RANDOM NUMBER GENERATOR INFLUENCE ON DIFFERENTIAL SEARCH PERFORMANCE

Wadee Alhalabi[1], Elena Niculina Dragoi[2], Areej Alsaeedi[1]
1 King Abdulaziz University Faculty of Computing and Information Technology, Department of Computer Science, Jeddah, Saudi Arabia
2 "Gheorghe Asachi" Technical University of Iasi, Faculty of Chemical Engineering and Environmental Protection, Department of Chemical Engineering, Blvd. D Mangeron, 700050, Iasi, Romania.

ABSTRACT
Differential Search is an optimization approach that simulates the Brownian like random walk movement of migration organisms. Its inner workings relay on Random Number Generators (RNG) to emulate the probability of the natural driven process and therefore, the properties and performance of the RNG used has an influence on the overall algorithm performance. This study focuses on assessing this influence and on the manner in which the best suited RNG must be selected in order to achieve the optimal results for different types of problems. The results showed that depending on the properties of the problem, different RNGs provide the best solutions. However, as the number of function evaluation rises in report to the problem dimensionality, the influence on the RNG becomes smaller.

KEY WORDS
Optimization, performance, random number generator, differential search.

1. Introduction

In the context of modelling and problem solving, optimization plays a key role. Solving real-world problems via optimization model has two steps: i) constructing an appropriate model and ii) solving the optimization model (which includes variables, constraints, and objective function)[1]. Most of the classic optimization algorithms are based on gradients methods, which have a series of problems related to the high possibility of getting trapped in local optima and to the fact that do not guarantee finding a global optimum solution [2]. Biologically inspired algorithms tend not to have these problems and, in the current work, the influence of different factors on a relatively new optimization algorithm Differential Search (DS) is studied.

DS is a bio-inspired population-based heuristic optimizer proposed by Civicioglu [3]. Its advantages consists in i) ability to search for solutions of multimodal functions; ii) has only two control parameters; iii) is easy to program [4]. The research reported in literature related to this algorithm indicates that it is robust and flexible and can generate high quality solutions within short calculation time [5].

As it can be observed, DS has a great potential. However, due to the fact that is a new algorithm and since its development in 2012 only a few years have passed, extensive studies related to its performance are required in order to assess its strong and weak points from different perspectives. Consequently, in this work, the influence of different random distributions on the efficiency of the DS algorithm is studied in detail. Different types of distributions, in combination with two problem dimensionalities and two types of problems (uni-modal and modal) are tested in order to determine if the characteristics of the problem being solved and the RNG distribution influence significantly the behaviour of DS algorithm.

2. Differential Search

DS was developed to solve the problem of transforming the geocentric cartesian coordinates to geodetic coordinates. However, testing it on different problems showed that it is more efficient that other established approaches. This efficiency is determined by the inner workings of the algorithm, which is based on the Brownian like random walk movements.

The population of the DS is formed from random solutions to the problem being solved and it corresponds to an artificial superorganism migrating (Civicioglu, 2012). The role of this migration is to determine the global optimum of the problem.

The main steps of the algorithm are initialization, evaluation, and migration. Migration is repeated until the stop criteria is reached and includes intermediary steps like: donor selection, stopover site determination, individual selection, stopover site evaluation, and superorganism update.

Initialization is the step in which the initial values of the superorganism are generated based on random number generated values. After that, each individual forming the superorganism is evaluated using a fertility function, equivalent to the fitness function used in the EAs.

The next step consists in selecting the donor using a shuffling function and randomly selected individuals move towards the donor in order to discover new stopover sites.

The new stopover site is evaluated and its fertility is compared to the source of discovery, the superorganism moving to that stopover site if it is more fertile.

The steps of the algorithm described above are presented in Figure 1, where, in order to show the extensive use of random number generators (RNGs), the steps using random numbers are clearly evidentiated in green colour.

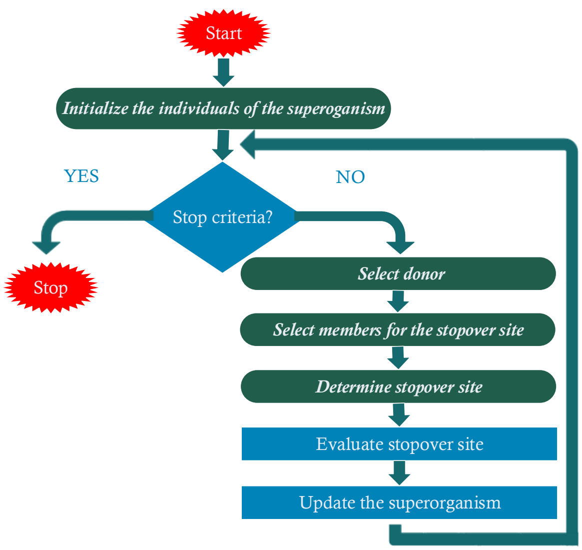

Figure 1. Schema of the DS algorithm

3. Random Number Generators

In this study, the applied RNGs are based on Binomial, Bernoulli, Beta, ChiSquare, Rayleigh and Weibull distributions.

The binomial distribution with parameters n and p is the discrete probability distribution of the number of successes in a sequence of n independent yes/no experiments, each of which yields success with probability p.

Bernoulli is a probability distribution of a random variable which takes the value 1 with success probability of p and the value 0 with failure probability of q=1-p (It is a special case of the binomial distribution).

Beta is a family of continuous probability distributions defined on the interval [0, 1] parameterized by two positive shape parameters, denoted by α and β, that appear as exponents of the random variable and control the shape of the distribution

The sampling distribution (if the null hypothesis is true) can be made to approximate a chi-square distribution as closely as desired by making the sample size large enough. The chi-squared test is used to determine whether there is a significant difference between the expected frequencies and the observed frequencies in one or more categories.

Rayleigh distributions is a continuous probability distribution and represents distribution of the magnitude of a two-dimensional random vector with independent coordinates.

Weibull is a continuous probability distribution which is interpolated between the exponential distribution and the Rayleigh distribution.

4. Case study

In order to assess the influence of the different RNGs on the DE algorithm performance, a set of benchmark functions were used. These benchmark functions are selected from the CEC 2013 special session on real parameter optimization [6]. The CEC-2013 testbed consist of 3 categories of test functions (28 numerical test functions) which are uni-modal functions, multi-modal functions and composition functions. From this set, four benchmark functions were selected: 2 uni-modal (sphere and rotated high conditional elliptic function) and 2 are multimodal (rotated rosenbrock's function and rotated schaffer's function).

The equations for these functions are presented below, where f1 indicates the Sphere function, f2 the Rotated High Conditioned Elliptic function, f3 the Rotated Rosenbrok's function and f4 the Rotated Schaffer F7 function

$$f_1(x) = \sum_{i=1}^{D} z_i^2 \text{ , } z=x-o \qquad (1)$$

$$f_2(x) = \sum_{i=1}^{D} (10^6)^{\frac{i-1}{D-1}} z_i^2, \; z=T_{osz}(M_1(x-o)) \qquad (2)$$

$$f_3(x) = \sum_{i=1}^{D} \left(100\left(z_i^2 - z_{i+1}\right)^2 + (z_i - 1)^2\right),$$

$$z = M_1\left(\frac{2.048(x-o)}{100}\right) + 1 \qquad (3)$$

$$f_4(x) = \left(\frac{1}{D-1} \sum_{i=1}^{D} \left(\sqrt{z_i} + \sqrt{z_i}\sin(50 z_i^{0.2})^2\right)\right)^2,$$

$$z_i = \sqrt{y_i^2 + y_{i+1}^2} \text{ for i=1..D}$$

$$y = \Lambda^{10} M_2 T_{asy}^{0.5}(M_1(x-o)) \qquad (4)$$

5. Results and discussion

For each function and RNG selected a number of 50 simulations were performed. In order to test the influence of RNG the problem dimensionality was set to 10 and 50, and number of functions evaluations to 100000. In addition, the population dimensionality is set to 20. These values were selected based on the consideration that the majority of real-word problems are not large or very large, 50 parameters for optimization in engineering systems representing a large group of problems.

The errors for all the combinations RNG-Function for all the simulation results obtained are listed in Table1 when dimensionality is 10 and Table 2 when dimensionality is 50. The optimal value for all the functions is 0 and therefore, the best solution is considered the one closest to 0.

for F3 and F4, while for F2, the best solution is determined when using a binomial distribution. F3 and F4 are both multi-modal and non-separable, fact that tends to point out that for this type of functions, the DS algorithm has the best performance. For F1, compared with the other functions, all the distributions provided the worst solutions. This indicates that, for this separable function and for the selected setting, DS is not able to determine acceptable solutions.

Civioglu tested the DS algorithm using a series of function, including F1 but with a dimensionality equal to 30 and a number of function evaluation equal to 200000.

Table 1. Simulation results for all the combinations RNG-Function when dimensionality is 10

Function	RNG	Best	Worst	Average	Median	Standard Deviation
F1	Bernoulli	1.12E-25	2.85E-20	1.77E-21	1.25E-24	5.74E-21
	Beta	2.75E+00	4.10E+03	4.15E+02	8.30E+01	7.25E+02
	Binomial	5.02E-26	3.10E-25	1.01E-25	8.80E-26	5.51E-26
	ChiSquare	4.91E-26	2.96E-25	1.14E-25	8.93E-26	7.24E-26
	Rayleigh	1.30E-01	1.92E+02	1.71E+01	4.73E+00	3.69E+01
	Weibull	6.76E-26	3E-25	1.08E-25	9.31E-26	5.94E-26
F2	Bernoulli	1.97E-29	1.06E-24	1.69E-25	5.13E-26	2.28E-25
	Beta	1.41E-02	2.86E+05	1.18E+04	8.67E+02	4.35E+04
	Binomial	5.48E-10	8.66E+01	2.25E+00	1.14E-03	1.22E+01
	ChiSquare	3.19E-09	1.39E+01	4.98E-01	1.31E-03	2.05E+00
	Rayleigh	5.03E-15	9.43E+02	3.25E+01	5.78E-08	1.48E+02
	Weibull	1.97E-29	2.85E-24	4.48E-25	1.85E-25	5.98E-25
F3	Bernoulli	2.47E-29	4.04E-19	1.02E-20	7.35E-26	5.76E-20
	Beta	5.62E-09	3.74E-01	1.12E-02	1.32E-04	5.32E-02
	Binomial	5.98E-28	1.03E-25	4.49E-26	4.43E-26	3.04E-26
	ChiSquare	3.58E-29	9.63E-26	3.68E-26	3.66E-26	3.04E-26
	Rayleigh	2.91E-21	3.29E-06	1.31E-07	1.84E-11	5.37E-07
	Weibull	1.09E-28	1.02E-25	3.83E-26	2.1E-26	3.64E-26
F4	Bernoulli	0.00E+00	1.74E-07	3.47E-09	0.00E+00	2.43E-08
	Beta	2.79E-01	1.03E+01	2.74E+00	2.52E+00	1.94E+00
	Binomial	0.00E+00	2.52E-07	4.87E-08	0.00E+00	8.43E-08
	ChiSquare	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
	Rayleigh	4.53E-04	2.24E+00	1.42E-01	1.13E-02	4.39E-01
	Weibull	0.00E+00	3.71E-07	4.64E-08	0.00E+00	9.45E-08

When problem dimensionality is 10, for F1, the best solution was obtained with Bernoulli distribution and for F2, the best solutions were provided by Bernoulli and Weibull distributions. For F3, the best solutions were generated by Bernoulli and ChiSquared distributions, and for F4, 4 distributions provided the same best solution. This points out to the fact that when sufficient function evaluations (FES) are used, the influence of the RNG becomes smaller.

On the other hand, when the problem dimensionality if 50, by comparing the best solutions obtained for F1, it can be observed that the best solution is obtained with the Weibull distribution. Similar findings are also observed

In these conditions (FES=66666*Dimensionality), the reported solution for DS was in the vicinity of the global optimum. When dimensionality is 50 and FES=2000*Dimensionality, F1 seems to pose some problems for DS, especially when not the appropriate RNG is used.

Regarding the type of distribution, for both dimensionality values, the Beta distribution provides the worst solutions in case of all four functions. These results indicate that the combination of the DS algorithm with this distribution is not favorable. The comparison of convergence speed for F1 and dimensionality is 50 when Beta and Weibull distributions are used (Figure 2). In

Figure 2, for the integrations between 2500 and 5000, as the difference between the two distributions becomes higher, the selection is shown using two y axis, on the left y axis being the Weibull distribution and on the right axis the Beta distribution.

the number of FES used. Consequently, it requires a higher number of FES to reach the vicinity of the global optimum. Although for some problems using a very high FES (in rapport to the dimensionality) is not an issue (as in the case simple benchmark functions), for the majority

Table 2. Simulation results for all the combinations RNG-Function when dimensionality is 50

Function	RNG	Best	Worst	Average	Median	Standard Deviation
F1	Bernoulli	4.06E+01	1.75E+03	3.74E+02	2.88E+02	3.27E+02
	Beta	6.98E+04	2.27E+05	1.46E+05	1.40E+05	3.87E+04
	Binomial	2.72E+03	8.19E+03	4.65E+03	4.43E+03	1.39E+03
	ChiSquare	5.02E+03	2.25E+04	1.42E+04	1.44E+04	4.25E+03
	Rayleigh	1.46E+02	5.87E+03	1.40E+03	1.07E+03	1.08E+03
	Weibull	2.07E-09	4.38E-05	1.06E-06	7.21E-08	6.12E-06
F2	Bernoulli	2.69E-23	2.16E-13	8.32E-15	1.78E-20	3.61E-14
	Beta	5.82E+01	4.82E+07	3.65E+06	1.21E+06	8.45E+06
	Binomial	7.57E-08	2.75E+01	7.44E-01	2.38E-03	3.88E+00
	ChiSquare	1.64E-08	1.17E+02	2.46E+00	3.98E-03	1.64E+01
	Rayleigh	3.31E-18	4.86E+03	1.18E+02	1.01E-08	6.90E+02
	Weibull	1.01E-23	1.59E-22	3.66E-23	2.98E-23	2.59E-23
F3	Bernoulli	4.15E-28	1.13E-15	2.31E-17	8.95E-26	1.58E-16
	Beta	7.81E-03	2.72E+01	3.81E+00	1.59E+00	6.13E+00
	Binomial	2.17E-13	2.27E-05	7.94E-07	2.89E-09	3.62E-06
	ChiSquare	2.19E-13	1.61E-05	4.59E-07	8.01E-09	2.29E-06
	Rayleigh	2.09E-23	8.18E-03	2.05E-04	1.72E-14	1.15E-03
	Weibull	8.08E-28	1.01E-25	3.68E-26	3.26E-26	3.09E-26
F4	Bernoulli	3.30E-06	6.31E-06	4.66E-06	4.77E-06	7.27E-07
	Beta	3.20E+01	6.99E+02	1.26E+02	1.25E+02	1.00E+02
	Binomial	9.09E-02	3.09E+00	1.08E+00	7.56E-01	8.14E-01
	ChiSquare	2.32E-02	4.65E+00	1.15E+00	7.85E-01	9.33E-01
	Rayleigh	4.71E-05	2.89E+01	3.04E+00	5.25E-02	6.47E+00
	Weibull	4.07E-06	7.99E-06	5.43E-06	5.40E-06	7.58E-07

Figure 2 indicates that the problem with the Beta based RNG consists in the fact that the evolution of the solution toward the global optimum is slow in report to

of problems (especially related to real life systems), using a high FES is computationally expensive and unpractical when the time for optimization is high (for example in

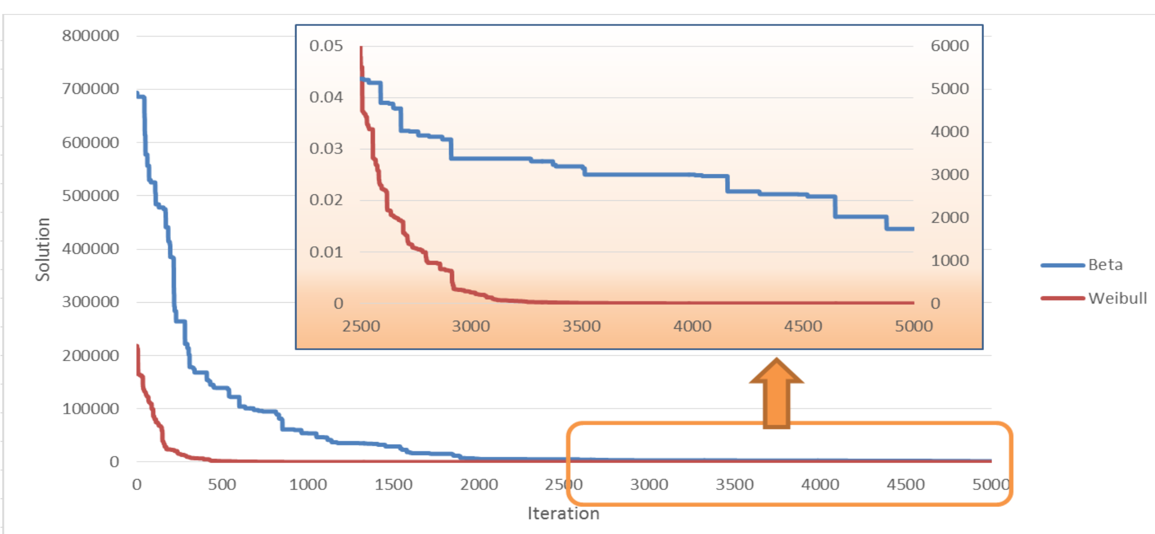

Figure 2. Convergence speed of F1 when dimensionality is 50 for Beta and Weibull distributions

control loops) and the resources available limited.

6. Conclusions

In this work, the performance of the Differential Search algorithm was studied and the influence of different distributions (Bernoulli, Binomial, Beta, ChiSquared, Weibull, and Rayleigh) for random number generators determined. The results obtained showed that when the number of function evaluations in rapport with the problem dimensionality is high, then the influence of the RNG becomes smaller, as there are sufficient iterations for the problem to converge on the global optimum.

On the other hand, when the number of function evaluations is smaller in rapport with the problem dimensionality, the type of RNG used has a big influence on performance, as the convergence speed should be fast enough in order for the solution to be near the global optimum.

Acknowledgement

According to the instructions, the acknowledgment will be introduced after revision.

References

[1] A. Levy, B. College, & B. Maine, *The Basiscs of Practical Optimization*. (Philadelphia: Society for Industrial and Applied Mathematics (SIAM), 2009).

[2] R.Angira, & B.V. Babu, Performance of modified differential evolution for optimal design of complex and non-linear chemical processes, *Journal of Experimental & Theoretical Artificial Intelligence*, 18(4), 2006, 501-512.

[3] P.Civicioglu, Transforming geocentric cartesian coordinates to geodetic coordinates by using differential search algorithm, *Computers & Geosciences*, 46, 2012, 229-247.

[4] L. Gan, & H. Duan, Biological image processing via Chaotic Differential Search and lateral inhibition. *Optik - International Journal for Light and Electron Optics*, 125(9), 2014, 2070-2075.

[5] Y. Amrane, M. Boudour, & M. Belazzoug, A new Optimal reactive power planning based on Differential Search Algorithm. *International Journal of Electrical Power & Energy Systems*, 64, 2014, 551-561.

[6] J.J. Liang, , B. Y.Qu, P. N.Suganthan, A. G Hernandez-Diaz, Problem definitions and evaluation criteria for the cec 2013 special session on real-parameter optimization. *Computational Intelligence Laboratory, Zhengzhou University, Zhengzhou, China and Nanyang Technological University, Singapore*, Technical Report, 2013. No. 201212.