Noname manuscript No.
(will be inserted by the editor)

Parameters Study of Sharing Mechanism on Genetic Algorithm for Multimodal Problem

Zhang Huiyao[1] · Atsushi Yokoyama [1,*]

Received: date / Accepted: date

Abstract Sharing mechanism is widely used in genetic algorithm (GA) to solve multimodal problems. In this paper, the influenece of parameters of share function on the search performance of GA is studied. There are three key factors in the share function, which are metric space, distance constant and index constant. First, the metric space is studied based on the peroformance of GA; second the distance constant is determined based on optimal metric space; finaly,According to the metric space and distance constant, the index constant is decided. According to the five criterias of the experiment, the optimal combination of these three parameters are determined to maxmize the performance of GAs.

Keywords Genetic Algorithm · Niche · Metric Space · Multimodal

1 Introduction

Genetic algorithms (GAs) are first introduced in 1975 by Holland[1], because of its powerful search ability, which was widely been used to solve many search problems. Compared with other random search algorithm, for example, particle swarm optimization, ant colony optimization, and simulated annealing algorithm[2], GAs are not easily trapped in local optima,Traditional GAs has been to successfully find the optimal value in the domain for unimodal problems, however, sometimes, not only the optimal point but also the secondary important point is needed. Because of its search rule, traditional GAs are failed to maintain the information for a multimodal problem.

Zhang Huiyao
Room 203,Bulding 3,Kyoto Institue of Technology
Matsugasaki,Sakyo-ku,Kyoto,606-8585,JAPAN
E-mail: zhanghy1012@gmail.com

S. Author
second address

1.1 Multimodal Optimization

In order to maintain the diversity of the population for a multimodal problems, it is significant to maximize the number of niche in the population. There are several methods to achieve this target,which are Crowding, Deterministic Crowding and Sharing methods. share Function[3]

1.2 Genetic Algorithm Procedure

According to Darwin natural evolution theory, the one who fits the environment most well are more possilbe to survive and reproduction. GA simulates the process of natural evolution which includes selection,crossover and mutation. Because of the powerful search ability of GA,it has been widely used in many fields for multimodal problems. The procedure diagram of GA as shown in figure 1

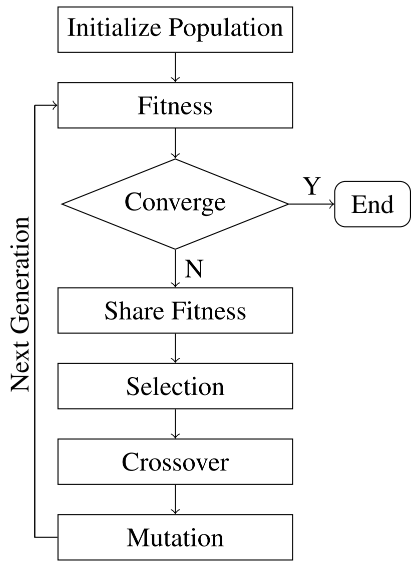

Fig. 1: GA Procedure with Share Function

1.3 Parent Selection

Selection is the most important step of GA algorithm which decides the diversity of the population. To improve the search ability and reduce the search cost, many selection methods [4] has been invented,the selection schemes can be divided into four classes which are proportionate reproduction, ranking selection, tournament selection and Genitor(or "steady state") selection.In this paper, to maintain a stable subpopulation for multimodel problems,the stochastic remainer algorithm is used for experiment.

1.4 Parents Crossover

According to the selected parents, crossover is used to generate the offspring, combining genes from parents to reproduce new members. There are so many research [5, 6] about the crossover operation in GAs to maximize the divistiy of the population. In this paper, one-point crossover strategy is implemented.

1.5 Offspring Mutation

Mutation simulates the situation genes on the chromosomes which received from parents randomly changed. The process are used to provide new evolution direction and information for the population, the parameters has during this process has been well researched by Schaffer[7].

1.6 Sharing

Sharing mechanism is first introduced by Holland [8] to reduce the fitness of similar individual in the population and increase the diversity of the population. For Sharing mechanism, the most important thing is how to identify the number of niches, Miller comed up dynamic niche sharing method, and lin developed the method to identify the number of niches in the population.[9]

1.7 Shareing Issue

The problem within sharing mechanism is how to identiy the range of parameters in the formula. Given the number of peaks in population,Deb and Goldberg comed up method to set up the appropriate value for σ_{sh}, however,in many problems, the number of peaks is difficult to identify, and more often, the fitness is discrete. So it is important to determine these parameters properly.

$$sh(d_{i,j}) = \begin{cases} 1 - \left(\frac{d_{i,j}}{\sigma_{sh}}\right)^{\alpha_{sh}} & \text{if } d_{i,j} < \sigma_{sh} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

2 Research Method

According to the following share function

$$\mathrm{sh(d)} = 1 - \left(\frac{\mathrm{d}}{\sigma_{\mathrm{share}}}\right)^{\alpha}, \quad \mathrm{d} < \sigma_{\mathrm{share}}$$
$$= 0, \quad \text{otherwise}$$

where d denotes distance between two strings, α is a contant number, σ_{share} is also a constant number.

define share fitness f_i' as the following

where m_i' indicates niche count,

$$d_{ij} = d(x_i, x_j) \tag{2}$$

There are three factors which influence the performance of GAs, which are metric space and two constant numbers. Metric space is one of the most important factors which influence the effect of GA. Suppose x_i and x_j denotes two different string in the population. Given the metric space (M,d), metric function d is:

$$d : M \times M \rightarrow R$$

This can be divided into the following three steps, the first is the normalization of metric space.

1: Metric Normalization

In order to discuss the influence of different parameters, the normalization of metric distance is necessary, suppose d_{max} means the longest distance between two strings in the population, So the distance can be normalized by the following formula:

$$d_{norm} = \frac{d_i}{d_{max}} \tag{3}$$

After normalization the distance between any two points is:

$$0 \leq d_{ij} \leq 1 \tag{4}$$

2: Define Tensor

Define D and D^* as the basic metric space and dual metric space, which can be used as the basic building block to define all the metric space, and a tensor with r basic metric space and q dual metric space is defined as by the formula:

$$D = |x_i - x_j| \tag{5}$$

$$D^* = \sqrt{|x_i - x_j|} \tag{6}$$

$$t(r,s) := \underbrace{D \bigotimes \cdots \bigotimes D}_{r} \times \underbrace{D^* \bigotimes \cdots \bigotimes D^*}_{s} \rightarrow R \tag{7}$$

where R stands for real number.

3: constant d_{share} in share formula

Given the best tensor condition, increase the value of d_{share} to check the influence of d on the search ability of GAs.take the step length $d_{step} = 0.5$, after the nomalization of metric space, the distance d_{max} between two points in M satisfiy $d_{max} \leq 1$. According to the share function theory, d_{share} needs great then d_{max}. So the range of is $\{d_{share} | d > 1\}$

3 Experiment and Results

3.1 Experiment Method

1. GA

In order to evaluate the impact of the parameters, the mutation process is ignored. The parameters of GA is shown in the table 1

Table 1: GA-parameters

parameter	value
population size	50
encoding method	binary encoding
encoding length	16
selection strategy	roulette wheel
crossover strategy	one-point
mutation strategy	None

The ith individual in the population converted from binary to decimal by the following formula:

$$x_i = \frac{\sum_{j=1}^{16} 2^{j-1} g_i^j}{2^{15}} \tag{8}$$

2. Evaluation indexes

To evaluate the performance of GA algorithm, four index are come up. The first evaluation criterion is apparent reliability denoted as R, which is calulate by the following formula

$$R = \frac{n}{\mathrm{N}} \tag{9}$$

n stands for the number of GA finds at least one optimal point, N stands for the number of the runtime of GA. The second criterion is the normalized cost of GA,denoted as C_n,which is calculated by the following equation:

$$C_n = N_g P / R \tag{10}$$

where P is population size. The third criterion is population richness, denoted as P_r, which is calculated by the following formula:

$$P_r = N/P$$

Population richness is used to denote how many optimal points these GA maintains. The fourth criterion is the number of best output, denoted by O_n. The last criterion is the number of niche, denoted by N_n, the number of niche can be used to represents the ability of the GA to maintain subpopluation.

3.2 Result

Example 1:

$$f_1(x) = \sin^6(5.1\pi x + .5) \tag{11}$$

The experiment result in the table 2, with d_share equals 2 and the value of σ is 1.

Table 2: Tensor Parameter

tensor	R	C	O_n	P_r	N_n
$t(1,0)$	0.99	2020	5.1	11.84	5.12
$t(0,1)$	0.97	2051	3.83	7.32	4.0
$t(1,1)$	0.82	2439	1.9	3.9	2.46
$t(0,2)$	0.86	2339	1.94	3.77	2.34
$t(1,2)$	0.94	2127	2.06	3.58	2.34
$t(2,2)$	0.82	2439	1.84	3.24	2.30
$t(0,3)$	0.89	2234	1.89	3.46	2.25
$t(1,3)$	0.89	2272	1.72	3.38	2.22
$t(2,3)$	0.90	2222	1.86	3.40	2.20
$t(3,3)$	0.94	2127	1.96	3.24	2.10

According to the table 2, with the sum of $r+s$ increases, the population richness P_r, The number of best output O_n and the number of niches N_n decrease,and when $r=1$ and $q=0$,apparent reliability R obtain the maximaize value,GA obtains the best result. as shown in Fig. 2

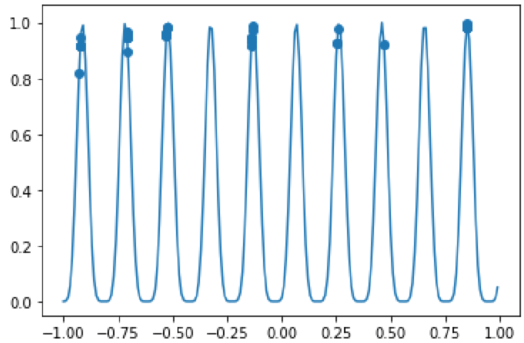

Fig. 2: Result 1

2. constant influence under tensor $t(1,0)$
Because $t(1,0)$ obtain the best search result,

Table 3: d_{share} parameter

d_{share}	R	C	O_n	P_r	N_n
1.1	0.98	2040	4.08	12.24	5.42
1.6	0.98	2040	5.10	13.36	5.52
2.1	0.98	2040	5.18	12.06	5.12
2.6	1.0	2000	5.16	10.2	4.66
3.1	0.98	2040	4.70	9.32	4.44
3.6	0.96	2083	4.5	8.22	4.12
4.1	0.98	2040	4.86	8.44	4.18

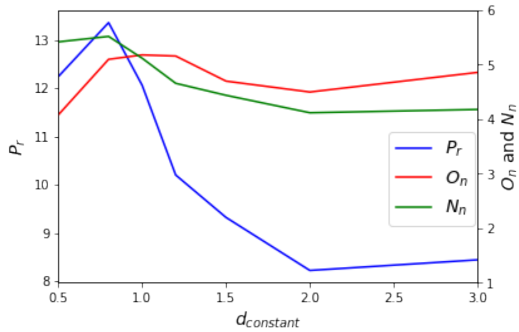

Fig. 3: metric denominator

According to Tab 3, we can see that when $d_{share} = 1.6$, the performance of GA is best; Combined with Fig.3, when

$$1.1 \le d_{share} \le 2.1$$

The GA performed the best.
3. power index influence

Table 4: σ parameter

σ	R	C	O_n	P_r	N_n
0.5	0.98	2040	4.64	13.02	5.08
0.8	0.92	2173	4.56	13.08	5.22
1.0	0.98	2040	5.10	13.36	5.52
1.2	0.98	2040	5.32	11.72	5.36
1.5	0.96	2083	4.80	9.80	4.68
2.0	0.96	2083	3.94	7.50	4.10
3.0	0.96	2083	2.08	3.58	2.34

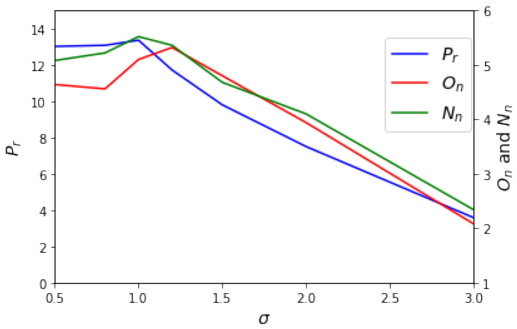

Fig. 4: σ parameter

Example 2:

$$f_1(x) = \cos^6(5.1\pi x + .5)$$

Table 5: Metric Space Example

Metric	R	C	O_n	P_r	N_n
$t(1,0)$	1.00	2000	4.98	11.6	4.94
$t(0,1)$	0.96	2083	3.74	7.82	4.40
$t(1,1)$	0.94	2127	1.76	4.56	2.86
$t(0,2)$	0.88	2272	1.72	3.60	2.30
$t(1,2)$	0.92	2173	1.6	3.74	2.36
$t(2,2)$	0.94	2127	1.82	3.62	2.22
$t(0,3)$	0.86	2325	1.76	3.52	2.22
$t(1,3)$	0.92	2173	1.68	3.54	2.26
$t(2,3)$	0.94	2127	1.86	3.54	2.20
$t(3,3)$	0.94	2127	2.00	3.32	2.18

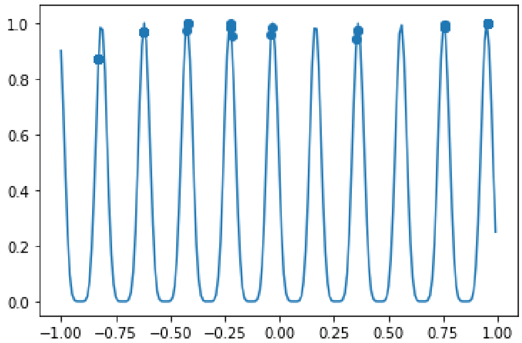

Fig. 5: Result 2

Table 6: Metric Space

l_{step}	R	C	O_n	P_r	N_n
1.1	0.98	2040	4.08	12.24	5.42
1.6	0.98	2040	5.10	13.36	5.52
2.1	0.98	2040	5.18	12.06	5.12
2.6	1.0	2000	5.16	10.2	4.66
3.1	0.98	2040	4.70	9.32	4.44
3.6	0.96	2083	4.5	8.22	4.12
4.1	0.98	2040	4.86	8.44	4.18

Four evaluation criteria are used. The first is normalized cost per genetic search,

$$C_n$$

cost is determinted by the following formula.

4 Conclusion

In this paper, three variables of sharing function tensor,d_{share},σ was analysed. The experiment result is measured by five indexes, normalized search cost, population richness, the number of niche number, the number of optima.

5 Acknowledgements

This is work was supported by China Schloarship Council(CSC).

References

1. Jeffrey R Sampson. Adaptation in natural and artificial systems (john h. holland), 1976.

2. Zelda B Zabinsky. Random search algorithms. *Wiley Encyclopedia of Operations Research and Management Science*, 2010.
3. David E Goldberg, Jon Richardson, et al. Genetic algorithms with sharing for multimodal function optimization. In *Genetic algorithms and their applications: Proceedings of the Second International Conference on Genetic Algorithms*, pages 41–49. Hillsdale, NJ: Lawrence Erlbaum, 1987.
4. David E Goldberg and Kalyanmoy Deb. A comparative analysis of selection schemes used in genetic algorithms. In *Foundations of genetic algorithms*, volume 1, pages 69–93. Elsevier, 1991.
5. Zakir H Ahmed. Genetic algorithm for the traveling salesman problem using sequential constructive crossover operator. *International Journal of Biometrics & Bioinformatics (IJBB)*, 3(6):96, 2010.
6. Antonio J Nebro, Juan J Durillo, Francisco Luna, Bernabé Dorronsoro, and Enrique Alba. Mocell: A cellular genetic algorithm for multiobjective optimization. *International Journal of Intelligent Systems*, 24(7):726–746, 2009.
7. JDAR Schaffer. A study of control parameters affecting online performance of genetic algorithms for function optimization. *San Meteo, California*, 1989.
8. Brad L Miller and Michael J Shaw. Genetic algorithms with dynamic niche sharing for multimodal function optimization. In *Proceedings of IEEE international conference on evolutionary computation*, pages 786–791. IEEE, 1996.
9. Chyi-Yeu Lin and Wen-Hong Wu. Niche identification techniques in multimodal genetic search with sharing scheme. *Advances in Engineering Software*, 33(11-12):779–791, 2002.