Neutral theory based mutation operator for Differential Evolutionary algorithms to enhance population diversity

C Wang*, Y C Liu, M Y Xu, X L Liang, Q J Zhang, H H Guo, Y Wei and Y Chen

Marine Engineering College, Dalian Maritime University, Dalian, Liaoning, People’s Republic of China

*Corresponding author: chuanwang0101@163.com

Abstract. As an easily used and powerful heuristic search technique based on population, Differential Evolution (DE) algorithm has been widely applied for various global optimization and real engineering problems. Nevertheless, as with other Evolutionary Algorithms (EA), DE could not avoid from premature convergence due to over concentrated population, which could be called losing population diversity. In order to enhance its performance, we propose a Neutral Mutation (NM) operator for DE algorithm. This novel operator is inspired by neutral theory of molecular evolution, which claims that most mutations at the molecular level are neutral. That is to say, most variations observed are with neither advantage nor disadvantage fitness. Thus, they would not affect an organism’s ability to survive and reproduce. The NM operator maintains slightly deleterious trial vectors, which we called neutral or nearly neutral, with a certain probability in the conventional selection operator of DE. Besides, some of these trial vectors have a chance to be neutrally mutated within the search domain randomly. As a result, the population is diversified with costing negligible Function Evaluations (FEs). Comprehensive experimental results demonstrate that the presented NM operator could improve population diversity to some extent, especially when the population is not divergent at all.

1. Introduction

Since 1997 firstly proposed, Differential Evolution (DE) algorithm has been used on widespread problems such as pattern recognition[1], data mining[2], power system[3], financial market and combinatorial optimization, due to the powerful searching capability and satisfying performance.

Although many achievements have been obtained on DE algorithms, a few weaknesses are still not overcome yet. One of these weaknesses is that the population is gradually losing diversity and gathering in a certain area during optimizing process. Like other population-based methods, successes of the methods depend on their abilities to balance exploration and exploitation. If all the individuals could not find any better solution in fitness landscape, i.e., the population meets a stasis, population diversity decreases to a low level. As a result, global optimum would be rarely found. Obviously, if a modified method could improve population diversity without disruptive effects, then the final performance might be ameliorated. Rare works have improved DE from this aspect. Instead, scholars pay their attentions to adaptive different mutation strategies, control parameters tuning, hybrid framework and other areas. Note that, according to Piotrowski, experimental results on a very simple version of a continuous differentiable and scalable problem show that for some problems the random sampling may indeed perform better than heuristic optimization algorithms. That means that...
population diversity is a key factor to find optimum for some problems, which the heuristic optimization algorithms are not suitable to deal with. Perhaps current DE algorithms already have very strong global and local searching capability, but lack of enough capability to enhance population diversity. Neutral mutation may be a possible solution to this problem.

Inspired by phenomenon in biology, random changes in the DNA do not affect the phenotype due to neutral theory of molecular evolution is introduced to enhance population diversity. According to the theory, mutations do not have significantly better or worse fitness values. However, it might be helpful to diversify population, resulting in increasing potential exploration capability. That is the reason why we present Neutral Mutation (NM) operator. This theory strongly inspires us to develop a neutral mutation and selection operator to mimic protein polymorphism at molecular level, resulting in sustainably divergent population.

The main contribution of this paper is that, we provide a possible way to improve population diversity when the population meets stagnation. In the proposed NM operator, slightly deleterious trial vectors persist into the next generation with a certain probability. Although fitness values of these trial vectors are worse, the population is diversified. Experimental results demonstrate that presented NM operator could improve population diversity to some extent, especially when the population is not divergent at all. Note that, this NM operator is different from another similar operator, which directly generates neutral mutations into the next generation. It is impossible for DE algorithm to do that for problems with unknown fitness landscapes. Besides, the operator lacks of definition of neutrality which is very important for neutral theory of molecular level.

This paper is organized as follows: In the next section, the authors propose the NM operator. Experimental results on test suit are displayed in the next section. Finally, we summarized this article.

2. Neutral mutation operator for DE algorithm
In this part, we will introduce the NM operator is proposed. Besides, two important control parameters of the presented NM operator are investigated. The authors propose a novel mutation operator based on neutral theory of molecular evolution. First, if the trial vector $U_{i,G}$ has better fitness value than its corresponding target vector $X_{i,G}$, the target vector $X_{i,G+1}$ in $G + 1$ generation will be replaced by the better trial vector $U_{i,G}$. This operation could be classified as positive selection compared with biological molecular evolution. Meanwhile, readers should notice that worse trial vector generated by DE is discarded by the above operation, although the detail pseudo codes are omitted. Accordingly, this part could also be classified as negative selection compared with biological molecular evolution. As we claimed, positive selection is rare and negative selection is ubiquitous. In other words, it is difficult to generate trial vectors with better fitness values in the next generation, well it is very common to delete most generated worse trial vectors. That is the reason why DE algorithm will meet stagnation in later searching stage. On the contrary, biological proteins keep phenotypic polymorphism at molecular level. That strongly inspires the authors to design a neutral or nearly neutral mutation operator for DE algorithm to keep population divergent, resulting in higher possibility to find satisfying optima.

In the proposed NM operator, two control parameters, which are neutral mutation rate $\mu$ and selection coefficient $s$, should be investigated. In genetics, the mutation rate is the frequency of new mutations in a single gene or organism over a various amount of time. In this paper, neutral mutation rate indicates $\mu$ the frequency of neutral mutation in vector of DE algorithm. Unlike the mutation operation in GA, we do not directly insert a mutation operator in DE algorithm, because it will destroy the searching process according to our previous experiments. Although neutral mutations at molecular level, such as mutation rate of haemoglobin molecules among different groups of animals, appear at roughly fixed rate, according to the neutral theory of molecular evolution. The authors believe that it might be more appropriate to adjust this parameter, resulting in better performance for DE algorithm. Thus, a dynamic neutral mutation rate tuning strategy is designed as follows:
where, $X_{i,G}$ and $X_{j,G}$ are two randomly selected target vectors from the population. If they are equal, the population may lose its diversity and reach stasis. So, it is reasonable to set the neutral mutation rate as 1 to immediately renew an individual. $\mu_s$ and $\mu_e$ represent mutation rate at start and end point, respectively. FEs denotes function evaluations, and Max_FEs means the maximum number of FEs. $\mu_s$ should be smaller than $\mu_e$. This nonlinear adaption mechanism may guarantee that DE has greater chance to diversify the population at later stage, without disturbing the exploration process in early stage.

Another parameter selection coefficient $s$ scales the definition of neutrality. In population genetics, suppose that there are two genotypes $A$ and $B$ in a population with relative fitnesses $f_A$ and $f_B$, respectively. Then, choosing genotype $A$ as our point of reference, we have $f_A = 1$, and $f_B = 1 + s$, where $s$ measures the fitness advantage (positive $s$) or disadvantage (negative $s$) of $B$. According to Kimura’s definition of neutrality, mutations with $|2Ns| < 1$ or $|s| \leq 1/(2N)$ are defined as neutral, where $N$ is the effective population size and $s$ is the selection coefficient for the mutant heterozygotes. That is to say, it could be named as neutrality that an offspring has a slightly more advantage or disadvantage fitness value than its parent because $s$ is close to 0. In this work, we propose a neutral trial vector remain scheme formulated as follows:

$$X_{i,G+1} = \begin{cases} 
X_{i,G}, & \text{if } f(X_{i,G}) - f(X_{best}) < s \\
X_{s,G}, & \text{otherwise}
\end{cases}$$

where, $X_{tmp}$ represents neutrally mutated target vector, which is randomly re-initialized. $X_{best}$ denotes the target vector with best fitness value so far. This operation will guarantee neutrally mutated vectors survive in the next generation, resulting in a more divergent population.

3. Experimental study on test functions

In this section, the authors are going to conduct comprehensive experiments to test the performance of proposed NM operator. The test suit of CEC 2013 is adopted. More details about the definition of benchmark problems can be found by visiting http://www3.ntu.edu.sg/home/epnsugan/.

In order to test the effectiveness of the proposed NM operator, in this section, we incorporate the proposed operator into some advanced DE variants. To make a fair comparison between the advanced DEs and their corresponding neutral-mutation-based DEs, all the control parameters are adopted as suggestions by their authors in the original literatures. The involved algorithms and parameter settings are shown in Tab. 1.

| Algorithm | Parameters |
|-----------|------------|
| SaDE[4]   | $F \sim N(0.5, 0.3)$, $Cr \sim N(C_{rank}, 0.1)$, $LP = 50$ |
| JADE[5]   | $c = 0.1$, $p = 0.05$, $\mu_F = 0.5$, $\mu_{CR} = 0.5$ |
| jDE [6]   | $F_l = 0.1$, $F_u = 0.9$, $\tau_1 = \tau_2 = 0.1$ |

For most evolutionary algorithms, the population is randomly initialized within the upper and lower limits of optimization problems. Thus, initial population is divergent. However, the individuals
become more and more concentrated within some narrow domains to find more precise solutions after generations. It is inevitable that the population will lose diversity. In some extreme cases, all the individuals will become the same, and no better solutions will be found. That is the reason why we proposed the neutral mutation operator to enhance population diversity. In this section, all methods with neutral mutation operator and their corresponding algorithms are tested with non-diversified initial population to simulate lost diversity. Every initial value of each target vector at each dimension is set as 0.

The mean errors and variances of all DE variants are reported in Tab. 2 when \( D = 30 \), respectively. In order to compare the significance between two algorithms, the Wilcoxon rank sum test is used. All results are averaged over 31 independent runs. The maximum number of FEs (Max_FEs) is \( 3000 \times D \). The population size \( p_s \) is set as 30 for \( D = 30 \).

|         | SaDE   | SaDE-NM | JADE   | JADE-NM | jDE    | jDE-NM |
|---------|--------|---------|--------|---------|--------|--------|
| \( f_1 \) | 7.05E+04  | + 2.96E+03 | 7.05E+04  | + 4.93E+02 | 7.05E+04  | + 9.37E+03 |
|         | ±2.95E-11 | ±8.56E+03 | ±2.95E-11 | ±1.62E+03 | ±2.95E-11 | ±1.21E+04 |
| \( f_2 \) | 7.61E+09  | + 3.52E+06 | 7.61E+09  | + 2.51E+06 | 7.61E+09  | + 2.80E+06 |
|         | ±2.90E-06 | ±1.69E+06 | ±2.90E-06 | ±7.67E+05 | ±2.90E-06 | ±1.91E+06 |
| \( f_3 \) | 1.44E+23  | + 1.44E+09 | 1.44E+23  | + 2.14E+10 | 1.44E+23  | + 7.92E+09 |
|         | ±5.11E+07 | ±2.87E+09 | ±5.11E+07 | ±1.01E+11 | ±5.11E+07 | ±1.56E+10 |
| \( f_4 \) | 2.81E+06  | + 3.11E+04 | 2.81E+06  | + 4.54E+04 | 2.81E+06  | + 3.18E+04 |
|         | ±1.89E-09 | ±5.89E+03 | ±1.89E-09 | ±6.39E+03 | ±1.89E-09 | ±2.15E+04 |
| \( f_5 \) | 1.04E+05  | + 1.46E+03 | 1.04E+05  | + 2.38E+02 | 1.04E+05  | + 6.49E+03 |
|         | ±4.43E-11 | ±5.02E+03 | ±4.43E-11 | ±8.38E+02 | ±4.43E-11 | ±1.85E+04 |
| \( f_6 \) | 2.64E+04  | + 1.53E+02 | 2.64E+04  | + 3.53E+01 | 2.64E+04  | + 1.30E+02 |
|         | ±7.39E-12 | ±3.12E+02 | ±7.39E-12 | ±2.40E+01 | ±7.39E-12 | ±3.02E+02 |
| \( f_7 \) | 3.59E+08  | + 1.06E+02 | 3.59E+08  | + 1.39E+02 | 3.59E+08  | + 2.92E+02 |
|         | ±2.42E-07 | ±5.73E+01 | ±2.42E-07 | ±5.59E+01 | ±2.42E-07 | ±1.24E+03 |
| \( f_8 \) | 2.18E+01  | + 2.10E+01 | 2.18E+01  | + 2.10E+01 | 2.18E+01  | + 2.09E+01 |
|         | ±3.46E-13 | ±5.14E-02 | ±3.46E-13 | ±5.03E-02 | ±3.46E-13 | ±4.54E-02 |
| \( f_9 \) | 6.25E+01  | + 3.11E+01 | 6.25E+01  | + 3.66E+01 | 6.25E+01  | + 3.69E+01 |
|         | ±3.46E-13 | ±5.15E+00 | ±3.46E-13 | ±4.18E+00 | ±3.46E-13 | ±4.24E+00 |
| \( f_{10} \) | 1.55E+04  | + 1.59E+02 | 1.55E+04  | + 1.57E+01 | 1.55E+04  | + 1.31E+02 |
|         | ±5.54E-12 | ±4.21E+02 | ±5.54E-12 | ±6.49E+01 | ±5.54E-12 | ±3.00E+02 |
| \( f_{11} \) | 1.30E+03  | + 8.43E+01 | 1.30E+03  | + 1.66E+02 | 1.30E+03  | + 1.97E+02 |
|         | ±6.93E-13 | ±1.37E+02 | ±6.93E-13 | ±6.75E+01 | ±6.93E-13 | ±1.86E+02 |
| \( f_{12} \) | 1.25E+03  | + 1.69E+02 | 1.25E+03  | + 3.87E+02 | 1.25E+03  | + 1.85E+02 |
|         | ±3.46E-13 | ±5.61E+01 | ±3.46E-13 | ±7.92E+01 | ±3.46E-13 | ±1.07E+02 |
| \( f_{13} \) | 1.33E+03  | + 2.90E+02 | 1.33E+03  | + 3.92E+02 | 1.33E+03  | + 2.51E+02 |
\[ f_{14} \quad \pm 4.62E-13 \quad \pm 1.20E+02 \quad \pm 4.62E-13 \quad \pm 8.22E+01 \quad \pm 4.62E-13 \quad \pm 1.03E+02 \]

\[ w/l \quad 25/3/0 \quad - \quad 25/3/0 \quad - \quad 25/3/0 \quad - \]

\[ \text{Figure 1. Some median convergence curves of compared methods for } D = 30. \]

The main contribution of this paper is that we provide a new way to improve population diversity when the population meets stagnation. Perhaps for current evolutionary algorithms exploration and exploitation capability are already satisfying, but the ability of maintaining population diversity is not
strong enough. It should be noticed that real biological systems at molecular level could keep protein polymorphism, resulting in sustainable evolution. We are not tolerant enough of slightly deleterious offspring, which might be a potential promising solution.

In the future, we will develop a more powerful scheme of neutral mutation operator, which imitates biological evolution better, to diversify population as much as possible. Self-adaptation strategy for $\mu$ and $s$ is also need to be designed. Furthermore, we would like to apply DE variants with neutral mutation operator to some engineering problems, such as parameters identification of shipboard PMSM and PID parameters optimization of frequency converter on shipboard.

Acknowledgments
This work is supported by Natural Science Foundation of China under Contract No. 51709027; Doctoral Scientific Research Foundation of Liaoning Province under Contract No. 20170520265; Natural Science Foundation of Liaoning Province, China under Contract No. 2014025006; Education Department General Project of Liaoning Province, China under Contract No. L2014209.

References
[1] Das S and Sil S 2010 Kernel-induced fuzzy clustering of image pixels with an improved differential evolution algorithm. Inf. Sci. 180 1237-1256.
[2] Das S, Abraham A and Konar A 2008 Automatic clustering using an improved differential evolution algorithm. IEEE Tran. Syst. Man Cy.-S. 38 218-237.
[3] Wang Y, Li B and Weise T 2010 Estimation of distribution and differential evolution cooperation for large scale economic load dispatch optimization of power systems. Inf. Sci. 180 2405-2420.
[4] Qin A K, Huang V L and Suganthan P N 2009 Differential Evolution Algorithm With Strategy Adaptation for Global Numerical Optimization. IEEE Trans. Evol. Comput. 13 398-417.
[5] Zhang J Q and Sanderson A C 2009 JADE: Adaptive Differential Evolution With Optional External Archive. IEEE Trans. Evol. Comput. 13 945-958.
[6] Brest J, Greiner S, Boskovic B, Mernik M and Zumer V 2006 Self-adapting control parameters in differential evolution: A comparative study on numerical benchmark problems. IEEE Trans. Evol. Comput. 10 646-657.