AncDE with gaussian distribution for numerical optimization problem

Siti Khadijah Mohd Salleh a,1,*, Siti Azirah Asmai a,2, Zuraida Abal Abas a,3, Abdul Samad Shihghatullah b,4, Diarmuid O’Donoghue c,5

a Department of Information and Communication Technology, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia
b Faculty of Business & Information Science, UCSI University Kuala Lumpur, Kuala Lumpur, Malaysia
c Department of Computer Science, Maynooth University, Co. Kildare, Ireland

1 sitikhadijah@puo.edu.my; 2 azirah@utem.edu.my; 3 zuraidaa@utem.edu.my; 4 AbdulSamad@ucsiuniversity.edu.my; 5 Diarmuid.ODonoghue@mu.ie

* corresponding author

ARTICLE INFO

ABSTRACT

This work is introducing an enhanced Differential Evolution (DE) called AncDE. This proposed algorithm is using an additional population from the current generation and located it as ancestor. There are two parameter controllers to manage the selection of ancestor vector; aup for selection frequency and arp for age of selection. In this work we were applying Gaussian distribution on aup and we tested it on CEC 2015 Numerical Optimization Problem. Standard Differential Evolution will act as the benchmark. The result shows that AncDE with Gaussian approach has produced better result than standard DE.

This is an open access article under the CC–BY-SA license.

1. Introduction

Differential Evolution (DE) was first introduced by Storn and Price [1] in 1995, and has finished third at the First International Contest on Evolutionary Optimization (1st ICEO), in Nagoya, Japan. DE has become more popular among researchers since DE turned out to be the best evolutionary algorithm for solving the real-valued test function suite of the 1st ICEO. Since that, DE has gain much interest from researches due to its problem solving capability but yet having simple code of instructions [2]. However, as any other Evolutionary Algorithms (EAs), DE will stuck in local solution [3] and pretty much rely on parameter controller strategy for mutation and crossover stage [4]. Similar to others EAs, DE employs evolution stages; mutation, crossover, and selection in every generation to generate population to reach global optimum.

During the mutation stage, a mutant vector is generated. Since DE having more than 5 distinct mutation strategies which give different result for different types of problems depending on the most suitable strategy. Therefore, several enhancements have been proposed on mutation strategy to provide more reliable performance as well as be able to optimize the result. JADE as for example used additional population to provide extra information of previous information [5]. Tanaka et al. [6] has improvised JADE by introduced SHADE which uses a historical memory of successful control parameter settings to guide the selection of future control parameter values. Meanwhile mDE-bES was proposed by Ali et al. [7] applied multi-population to boost the diversity for large scale optimization. Furthermore, DPDE was introduced by Zhong et al. used dual population to control the exploration and exploitation capabilities of DE as stated in [8]. Another strategy for parameter control is using probability distribution lead by jDE proposed by Brest et al. [9], using probability to change the value of F and CR to a random
value within the predetermined range. Another proposed strategy called MOSADE inspired by jDE using probability distribution for F and CR with the values are re-sampled at every generation [10]. From here, we note that probability may give positive impact on DE. Yang et al. [11] has introduced NSDE based on generalization of neighborhood search strategy. NSDE use both Gaussian and Cauchy to control the size and shape of the neighborhood. GPDE is another proposed variant of DE is using Gaussian during mutation stage to produce new trial vector [12]. This current work has encouraged us to apply Gaussian as parameter controller on our proposed method called AncDE. In this work, we will use Gaussian on only controller adap during mutation stage. Meanwhile, value of F and CR are initially decided. The reason behind this is to reduce the human intrusion since AncDE has 4 parameters not including size of population NP. We test this proposed work on CEC2015 Numerical Optimization Problem and compared it with standard DE as well as with standard AncDE.

The rest of this paper is organized as follows: in the second section, it will give a brief explanation of standard Differential Evolution. A complete elaboration of AncDE can be found in third section. AncDE with Gaussian distribution is in fourth section. The parameter setting as well as the result of experiment will be presented in the fifth section. Section 6 will draw the conclusion of this paper.

2. Method

2.1. Differential Evolution

Differential Evolution will start with a random population generated according to stated size of NP, \( x_i = [x_{i1}, x_{i2}, x_{i3}, \ldots, x_{iD}] \). Vectors \( x \) will randomly select from the population with \( x_{iG} \) as the target vector during mutation stage to produce a donor vector \( v_i \). DE has five popular variants for mutation as follows:

1) DE/rand/1
\[
v_i = x_{r1} + F \times (x_{r2} - x_{r3})
\]

2) DE/best/1
\[
v_i = x_{best} + F \times (x_{r1} - x_{r2})
\]

3) DE/current-to-best/1
\[
v_i = x_i + F \times (x_{best} - x_{r1}) + F \times (x_{r2} - x_{r3})
\]

4) DE/best/2
\[
v_i = x_i + F \times (x_{r1} - x_{r2}) + F \times (x_{r3} - x_{r4})
\]

5) DE/rand/2
\[
v_i = x_{r1} + F \times (x_{r2} - x_{r3}) + F \times (x_{r4} - x_{r5})
\]

with \( r1, r2, r3, r4 \) and \( r5 \in [1, NP] \) and mutually different. \( F \) is a positive real constant for scaling the difference vector. \( x_{best} \) is the best individual vector with best fitness of the population for current iteration. During the crossover stage, a trial vector is produced with the combination of target and donor vector. Crossover has two option either binomial or exponential, though binomial is widely used as the following scheme:

\[
u_{ij} = \begin{cases} v_{ij}, & \text{if } (rand)[0,1] \leq CR \text{ or } (j = i) \text{ and} \\ x_{ij}, & \text{otherwise} \end{cases}
\]
where CR crossover rate that is any random positive value in range [0,1], rand is a random number between [0,1] and jrand is any random value from {1, 2, ... D} where D is the size of the dimension to ensure the trial vector is different than target vector at least in one dimension. Selection stage will determine either the new trial vector is selected or the current target vector into new population as describe in (7). All stages will repeat until it satisfies the current criteria.

\[ x_{t+1} = \begin{cases} u_{t,j} & \text{if } f(u_{t,j}) \leq f(x_{t,j}) \\ x_{t,j} & \text{otherwise} \end{cases} \]  

(7)

2.2. Ancestor differential evolution - AncDE

In regards to the archive algorithms performance in multi-objective optimization as well to favour the diversity, there are four (4) causes of improvement in natural evolution: selection, mutation, migration and genetic drift. Lolli et al. [13] controversially proposed a 5th cause, that of ancestor-based genetic repair. Building on previous work [14] and [15], this thesis attempts to improve that quality and reliability of results produced by an ancestor driven extension to the DE algorithm. Therefore, we used the ancestor vector as an archive population to extend the diversity in vectors selection. Although MOEAs are popular within multi objective optimization problems, however the approach of how these algorithms solve the problem would have similarities with our proposed Ancestor Differential Evolution (AncDE) that has been using for single objective problems. Because of the similarities between AncDE and a certain MOEAs, hence we took it as our closed referenced to justify AncDE structure. Nevertheless we did not make any statistic comparison between AncDE and MOEAs in this particular test. This is because AncDE is a new approach with ancestor implementation. Therefore, we are focusing into making AncDE more capable to solve problem by applying Gaussian distribution in one of the controllers in mutation stage in this test.

MOEAs promotes diversity by extending their population of solution, thus it will provide proper balance between exploration and exploitation. Since the population size might affect the algorithm performance. A smaller size of population would convergence fast but will cause premature convergence and stagnation. In contrast, a big size population would avoid the premature convergence but it will cost run time [16]. Therefore by having extended population from current population would solve the issue of premature convergence and stagnation but will also provide diversity. One of the most popular MOEA called Pareto Archived Evolution Strategy (PAES) has extended the population by using reference archive of previously found solution in order to identify the approximate dominance ranking of the current and candidate solution vector [17]. Another MOEA algorithm with similar concept was been introduced called the Two-Archive Algorithm, where the first archive (convergence archive) for candidates that non-dominated member in the population and second archive (diversity) for dominated candidates [18]. Meanwhile, Knowles and Corne [19] stated the drawback of EA is this algorithm usually do not keep track of the search history and will evaluated solutions that already visited, thus it caused the population loss diversity value. Therefore, they propose EA with solution archives and bounding extension for detecting branches in the archive only has inferior solution.

AncDE [20] was first inspired by Hatton and Donoghue [21][15] to implement it into DE. The idea is to modify the current generation by using the ancestral template although the evolution process may slowly converge however the result is promising. Taking from that point, AncDE introduced an archive population as repository of selected trial vector from current population. However, the selection of nominated trial vector only can be done if it satisfies a controller called arp. AncDE having arp in order to manage the ancestor replacement thus control the relative age of ancestor vector. This is to avoid of having ancestral cache that too similar of the current population. By having second population has the possibility to increase the solution diversity, therefore, the ancestral vector will be selected during the mutation stage for difference vector calculation. Random ancestral vector is selected from the ancestral with control from aup controller in favour to control the frequency of selected ancestral vector. The selected ancestral vector then calculated with vector from current population.
\[
v_{ij} = \begin{cases} 
    (x_{\text{best}} + F \times (x_{\text{anc}} - x_{\text{best}})), & \text{if } \text{rand}_{ij} < \text{aup}_{ij} \\
    (x_{\text{best}} + F \times (x_i - x_{\text{best}})), & \text{otherwise}
\end{cases}
\]  

(8)

where \(x_{\text{base}}\) is a vector from current population, \(x_{\text{anc}}\) is randomly select from the ancestral population, \(\text{rand}_{ij}\) is any positive random value between \([0, 1]\). If the random value is smaller than \(\text{aup}\), then AncDE will implement standard variant of DE. The rest of AncDE process are similar with standard DE for both crossover with binomial and selection.

### 2.3. AncDE with gaussian distribution

DE has received different attempts of enhancement on mutation strategy, however defining the right value for parameter controller is crucial, and frequently it depending on the user experience in previous problem-task [9]. Eiben et al. [22] has classified parameter controller into 3 category:

1) **Deterministic parameter control**: takes place when the value of parameter is altered by some deterministic rules.

2) **Adaptive parameter control**: is apply in place when there is some form of feedback from the search that is used to determine the direction and/or magnitude of the change to the strategy parameter.

3) **Self-Adaptive Parameter Control**: The idea of the “evolution of evolution” can be used to implement the self-adaptation of parameters. Where the parameters to be adapted are encoded into the vector (individual) and undergo mutation and recombination. The better values of these encoded parameters lead to better individuals, which in turn are more likely to survive and produce offspring and hence propagate these better parameter values.

Meanwhile Eiben and Schut [23] stated most of the work related to parameter setting is focused on variation operators (e.g. mutation) and population size. Chiang et al. [24] in their review has distinguish parameter control in DE into three aspects: (a) the number of candidate parameter values, (b) the number of parameter values used in a single generation, and (c) the source of considered information. For example, ACGDE [25] and FADE [26] falls under (a) category. On the other hand jDE [9] and SaDE [27] are under continued/multiple/population. Not to mention in the previous section examples of methods that using probability as parameter controller. Efren et al. [28] has took another step by determine the performance of modified parameter controller using probability and analyse the behaviour of the parameter. However, it quite difficult to classify DE methods for strategy selection and parameter adaptation, but Das et al. [8] elaborate both under headline: DE methods with both strategy and control parameter adaptation, DE with only control parameter (F and Cr) adaptation, and DE with population size control. Gaussian distribution in the other hand, has been widely used in evolutionary algorithms (EAs) specifically in Differential Evolution. Lu et al. [29] has introduced random-type Gaussian barebones differential evolution (MGBDE) to improve the mutation strategy by employing the population diversity and global search ability. This work is enhancement of the previous GBDE introduced by Wang et al. [30]. Mutation strategy in MGBDE is been randomly determine either Gaussian mutation or DE variant; DE/best/1 during population initialization. Another finding using Gaussian for improvised DE is Gaussian Process assisted Differential Evolution (GPDE) by Su [31], where it has indicate that Gaussian Process (GP) is a newly developed machine learning and it will provide trial vector for the whole DE process. Based on the previous literature, we are motivated to improvise AncDE with only focusing on one of the introduced controller called \(\text{aup}\).

As mention in the previous section, AncDE has add two additional controller which increase the burden of defining best value for each controller. In the previous work, AncDE has run test on various types of variants [32], which the result shows AncDE has performed better using variant with best vector from each iteration. From the result from previous test we take another step on AncDE by proposed an improvised AncDE with Gaussian distribution (normal distribution) to determine the optimal value for only on \(\text{aup}\). We call this proposed algorithm as NAncDE (9).

\[
v_{ij} = \begin{cases} 
    (x_{\text{best}} + F \times (x_{\text{anc}} - x_{\text{best}})), & \text{if } \text{rand}_{ij} < \text{aup}\_ij \\
    (x_{\text{best}} + F \times (x_i - x_{\text{best}})), & \text{otherwise}
\end{cases}
\]  

(9)

Mohd Salleh et al. (AncDE with gaussian distribution for numerical optimization problem)
where \( x_{best} \) is the best vector of the generation, \( x_{alc} \) is random vector from ancestral archive, \( aupN \) is controller \( aup \) that received value from Gaussian distribution. Gaussian distribution will generate a random value with the combination of mean \( \mu \) and standard deviation \( \sigma \). Mean and standard deviation are derived from the previous test on defining the most appropriate value for only \( aup; 0.15, 0.3 \) and \( 0.5 \) as suggested in Sawant et al. [20]. For \( aup \), Gaussian distribution will provide the value for each iteration (10), and NAncDE will follow the standard AncDE on selecting variant as in (9). Algorithm 1 show the procedures stated in NAncDE (Fig. 1).

Algorithm 1. AncDE with Gaussian distribution procedures

1. Initialization phase
   a. Initialize the population
   b. Make a copy of ancestor population
2. Mutation phase
   a. If random_value1 < \( aup = N(\mu, \sigma) \)
      Calculate the difference vector using ancestor vector
   b. Else using standard DE/best/1
3. Crossover (binomial) phase
   Generate trial vector
4. Selection phase
   a. If random_value2 < \( arp \) & trial vector is smaller than target vector
      i. Select trial vector into new offspring
      ii. Select trial vector into ancestor population
   b. Else select target vector into new offspring

Fig. 1. AncDE with Gaussian distribution procedures

3. Results and Discussion

3.1. Parameter setting

All the algorithms; NAncDE, AncDE and DE are tested using a set of standard benchmark functions from special session and competition on real numerical optimization held by IEEE CEC2015 Bound Constraint Single-Objective Computationally Expensive Numerical Optimization Problem [33]. These functions consist of 15 different set of problem under category unimodal (Problem 1 and 2), simple multimodal (Problem 3, 4 and 5), hybrid function (Problem 6, 7, and 8), and composition function (Problem 9, 10, 11, 12, 13, 14 and 15). CEC2015 is divided into 2 dimensions; 10 and 30.

For dimension 10, standard DE will follow the same setting as suggested in Sawant et al. [20] where \( NP = 55, F = 0.55, CR = 0.95 \) and Range = \(+75\). Meanwhile, AncDE are \( NP = 12, F = 0.6, CR = 0.75 \), Range = \(+75\), \( arp = 0.15 \) and \( aup = 0.3 \). Followed by NAncDE with \( NP = 15, F = 0.55, aup = \) Gaussian distribution (mean = 0.317, standard deviation = 0.1756), \( arp = 0.15 \).

Parameter setting for dimension 30 are as follows; AncDE setting \( NP = 25, F = 0.6, CR = 0.6 \), Range = \(+75\), \( arp = 0.15 \) and \( aup = 0.3 \). DE setting \( NP = 55, F = 0.55, CR = 0.95 \) and Range = \(+75\). NAncDE setting \( NP = 25, F = 0.6, CR = 0.6 \), Range = \(+75\), \( aup = \) Gaussian distribution (mean = 0.317, standard deviation = 0.1756), \( arp = 0.15 \). Standard DE is using DE/best/1 variant, and all the three algorithms are using binomial crossover for both 10 and 30 dimensions. All the simulations were done on a 2.4 GHz Intel Core i5 processor with 4GB RAM DDR3.

3.2. Results

Table 1 shows result of mean and standard deviation of DE, AncDE and NAncDE for dimension 10. The result of dimension 30 as shown in Table 2. NAncDE has outperformed other algorithms over 7
functions in Table 1 (Problem 1, 2, 3, 4, 9, 10, and 12). NAncDE is considered equal with AncDE on the other 7 functions (Problem 5, 6, 7, 8, 11, 13, and 14). On the other hand, DE has performed well only on function 15.

Table 1. Dimension 10 between NAncDE, AncDE and DE

<table>
<thead>
<tr>
<th>Function</th>
<th>NAncDE</th>
<th>AncDE</th>
<th>DE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MEAN</td>
<td>STD</td>
<td>MEAN</td>
</tr>
<tr>
<td>F1</td>
<td>2.22E+09</td>
<td>4.19E+09</td>
<td>2.43E+09</td>
</tr>
<tr>
<td>F2</td>
<td>1.50E+07</td>
<td>1.66E+08</td>
<td>1.55E+07</td>
</tr>
<tr>
<td>F3</td>
<td>3.09E+02</td>
<td>3.08E+00</td>
<td>3.09E+02</td>
</tr>
<tr>
<td>F4</td>
<td>1.98E+03</td>
<td>7.04E+02</td>
<td>2.01E+03</td>
</tr>
<tr>
<td>F5</td>
<td>5.04E+02</td>
<td>1.34E+00</td>
<td>5.04E+02</td>
</tr>
<tr>
<td>F6</td>
<td>6.02E+02</td>
<td>1.88E+00</td>
<td>6.02E+02</td>
</tr>
<tr>
<td>F7</td>
<td>7.17E+02</td>
<td>2.69E+01</td>
<td>7.16E+02</td>
</tr>
<tr>
<td>F8</td>
<td>3.71E+04</td>
<td>3.52E+05</td>
<td>3.10E+04</td>
</tr>
<tr>
<td>F9</td>
<td>9.04E+02</td>
<td>3.21E-01</td>
<td>9.04E+02</td>
</tr>
<tr>
<td>F10</td>
<td>6.52E+06</td>
<td>6.23E+07</td>
<td>6.76E+06</td>
</tr>
<tr>
<td>F11</td>
<td>1.13E+03</td>
<td>1.31E+02</td>
<td>1.13E+03</td>
</tr>
<tr>
<td>F12</td>
<td>1.80E+03</td>
<td>1.98E+04</td>
<td>1.95E+03</td>
</tr>
<tr>
<td>F13</td>
<td>1.76E+03</td>
<td>4.19E+02</td>
<td>1.75E+03</td>
</tr>
<tr>
<td>F14</td>
<td>1.61E+03</td>
<td>2.31E+01</td>
<td>1.61E+03</td>
</tr>
<tr>
<td>F15</td>
<td>1.93E+03</td>
<td>1.72E+02</td>
<td>1.92E+03</td>
</tr>
</tbody>
</table>

Table 2 revealed the achievement of NAncDE for 30D. NAncDE has performed better than other methods on 7 function (Problem 3, 4, 6, 7, 11, 12, and 15). We noted that NAncDE has maintained it stability for Problem 3, 4 and 12. While AncDE has performed well on 3 functions (Problem 1, 8, and 13). However, DE has done well on the other 5 functions (Problem 2, 5, 9, 10, and 14). From the result, NAncDE has done pretty well job for composition functions category for both 10D and 30D environment.

Table 2. Dimension 30 between NAncDE, AncDE and DE

<table>
<thead>
<tr>
<th>Function</th>
<th>NAncDE</th>
<th>AncDE</th>
<th>DE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MEAN</td>
<td>STD</td>
<td>MEAN</td>
</tr>
<tr>
<td>F1</td>
<td>1.44E+10</td>
<td>2.05E+10</td>
<td>1.41E+10</td>
</tr>
<tr>
<td>F2</td>
<td>5.36E+05</td>
<td>5.73E+06</td>
<td>5.43E+05</td>
</tr>
<tr>
<td>F3</td>
<td>3.33E+02</td>
<td>7.78E+00</td>
<td>3.34E+02</td>
</tr>
<tr>
<td>F4</td>
<td>7.66E+03</td>
<td>1.01E+03</td>
<td>7.70E+03</td>
</tr>
<tr>
<td>F5</td>
<td>5.05E+02</td>
<td>9.99E+01</td>
<td>5.05E+02</td>
</tr>
<tr>
<td>F6</td>
<td>6.02E+02</td>
<td>9.88E+00</td>
<td>6.02E+02</td>
</tr>
<tr>
<td>F7</td>
<td>7.29E+02</td>
<td>4.52E+01</td>
<td>7.29E+02</td>
</tr>
<tr>
<td>F8</td>
<td>5.51E+06</td>
<td>2.21E+07</td>
<td>5.39E+06</td>
</tr>
<tr>
<td>F9</td>
<td>9.14E+02</td>
<td>2.62E+01</td>
<td>9.14E+02</td>
</tr>
<tr>
<td>F10</td>
<td>5.43E+07</td>
<td>8.51E+07</td>
<td>5.42E+07</td>
</tr>
<tr>
<td>F11</td>
<td>1.25E+03</td>
<td>1.79E+02</td>
<td>1.25E+03</td>
</tr>
<tr>
<td>F12</td>
<td>1.83E+04</td>
<td>1.88E+05</td>
<td>2.15E+04</td>
</tr>
<tr>
<td>F13</td>
<td>1.88E+03</td>
<td>4.14E+02</td>
<td>1.88E+03</td>
</tr>
<tr>
<td>F14</td>
<td>1.74E+03</td>
<td>1.08E+02</td>
<td>1.74E+03</td>
</tr>
<tr>
<td>F15</td>
<td>2.72E+03</td>
<td>3.64E+02</td>
<td>2.72E+03</td>
</tr>
</tbody>
</table>

A further illustration in Fig. 2 for 10D and Fig. 3 for 30D. In Fig. 2, as we can see NAncDE has the closest gap to the solution for Problem 1, 2, 3, 4, 9, 10, and 12 with AncDE stand close behind. Fig. 3

Mohd Salleh et al. (AncDE with gaussian distribution for numerical optimization problem)
also shows the NAncDE converges for Problem 3, 4, 6, 7, 11, 12, and 15 where NAncDE has performed better than other algorithms.

From both results, we believe that we have fulfil our hypothesis; Gaussian probability has improved AncDE in determining value for controller \( a_{up} \). Although, we noted there were many studies has done on parameter controller for F, CR and NP, however, for this particular work we only want to focus on \( a_{up} \). This is because, both \( arp \) and \( a_{up} \) are not a part of standard DE at the first place. These two parameters are only been introduced for the sake of controlling the ancestral cache, hence we are considering to manage and handle \( a_{up} \) first. We do consider to tackle parameter value for \( arp \), F, CR and NP in our next project.

**Fig. 2.** Boxplot for D10 between NAncDE, AncDE and DE
4. Conclusion

Based on the result of this experiment, we may conclude having Gaussian distribution in AncDE to determine the value for aup controller has produced a positive result. Gaussian distribution approach not only gave advantage on AncDE, at once maintaining its ability compared to standard Differential Evolution. Furthermore, this approach has lessened the burden of verifying suitable value for parameter controller, which proven our hypothesis. With this positive result, we might be continuing to improve NAnAncDE in other aspect such as population size or other parameter controller in next work. We do take some consideration on applying NAnAncDE in different field, because currently we only focusing in single objective numerical problems.
Acknowledgment

This work also was fully supported by Faculty of Information and Communications Technology, UTEM. We gratefully acknowledge support from Mr. Donagh Hatton, Mr. Sawant Rushikesh and Computer Science Department, Maynooth University for help with AncDE procedures and resources.

References


