Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Biogeography-Based Optimization Combined with Evolutionary Strategy and Immigration Refusal

Dawei Du, Dan Simon, and Mehmet Ergezer
Department of Electrical and Computer Engineering
Cleveland State University
Cleveland, Ohio, USA
d.du@csuohio.edu, d.j.simon@csuohio.edu, m.ergezer@csuohio.edu

***Abstract*—Biogeography-based optimization (BBO) is a recently developed heuristic algorithm which has shown impressive performance on many well known benchmarks. In order to improve BBO, this paper incorporates distinctive features from other successful heuristic algorithms into BBO. In this paper, features from evolutionary strategy (ES) are used for BBO modification. Also, a new immigration refusal approach is added to BBO. After the modification of BBO, F-tests and T-tests are used to demonstrate the differences between different implementations of BBOs.**

***Index Terms*—Biogeography-based optimization, heuristic algorithm, evolutionary strategy, immigration refusal.**

I. Introduction

The research area of heuristic optimization algorithms has been attracting researchers for many years. In the past 50 years, algorithms have been brought out one after another [1]. Some of them, like genetic algorithms (GAs) and evolutionary strategy (ES), have solved many problems which are difficult to solve using more traditional optimization algorithms. The importance of heuristic algorithms are generally recognized by the engineering research community. More and more researchers turn to heuristic algorithms for different kinds of hard-to-solve problems.

In [2], biogeography-based optimization (BBO), a heuristic algorithm, was first published. After tests on many benchmarks, and comparisons with many other widely used heuristic algorithms like GAs, stud GAs, and others, BBO outperformed most of the other algorithms on most of the benchmarks. This shows that BBO is an algorithm that has much promise and merits further development and investigation.

BBO has shown its ability to solve optimization problems. However, in order to improve this advantage relative to other heuristic algorithms, it is necessary to improve BBO. The goal of this paper is to improve the performance of BBO by adding some features from other algorithms. One feature is borrowed from ES, and the other one is called immigration refusal.

This paper is structured as follows. In the second section, we mainly talk about the BBO algorithm and its performance. In the third section, we discuss how to improve the performance of BBO. The fourth section includes the simulation of the original BBO and modified BBOs. In the fifth section, statistical analysis methods are used to distinguish the performance differences between the original BBO and modified BBOs.

This work was supported by NSF Grant 0826124 in the CMMI Division of the Engineering Directorate.

II. Biogeography-Based Optimization

As its name implies, BBO is based on the science of biogeography. Biogeography is the study of the distribution of animals and plants over time and space. Its aim is to elucidate the reason of the changing distribution of all species in different environments over time. As early as the 19th century, biogeography was first studied by Alfred Wallace [3] and Charles Darwin [4]. After that, more and more researchers began to pay attention to this area.

The environment of BBO corresponds to an archipelago, where every possible solution to the optimization problem is an island. Each solution feature is called a suitability index variable (SIV). The goodness of each solution is called its habitat suitability index (HSI), where a high HSI of an island means good performance on the optimization problem, and a low HSI means bad performance on the optimization problem. Improving the population is the way to solve problems in heuristic algorithms. The method to generate the next generation in BBO is by immigrating solution features to other islands, and receiving solution features by emigration from other islands. Then mutation is performed for the whole population in a manner similar to mutation in GAs.

The basic procedure of BBO is as follows:

1) Define the island modification probability, mutation probability, and elitism parameter. Island modification probability is similar to crossover probability in GAs. Mutation probability and elitism parameter are the same as in GAs.
2) Initialize the population (n islands).
3) Calculate the immigration rate and emigration rate for each island. Good solutions have high emigration rates and low immigration rates. Bad solutions have low emigration rates and high immigration rates.
4) Probabilistically choose the immigration islands based on the immigration rates. Use roulette wheel selection based on the emigration rates to select the emigrating islands.
5) Migrate randomly selected SIVs based on the selected islands in the previous step.

978-1-4244-2794-9/09/$25.00

6) Probabilistically perform mutation based on the mutation probability for each island.
7) Calculate the fitness of each individual island.
8) If the termination criterion is not met, go to step 3; otherwise, terminate.

In 2008, biogeography was applied to engineering optimization [2] for the first time. Fourteen benchmarks were used to test the performance of various heuristic algorithms. The algorithms that were tested include:

1) ACO - ant colony optimization
2) BBO - biogeography-based optimization
3) DE - differential evolution
4) ES - evolutionary strategy
5) GA - genetic algorithm
6) PBIL - population-based incremental learning
7) PSO - particle swarm optimization
8) SGA - stud genetic algorithm

TABLE I
THE PERFORMANCE OF THE HEURISTIC ALGORITHMS. THE BEST PERFORMANCE OF ALL ALGORITHMS ON EACH BENCHMARK IS NORMALIZED TO 100 [2].

	ACO	BBO	DE	ES	GA	PBIL	PSO	SGA
Ackley	205	**100**	178	220	224	325	262	114
Fletcher	1711	109	527	544	632	1947	1451	**100**
Griewank	240	181	576	1081	404	4665	2241	**100**
Penalty #1	**100**	3660	2.67E5	5.47E7	6198	1.65E10	4.05E7	1090
Penalty #2	**100**	4651	3.42E7	4.69E8	8.79E5	2.60E10	1.13E9	4878
Quartic	1.64E4	432	4847	2.50E4	4378	1.57E5	3.51E4	**100**
Rastrigin	541	**100**	502	564	466	798	544	123
Rosenbrock	2012	**100**	418	615	443	2696	558	103
Schwefel 1.2	391	174	1344	1209	186	2091	1742	**100**
Schwefel 2.21	259	109	571	381	249	597	307	**100**
Schwefel 2.22	779	**100**	374	560	468	1297	670	142
Schwefel 2.26	**100**	119	215	174	161	231	188	104
Sphere	1721	115	278	111	751	5196	1445	**100**
Step	279	106	585	1155	530	5595	1580	**100**

The results shown in Table I are the best results after 100 Monte Carlo simulations of each algorithm, where the best performances are normalized to 100 in each row. In Table I, BBO has good performance compared to the other algorithms. Though BBO is outperformed on ten benchmarks, its performances are usually only slightly worse than the winners. BBO performs the best on four benchmarks, the second best on seven benchmarks, and the third best on the other three benchmarks.

III. BBO MODIFICATION

In this section, we propose two modifications to BBO in order to improve its performance. A hybrid evolutionary algorithm is an attempt to combine two or more evolutionary algorithms. This can get the best from the algorithms that are combined together. Each heuristic algorithm has its own advantages with respect to robustness, performance in noisy environments, performance in the presence of uncertain parameters, or performance on different types of problems. At the same time, no algorithm can avoid marginal performance on certain problems. Hybrid algorithms can combine the advantages of each algorithm and avoid their disadvantages.

For example, in [5], ACO, GA and local search (LS) are combined to solve the quadratic assignment problem (QAP). ACO is used to construct a good initial population and provide feedback to the GA. With a well-constructed initial population instead of a random one, the GA is used to solve the QAP. After obtaining the solution using GA and ACO, LS can be used to improve this solution. When these three algorithms are combined, their advantages are combined too. Reference [5] demonstrates that the hybrid evolutionary algorithm gives outstanding performance compared with the algorithms acting separately.

There are also many other examples of hybrid heuristic algorithms. For example, PSO and incremental evolution strategy (PIES) have been combined to solve function optimization problems in Chapter 5 of [6]. The combination of GA and bacterial foraging is another hybrid evolutionary algorithm for solving function optimization problems in Chapter 8 of [6]. Hybrid heuristic algorithms can perform significantly better than a single heuristic algorithm. Because of this advantage, more and more attention has focused on the hybrid evolutionary algorithm field in recent years.

In order to make an improvement to the performance of BBO like the examples mentioned above, in this paper, the aim is to add the features of ES and immigration refusal to BBO to get better performance. In section III-A, we will discuss ES. In section III-B, we will introduce immigration refusal. In section III-C, we will talk about how to use the features borrowed from ES and immigration refusal to modify BBO.

A. Evolutionary Strategy

Evolutionary strategy was created by students at the Technical University of Berlin in the 1960s and 1970s [7]. It is one of the classic optimization techniques among heuristic methods. The basic procedure of the evolutionary strategy algorithm can be described as follows [7]:

1) Define α as the number of parents and β as the number of children.
2) Initialize the population of α individuals.
3) Perform recombination using the α parents to form β children.
4) Perform mutation on all the children.
5) Evaluate K population members, where $K \in [\alpha, \alpha+\beta]$.
6) Out of the K individuals in the previous step, select α individuals for the new population.
7) If the termination criterion is not met, go to step 3; otherwise, terminate.

In step 5, we can evaluate either α + β population members or just the β children. If α = β, and we evaluate all α + β individuals, the probability of getting fitter individuals for the next generation is increased dramatically. Likewise, if this method is used in BBO, the chance of finding the best island can be increased. If we set α = β, and K = α + β, the fitness values of the α parents have already been calculated in the previous generation, so the burden of cost function evaluation does not increase relative to the standard BBO algorithm.

In many realistic problems, the cost function evaluation of the population is computationally expensive. So if we use the α + β option of ES, we significantly increase the probability of finding the best island without introducing more calculation. That is the reason this feature from ES is added to BBO.

B. Immigration refusal

In BBO, the emigration rate and immigration rate are used to determine to where to emigrate and from where to get immigration. Fig. 1 shows the relationships between fitness of habitats (number of species), emigration rate μ and immigration rate λ. E is the possible maximum value of emigration rate to the habitat, and I is the possible maximum value for immigration rate. S is the number of species in this habitat, which corresponds to fitness. S_{max} is the maximum number of species the habitat can support. S_0 is the equilibrium value; when $S = S_0$, the emigration rate μ is equal to the immigration rate λ.

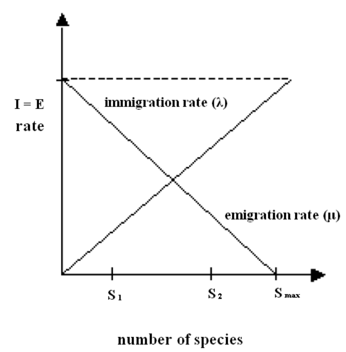

Fig. 1. The relationship of fitness of habitats (number of species), emigration rate μ and immigration rate λ [2]

From Fig. 1, it is clear that the island which has good performance like S_2 has a high emigration rate and a low immigration rate. On the other hand, the island which has poor performance like S_1 has a high immigration rate and a low emigration rate.

In the original BBO, where to emigrate and from where to receive immigration are based on the emigration rate and immigration rate. If the island has a high emigration rate, the probability of emigrating to other islands is high. On the other hand, the probability of immigration from other islands is low. But the low probability does not mean that immigration will never happen. Once in a while a highly fit solution will immigrate solution features from a low-fitness solution. This may ruin the high fitness of the island which receives the immigrants. So when the solution features from an island which has low fitness try to emigrate to other islands, the other islands should carefully consider whether or not to accept these immigrants. That is, if the emigration rate of the island which sends the solution feature is less than some threshold, and its fitness is also less than that of the immigrating island, the immigrating island will refuse the immigrating solution features. This idea, called immigration refusal, is what we add to BBO.

C. Modified BBO Algorithms

First, we borrow the feature from ES. That is, for every generation, we evaluate α + β individuals, where $\alpha = \beta$. Second, we add the immigration refusal approach to BBO. This will decrease the potential harm from low fitness islands. With the combination of these two features, we have modifications to BBO.

1) Original BBO
2) BBO with features borrowed from ES (BBO/ES)
3) BBO with immigration refusal (BBO/RE)
4) BBO with features borrowed from ES and immigration refusal (BBO/ES/RE)

The basic procedure of BBO/ES is as follows:

1) – 5) are the same as 1) – 5) in original BBO described in Section II.
6) Probabilistically perform mutation based on the mutation probability for each child island.
7) Calculate the fitness of each individual island, including both parent and child islands. Store them for the use in the next generation.
8) Based on the features borrowed from ES, select the best n islands from the n parents and n children as the population for the next generation.
9) If the termination criterion is not met, go to step 3; otherwise, terminate.

The basic procedure of BBO/RE is as follows:

1) – 4) are the same as 1) – 4) in original BBO described in Section II.
5) Migrate randomly selected SIVs based on the selected islands in the previous step. When receiving immigration from other islands, use the immigration refusal idea to decide whether or not to accept the immigration.
6) Probabilistically perform mutation based on the mutation probability for each child island.
7) Calculate the fitness of each individual island.
8) If the termination criterion is not met, go to step 3; otherwise, terminate.

The basic procedure of BBO/ES/RE is as follows:

1) – 5) are the same as 1) – 5) in BBO/RE.
6) – 9) are the same as 6) – 9) in BBO/ES.

IV. Simulation

A. Simulation parameters

The simulation parameters that we used are as follows:

1) Number of Monte Carlo simulations: 100
2) Number of islands: 100

3) Number of SIVs per island: 20
4) Generations per Monte Carlo simulation: 100
5) Island modification probability: 1
6) Mutation probability: 0.005
7) Elitism parameter: 1

Note that the number of islands is the population size, the number of SIVs per island is the problem dimension, and the elitism parameter is the number of elite islands saved for the next generation. The threshold of immigration refusal is 0.5. When the emigration rate of the immigration island is larger than 0.5, where emigration rate is normalized from 0 to 1, immigration refusal does not apply. When the emigration rate is less than 0.5, the island only accepts immigration if it comes from an island which has better fitness.

B. Performances of BBOs

Table II shows the performance of the four different BBOs. See [2] for a description of the 14 benchmarks used in this study.

TABLE II
THE BEST PERFORMANCE OF DIFFERENT BBOS AFTER 100 MONTE CARLO SIMULATIONS

	BBO	BBO/ES	BBO/RE	BBO/ES/RE
Ackley	3.56	**1.34**	3.03	1.42
Fletcher	9570.10	4503.96	6216.63	**2248.52**
Griewank	1.40	**1.04**	1.42	1.07
Penalty #1	1.05	0.04	1.10	**0.03**
Penalty #2	4.07	**0.46**	4.56	0.51
Quartic	3.68E-04	**4.81E-06**	2.22E-04	6.33E-06
Rastrigin	1.93	**0.00**	4.04	**0.00**
Rosenbrock	17.83	**12.80**	21.41	13.44
Schwefel 1.2	51.41	**9.52**	28.69	12.10
Schwefel 2.21	680.93	**654.65**	866.16	889.69
Schwefel 2.22	0.80	**0.10**	0.70	**0.10**
Schwefel 2.26	10.70	**8.40**	10.50	9.30
Sphere	0.16	**0.00**	0.12	0.01
Step	62.00	**7.00**	39.00	**7.00**

The differences in performance between the BBOs can be summarized as follows:

1) BBO/ES vs. BBO: With the features borrowed from ES, BBO/ES has a better performance than BBO. This is especially true for the Penalty #1, Penalty #2, Quartic, Rastrigin and Sphere functions, where we see that BBO/ES is more than ten times better than BBO.
2) BBO/RE vs. BBO: BBO/RE outperforms BBO eight times, and BBO performs better than BBO/RE six times, so it is not clear whether or not BBO/RE is better than BBO.
3) BBO/ES/RE vs. BBO: BBO/ES/RE performs better than BBO every time. The improved performance is especially large for the Penalty #1, Quartic, Rastrigin and Sphere functions, where we can see that BBO/ES/RE is more than 10 times better than BBO.
4) BBO/ES outperforms both BBO and BBO/RE every time. BBO/ES/RE outperforms BBO and BBO/RE almost every time, losing only one time. In other words, the features borrowed from ES have a strong effect on BBO, and increase the performance of BBO a lot.
5) BBO/ES/RE outperforms BBO/ES two times, has the same performance three times, and has worse performance nine times.

From the values shown in Table II, the features borrowed from ES give BBO a large improvement, but the effect of the immigration refusal does not have a large impact. So tuning the parameters of immigration refusal is an area for future research.

V. RESULTS ANALYSIS

From Table II, we can see the differences between different kinds of BBOs. But without any statistical analysis, conclusions drawn from Table II are only a personal judgement. In this section, there are two statistical methods used to analyze the differences between different BBOs: F-tests and T-tests.

A. F-tests

The F-test is a statistical test for several groups of numbers, where the number of groups is greater than two. The F-test procedure can be summarized as follows [8]:

1) G is the number of groups of data, where each group is distinguished by some set of independent variables. N is the number of experiments per group.
2) $$\overline{X}_g = \frac{1}{N}\sum_{i=1}^{N} X_{gi}$$
$\overline{X}_g$ is the average value of the dependent variable for group g. X_{gi} is the dependent variable of the i-th experiment for group g.
3) $$\overline{X} = \frac{1}{NG}\sum_{g=1}^{G}\sum_{i=1}^{N} X_{gi}$$
$\overline{X}$ is the average value of the entire population including all groups.
4) $$S_w = \frac{1}{G}\sum_{g=1}^{G}\frac{1}{N-1}\sum_{i=1}^{N}(X_{gi} - \overline{X}_g)^2$$
S_w is the within-group variance.
5) $$S_b = \frac{1}{G-1}\sum_{g=1}^{G}(\overline{X} - \overline{X}_g)^2$$
S_b is the between-group variance.

6) The F-test value is equal to S_b/S_w

The F-test value can be used as follows in order to determine if the differences between the groups of data are statistically significant.

1) The user chooses a probability P. This is the probability that the groups of data are from the same distribution. As an example, if the user wants to have a 99% confidence that there is a statistically significant difference between the groups of data, then P=0.01.
2) Compute the F-test value from the procedure above.
3) The numerator degrees of freedom is defined as $G-1$, and the denominator degrees of freedom is defined as $NG-G$.
4) According to the P value, numerator degrees of freedom, and denominator degrees of freedom, we can find the F-test threshold from tables such as those given in [9].
5) If the F-test value is bigger than the threshold, that means there is a probability P or less that results of the groups are from the same distribution.

In this paper, there are fourteen benchmarks used to test different BBOs. For each benchmark, each of the four BBOs is run for 100 Monte Carlo simulations. That means for each benchmark, each BBO converges to 100 different minimum values.

Table III shows the F-test values for each benchmark, along with the F-test thresholds for the 95%, 99%, and 99.9% confidence levels. If the F-test value is greater than the threshold for some P, that means the BBO differences are statistically significant with a probability of $1-P$ or greater.

TABLE III
F-TESTS VALUES AND THRESHOLDS

	Threshhold			
	P= 0.001	P= 0.01	P= 0.05	F-test value
Ackley	5.51	3.82	2.62	6.23
Fletcher				0.29
Griewank				3.74
Penalty #1				3.65
Penalty #2				0.16
Quartic				1.64
Rastrigin				5.40
Rosenbrock				0.20
Schwefel 1.2				2.90
Schwefel 2.21				0.05
Schwefel 2.22				3.14
Schwefel 2.26				0.29
Sphere				4.04
Step				3.68

The F-tests results can be summarized as follows:

1) When P = 0.05, on Ackley, Griewank, Penalty #1, Rastrigin, Schwefel 1.2, Schwefel 2.22, Sphere, and Step, the F-test values are larger than the threshold. So with 95% (or greater) confidence, on these benchmarks, the differences between the different BBOs are statistically significant and are not from the same distribution.
2) When P = 0.01, on Ackley, Rastrigin, and Sphere, the F-test values are bigger than the threshold. So with 99% (or greater) confidence, on these benchmarks, the differences between the different BBOs are statistically significant and are not from the same distribution.
3) When P = 0.001, only on Ackley, is the F-test value bigger than the threshold. So with 99.9% (or greater) confidence, on this benchmark, the differences between the different BBOs are statistically significant and are not from the same distribution.

B. T-tests

The F-tests give us a certain level of confidence that the results from different BBOs are from different distributions. But the F-tests can not be used to determine which of the four BBOs caused the differences between the four sets of data. We need to use another method in order to isolate the differences in pairs between each type of BBO.

The method we use is the T-test. In 1908, the T-test was invented by William Sealy Gosset, and his pen name was "Student". That is why the T-test is also called the Student's T-test [8].

The T-test procedure can be summarized as follows [8]:

1) N_1 is the number of dependent values in group 1. N_2 is the number of dependent values in group 2.
2) $$\overline{X}_1 = \frac{1}{N_1}\sum_{i=1}^{N_1} X_{1i}$$ $$\overline{X}_2 = \frac{1}{N_2}\sum_{i=1}^{N_2} X_{2i}$$ $\overline{X}_1$ is the average value of group 1 data. X_{1i} represents the i-th datum of group 1. $\overline{X}_2$ is the average value of group 2 data. X_{2i} represents the i-th datum of group 2.
3) Calculate the standard deviations of group 1 and group 2, denoted as S_1 and S_2 .
4) $$S_t = \sqrt{S_1^2 + S_2^2}$$
5) The T-test value is defined as $(\overline{X}_1 - \overline{X}_2)/S_t$

The T-test value can be used as follows in order to determine if the differences between the groups of data are statistically significant [8].

1) Calculate the T-test value as shown above.
2) Calculate the degree of freedom as N_1+N_2-2.
3) Use the T-test value and the degree of freedom to find the P value according to [10]. This is the probability

that the two sets of data come from the same distribution.

Using the calculation steps described above, T-test values for each benchmark can be calculated. For each benchmark, we use three pairs of BBOs to check for statistically significant differences. The pairs that we use are BBO and BBO/ES, BBO and BBO/RE, and BBO and BBO/ES/RE. Using these three pairs, we can find out how significant the differences are between the original BBO and each modified BBO. The T-test values are summarized in Table IV.

TABLE IV
T-TEST VALUES AND P VALUES

	BBO and BBO/ES		BBO and BBO/RE		BBO and BBO/ES/RE	
	P	T-test	P	T-test	P	T-test
Ackley	9.34E-04	3.15	0.25	0.66	1.73E-03	2.96
Fletcher	0.30	0.52	0.39	0.27	0.25	0.67
Griewank	0.01	2.23	0.31	0.49	0.01	2.27
Penalty #1	0.01	2.25	0.20	0.85	0.01	2.26
Penalty #2	0.35	0.37	0.38	0.32	0.35	0.37
Quartic	0.07	1.50	0.32	0.46	0.07	1.50
Rastrigin	1.86E-03	2.94	0.37	0.32	2.27E-03	2.87
Rosenbrock	0.27	0.63	0.41	0.22	0.47	0.73
Schwefel 1.2	0.01	2.37	0.34	0.42	0.01	2.30
Schwefel 2.21	0.47	0.07	0.36	0.36	0.45	0.12
Schwefel 2.22	8.44E-03	2.41	0.34	0.42	9.52E-03	2.36
Schwefel 2.26	0.24	0.71	0.34	0.41	0.20	0.83
Sphere	3.46E-03	2.73	0.32	0.46	3.27E-03	2.75
Step	7.37E-03	2.46	0.28	0.60	6.07E-03	2.53

1) BBO vs. BBO/ES
Only four P values are larger than 0.25. There are ten P values smaller than 0.25. Based on this result, the probability that the results of BBO and BBO/ES are from the same distribution is low.
2) BBO vs. BBO/RE
Only one P value is less than 0.25. It is therefore hard to say that the results of BBO and BBO/RE are from different distributions.
3) BBO vs. BBO/ES/RE
Four P values are larger than 0.25. This result is similar to that of BBO vs. BBO/ES, so the probability that the results of BBO and BBO/ES/RE are from the same distribution is low.

Based on the T-test results, we conclude that using the features from ES has a big effect on BBO, but the effect of using immigration refusal is not that large.

VI. Conclusion

In the first part of the paper, BBO was introduced along with its initial simulation results comparing it with other widely used heuristic algorithms. In order to enhance the performance of BBO, features borrowed from ES and immigration refusal were added to BBO. We use the notation BBO/ES for the combination of BBO and ES. We use the notation BBO/ES/RE for the combination of BBO, ES, and immigration refusal. Fourteen benchmarks were used to test the new BBOs. The modified BBOs performed significantly better than the original BBO. The features borrowed from ES had an especially significant effect as seen in Tables II – IV.

After the benchmark tests, F-tests were used to quantify the confidence level that the results of the different BBOs were from different distributions. According to the F-tests, the probability that the results are from different distributions is high (more than 95%) for over half of the benchmarks.

The T-tests showed that for most of the benchmarks, the BBO/ES and BBO/ES/RE results are from different distributions than the BBO results. According to the simulation results in Section V, and with the evidence from the T-tests, BBO/ES and BBO/ES/RE have performance that is statistically significantly better than the original BBO. In this paper, immigration refusal does not have an equivalent enhancement compared to the use of features from ES.

For future work, first of all, we will tune the parameters in immigration refusal, then test the modified BBO using different values of immigration refusal parameters. Second, we will tune the parameters in the feature borrowed from ES, making sure that it can work in the most efficient way. Our third aim is to extend this work to a more comprehensive set of benchmarks. The combination of features from other heuristic algorithms with BBO is also a direction that should be considered. Another direction for future work is how to apply BBO to real applications like the traveling salesman problem, power distribution problem, or message sending problem. Finally, future analysis of our BBO modifications along the lines of [11] can give us theoretical insight into their advantages and disadvantages.

References

[1] J. Pearl, Heuristics: Intelligent Search Strategies for Computer Problem Solving, Addison-Wesley, 1983.
[2] D. Simon, "Biogeography-Based Optimization," IEEE Transactions on Evolutionary Computation, Vol. 12, No. 6, pp. 702 – 713, December 2008.
[3] A. Wallace, "The Geographical Distribution of Animals (Two Volumes)," Adamant Media Corporation, 2005.
[4] C. Darwin, The Origin of Species, Gramercy, 1995.
[5] L. Tseng and S. Liang, "A hybrid metaheuristic for the quadratic assignment problem," Computational Optimization and Applications, Vol. 34, No. 1, pp. 85 – 113, 2006.
[6] C. Grosan, A. Abraham and H. Ishibuchi, Hybrid Evolutionary Algorithms: Methodologies, Architectures, and Reviews, Springer Berlin/Heidelberg, 2007.
[7] H. Beyer and H. Schwefel, Natural Computing, Kluwer Academic Publishers, 2002.
[8] J. Kennedy and R. Eberhart, Swarm Intelligence, Morgan Kaufmann Publishers, 2001.
[9] C. Wild and G. Seber, Chance Encounters: A first course in data analysis and inference, John Wiley & Sons, 2000.
[10] G. Hill, "Algorithm 395: Students t-distribution," Communications of the ACM, Vol. 13, pp. 617 – 619, October 1970.
[11] D. Simon, M. Ergezer and D. Du, "Markov Analysis of Biogeography-Based Optimization," unpublished. Available online at http://academic.csuohio.edu/simond/bbo/markov.