Biogeography Migration Algorithm for Traveling Salesman Problem

Hongwei Mo[1,2],Lifang Xu[3]

[1] National Defence Key Laboratory of Autonomous Underwater Vehicle Technology,
[2] Automation College, Harbin Engineering University,
150001 Harbin, China,
honwei2004@126.com,
[3] Engineering Training Center, Harbin Engineering University,
150001 Harbin, China,
mxlfang@163.com

Abstract. Biogeography-based optimization algorithm(BBO) is a new kind of optimization algorithm based on Biogeography. It is designed based on the migration strategy of animals to solve the problem of optimization. In this paper, a new algorithm-Biogeography Migration Algorithm for Traveling Salesman Problem(TSPBMA) is presented. Migration operator is designed. It is tested on four classical TSP problems. The comparison results with the other nature inspired optimization algorithms show that TSPBMA is a very effective for TSP combination optimization. It provides a new way for this kinds of problem.

Keywords: Biogeography; Biogeography-based optimization; Biogeography migration algorithm; Traveling salesman problem.

1   Introduction

In recent years, we have seen that many algorithms inspired by natural phenomenon or mechanisms. In this paper, we mainly focused on such a new kind of algorithm, which is called biogeography based optimization, which is inspired by the science of biogeography. It is very interesting in that it mimics the migration process of animals to design method for solving engineering problems, especially optimization. The science of biogeography can be traced to the work of nineteenth century naturalists such as Alfred Wallace [1] and Charles Darwin [2]. In the early 1960s, Robert MacArthur and Edward Wilson began working together on mathematical models of biogeography. Since their distinct work, biogeography has become a major area of research[3].Mathematical models of biogeography describe how species migrate from one island to another, how new species arise, and how species become extinct. The term "island" here is used descriptively rather than literally. That is, an island is any habitat that is geographically isolated from other habitats.

In view of this, Simon presented the first paper on biogeography inspired algorithm for engineering[4], which is called biogeography based optimization(BBO). In his creative work, he merged the burgeoning field of biogeography with engineering in order to see how the two disciplines can be of mutual benefit. Although the idea of application of biogeography to engineering is similar to those nature inspired algorithms mentioned above, it has completely different mechanisms and

process from those ones. It is again to prove the great power of nature. In the past two years, Simon and the other authors had published several papers about BBO. In the first paper on BBO, Simon introduced the main idea of how to use biogeography to design an optimization algorithm and gave us the basic definitions, steps of algorithms. The experiments results showed that BBO is indeed effective in solving these problems. In[5], Simon simplifies the original BBO in order to analyze its theory. They present a simplified version of BBO and then analyze its population using probability theory. Their analysis provides approximate values for the expected number of generations before the population's best solution improves. In[6], they develop a Markov analysis of BBO, including the option of elitism. Their analysis gives the probability of BBO convergence to each possible population distribution for a given problem. Analytical comparisons on three simple problems show that with high mutation rates the performance of GAs and BBO is similar, but with low mutation rates BBO outperforms GAs. In[7],Simon et al. propose a novel variation to biogeography-based optimization (BBO), which employs opposition-based learning (OBL) alongside BBO's migration rates to create oppositional BBO (O B BO). They mathematically prove that it has the highest expected probability of being closer to the problem solution among all OBL methods. In [8],in order to improve BBO, Du et al. incorporate distinctive features from other successful heuristic algorithms into BBO. F-tests and T-tests are used to demonstrate the differences between different implementations of BBOs. In[9], Ma et al. generalize the equilibrium species count in biogeography theory, explores the behavior of six different migration models in BBO. Their study shows that sinusoidal migration curves provide the best performance among the six different models. In[10], Bhattacharya et al. use BBO to solve the problem of economic load dispatch problem. In[11], the BBO is combined with quantum to produce a new kind of hybrid algorithm.BBO is not used for TSP problems since it was presented. In this paper, we use the idea of BBO to solve TSP.

The paper is organized as follows. Section II reviews the ideas of biogeography. Section III introduces the model and algorithm of TSPBMA. Section IV provides some simulation results of TSPBMA compared with other optimization algorithms for TSP. Section V presents some concluding remarks and suggestions for further work.

2   Biogeography and TSPBMA

2.1   Biogeography

In geography, geographical areas that are well suited as residences for biological species are said to have a high habitat suitability index (HSI). Biogeography is nature's way of distributing species, and is analogous to general problem solutions. Suppose that we are presented with a problem and some candidate solutions. A good solution is analogous to an island with a high HSI, and a poor solution represents an island with a low HSI. High HSI solutions resist change more than low HSI solutions. By the same token, high HSI solutions tend to share their features with low HSI solutions. The shared features remain in the high HSI solutions, while at the same time appearing as new features in the low HSI solutions. Poor solutions accept a lot of

new features from good solutions. This addition of new features to low HSI solutions may raise the quality of those solutions. The immigration and emigration curves shown in Fig. 1 as straight lines are a simple model. It illustrates a model of species abundance in a single habitat [12]. The immigration rate λ and the emigration rate μ are functions of the number of species in the habitat. The equilibrium number of species is S_0, at which point the immigration and emigration rates are equal [3][13].
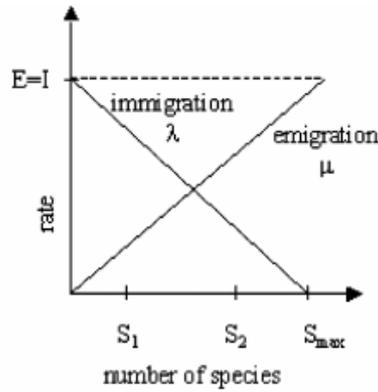

Fig. 1. The model of immigration rate and emigration rate of biology[4].

2.2 TSPBMA Model

Now, consider the probability P_s that the habitat contains exactly S species. P_s changes from time t to time $t + \Delta t$ as follows[1]:

$$P_s(t + \Delta t) = P_s(t)(1 - \lambda_s \Delta t - \mu_s \Delta t) + P_{s-1}\lambda_{s-1}\Delta t + P_s \mu_{s+1}\Delta t \qquad (1)$$

where λ_s and μ_s are the immigration and emigration rates when there are S species in the habitat. This equation holds because in order to have S species at time , one of the following conditions must hold:

   1) There were S species at time t, and no immigration or emigration occurred between t and $t + \Delta t$;

   2) There were $S - 1$ species at time t, and one species immigrated;

   3) There were $S + 1$ species at time , and one species emigrated.

Δt is small enough so that the probability of more than one immigration or emigration can be ignored.Taking the limit of (1) as $\Delta t \to 0$ gives equation (2) shown at the bottom of the page.

   Define $n = S_{max}$ and $P = [P_0, P_1 ... P_n]^T$, for notational simplicity. Now, the P_s equations is   arranged (for $S = 0, 1, ... n$)) into the single matrix equation

$$\dot{P} = AP \tag{2}$$

where the matrix A is given as (6). For the straight line curves shown in Fig. 1, we have

$$\mu_k = \frac{Ek}{n} \tag{3}$$

$$\lambda_k = I\left(1 - \frac{k}{n}\right) \tag{4}$$

$$\dot{P} = \begin{cases} -(\lambda_s + \mu_s)P_s + \mu_{s+1}P_{s+1}, S = 0 \\ -(\lambda_s + \mu_s)P_s + \lambda_{s-1}P_{s-1} + \mu_{s+1}P_{s+1}, 1 \leq S < S_{max} - 1 \\ -(\lambda_s + \mu_s)P_s + \lambda_{s-1}P_{s-1}, S = S_{max} \end{cases} \tag{5}$$

$$A = E\begin{bmatrix} -(\lambda_0 + \mu_0) & \mu_1 & 0 & ... & 0 \\ \lambda_0 & -(\lambda_1 + \mu_1) & \mu_2 & ... & ... \\ ... & ... & ... & ... & ... \\ ... & ... & \lambda_{n-2} & -(\lambda_{n-1} + \mu_{n-1}) & \mu_n \\ 0 & ... & 0 & \lambda_{n-1} & -(\lambda_n + \mu_n) \end{bmatrix} \tag{6}$$

It is easy to prove that the steady-state value for the probability of the number of each species is given by

$$P_k = \begin{cases} P_0 = \dfrac{1}{1 + \sum\limits_{l=1}^{n} \dfrac{\lambda_0 \lambda_1 ... \lambda_{l-1}}{\mu_1 \mu_2 ... \mu_l}}, k = 0 \\ P_k = \dfrac{\lambda_0 \lambda_1 ... \lambda_{k-1}}{\mu_1 \mu_2 ... \mu_k\left(1 + \sum\limits_{l=1}^{n} \dfrac{\lambda_0 \lambda_1 ... \lambda_{l-1}}{\mu_1 \mu_2 ... \mu_l}\right)}, 1 \leq k \leq n \end{cases} \tag{7}$$

2.3 TSPBMA Algorithm

In BBO,each individual is considered as a "habitat" with a habitat suitability index (HSI), which is similar to the fitness of EAs, to measure the individual. A good solution is analogous to an island with a high HSI, and a poor solution indicates an island with a low HSI. High HSI solutions tend to share their features with low HSI solutions. Low HSI solutions accept a lot of new features from high HSI solutions.In TSPBMA, each individual has its own immigration rate λ and emigration rate μ. A good solution has higher μ and lower λ, vice versa. The immigration rate and the emigration rate are functions of the number of species in the habitat.

They can be calculated as (3) and (4).Note that Eqns. 2 and 3 are just one method for calculating λ and μ. There are other different options to assign them based on different specie models [1].

The main steps of TSPBMA are shown as follows:

```
Initialize  m_max , generation number N, generate the
initial path habitats H^n randomly;
Evaluate the HSI for each path individual in H^n;
Sort the population from best to worst based on cost;
Initialize the generation counter t = 1;
While the halting criterion is not satisfied do
     For each path individual, map the HSI to the number
     of species;
      Calculate  the  immigration  rate  λ_i  and  the
       emigration rate μ_i for each path H(i);
      Calculate P_s ;
        If  rand< λ_i and  μ_i <sum( μ_i );
         Select a habitat H(i) ;
         Crossover( H(i) , H(i+1) );
         End if;
      End for
Calculate mutation rate m_i;
Mutate each habitat in H^n with mutation rate m_i ;
Evaluate H^n  update HSI;
Sort H^n according to cost;
Keep the first two best individuals;
t = t + 1;
End while;
```

It begins by computing the immigration and emigration rates of each path habitat
Then, path habitat modification is performed on each habitat. The emigration and
immigration rates of each solution are used to probabilistically share information
between habitats. In TSPBMA, it is carried out by the operation of selection and
crossover. According to the model curve of λ and μ in Fig.1, we suppose that
the better solution should be with lower λ and higher μ than the other emigration
rates μ of the other solutions to decide whether $H(i)$ should be selected or not. If
its μ value is bigger than the sum of the other habitats, then it must be better solution
than the other ones. Based on the two selection steps, if $H(i)$ is selected, it is used
to crossover with its neighbor $H(i+1)$ to share its good information and produce
more diverse habitats. The crossover used here is regular cycle crossover. This step is
different from that of GA because two parents are selected to crossover to produce
offspring in the population of GA. And the crossover operation is separated from
selection operation. In TSPBMA, only one habitat will be selected based on its

λ and μ. And crossover and selection are in one step, that is, they cannot be separated here. This step mimics the migration phenomena happening in biology. During the migration process, information will be shared among different habitats when animals migrate from one habitat to another. We use inversion mutation on both poor solutions and good solutions.

The mutation rate m_i is the mutation rate that is calculated as

$$m_i = m_{max}(1 - \frac{p_i}{p_{max}})$$ (8)

where m_{max} is a user-defined parameter, and $P_{max} = \text{argmax} \, P_i$, $i = 1,...NP$. With the migration operator, TSPBMA can share the information among solutions. Especially, poor solutions tend to accept more useful information from good solutions. This makes TSPBMA be good at exploiting the information of the current population. Additionally, the mutation operator tends to increase the diversity of the population.

3   Benchmark Results

In order to explore the benefits of BBO, we compared its performance on some classical TSPs with five other population-based optimization methods, including ACO,GA,PSO,IA,Fish Swarm(FS)[14]. The benchmarks are Oliver30, Eil50,and Eil75 in TSPLIB.
The parameters of TSPBMA are: habitat modification probability=1, maximum immigration and migration rates=1 for each habitat.
We did some rough tuning on each of the optimization algorithms to get reasonable performance, but we did not make any special efforts to fine-tune the algorithms. For ACO, we used the following parameters: initial pheromone value $\tau_0 = 1E - 5$,pheromone update constant $Q = 20$, exploration constant $q_0 = 1$, global pheromone decay rate $\rho_g = 0.9$, local pheromone decay rate $\rho_t = 0.5$, pheromone sensitivity $\alpha = 2$, and visibility sensitivity $\beta = 6$. For the GA, we used roulette wheel selection, single point crossover with a crossover probability of 0.3, and a mutation probability of 0.1. For PSO, we used initial and ending inertia weight 0.9 and 0.3 respectively, and a social constant 0.7 for swarm interaction. For IA, we used single point crossover with a crossover probability of 0.7, and a mutation probability of 0.07. We use try number 100, sense distance 6, crowd factor $\sigma = 0.5$ for FS.

In our test, the distance is considered only as the cost, to find the best path $T = (t_1, t_2,..., t_n)$ is the same as to make following targeting function minimum

$$f(T) = \sum_{i=1}^{n-1} d(t_i, t_{i+1}) + d(t_n, t_1)$$ (9)

t_i is the numbering of the city which is a natural number between 1 and N, $d(t_i, t_j)$ denotes the distance from city i to city j and for symmetrical TSP $d(t_i, t_j) = d(t_j, t_i)$. In our TSPBMA, we sort the habitats according to the cost $f(T)$, and recalculate their immigration rate λ and emigration rate μ.

Each algorithm had a population size of 50, an elitism parameter of 2, and ran 1000 generations and 20 times to get average results. Fig.2 to Fig.5 show the results of the simulations. For the space limited, we only give the results of TSPBMA on the KroA100 cities and also the comparison results with the other algorithms.
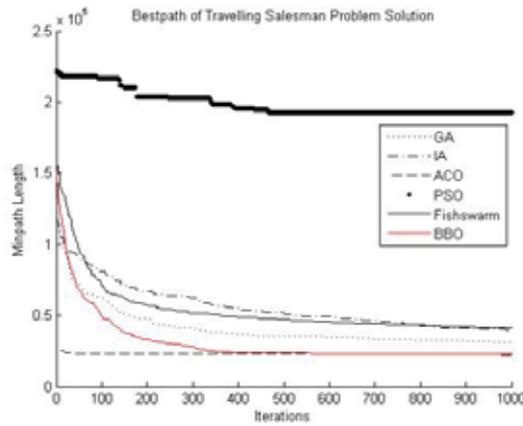

Fig. 2. The comparison results of TSPBMA with the other algorithms on KroA 130 cities

In Fig2, the comparison results of TSPBMA with the other five algorithms are shown. We can see that TSPBMA is much better than classical GA, IA,PSO and FS both in solution quality and converging speed. It can find the best result at 400 generation. But it is worse than ACO in the converging speed.
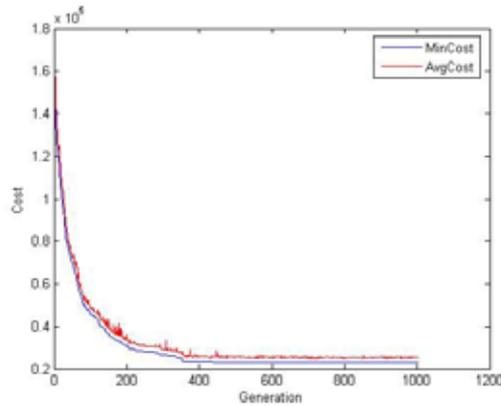

Fig.3. The average and minimum length of TSPBMA for KroA100

In Fig3, we give the average and min length gained by TSPBMA for KroA100 cities. In fact, for the other several problems, it has similar performance. The average results changed slowly with generations. It means that TSPBMA is stable in the searching process and can find the relative good path in each generation without too many trials in population.
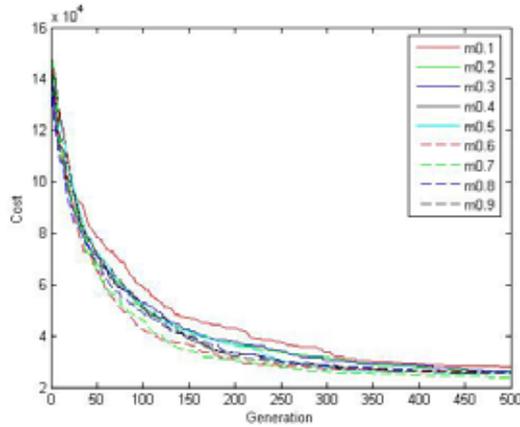
Fig.4. The performance with the change of mutation rate of TSPBMA

The only parameter defined by user is the mutation rate(m_{max}). And we get good results for m_{max} =0.3-0.9. So it is not sensitive to this parameter. It can be seen in Fig.4. This is a good performance that many other nature inspired algorithms do not own.
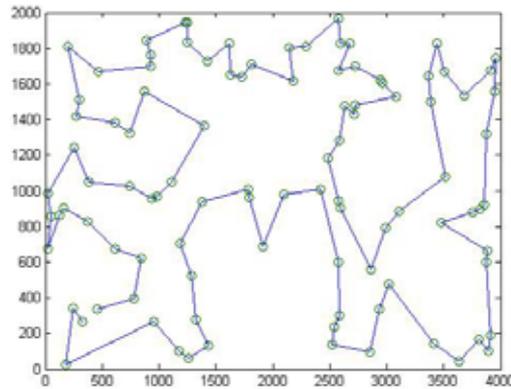
Fig.5. The best path found by TSPBMA for KroA100 cities

In Fig5, the best path of KroA100 cities is shown.

Table 1. TSPBMA results compared with other methods

Benchmark	TSPBMA	ACO	GA	IA	FS	PSO	Best Result
Oliver30	420	420	425	442	430	520	420
	422	422	426	453	442	552	
Eil50	425	425	N/A	453	442	530	425
	425	424	428	464	451	554	
Eil75	535	535	N/A	576	561	675	535
	536	535	545	583	572	684	
KroA100	21282	21282	N/A	22322	21923	64919	21282
	21282	21282	21761	22435	22067	66635	

The problems provided in TSPLIB for algorithm testing are in two expression forms. One is integer distance and another is real number distance between two cities. We use integer one here. The results in the parenthesis refer to average generations of 20 times running. N/A means the result cannot be gained.

In Table1, the comparison results of TSPBMA with the ACO, GA, IA,PSO and FS are shown. We report the best integer tour length. Results using GA are from[15]. The best results for these problems are included in TSPLIB. In total, the behavior of TSPBMA with respect to stability, convergence, equilibria, and other issues are better than the other nature inspired algorithms.

4   Conclusion

We have shown how biogeography, the study of the geographical distribution of biological species, can be used to derive algorithms for TSP combination optimization. This new algorithms is called TSPBMA. We have applied TSPBMA to benchmarks oliver30,eil50,eil75, KroA130 cities. And the results showed that it has better performance than most other classical nature-inspired methods, including converging speed, less parameters and robust to parameters. The results show that TSPBMA is a good method and has great potential in solving combination optimization problems including TSP. It provides a new way for researchers to solve similar problems in future. And it also proves that the idea of biogeography can be used to solve real engineering problem, such as TSP. But when compared with the proposed method, ACO was as successful as it in terms of solution quality and significantly better than it in terms of speed of convergence. In future, we will focus on improving our algorithm's performance including converging speed and solution quality on combination optimization problems and solving more complex ones in further.

References

1. Wallace, A.: The Geographical Distribution of Animals (Two Vol-umes). Adamant Media Corporation, Boston, MA (2005)
2. Darwin, C.: The Origin of Species. Gramercy , New York (1995)
3. I. Hanski and M. Gilpin, Metapopulation Biology. New York: Academic(1997)
4. Simon, D. Biogeography-based optimization. IEEE Transactions on Evolutionary Computation 12, pp. 702 - 713 (2008)
5. Simon, D.: A Probabilistic Analysis of a Simplified Biogeography-Based Optimization Algorithm. Http://academic.csuohio.edu/simond/bbo/simplified/bbosimplified.pdf
6. Simon, D., Ergezer, M., Du, D.: Markov Analysis of Biogeography Based Optimization. http://embeddedlab.csuohio.edu/BBO
7. Ergezer, M., Simon, D., Du, D.W.: Oppositional Biogeography-Based Optimization. IEEE Conference on Systems, Man, and Cybernetics, San Antonio, TX, pp. 1035-1040, (2009)
8. Du, D., Simon, D., Ergezer, M.:Biogeography-Based Optimization Combined with Evolutionary Strategy and Immigration Refusal 2009 IEEE International Conference on Systems, Man, and Cybernetics. San Antonio, TX, pp. 1023-1028(2009)
9. Ma, H., Chen, X.:Equilibrium Species Counts and Migration Model Tradeoffs for Biogeography-Based Optimization.48th IEEE Conference on Decision and Control.(2009)
10. Bhattacharya, A. , Chattopadhyay, P.K.: Solving complex economic load dispatch problems using biogeography-based optimization, Expert Systems with Applications, 37, pp.3605–3615(2010)
11. Li ,X., Guo, T.:Quantum and Biogeography based Optimization for a Class of Combinatorial Optimization, 2009 World Summit on Genetic and Evolutionary Computation pp.969-972(2009)
12. Wesche, T., Goertler, G., Hubert, W. :Modified habitat suitability index model for brown trout in southeastern Wyoming. North Amer. J. Fisheries Manage., vol. 7, pp. 232–237, (1987)
13. MacArthur, R, Wilson, E.:The Theory of Biogeography. Princeton Univ. Press, Princeton, NJ (1967)
14. Li.,X.L, Shao,Z.J.Qian,J.X: An Optimizing Method Based on Autonomous Animats: Fish-Swarm Algorithm. Theory and Practice of System Engineering,.11, pp.33-49 (2002)
15. Wang,Q.,Xiong, L., and Liu,H.Y.:Improved Particle Swarm Algorithm for TSP Based on the Information Communication and Dynamic Work Allocation. Asian Journal of Information Technology,5(11), pp.1191-1196(2006)