Journal of Computational Science & Engineering 10 (2014) 383-390

Journal of Computational Science & Engineering

Available at http://www.asocse.org ISSN 1710-4068

An Estimation of Optimization Problem Hardness Based on the Optimal Contraction Theorem

[1]Kun Li*, [2] Ming Li, [3] Hao Chen

[1] *College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, 211106, Jiangxi China*

[2] *Key Laboratory of Nondestructive Testing, Nanchang Hangkong University, 330063, Jiangxi China*

[3] *Key Laboratory of Nondestructive Testing, Nanchang Hangkong University, 330063, Jiangxi China*

Article Information

Article history:
Received 11 December 2013
Revised 28 January 2014
Accepted 26 January 2014
Available online 29 January 2014

Keywords:
Optimal Contraction Theorem
Optimization Hardness
Particle Swarm Optimizer
Genetic Algorithm

Abstract

On the basis of the optimal contraction theorem, this paper analysis the causes of the Optimization Hardness, proposes new indicators of optimization hardness−the effective ratio of exploration and the effective ratio of exploitation, and make the efficiency ratio curve of exploration (REr Curve) and exploitation(REi Curve). Furthermore, a new method was proposed to estimate the optimal ratio of the optimization problem. Finally, the method was confirmed by eight test functions, and compared with other problem hardness indicators and test results of Particle Swarm Optimizer and Genetic Algorithm.

1. Introduction

Optimization problems exist in many areas as artificial intelligence, automatic control and management science. A lot of optimization algorithms are designed to solve a wide variety of optimization problems, such as the steepest ascent method, hill climbing method, branch and bound method, simulated annealing algorithm, genetic algorithms and particle swarm algorithm. However, the researchers don't have enough Information for selecting appropriate algorithm for certain problems. Then, as the No Free Launch theorem [1] assured that there is no super algorithm can solve all sorts of optimization problems. Therefore, methods of detecting optimization

* Corresponding author: likuncba@126.com

2014 © The American Computational Science Society. All rights reserved.

problems' features should be explored.

There are three sorts of traditional methods of problem hardness estimate methods [2] as Fitness Distance Correlation (FDC), Correlation Length (CL) and Epistasis Measure (EM), and innovated analysis methods such like Schema Fitness Landscape [3] (SFL), Information Theoretic Epistasis Metrics [4,5] (ITEM), Landscape State Machine [6,7] (LSM), Information Landscape [8] (IL), Spatial Correlation [9] (SC) and so on. These methods can be classified into two categories: one category focus on the character of search scope, the other category focus on the structure of single feasible corner point. The FDC and SC methods assume that the relationship between fitness value and Euclid distance of feasible solutions should indicate the problem's hardness level. The CL method based on the Fitness Landscape (FL) model is proposed to describe the stochastic correlation between different feasible solutions. The LSM model is a finite state machine, which simulates the transition process between feasible solutions caused by optimization algorithm. The traditional EM based on the Schema theorem, include two indicators——Epistasis Variance and Epistasis Correlation, which calculate the influence of each bit in the binary string of feasible solution. The new EM is expanded to real code Genetic Algorithm and already utilized to analysis the correlation between different decision variables. The excellent performance of these theories and methods mentioned above is limited in a small and finite problem set. The Optimal Contraction Theorem [10] (OCT) provide a series of optimization problem hardness indicators that can overcome the disadvantages of the methods mentioned above, but these indicators can not be calculated. This paper proposes a new estimate method of the OCT's optimization hardness indicator.

The Section 2 will introduce the OCT's basic concepts and the Hardness Indicators of OCT. The In Section 3, definitions of efficiency of exploration and exploitation are given, according to the OCT framework. Then, the estimation of the hardness indicator based on the definitions are proposed and discussed. In Section 4, some numerical experiments are made in purpose of analysis and evaluate the performance of the estimate method. Conclusion and prospect are made in Section 5.

2. Optimal Contraction Theorem and Hardness Indicator

According to the optimal contraction theorem, the tradeoff between exploration and exploitation [11] always exists in the process that optimizer searching the global optimum in the search scope. The performance of the optimizer depends on how suitable the tradeoff between Ei & Er is for the Problem's characters. In OCT, the exploration and exploitation are redefined using the language of probability theory, and the definition is cited as follows in purpose of describe the understructure of this paper.

Cited definition 1—Exploration (Er):

A sampling behavior is exploration iff its sampling point is generated independent of the information acquired by historical sampling points. Mathematically, a sampling X_n at time n is exploration iff

$$P(X_n = x | X_{n-1} \quad \cdots \quad X_0) = P(X_n = x) .$$

Cited definition 2—Exploitation (Ei):

A sampling behavior is exploitation iff the generation of its sampling point depends upon the information acquired by historical sampling points. Mathematically, a sampling X_n at time n is exploitation iff

$$P(X_n = x | X_{n-1} \quad \cdots \quad X_0) \neq P(X_n = x) \quad .$$

The search process mentioned above can

2014 © The American Computational Science Society. All rights reserved.

be divided into several stages, and each stage is a subset of the whole search space β. There are two special stages can be used as the understructure of problem hardness indicator. In OCT, The definitions of the special stages are as follows and the Figure 1 depicts the:

Cited definition 3—Optimal Field (OF):

A subspace D^* in the whole search space B is optimal field iff both of the following conditions are satisfied.

1) Arbitrary point in D^*, denoted by x, satisfies $f(x) > f(x')$, where x' is a local optimum which is only inferior to global optima among all local optima.

2) Arbitrary point in B satisfying the inequality $f(x) > f(x')$ is involved in D^*.

OF can be denoted by the level set $\left\{x \middle| \forall x \in B : f(x) > f(x^*)\right\}$

Cited definition 4—Optimal Supreme Cap (OSC):

For each isolated global optimum (denoted by x^*), its optimal supreme cap $C^*(x^*)$ is the section of function landscape that includes all points for which there exists a steepest ascent path to reach the global optimum. Particularly, if many global optima gather together to form a continuous field, they will be regarded as an ensemble, and the unique OSC corresponding to the ensemble contains all points that can reach at least one optimum in the cap along a steepest ascent path. OSC is similar to the concept of the attraction basin of local optima for minimization problems.

3. The Estimation of Optimal Feature Factor

First of all, the Optimization problem discussed in this paper can be formulated as follows:

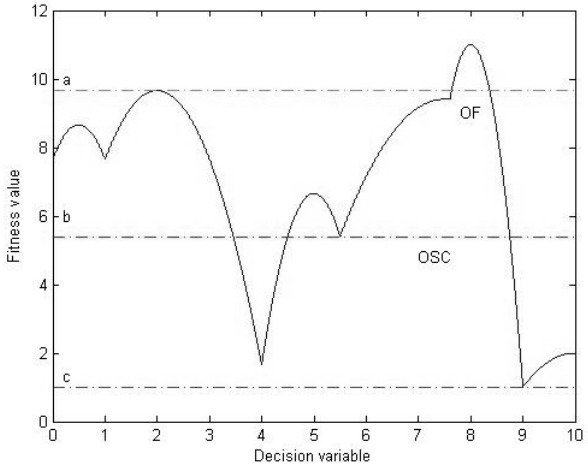

Fig. 1 Optimal Field and Optimal Supreme Cap

$$\begin{cases} \max & f(X) \\ s.t. & X \subseteq \beta \end{cases}$$

Where β denotes the whole search space, $X = [x_1 \quad x_2 \quad \cdots \quad x_n]$ denotes a point in whole search space β, $\sigma(\beta)$ is a measure of search space β.

3.1 The Efficiency of Exploration and Exploitation

According to the OCT, The shrinkage and dilation of contraction search scope may coexist in the searching process, but the continuous shrinkage is definitely the concerned point. For the convergent optimizer, the contraction search scope shrinks to the global optimum at last. With the shrinkage process, the floor level of fitness value of the current population is raising. Howbeit, the change of the fitness floor level can be traced and utilized as the tool of estimating strict optimal ratio and optimal ratio.

If the current population's fitness floor level can denoted by Y, and a point is in the current search scope, while the fitness value of this point is higher than Y. Hence, the whole search space β can be divided into two subsets, the subset that all fitness values are higher than Y, is denoted by S, and S is also the current search scope. On the basis of the cited definition

2014 © The American Computational Science Society. All rights reserved.

1, the ratio of the measure of the subset S to that of the whole search space β can be the indicator of the efficiency of optimizer's exploration behavior.

Definition 1——exploration effective ratio: If the floor level of fitness value of current search scope can be denoted by Y, the subset S consist of all points in the whole space that their fitness value is higher than Y. The ratio of the measure of the subset S (denoted by $\sigma(S)$) to that of the whole search space β (denoted by $\sigma(\beta)$), is defined as exploitation effective ratio, and denoted by $R_{Er} = \sigma(S)/\sigma(\beta)$.

As Figure 1 shows, if Y=a, S equals to the optimal field, and the exploration effective ratio equals to the strict optimal ratio. If Y=b, S equals to the optimal supreme cap, and the exploration effective ratio equals to the optimal ratio. Therefore, the strict optimal ratio and optimal ratio is the special condition of the exploration effective ratio.

In most cases, the elements of subsets S are separated into several unconnected regions. These isolated regions can be divided into two parts: the region contains the global optimal solution (denoted by S') or not. Hence, the intersection between set S and optimal supreme cap definitely include the global optimum. But, the relationship of set *S*, set *S'*, OSC and OF are complicated. As shown in Figure 1,

if Y<c, $OF \subseteq OSC \subset S' = S$;

if c<Y<b, $OF \subseteq OSC \subset S' \subseteq S$;

if b<Y<a, $OF \subseteq S' \subset OSC \subseteq S$;

if Y>a, $S' = S \subset OF \subseteq OSC$.

According to the definition of exploitation in optimal contraction theorem, set *S'* contains the information that lead the optimizer to global optimum. The measurement of set *S'* indicates the amount of positive information in the whole search space. The definition of exploitation effective ratio is as follows:

Definition 2——exploitation effective ratio: If the floor level of fitness value of current search scope can be denoted by Y, the subset S' consists of points in the whole search space β satisfy two following conditions:

(1) The fitness value of this point is higher than *Y*;

(2) There is one neighborhood of this point exist, which include at least one global optimum;

The ratio of the measure of the subset *S'* (denoted by $\sigma(S')$) to that of the whole search space β (denoted by $\sigma(\beta)$), is defined as exploitation effective ratio, and denoted by $R_{Er} = \sigma(S')/\sigma(\beta)$.

The exploitation effective ratio is less than or equal to the exploration effective ratio, the detailed relationship between the two indicators is shown in Figure 1. If $Y<c$, $R_{Ei}=R_{Er}$; if $c<Y<a$, $R_{Ei}<R_{Er}$; if $Y>a$, $R_{Ei}=R_{Er}$.

3.2 The Curve of Er and Ei Effective Ratio

Obviously, *Y* is varying in the search process, and the value of *Er* and *Ei* Effective Ratio will also change accordingly. Then, the relationship between the value of *Y* and *Er* and *Ei* Effective Ratio can be described by two curves——the R_{Er} and R_{Ei} curves. Clearly, the variation range of two curves is restricted into interval [0, 1].

Compared Figure 2 with Figure 1, the R_{Er} and R_{Ei} curve illustrate some characters of the whole search space: (1) The point where R_{Er} curve cuts R_{Ei} curve equals to the strict optimal ratio of the problem; (2) Every sharply decrease of the R_{Ei} curve indicate that at least one isolation region is separated from set S', which do not include global optimum, and the last sharply decrease of the R_{Ei} curve equals to the optimal ratio. Figure 1 can vividly depict that the sharply decrease of

2014 © The American Computational Science Society. All rights reserved.

the R_{Ei} curve can show the separation of every isolate region clearly, and the decrease range can indicate the scale of the separated region. For example, there are three sharply decrease show in Figure 2, and each of them correspond to one isolate region shown in Figure 1. Therefore, the R_{Er} and R_{Ei} curve can be utilized as a tool of estimating optimal ratio and strict optimal ratio.

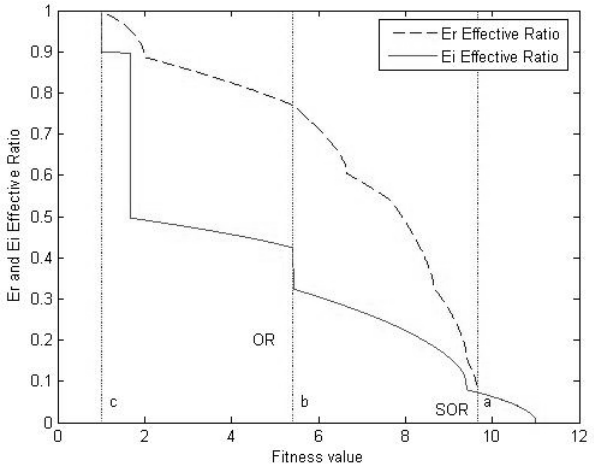

Fig. 2 the Curve of Er and Ei Effective Ratio

Nevertheless, the key point of this estimate method is locate the sharply decrease point on the R_{Ei} curve. Get the differential sequence of the R_{Ei} curve is the first step of estimation, then, a proper threshold value is need to filter the differential sequence, and a point can stay in the sequence only if its value is higher than the threshold value. What should be concerned in the filtered differential sequence is not the value of the last stayed point but the position of that point. If the position of the last stayed point can be denoted by *i*, the value of the point *i+1* on the R_{Ei} curve is the estimation value of optimal ratio. Figure 3 shows the stayed sequence of the R_{Ei} curve showed in Figure 2. Yet, getting the R_{Er} and R_{Ei} curve by calculating each point in the whole search space is unacceptable. According to the statistic knowledge, the statistic feature of uniform random sample set can be utilized to estimate of the ensemble. Hence, in the estimate method of optimal ratio mentioned above, a uniform random sample set P should be introduced to calculate the approximated R_{Er} and R_{Ei} curve, and instead of the whole search space.

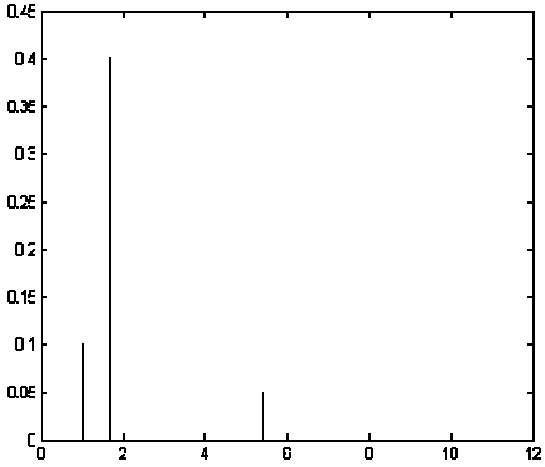

Fig. 3 the processed difference sequence

The use of approximated R_{Er} and R_{Ei} curve may bring two disadvantages. First, the approximation of global optimal may bring some inaccuracy. Then, the OSC is too small to detect in the sample set P, the estimate method of optimal ratio mentioned above can not work totally. So, the minimum resolution of the estimation method can be expressed by the ratio of the measurement of sample set P (denoted by $\sigma(P)$) to that of the whole search space (denoted by $\sigma(\beta)$) ,which denoted by $\sigma(P)/\sigma(\beta)$. The estimate method proposed in this paper can not work, while the optimal ratio is smaller than $\sigma(P)/\sigma(\beta)$.

4. Numerical Experiments and Result Analysis

4.1 Test Functions

(1) Test Function F1:

$$f_1 = \sum_{i=1}^{N} x_i^2, \quad x_i \in [-100 \quad 100]$$

2014 © The American Computational Science Society. All rights reserved.

Sphere function, global optimum $f_{min}(0)=0$.

(2) Test Function F2:

$$f_2 = 1 + \frac{1}{4000}\sum_{i=1}^{N}(x_i)^2 - \prod_{i=1}^{N}\cos\left(\frac{x_i}{\sqrt{i}}\right) \quad x_i \in [-50 \quad 50]$$

Griewangk function, global optimum $f_{min}(0)=0$.

(3) Test Function F3:

$$f_3 = \sum_{i=1}^{N-1}\left[100\left(x_{i+1} - x_i^2\right) + \left(x_i - 1\right)^2\right] \quad x_i \in [-2.048 \quad 2.048]$$

Rosenbrock function, global optimum $f_{min}(1)=0$.

(4) Test Function F4:

$$f_4 = 0.5 + \frac{\sin^2\left(\sqrt{\sum_{i=1}^{N} x_i^2}\right) - 0.5}{\left[1 + 0.001\left(\sum_{i=1}^{N} x_i^2\right)\right]^2} \quad x_i \in [-100 \quad 100]$$

Schaffer F6 function, global optimum $f_{min}(0)=0$.

(5) Test Function F5:

$$f_5 = -\sum_{i=1}^{N} x_i \sin\left(\sqrt{|x_i|}\right) \quad x_i \in [-500 \quad 500]$$

Schwefel function, global optimum $f_{min}(420.9687,420.9687)=-837.9658$.

(6) Test Function F6:

$$f_6 = \prod_{i=1}^{N}\left(\sum_{j=1}^{5} j\cos\left((j+1)x_i + j\right)\right) \quad x_i \in [-10 \quad 10]$$

Shubert function, there are 18 global optimums $f_{min}=-186.73$.

(7) Test Function F7:

$$f_7 = -e^{\frac{1}{N}\left(\sum_{i=1}^{N}\cos(2\pi x_i)\right)} - 20e^{-0.2\sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}} + 20 + e \quad x_i \in [-30 \quad 30]$$

Ackley's path function, global optimum $f_{min}(0)=0$.

(8) Test Function F8:

$$f_8 = 1 + \prod_{i=1}^{N}\sin\left(x_i\right) + \frac{1}{0.05 + \sum_{i=1}^{N}\left(x_i - 4\right)^{\alpha}} \quad \alpha \geq 2 \quad N \leq 10 \quad x_i \in [0 \quad 3$$

Need in the Haystack (type II) function, global optimum $f_{max}(4)=21.5731$.

4.2 Test Results and Discussions

Table 1 shows three indicators of the eight test functions and the test results of PSO and GA. The size of PSO and GA population is 500, the max generation is 1000.The GA is real coded, roulette wheel selection, and crossover rate is *0.8*, mutation rate is *0.1*; PSO learning factor $C_1=C_2=1.4$, and $V_{max}=0.5$. Second is the count unit of convergence time in Table 1. The optimization hardness sort of these test functions from top to bottom is F_5, F_4, F_2, F_6, F_7, F_8, F_1, F_3 ,depend on PSO, or F_5, F_8, F_7, F_6, F_4, F_2, F_3, F_1,depend on GA. Figure 4 shows the R_{Er} and R_{Ei} curves of the test functions section 4.1. The the R_{Er} and R_{Ei} curves of test function F1are coincidence, because of it is a unimodal function.

The bar graph in Figure 5 can show the differences between there optimization problem hardness indicators clearly, particularly the estimate value of optimal ratio in the bar graph is normalized after logarithm, and signify the minimal hardness function by the longest bar. Based on correlation length value, F_8 is the most difficult problem, and F_4 is the most difficult problem according to fitness distance correlation value and the estimate value of optimal ratio. The correlation length matches the GA results in Table 1 best, but the fitness distance correlation and estimation of optimal ratio match the PSO and GA results well both. The test results prove that the estimation method of optimal ratio is available.

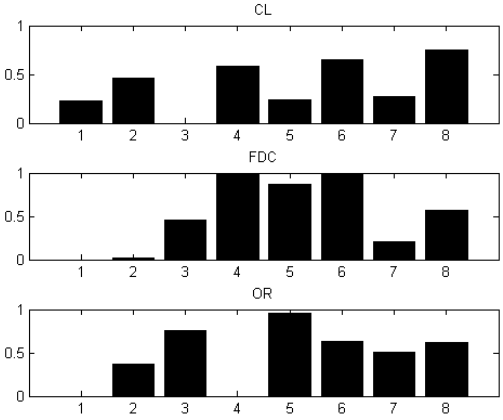

Fig. 5 the comparison of there hardness indicators

2014 © The American Computational Science Society. All rights reserved.

Table 1 the comparison between hardness indicators and test results of SGA and PSO (Dimension N=2)

Test Function	CL	FDC	OR	PSO		SGA	
			(estimate)	Conver. ratio	Conver. time	Conver. ratio	Conver. time
F1	0.4351	-0.9577	1	1	0.058	1	0.734
F2	0.3045	-0.9390	2.7e-3	0.94	0.73	0.81	1.214
F3	0.5631	-0.5194	5.95e-6	0.97	0.037	1	0.048
F4	0.2330	0.0015	0	0.63	0.764	0.56	0.577
F5	0.4274	-0.1257	2.5e-7	0.02	0.083	0	—
F6	0.1977	-0.0005	4.5e-5	1	0.474	0.23	0.777
F7	0.4124	-0.7639	3.28e-4	1	0.163	0.19	0.883
F8	0.1424	-0.4076	4.76e-5	1	0.095	0.09	0.96

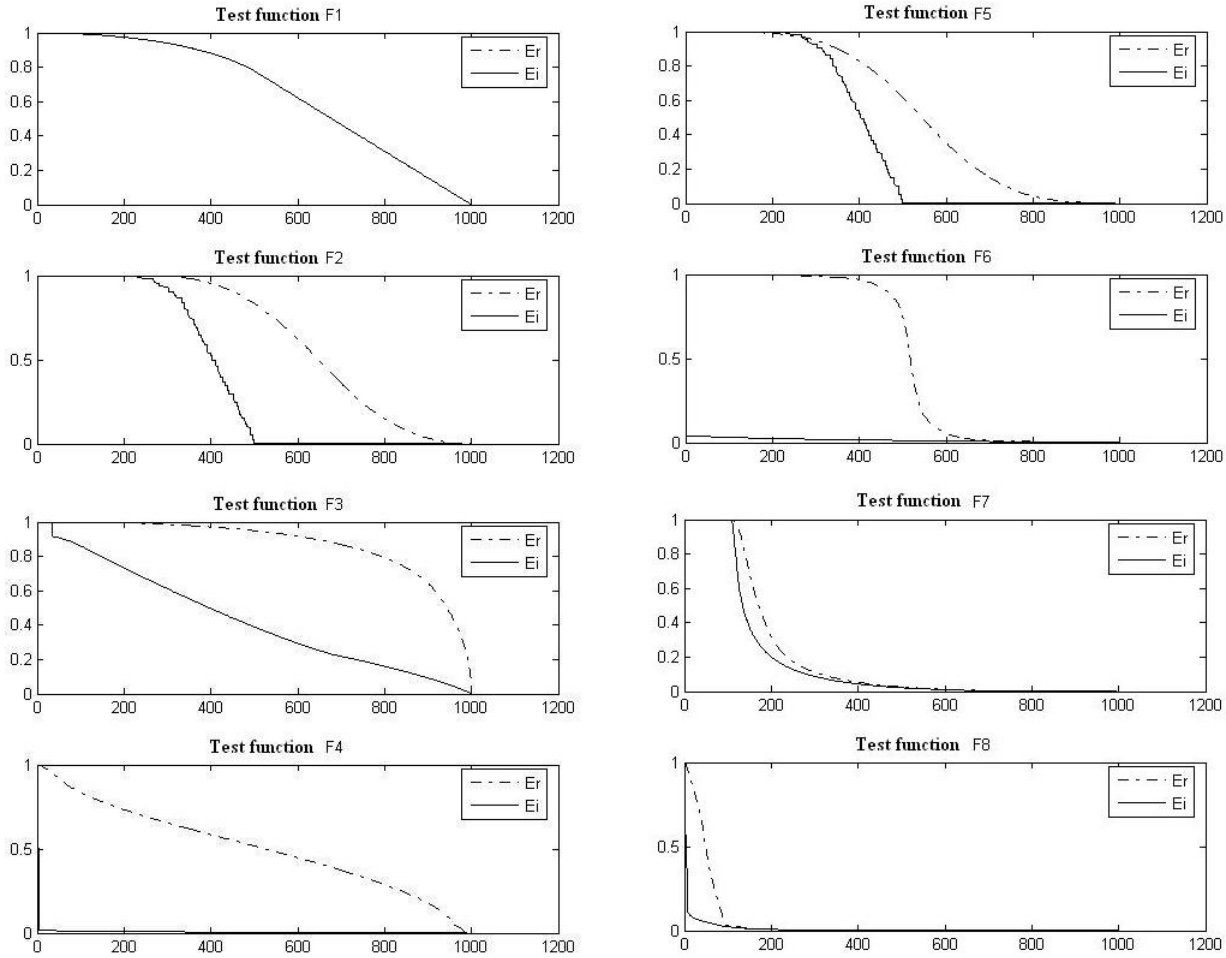

Fig. 4 the Er and Ei curve of test functions

5. Conclusion and Prospect

In this paper, the concepts of exploration and exploitation effective ratio are proposed based on the optimal contraction theorem. The R_{Er} and R_{Ei} curves are constructed and utilized as the tool of estimating optimal ratio. Furthermore, the estimate value of optimal ratio matches the fitness distance correlation and the test results of two optimizers well. Yet, the estimation method still has its disadvantages,

2014 © The American Computational Science Society. All rights reserved.

such as the degree of accuracy of the method is restricted by the size of sample set P. Beside used as the tool of estimating optimal ratio, the R_{Er} and R_{Ei} curves that contain a lot of information of the optimization problem should have extended application in the field of optimization hardness research.

References

[1] D.H. Wolpert, W.G. Macready, IEEE Trans on Evolutionary Computation, 1 (1997) 67-82.

[2] B. Naudts, L. Kallel, IEEE Trans on Evolutionary Computation, 4 (2000) 1-15.

[3] J.W. Li, Research on Fitness Landscapes of Genetic Algorithms and GA-hardness, University of Tianjin (2003).

[4] D.I. Seo, Y.H. Kim, B.R. Moon, Genetic and Evolutionary Computation-Gecco Proceedings, (2003) 1345-1356.

[5] D.I. Seo, S.S. Choi, B.R. Moon, Applications of Evolutionary Computing Lecture Notes in Computer Science, 3103 (2004) 150-161.

[6] D. Corne, M. Oates, D. Kell, Applications of Evolutionary Computing Lecture Notes in Computer Science, 2611 (2003) 187-198.

[7] W. Rowe, D. Corne, J. Knowles, 2006 IEEE Congress on Evolutionary Computation, 1-6 (2006) 2929-2936.

[8] Y. Borenstein, R. Poli, Proceedings of the 2005 conference on Genetic and evolutionary computation, (2005) 1515-1522.

[9] M.S. Gibbs, H.R. Maier, G.C. Dandy, Engineering Optimization, 43 (2011) 349-376.

[10] J. Chen, B. Xin, P. Zhihong, L.H. Dou, IEEE Transactions on Systems Man and Cybernetics Part a-Systems and Humans, 39 (2009) 680-691.

[11] A. Bocanet, C. Ponsiglione, Vine, 42 (2012) 15-35.

2014 © The American Computational Science Society. All rights reserved.