in Natural and Artificial Systems," MIT Press, 1992.
- [2] E. Aarts and J. K. Lenstra, Local Search in Combinatorial Optimization, Princeton University Press, 2003.
- [3] D. E. Goldberg, "Genetic Algorithms in Search, Optimization, and Machine Learning," Addison-Wesley Publishing Company, Inc., 1989.
- [4] M.-K. Tsai, J.-M. Yang, Y.-F. Tsai, and C.-Y. Kao, "An Evolutionary Algorithm for Large Traveling Salesman Problems," IEEE Transactions on Systems, Man, and Cybernetics- Part B. Cybernetics., Vol. 34, No.4, pp. 1718-1729, 2004.
- [5] K. Maekawa, N. Mori, H. Tamaki, H. Kita and Y. Nishikawa, "A Genetic Solution for the Traveling Salesman Problem by means of a Thermodynamical Selection Rule", Proceedings of the 1996 IEEE Congress on Evolutionary Computation (ICEC' 96), pp. 529-534, 1996.
- [6] K. Mori, M. Tsukiyama, and T. Fukuda, "Application of an Immune Algorithm to Multi-Optimization Problem," T. IEE Japan, Vol.117-C, No.5, pp.593-598, 1997 (in Japanese).
- [7] J. Lu and M. Xie: "Immune-Genetic Algorithm for Traveling Salesman Problem," TRAVELING SALESMAN PROBLEM, THEORY AND APPLICATIONS, Edited by Donald Davendra, Published by InTech, pp. 81-96, 2010.
- [8] S. Ono and S. Nakayama, "Multiple Solution Search Based on Hybridization of Real-Coded Evolutionary Algorithm and Quasi-Newton Method," Proceeding of 2007 IEEE Congress on Evolutionary Computation (CEC 2011), pp. 1133-1140, 2007.
- [9] R. Takahashi, "Verification of Thermo-dynamical Genetic Algorithm to Solve the Function Optimization Problem through Diversity Measurement," Proceedings of the 2016 IEEE Congress on Evolutionary Computation (IEEE CEC 2016), pp. 168-177, 2016.
- [10] K. A. De Jong, "An Analysis of the Behavior of a Class of Genetic Adaptive Systems," Doctoral Dissertation, University of Michigan, January 1975.
- [11] R. L. Haupt and S. E. Haupt, "Practical Genetic Algorithms," John Wiley & Sons, Inc., 2004.
- [12] R. Takahashi, "Quantitative Evaluation of Iterative Extended Changing Crossover Operators to Solve the Traveling Salesman Problem," in Proceedings of 2014 10th International Conference on Natural Computation (ICNC), pp. 235-244, IEEE, 2014.
- [13] M. Dorigo and L. M. Gambardella, "Ant Colony System: A Cooperative Learning Approach to the Traveling Salesman Problem," IEEE Transactions on Evolutionary Computation, 1(1), pp. 53-66, 1997.
- [14] R. Takahashi, "Solving the Traveling Salesman Problem through Iterative Extended Changing Crossover Operators," Proceedings of Tenth International Conference on Machine Learning and Applications, pp. 253-258, IEEE Man, Machine and Cybernetics Society, 2011.
- [15] S. Kobayashi, "The Frontier s of Real-Coded Genetic Algorithms, " Transaction of Japanese Society for Artificial Intelligence, Vol.24, No.1, pp. 128-143, 2009.
- [16] L. J. Eshelman and J. D. Schaffer, "Real Coded Genetic Algorithms and Interval-Schemata," Foundations of Genetic Algorithms 2, pp. 187-202, Morgan Kaufmann, 1993.
- [17] M. Takahashi and H. Kita, "A Crossover Operator Using Independent Component Analysis for Real-Coded Genetic Algorithms," Proceedings of the 2001 IEEE Congress on Evolutionary Computation, pp.643-649, 2001.
- [18] S. Tsutsui, M. Yamamura, and T. Higuchi, "Multi-parent Recombination with Simplex Crossover in Real Coded Genetic Algorithms," Proceeding of Genetic and Evolutionary Computation Conference, pp.657-664, ACM, 1999.
- [19] H. Kita, "Up-to-date Genetic algorithm to solve the combinatorial optimization problem," Tutorial on Information Processing based on Probabilistic Algorithm, Workshop on SMAPIP, 2003 (in Japanese).
- [20] T. Sato and M. Hagiwara, "Bee System: Finding Solution by a Concentrated Search," T. IEE Japan, Vol.118-C, No.5, pp.721-726, 1998 (in Japanese).
- [21] J. M. Dieterich and B. Harkte, "Empirical Review of Standard Benchmark Functions Using Evolutionary Global Optimization," Applied Mathematics, Vol. 3, pp. 1552-1564, 2012.
- [22] X.-S. Yang, "Test Problems in Optimization, in: Engineering Optimization: An Introduction with Metaheuristic Applications (Eds. X.-S. Yang)", John Wiley & Sons, 2010.
- [23] O. Abdul-Rahman, M. Munetomo and K. Akama, "An Improved Binary-Real Coded Genetic Algorithm for Real Parameter Optimization, " Proceedings of 2011 Third World Congress on Nature and Biologically Inspired Computing, pp.149-156, 2011.
- [24] C. Darwin, "On Origin of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life," 1859 (translated into Japanese by R. Yasugi, published by Iwanami, 1996).
- [25] G.T. Parks and P. J. Clarkson " Erratum: A Species Conserving Genetic Algorithm for Multimodal Function Optimization, " Evolutionary Computation, Vol.10, No.3, pp. 207-234, 2003.
- [26] N. Mori, J. Yoshida, H. Tamaki, H. Kita and Y. Nishikawa, " A Thermodynamical Selection Rule for the Genetic Algorithm, " Proceeding of ICEC'95, pp.188-192, 1995.
- [27] J. J. Grefenstette, "Optimization of Control Parameters for Genetic Algorithms," IEEE Trans. on Systems, Man, and Cybernetics, vol.SMC-16, no.1, pp.122-128,1986.
- [28] H. -P. Schwefel, I. Wegener and K. Weinert, "Advances in Computational Intelligence," Springer-Verlag Berlin Heiderberg, 2003.