BADE: A New Strategy Based on BA Topological Structure to Improve DE Algorithm

Yifeng wang

Abstract—**Differential Evolution algorithm is an efficient meta-heuristic algorithm, which is highly attractive in the face of practical problems. In the original DE algorithm and its corresponding improved algorithm, all the particles in the process of mutation are treated equally and chosen randomly, while the influence of the topological relation between individuals on the mutation and its subsequent optimization is neglected. In the BADE strategy proposed in this paper, the connection between all individuals will follow the BA scale-free network. In the process of mutation of each individual in each generation of DE, the three individuals required to produce new individuals will no longer be generated by the random equal probability in the rest of the individuals except the ontology; in contrast, only the remaining individuals connected to the ontology are likely to be selected as the mutation factors of the new individuals in each generation process of a new individual. Compared with previous DE algorithm and its typical improved variants, corresponding improved algorithms using BADE strategy, converge more rapidly and more reliably in most test functions. In other test cases this new strategy is also quite competitive.**

I. INTRODUCTION

To solve complex computations, researchers have sought out inspiration for many years from nature as the algorithm prototype. The problem of optimization has long been accompanied by the development of nature. Over the millions of years that nature has evolved and developed, all living organisms have had to adjust their body structures to adapt to their surroundings by mutation and selection, then optimizing the offspring of the entire population and better their propagation. Thus, the evolutionary algorithm, a very calculated and important method, was invented for solving complex search and optimization matter, to handle optimization problems.

The Differential Evolution algorithm (For brevity, we will refer to it as DE algorithm in the following paragraphs), tracing back to 1995, is currently very competitive in the field of evolutionary computation. DE algorithm shows excellent performance for many optimization problems in the real world, and the optimization ability of DE algorithm is largely dependent on the strategy of mutation generated by new individuals in each generation.

Due to the excellent performance and results of the algorithm, its modified versions have been proposed in succession over the course of the next few years.

The key of DE algorithm is how to find and use high quality information more quickly, which is also called the ability of exploration and exploration. When these improved algorithms focus on how to improve the performance of DE algorithm by improving mutation strategy, we focus on the effect of the topological structure of the population on the performance of the DE algorithm.

In nature, the mutated gene is randomly generated and it may occur to any individual of the population, or be attached to any gene of an individual.

But if any individual with the gene mutation has a closer relationship with another one in its population, the mutated gene is more likely to be passed on to its descendant(s). It is likely that, if any gene mutation happens to an isolated individual, regardless of its benefits, it will not be passed on to its descendants.

In accordance with this, we proposed the BADE strategy in which a relationship between the BA scale-free network and each individual of the original population is established. In the BA network, nodes with larger degrees correspond to individuals with greater influence, while nodes with lesser degrees correspond with individuals that have more distant relationships with the other individuals in the network. Only individuals with certain relationship between them, symbolized by an edge between them, can be involved in the generation of a new mutated individual.

The remainder of the paper consists of the following sections: Section II will describe the basic steps of the original DE algorithm and some of its typical variants. The concepts and attributions of some representative DE improvements will also be presented in this section. The BA topological structure will be described in Section III, and in this section, the BADE mutation strategy will be described in detail. The simulation results and their comparison with the original DE algorithm and its variants will be presented in Section IV. Finally, a selection of concluding observations will be presented in Section V.

II.

In this section, basic operations of the DE algorithm, along with a few fundamental terms and nouns, will be introduced. These words, terms, and expressions are efficient for future use of the DE algorithm and its variants.

the original DE algorithm

The DE algorithm is in accordance with the general processes of the evolutionary algorithm. The evolutionary algorithm uses the iteration process that is exemplified by the development or growth of a population. According to evolutionary law, the population is evolving towards an ultimate goal.

The initial population $\{X_{i,0} = (x_{1,i,0}, x_{2,i,0}, ..., x_{D,i,0}) | i = 1, 2, ..., NP\}$ is generated on the basis of the limited random equation of $[x_{j,down} \leq x_{j,i,0} \leq x_{j,up}]$ and is in equal probability, where D denotes the dimension of the problem and NP represents the individual quantity in the population. After initialization, DE algorithm will enter a mutation, crossover and selection evolutionary cycle.

The key to the DE algorithm lies in generating trial parameter vectors. The variants and cross section are both used for generating new trial vectors, while a part of them will be selected to decide which of those vectors will survive and become the next population.

1). Mutation

In the generation g, a variation vector $V_{i,g}$ will be randomly created based on the existing population

$$\{X_{i,g} \mid i = 1,2...,NP\} \qquad (1)$$

$$V_{i,g} = X_{r_1,g} + F_i\left(X_{r_2g} - X_{r_3g}\right) \qquad (2)$$

Here, i denotes the mark of the individual in the population and G denotes the generation of the population.

Meanwhile, r_1, r_2, r_3, and i are different from each other, as well as r_1, r_2, and r_3 are generated randomly from $[1, NP]$. F is a real number for deciding the multiples of the product for the differences between X_{r_2g} and X_{r3g}.

If a trial vector's value is out of the value range of the definitional domain, then this variant will automatically be regarded as the boundary value of the neighboring definitional domain.

2). Crossover

The target vector is formed through the combination of the variation vector and original vector; strategy for generating the trial vector is as follows:

$$U_{i,g+1} = \left(U_{1_i,G+1}, ..., U_{D_i,G+1}\right) \qquad (3)$$

Where

$$u_{ji,G+1} = v_{ji,G+1}, if(randb(j) \le CR)or(j = rnbr(i)) \qquad (4)$$

$$u_{ji,G+1} = x_{ji,G+1}, if(randb(j) \ge CR)or(j \ne rnbr(i)) \qquad (5)$$

$$(j = 1,2,...,D)$$

$randb(j)$ is a real number that is randomly generated in the j-th evolutionary process and randomly distributed between 0 to 1. CR is a cross constant that lies between 0 and 1 and is usually decided by the user. $rnbr(i)$ is a random integer between 1 and D and is used for ensuring at least one element from UiG+1 to ViG+1 can be obtained. Otherwise, the population will not produce changes without the creation of new trial vectors.

3). Selection

Hereby Greedy Selection is used for $j = 1,2,...,D$, and $X_{i,G+1}$ won't be replaced with $U_{i,G+1}$ unless the trial vector $U_{i,G+1}$ generates a better value than $X_{i,G}$ in response to the trial function; if it does not, the original $X_{i,G+1}$ will be retained.

Some representative variants of the DE algorithm

After the DE algorithm is proposed, in order to improve the algorithm's ability of exploration and exploration, there are many representative of the improved algorithm has been proposed. In order to better introduce the BADE strategy proposed in this paper, we only select two algorithms which

are very representative in the process of exploration and exploration because of limited space.

1). DE/best/1"

$$V_{i,g} = X_{best,g} + F_i\left(X_{r_1g} - X_{r_2g}\right) \qquad (6)$$

DE/best/1 is a rather well-known variant based on the original DE algorithm that was put forward by Storn and Price. It contrasts the original DE algorithm in that each of its variants is based on the most excellent individuals found in every generation of the population.

Experimental results show that the DE / best / 1 algorithm's exploration ability are of deficient capabilities and are easy to obtain locally optimal solutions while solving multimodal functions. However, it is embodied with outstanding exploitation capability and faster convergence rate when faced with a unimodal function.

2). "jDE"

The jDE is an adaptive algorithm that is based on the original DE and developed by Brest et al. Similar to most advanced strategies, each generation of the jDE's populations are fixed, however, the control variables F and CR are related to each individual. In the initial stage, F is set for each of the individuals as 0.5, CR as 0.9, while the new F and CR that are created by jDE will average out and randomly distribute the definition domains of $[0.1,1]$ and $[0,1]$ (the probability of the occurrence of these two is 0.1). This is due to fact that the author of jDE believes that the better parameters are easier to generate individuals with more adaptability for survival while also passing these values on to the next generation.

The experimental result shows that the jDE has better effects than the traditional algorithms DE / rand / 1bin, as well as is better than FEP, CEP, LEP, and FADE. An advanced jDE algorithm named jDE-2 is later created with the basis of the jDE and it shows equal competitiveness with SaDE in most test functions

III. PREPARE YOUR PAPER BEFORE STYLING

In this section, basic concept of BA scale-free network and procedures of constructing a BA network will be described. Then based on this, the BADE strategy is introduced in this section.

BA scale-free network

Complex network theory abstracts the actual complex systems into entities consisting of nodes and lines, in which a node represents the studied individual and a line represents the interaction between these individuals. In 1999, Barabási and Albert explored several large-scale network databases such as WWW. The network whose node degrees are according with power-law distribution is called scale-free network. After a long evolution, most nodes in the network will be left with few connecting edges, and few nodes will be left with many edges. Growth and preferential connection in the network play an important role in the process of evolution. So they put forward the famous network evolution model, namely the BA model.

In the whole process of network growth, new nodes will be added to the system constantly. Two basic assumptions for the BA scale-free model: (1) network growth: starts from M0

nodes, and each node will be added to the network after a time step, pick m (m≤m0) nodes in M0 nodes and connect them with the new nodes; (2) preferential connection: the probability for an existing node to be connected to the node i is $\prod i$ and its relationship with the node degree are in accordance with the following equation:

$$\prod i = k_i / \sum k_j \qquad ()$$

After time step t, the algorithm will generate a network consisting of $N = t + m_0$ nodes and m_t edges. This model illustrated that network node is increasing, but connection between the newly added node and existing node(s) is by no means random, it is proportional to the degrees of existing nodes.

Consequently, a network with features of some nodes have a large number of connected edges and most nodes have few connected edges will be formed.

Degree of BA network model follows a power law distribution, and this is proved to exist widely in the real networks.

BADE strategy

In nature, information exchange is completed through interaction. In the DE algorithm, it is completed by variation. Previous improvements to DE algorithm were done by optimizing the process of variation, thus identifying better information in the gene pool of the population (DE / best / 1) and making effective usage of such choice information (jDE).

It is not difficult to find that these improved algorithms are based on the graphs of a all-connected stochastic network (a), that is, In the population, different individuals are connected with each other, specifically in the variation process $V_{i,g} = X_{r_1,g} + F_i\left(X_{r_2,g} - X_{r_3g}\right)$. After $X_{r_1,g}$ is randomly selected, $X_{r2,g}$, and $X_{r3,g}$ are generated at random, and the constraint condition is that these three variables are different from each other.

Here we put forward the BADE strategy which, on the basis of original DE algorithm, correlates all individuals in the DE population and nodes of BA scale-free network one-to-one(b), so that only individuals in the DE population with corresponding connected edges in the BA network can enjoy information exchange, also known as the variation.

Specifically, for the i th individual in the population, the mutation process is $V_{i,g} = X_{r_1,g} + F_i\left(X_{r_2g} - X_{r_3g}\right)$, that is, on the basis of $X_{r_1,g}$ plus $X_{r2,g}$, and $X_{r3,g}$ as a multiple of trial vector, So we have to produce $X_{r_1,g}$ at first, In the process of generating Xr1g, in addition to the original requirements r_1 and i are not the same, only with connected individuals are eligible to be randomly selected as $X_{r_1,g}$ when $X_{r_1,g}$ generated, empathy, only with $X_{r_1,g}$ connected individuals have chance to be randomly selected as $X_{r2,g}$, and $X_{r3,g}$, where i,r_1,r_2,r_3 also meet different each other.
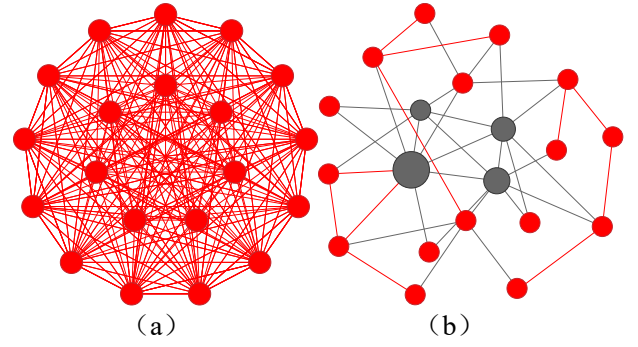

Figure 1. Example of a figure caption. (figure caption)

The advantage of replacing a whole all-connected stochastic network with a BA scale-free network is that, once an excellent gene is generated by mutation, since the position of each individual in the population won't be the same, the rate of diffusion of this information in the BA scale-free network will be improved. In other words, the excellent gene would be more helpful to the evolution of the entire population.

IV. USING THE TEMPLATE

In this section, DE, DE / best / 1, and jDE, which incorporate the BADE strategy, are compared with their initial algorithms. For fair comparison, the original DE algorithm (DE / rand / 1) and DE / best / 1 parameters, we set F = 0.5, CR = 0.9 recommended in [] .Our jDE algorithm follows the previous article's restrictions on jDE parameters.

As shown in Table I, these algorithms are compared in optimizing a set of 13 known functions. Among them, f1-f4 are continuous unimodal functions. F5 is the Rosenbrock function which is unimodal for D=2and 3, but it has multiple minima in higher dimensional cases. F6 is a discontinuous step function, and f7 is a noisy quartic function. F8-f13 are multimodal and the number of their local minima increases exponentially with the problem dimension.

In all the simulation tests, we set the population size NP=100 and D = 30, the number of the generations are set to be the same as that in [] for D = 30. In addition, all results reported in this section are calculated based on 50 independent experiments.

TABLE I.   TEST FUNCTION OF DIMENSION D. EACH OF THEM HAS A GLOBAL MINIMUM VALUE OF 0.

Test Function	*Initial Range*
$f_1(x) = \Sigma_{i=1}^{D} x_i^2$	$[-100,100]^D$
$f_2(x) = \Sigma_{i=1}^{D}\lvert x_i\rvert + \prod_{i=1}^{D}\lvert x_i\rvert$	$[-10,10]^D$
$f_3(x) = \sum_{i=1}^{D}(\sum_{j=1}^{i} x_j)^2$	$[-100,100]^D$
$f_4(x) = \max_i\{\lvert x_i\rvert\}$	$[-100,100]^D$
$f_5(x) = \sum_{i=1}^{D-1}[100(x_{i+1} - x_i)^2 + (x_{i-1})^2]$	$[-30,30]^D$
$f_6(x) = \sum_{i=1}^{D}[x_i + 0.5]^2$	$[-100,100]^D$

$f_7(x) = \sum_{i=1}^{D} i x_i^4 + rank[0,1]$		$[-1.28, 1.28]^D$	
$f_8(x) = \sum_{i=1}^{D} -x_i \sin\sqrt{|x_i|}$ $+ D \cdot 418.98288717243369$		$[-500, 500]^D$	
$f_9(x) = \sum_{i=1}^{D} [x_i^2 - 10\cos(2\pi x_i) + 10]$		$[-5.12, 5.12]^D$	
$f_{10}(x) = -20\exp(-0.2\sqrt{\dfrac{1}{D}\sum_{i=1}^{D} x_i^2})$ $-\exp(\dfrac{1}{D}\sum_{i=1}^{D}\cos(2\pi x_i)) + 20 + e$		$[-32, 32]^D$	
$f_{11}(x) = \dfrac{1}{4000}\sum_{i=1}^{D} x_i^2 - \prod_{i-1}^{D}\cos(\dfrac{x_i}{\sqrt{i}}) + 1$		$[-600, 600]^D$	

f11	500	3.36E-02	2.02E-02
	3000	3.19E-02	7.39E-03
f12	500	1.25E+00	3.75E-01
	1500	2.41E+00	8.53E-01
f13	500	2.82E+00	1.57E-01
	1500	2.46E+00	2.13E-01

$$f_{12}(x) = \frac{\pi}{D}\{10\sin^2(\pi y_1)$$
$$+ \sum_{i=1}^{D-1}(y_i - 1)^2[1 + 10\sin(\pi y_{i+1})]$$
$$+ (y_D - 1)^2\} + \sum_{i=1}^{D} u(x_i, 10, 100, 4)$$

$$f_{12}(x)\ y_1 = 1 + \frac{1}{4}(x_i + 1)\ \&\ u(x_i, a, k, m)$$

$$= \begin{cases} k(x_i - a)^m & x_i > a \\ 0 & -a \le x_i \le a \\ k(-x_i - a)^m & x_i < -a \end{cases}$$

Range: $[-50, 50]^D$

$$f_{13}(x) = 0.1\{\sin^2(3\pi x_1)$$
$$+ \sum_{i=1}^{D-1}(x_i - 1)^2[1 + \sin^2(3\pi x_{i+1})]$$
$$+ (x_D - 1)^2[1 + \sin^2(2\pi x_D)]\}$$
$$+ \sum_{i=1}^{D} u(x_i, 5, 100, 4)$$

Range: $[-50, 50]^D$

As shown in Table II, Table III and Table IV，the mean and the standard deviation of the best-so-far function values are summarized.

TABLE II.　A COMPARISON OF EXPERIMENTAL RESULTS FOR ORIGINAL DE (DE/RAND/1) ALGORITHM

	Gen	DE/rand/1	BA_DE/rand/1
f1	500	2.60E-04	**5.13E-10**
	1500	7.40E-21	**3.26E-40**
f2	500	1.15E-01	**3.51E-05**
	2000	4.94E-12	**5.81E-25**
f3	500	7.89E-02	**2.79E-07**

	5000	5.63E-64	**4.63E-117**
f4	5000	**1.01E-02**	1.03E-02
f5	3000	**2.90E-02**	3.80E-02
	20000	5.93E-30	5.93E-30
f6	100	1.90E+03	**1.29E+02**
	1500	2.71E-16	**1.18E-31**
f7	3000	8.15E+01	**2.14E+01**
f8	1000	1.78E+02	**1.31E+02**
	9000	8.29E+01	**5.92E+01**
f9	1000	1.86E+02	**4.52E+01**
	5000	8.24E+01	**1.69E+01**
f10	500	2.11E-02	**1.83E-05**
	2000	2.27E-12	**1.13E-15**
f11	500	8.83E-03	**3.96E-08**
	3000	0.00E+00	0.00E+00

	Gen	DE/best/1	BA_DE/best/1
f1	500	2.73E-72	5.97E-72
	1500	3.25E-227	5.21E-227
f2	500	3.38E-14	1.71E-14
	2000	1.58E-52	8.59E-58
f3	500	8.10E-54	1.59E-53
	5000	0.00E+00	0.00E+00
f4	5000	1.37E-09	1.37E-09
f5	3000	5.27E-24	1.76E-25
	20000	9.51E-28	9.51E-28
f6	100	2.88E-08	3.69E-08
	1500	7.16E-31	3.85E-31
f7	3000	5.03E+01	6.42E-01
f8	1000	2.85E+03	2.07E+03
	9000	3.12E+03	2.51E+03
f9	1000	7.79E+01	5.70E+01
	5000	7.79E+01	4.22E+01
f10	500	5.56E+00	3.38E+00
	2000	4.31E+00	2.96E+00
f12	500	2.72E-04	**5.36E-10**
	1500	5.03E-17	**3.38E-17**
f13	500	1.63E-03	**4.33E-09**
	1500	2.67E-16	**2.88E-17**

TABLE III.　A COMPARISON OF EXPERIMENTAL RESULTS FOR DE/BEST/1 ALGORITHM

TABLE IV.　A COMPARISON OF EXPERIMENTAL RESULTS FOR JDE ALGORITHM

	Gen	jDE	BA_jDE
f1	500	5.41E-08	**6.11E-09**
	1500	7.70E-33	**3.96E-35**
f2	500	1.63E-04	**6.48E-05**
	2000	1.59E-23	**7.69E-25**
f3	500	1.08E-05	**1.87E-06**
	5000	1.40E-104	**9.72E-111**
f4	5000	1.04E-14	**8.98E-15**
f5	3000	4.82E+00	**4.77E+00**

	20000	3.16E-29	**2.04E-29**
f6	100	5.96E+02	**2.28E+02**
	1500	9.78E-29	**4.01E-30**
f7	3000	8.18E+01	**3.02E+01**
f8	1000	0.00E+00	0.00E+00
	9000	0.00E+00	0.00E+00
f9	1000	5.03E+00	**2.53E+00**
	5000	0.00E+00	0.00E+00
f10	500	2.49E-04	**8.85E-05**
	2000	1.33E-15	1.33E-15
f11	500	2.01E-06	**2.90E-07**
	3000	0.00E+00	0.00E+00
f12	500	9.92E-08	**7.22E-09**
	1500	3.02E-17	3.02E-17
f13	500	9.35E-07	**1.35E-07**
	1500	2.88E-17	2.88E-17

Important of the rate and the reliability of BADE strategy can be made from the results presented in Fig2 and Table II, Table III and Table IV. First of all, these simulations illustrate the use of BADE strategy of the algorithm, whether it is the traditional DE algorithm, or DE / best / 1, or jDE, converge faster in most of the thirteen test function. They are either unimodal or multimodal, clean or noisy, continuous or discontinuous.

Second, it is important to compare the reliability of different algorithm. DE/best/1 is less satisfactory because of premature convergence. The improved DE/best/1 algorithm combined with BADE strategy is shown to work well, as expected, since it is based on the BA scale-free network rather than a whole all-connected stochastic network and thus more robust than usual greedy strategy.

V. CONCLUSION

There has been a great deal of effort to optimize the generation and use of quality information through improved mutation strategies, thereby improving the DE algorithm. In this paper, we propose a BADE strategy that focuses on the effect of the topological structure of a network of individuals on the DE algorithm.

Our proposed BADE strategy is applied to the original DE algorithm and the corresponding two DE algorithms (jDE and DE/best/1). Experimental results show that DE algorithm, DE / best / 1 algorithm and the jDE algorithm combined with of BADE strategy show better convergence than their original algorithms on the standard 13 test functions.

Due to limited space, we do not list the BADE strategy to other DE improved algorithm on the optimization effect. At the same time, the specific optimization mechanism of BADE strategy needs to be further studied.