Differential Evolution Enhanced by Combining Group Learning and Elite Learning

Guang-Xu Shen, Jian-Yu Li, Pei-Fa Sun, Sang-Woon Jeon, and Hu Jin*
Dept. Electrical and Electronic Engineering, Hanyang University, Ansan, South Korea
e-mail:{gxshen, lijianyu, pfsun, sangwoonjeon, hjin}@hanyang.ac.kr

***Abstract*—Differential evolution (DE) is fully validated as a feasible algorithm for solving optimization problems. Additionally, for the complex optimization problems with high dimension, the traditional DE suffers from slow convergence. This paper proposes an enhanced DE algorithm that combines group learning and elite learning. The proposed algorithm improves the global search capability while guaranteeing a certain convergence speed. Through extensive experiments we confirm the superior competitiveness of the proposed DE algorithm compared to the traditional ones.**

***Index Terms*—Differential evolution, group learning, elite learning, mutation strategy, optimization**

I. Introduction

In the past few decades, researchers from both academic and industry have shown considerable interest in differential evolution (DE) due to its advantages of being easily understandable and operationally simple [1]. Nowadays, DE has found extensive applications in diverse areas, such as industrial applications or routine daily problem-solving scenarios [2]. However, the optimization problems raised from industry and academic become more complicated [3]. With the gradual increase in the complexity of optimization problems, the challenges brought by high-dimension issues cannot be ignored, especially in terms of algorithm convergence speed [4].

To address such problems, several researchers have started to employ distributed evolutionary algorithms. Zhan *et al.* [5] introduced an adaptive distributed differential evolution algorithm (ADDE). This approach accelerates convergence by facilitating the simultaneous evolution of three populations possessing distinct characteristics. In [6], authors proposed a distributed genetic algorithm (DGA). This approach enhances the algorithm's convergence speed by leveraging multiple processors to execute the evolutionary process concurrently.

Notably, the impact of the mutation process within the DE algorithm on the overall convergence speed is more pronounced compared to other operators [7]. By refining the mutation process, individuals are guided to learn from superior individuals and better-performing groups. This enhancement effectively mitigates the risk of the algorithm becoming trapped in local optima, thereby significantly bolstering the algorithm's overall convergence speed. In response, this paper introduces an enhanced DE algorithm with group learning and elite learning named differential evolution enhanced by combining group learning and elite learning (GEDE).

II. Differential Evolution Enhanced by Combining Group Learning and Elite Learning

A. Differential Evolution (DE)

DE is a heuristic population-based intelligence algorithm originating from the genetic algorithms domain, wherein iterative retention of superior individuals culminates in the attainment of an optimal solution [8]. In DE, individuals of size N constitute a population, and an individual has dimension D, which can be expressed as

$$x_{i,g} = \{x_{i,g,1}, x_{i,g,2}, ..., x_{i,g,D}\}, \tag{1}$$

where i is the individual index from $[1, 2, ..., N]$; g is the number of generations; $x_{i,g}$ represents the ith individual in the gth generation.

Initial population is obtained by

$$x_{i,j,0} = L_j + \text{rand}(0,1) \cdot (U_j - L_j) \tag{2}$$

where U_j and L_j are the upper and lower bounds of the search range in dimension j, respectively; $\text{rand}(0,1)$ is a random number generated uniformly in $[0,1]$.

Subsequently, DE conducts a comprehensive traversal encompassing mutation, crossover, and selection operations.

a) Mutation: The target vectors are operated with the difference vectors according to specific rules to produce new individual vectors with global search capability in each generation. Several common mutation strategies are listed as follows.

- DE/rand/1:

$$v_{i,g} = x_{r1,g} + F \cdot (x_{r2,g} - x_{r3,g}) \tag{3}$$

- DE/best/1:

$$v_{i,g} = x_{best,g} + F \cdot (x_{r1,g} - x_{r2,g}) \tag{4}$$

- DE/rand-to-best/1:

$$v_{i,g} = x_{r1,g} + F \cdot (x_{best,g} - x_{r1,g}) + F \cdot (x_{r2,g} - x_{r3,g}) \tag{5}$$

- DE/current-to-rand/1:

$$v_{i,g} = x_{i,g} + F \cdot (x_{r1,g} - x_{i,g}) + F \cdot (x_{r2,g} - x_{r3,g}) \tag{6}$$

- DE/current-to-best/1:

$$v_{i,g} = x_{i,g} + F \cdot (x_{best,g} - x_{i,g}) + F \cdot (x_{r1,g} - x_{r2,g}) \tag{7}$$

- DE/rand/2:

$$v_{i,g} = x_{r1,g} + F \cdot (x_{r2,g} - x_{r3,g} + F \cdot (x_{r4,g} - x_{r5,g}) \tag{8}$$

- DE/best/2:

$$v_{i,g} = x_{best,g} + F \cdot (x_{r1,g} - x_{r2,g}) + F \cdot (x_{r3,g} - x_{r4,g}) \quad (9)$$

where $v_{i,g}$ is the gth generation of individual after mutation, which is the mutant vector or donor vector; $x_{best,g}$ is the gth generation of individual characterized by the best fitness value; indices $r1$, $r2$, $r3$, $r4$, and $r5$ correspond to unique and randomly selected individuals within the sampled population; F is the scale factor used to control the difference vector.

b) Crossover: This operation randomly crosses $v_{i,j,g}$ and $x_{i,j,g}$ according to a certain probability to make individuals more diverse. This probability is called crossover rate (CR) and is usually set between $[0.3, 0.9]$.

$$u_{i,j,g} = \begin{cases} v_{i,j,g}, & \text{if rand}(0,1) \leq \text{CR or } j = j_{\text{rand}} \\ x_{i,j,g}, & \text{otherwise} \end{cases} \quad (10)$$

where j_{rand} belongs to $\{1, 2, ..., D\}$ to ensure that after the crossover, there must be components of the mutant vector $v_{i,g}$.

c) Selection: By utilizing the fitness values, better individuals are selected between $u_{i,g}$ and $x_{i,g}$, i.e.,

$$x_{i,g+1} = \begin{cases} u_{i,g}, & \text{if } f(u_{i,g}) \leq f(x_{i,g}) \\ x_{i,g}, & \text{otherwise} \end{cases} \quad (11)$$

where $f(x)$ is the fitness value associated with the variable x.

B. Group Learning Approaches

A suitable choice of strategies for different mutations can enhance the traditional DE. Based on DE/rand-to-best/1, we modify the vector of optimal fitness individuals $x_{best,g}$ to take individuals ranked by 20% fitness values and calculate the mean denoted as $x_{mean,g}$ at each generation. The vector distance from $x_{r1,g}$ to $x_{mean,g}$ is controlled with F. The expression of the proposed operation is expressed as

$$v_{i,g} = x_{r1,g} + F \cdot (x_{mean,g} - x_{r1,g}) + F \cdot (x_{r2,g} - x_{r3,g}). \quad (12)$$

The proposed DE with Group Learning (GDE) may have better exploitation capability than the ones with randomly selected individuals. Also, GDE has better global search capability than the best individual influence scheme.

C. Elite Learning Approaches

The convergence speed of an algorithm is an essential metric when evaluating an algorithm, and our approach of incorporating elite learning into GDE aims to improve the convergence speed further. It is notable that the algorithm's complexity does not increase since the individuals in the population are already sorted according to their fitness values in the group learning operation.

Our proposed GEDE algorithm is to randomly select individuals with better fitness values than the target individuals $x_{i,g}$ as elite individuals denoted by $x_{elite,g}$. The expression of the proposed operation is expressed as

$$\begin{aligned} v_{i,g} = & x_{r1,g} + F \cdot (x_{mean,g} - x_{r1,g}) + F \cdot (x_{r2,g} - x_{r3,g}) \\ & + F_e \cdot (x_{elite,g} - x_{r1,g}). \end{aligned} \quad (13)$$

TABLE I: simulation parameters and values

Parameters	Values
Population size, N	100
Variable dimension, D	30
Scale factor, F	0.5
Crossover rate, CR	0.6
Enhanced scale factor, F_e	0.1

The elite learning part of GEDE has a better ability to detach from the local optimum compared to the best individual influence scheme, for which we set F_e to 0.1.

III. Experimental Results

A. Experimental Setup

Typically, the efficacy of an algorithm is assessed using designated test functions. In this study, we evaluated our algorithm using four specific functions: Rastrigin, Weierstrass, Griewank, and Happy Cat. Considering the fairness of the comparison algorithm, we use the $10^4 \times D$ maximal fitness evaluations (maxFEs), in which D is the dimension of the selected function.

The formulas for the selected test functions are listed as follows.

- Rastrigin Function:

$$f_1(x) = 10d + \sum_{i=1}^{d} \left[x_i^2 - 10 \cos(2\pi x_i) \right] \quad (14)$$

- Alpine N.1 Function:

$$f_2(x) = \sum_{i=1}^{d} |x_i sin(x_i) + 0.1 x_i| \quad (15)$$

- Griewank Function:

$$f_3(x) = \sum_{i=1}^{d} \frac{x_i^2}{4000} - \prod_{i=1}^{d} \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1 \quad (16)$$

- Happy Cat Function:

$$f_4(\mathbf{x}) = \left[\left(||\mathbf{x}||^2 - d \right)^2 \right]^{\alpha} + \frac{1}{d} \left(\frac{1}{2} ||\mathbf{x}||^2 + \sum_{i=1}^{d} x_i \right) + \frac{1}{2} \quad (17)$$

We proceed to compare the proposed algorithms, namely GDE and GEDE, against the DE of traditional mutation strategies such as DE/rand/1, DE/best/1, DE/rand-to-best/1, DE/current-to-rand/1, DE/current-to-best/1, DE/rand/2, DE/best/2. To ensure the robustness of our results, we evaluated the above test function 20 times for each algorithm and took the expectation as the result. Table I shows some of the primary parameters used in the experiments.

B. Comparative Analysis with Other DEs

We compare the algorithms named GDE and GEDE proposed in this paper with the traditional DE algorithms employing different mutation strategies. GEDE is an enhanced algorithm that integrates elite learning into GDE. For the results shown in Figs. 1-4, X-axis shows the fitness evaluations

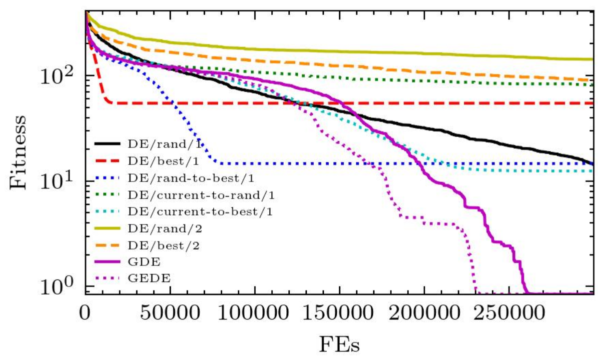

Fig. 1: Fitness values over FE. (Rastrigin)

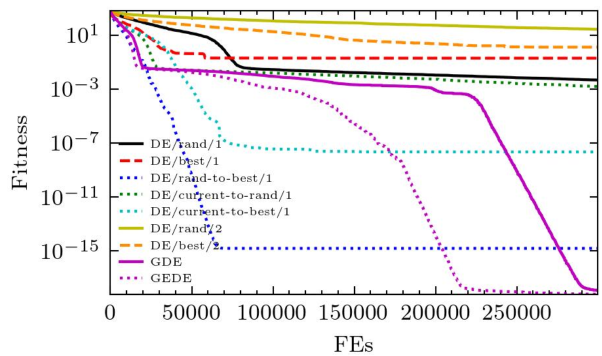

Fig. 2: Fitness values over FE. (Alpine N.1.)

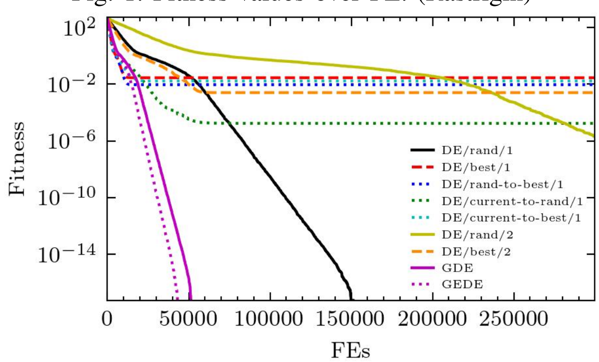

Fig. 3: Fitness values over FE. (Griewank)

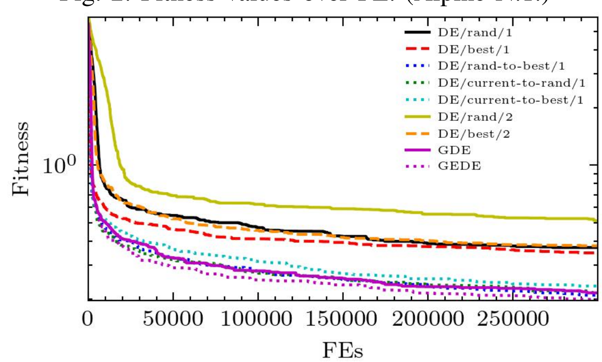

Fig. 4: Fitness values over FE. (Happy Cat)

(FEs) value of the number of fitness functions used, while Y-axis shows the corresponding fitness values. As the generations proceed, each algorithm converges to its optimal solution.

The performance of our proposed algorithm is depicted by the distinctive magenta line. Each algorithm uses the parameters in Table I. For the Griewank and Happy Cat functions, the proposed algorithms have the fastest convergence rate and reach the global optimal solution with the same maxFEs condition. Regarding the Rastrigin and Alpine N.1 functions, our proposed algorithms demonstrate the capability to attain the utmost precise optimal solution within the stipulated maxFEs.

IV. Conclusion

In this paper, for the mutation process of DE, we propose an enhanced differential evolution algorithm that combines group learning and elite learning. The GEDE algorithm is based on the randomly selected individual and the best individual influence scheme. While keeping the factors constant, comprehensive experiments show that GEDE has better global search capability and faster convergence than existing DEs. In future work, We plan to extend our algorithm to be matrix-based DE and apply it to a broader range of high-dimensional optimization problems.

Acknowledgment

This work is supported by the Institute of Information and Communications Technology Planning and Evaluation (IITP) Grant funded by the Korea Government (MSIT, Development of Technology for Securing and Supplying Radio Resources) under Grant 2021-0-00092.

References

[1] K. R. Opara and J. Arabas, “Differential evolution: A survey of theoretical analyses,” *Swarm and evolutionary computation*, vol. 44, pp. 546–558, Feb. 2019.

[2] M. Pant, H. Zaheer, L. Garcia-Hernandez, A. Abraham, *et al.*, “Differential evolution: A review of more than two decades of research,” *Engineering Applications of Artificial Intelligence*, vol. 90, p. 103479, Apr. 2020.

[3] W. Deng, J. Xu, Y. Song, and H. Zhao, “Differential evolution algorithm with wavelet basis function and optimal mutation strategy for complex optimization problem,” *Applied Soft Computing*, vol. 100, p. 106724, Mar. 2021.

[4] Z.-H. Zhan, J. Zhang, Y. Lin, J.-Y. Li, T. Huang, X.-Q. Guo, F.-F. Wei, S. Kwong, X.-Y. Zhang, and R. You, “Matrix-based evolutionary computation,” *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 6, pp. 315–328, Jan. 2021.

[5] Z.-H. Zhan, Z.-J. Wang, H. Jin, and J. Zhang, “Adaptive distributed differential evolution,” *IEEE transactions on cybernetics*, vol. 50, pp. 4633–4647, Oct. 2019.

[6] Z.-J. Wang, Z.-H. Zhan, and J. Zhang, “Solving the energy efficient coverage problem in wireless sensor networks: A distributed genetic algorithm approach with hierarchical fitness evaluation,” *Energies*, vol. 11, p. 3526, Dec. 2018.

[7] A. W. Mohamed and A. K. Mohamed, “Adaptive guided differential evolution algorithm with novel mutation for numerical optimization,” *International Journal of Machine Learning and Cybernetics*, vol. 10, pp. 253–277, Aug. 2019.

[8] R. Storn and K. Price, “Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces,” *Journal of global optimization*, vol. 11, pp. 341–359, Dec. 1997.