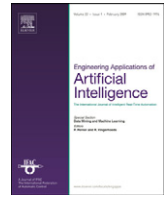




ELSEVIER

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Analysis of migration models of biogeography-based optimization using Markov theory

Haiping Ma^{a,*}, Dan Simon^b^a Department of Electrical Engineering, Shaoxing University, Shaoxing, Zhejiang, China^b Department of Electrical and Computer Engineering, Cleveland State University, Cleveland, Ohio, USA

ARTICLE INFO

Article history:

Received 26 November 2010

Received in revised form

28 March 2011

Accepted 30 April 2011

Available online 20 May 2011

Keywords:

Biogeography-based optimization

Evolutionary algorithms

Migration model

Markov chain

Population distribution

ABSTRACT

Biogeography-based optimization (BBO) is a new evolutionary algorithm inspired by biogeography, which involves the study of the migration of biological species between habitats. Previous work has shown that various migration models of BBO result in significant changes in performance. Sinusoidal migration models have been shown to provide the best performance so far. Motivated by biogeography theory and previous results, in this paper a generalized sinusoidal migration model curve is proposed. A previously derived BBO Markov model is used to analyze the effect of migration models on optimization performance, and new theoretical results which are confirmed with simulation results are obtained. The results show that the generalized sinusoidal migration model is significantly better than other models for simple but representative problems, including a unimodal one-max problem, a multimodal problem, and a deceptive problem. In addition, performance comparison is further investigated through 23 benchmark functions with a wide range of dimensions and diverse complexities, to verify the superiority of the generalized sinusoidal migration model.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Mathematical models of biogeography describe the immigration and emigration of species between habitats. Biogeography-based optimization (BBO) was first presented in 2008 (Simon, 2008) and is an extension of biogeography theory to evolutionary algorithms (EAs). BBO has demonstrated good performance on various unconstrained and constrained benchmark functions (Du et al., 2009; Ergezer et al., 2009; Ma and Simon, 2010). It has also been applied to real-world optimization problems, including sensor selection (Simon, 2008), groundwater detection (Kundra et al., 2009), satellite image classification (Panchal et al., 2009), and power system optimization (Rarick et al., 2009). See reference Gardner and Simon (2009) for a web-based BBO graphical user interface. Like other EAs (Ahn, 2006; Schwefel, 1995; Yao et al., 1999), BBO is based on the idea of probabilistically sharing information between candidate solutions based on their fitness values. In BBO, each solution is comprised of a set of features. Each solution immigrates features from other solutions based on its immigration rate, and emigrates features to other solutions based on its emigration rate. In the original BBO

paper (Simon, 2008), a linear migration model is used for the sake of simplicity. In Ma (2010) and Ma et al. (2009) more complicated and life-like migration models are presented to give better optimization results. These research provided empirical evidence of the potential benefit of alternative migration models of BBO. However, as with most other EAs, there are limited theoretical results for BBO.

Markov models have been a valuable theoretical tool to analyze EAs, including simple genetic algorithms (Davis and Principe, 1993; Nix and Vose, 1992; Reeves and Rowe, 2003; Suzuki, 1995, 1998) and simulated annealing (Lundy and Mees, 1986). Markov models have already been derived for BBO (Simon et al., 2009, 2010), along with Markov model comparisons between BBO and genetic algorithms (Simon et al., 2011). A Markov chain is a random process, which has a discrete set of possible state values s_i ($i=1,2, \dots, T$). The probability that the system transitions from state s_i to s_j is given by the probability p_{ij} , which is called a transition probability. The $T \times T$ matrix $P=[p_{ij}]$ is called the transition matrix. A Markov state in Simon et al. (2010) represents a BBO population distribution. Each state describes how many individuals at each point of the search-space are there in the population. Probability p_{ij} is the probability that the population transitions from the i th population distribution to the j th population distribution in one generation. Although the BBO Markov model is established and some useful results are obtained, there have not been any reports in the literature to

* Corresponding author.

E-mail addresses: Mahp@usx.edu.cn (H. Ma), d.j.simon@csuohio.edu (D. Simon).

analyze the influence of migration models based on Markov theory. This paper uses Markov models to study the effect of migration models on the performance of BBO.

Section 2 gives an introduction to BBO, provides its Markov chain model and presents a new generalized sinusoidal migration model. Section 3 compares various migration models based on the Markov chain model and simulation results. Some concluding remarks and directions for future work are provided in Section 4.

2. Markov chains for biogeography-based optimization

This section presents an overview of the BBO algorithm (Section 2.1), provides its Markov chain model (Section 2.2), and discusses different migration models of BBO (Section 2.3).

2.1. Biogeography-based optimization

In this section a general presentation of the BBO algorithm is given. Consider an optimization problem with a certain number of candidate solutions. A good solution is analogous to a habitat with a high habitat suitability index (HSI). This corresponds to a geographical area that is well suited for hosting biological species in biogeography. In optimization problems, HSI corresponds to a measure of the goodness of a BBO solution, which is also called fitness. In the following text the term fitness instead of HSI is used to be consistent with standard EA notation. A poor solution is like a habitat with a low fitness. High fitness solutions correspond to habitats with a large number of species, and low fitness solutions correspond to habitats with a small number of species. High fitness solutions are more likely to share their features with other solutions, and low fitness solutions are more likely to accept shared features from other solutions. This approach to solve general optimization problems is called biogeography-based optimization (BBO). Similar to all other EAs, BBO consists of two main steps: information sharing (which is implemented with migration in BBO) and mutation.

Migration is a probabilistic operator that improves a candidate solution y_k . The migration rates of each solution are used to probabilistically share features between solutions. For each solution y_k , the immigration rate λ_k is used to probabilistically decide whether or not to immigrate. If immigration is selected, then the emigrating solution y_j is selected probabilistically based on the emigration rate μ_j . Migration is denoted by

$$y_k(s) \leftarrow y_j(s) \tag{1}$$

where s is a solution feature, equivalent to a gene in GAs. Here, immigration rate λ and emigration rate μ are based on a particular migration model, such as the linear model presented in Simon (2008). Additional details about migration models are discussed in Section 2.2.

Mutation is a probabilistic operator that randomly modifies a solution feature. The purpose of mutation is to increase diversity among the population. For low fitness solutions, mutation gives them a chance of enhancing the quality of solutions, and for high fitness solutions, mutation is able to improve them even more than they already have.

A description of one generation of BBO is given in Table 1. Migration and mutation of the entire population take place before any of the solutions are replaced in the population, which requires the use of the temporary population z in the algorithm.

2.2. Markov chain model

In Simon et al. (2010) a BBO Markov chain model is derived. This subsection reviews this Markov model. A Markov model of

Table 1

One generation of the BBO algorithm. y is the entire population of candidate solutions, y_k is the k th candidate solution, and $y_k(s)$ is the s th feature of y_k .

```

For each solution  $y_k$ , define emigration rate  $\mu_k$  proportional to fitness
of  $y_k$ ,  $\mu_k \in [0,1]$ 
For each solution  $y_k$ , define immigration rate  $\lambda_k$  inversely proportional to
fitness of  $y_k$ ,  $\lambda_k \in [0,1]$ 
 $z \leftarrow y$ 
For each solution  $z_k$ 
  For each solution feature  $s$ 
    Use  $\lambda_k$  to probabilistically decide whether to immigrate to  $z_k$ 
    If immigrating then
      Use  $\{\mu_j\}$  to probabilistically select the emigrating solution  $y_j$ 
       $z_k(s) \leftarrow y_j(s)$ 
    End if
  Probabilistically decide whether to mutate  $z_k(s)$ 
Next solution feature
Next solution
 $y \leftarrow z$ 

```

BBO provides the probability p_{ij} of transitioning from the i th population distribution to the j th population distribution. In BBO, two main steps, migration and mutation, are significant, which indicate that the transition probability includes the migration probability and the mutation probability for one generation.

Consider a problem whose solutions are in a binary search space. The set of candidate solutions is the set of all bit strings x_i consisting of q bits each. Therefore, the cardinality of the search space is $n = 2^q$. Use N to denote the population size, and use v to denote the population vector, where the component v_i is the number of candidate solutions x_i in the population. Use y_k to denote the k th individual in the population, and use s to denote the s th feature of a solution. According to Simon et al. (2010), the migration probability during generation t , which results in an individual at generation $t + 1$, is the following:

$$\begin{aligned}
 Pr(y_{k,t+1}(s) = x_i(s)) &= Pr(\text{no immigration to } y_{k,t})Pr(y_{k,t+1}(s) = x_i(s) | \text{no immigration}) \\
 &\quad + Pr(\text{immigration to } y_{k,t})Pr(y_{k,t+1}(s) = x_i(s) | \text{immigration}) \\
 &= (1 - \lambda_{m(k)})1_0(x_{m(k)}(s) - x_i(s)) + \lambda_{m(k)} \frac{\sum_{j \in \zeta_i(s)} v_j \mu_j}{\sum_{j=1}^n v_j \mu_j} \tag{2}
 \end{aligned}$$

where 1_0 is the indicator function on the set $\{0\}$, and

$$y_k = x_{m(k)} \text{ for } k = 1, \dots, N \tag{3}$$

where $m(k)$ is defined as

$$m(k) = \min r \text{ such that } \sum_{i=1}^r v_i \geq k \tag{4}$$

The notation $\zeta_i(s)$ denotes the set of population indices j such that the s th bit of x_j is equal to the s th bit of x_i . That is

$$\zeta_i(s) = \{j : x_j(s) = x_i(s)\} \tag{5}$$

In fact, from (2) the total migration probability includes two parts: the probability that immigration did not occur and the probability that immigration occurred. When $y_k(s)$ does not change from generation t to generation $t + 1$, that is, the s th feature of y_k is not selected for immigration during generation t , then

$$y_k(s)_{t+1} = x_{m(k)}(s) \text{ (no immigration to } y_{k,t}) \tag{6}$$

When the s th feature of y_k is selected for immigration during generation t , the probability that $y_k(s)_{t+1}$ is equal to $x_i(s)$ is proportional to the combined emigration rates of all individuals

whose sth feature is equal to $x_i(s)$. This probability can be written as

$$Pr_{imm}(y_k(s)_{t+1} = x_i(s)) = \frac{\sum_{j \in \zeta_i(s)} V_j \mu_j}{\sum_{j=1}^n V_j \mu_j} \quad (7)$$

Eqs. (6) and (7) are combined with the fact that the probability of immigration to $y_k(s)$ is equal to λ_k to obtain (2).

For q bits in each solution, $P_{ki}(v)$ denotes the probability that immigration results in $y_k = x_i$, given that the population distribution is equal to v , which can be written as

$$P_{ki}(v) = Pr(y_{k,t+1} = x_i) = \prod_{s=1}^q \left[(1 - \lambda_{m(k)}) 1_0(x_{m(k)}(s) - x_i(s)) + \lambda_{m(k)} \frac{\sum_{j \in \zeta_i(s)} V_j \mu_j}{\sum_{j=1}^n V_j \mu_j} \right] \quad (8)$$

Note that the k th row of $P(v)$ corresponds to the k th iteration of the outer loop in Table 1 (there are N iterations of the outer loop in Table 1). The i th column of $P(v)$ corresponds to the probability of obtaining island x_i during each outer loop iteration; that is, $P_{ki}(v)$ means the probability of the i th outcome on the k th migration trial.

In (8) only migration is calculated. Mutation probability needs to be included after migration is completed. Use U to denote the $n \times n$ mutation matrix, where U_{ij} is the probability that x_j mutates to x_i . The probability that the k th immigration trial followed by mutation results in x_i is denoted as $P_{ki}^{(2)}(v)$. This can be written as

$$P_{ki}^{(2)}(v) = \sum_{j=1}^n U_{ij} P_{kj}(v) \quad (9)$$

$$P^{(2)}(v) = P(v)U^T$$

where the elements of $P(v)$ are given in (8). $P^{(2)}(v)$ contains the probabilities when both migration and mutation are considered. Define u as the population vector after migration and mutation are completed, where the component u_i is the number of solutions x_i in the population. Then the transition probability $Pr(u|v)$ where population vector u is obtained after one generation, given that the population vector is v at the beginning of the generation, can be obtained as

$$Pr(u|v) = \sum_{J \in Y} \prod_{k=1}^N \prod_{i=1}^n [P_{ki}^{(2)}(v)]^{J_{ki}}, \quad (10)$$

$$Y = \left\{ J \in \mathbb{R}^{N \times n} : J_{ki} \in \{0, 1\}, \sum_{i=1}^n J_{ki} = 1 \text{ for all } k, \sum_{k=1}^N J_{ki} = u_i \text{ for all } i \right\}$$

Eq. (10) can be used to find the transition matrix for BBO with migration and mutation. In order to find the probability that the BBO population transitions from v to u after one generation, find all of the J matrices that satisfy the condition of (10). For each of these J matrices, compute the product of products given in (10). Then add up all the product-of-products to obtain the desired probability. The Markov transition matrix Q is obtained by computing (10) for each possible v and each possible u . The element Q_{ij} will give the probability of transitioning from population vector v to u after one generation. Note the matrix Q is a $T \times T$ matrix, where T is the total number of possible populations, which can be calculated by several different methods, as discussed in Simon et al. (2010). Once the transition matrix Q is calculated, a wealth of Markov tools (Grinstead and Snell, 1997) can be applied to the transition matrix to find statistical properties of BBO, including the limiting probability of each possible BBO population, and population distributions of different BBO migration models. This is discussed further in Section 3.

2.3. Migration models

According to different mathematical models of biogeography (Lomolino et al., 2009; Whittaker, 1998), various migration model

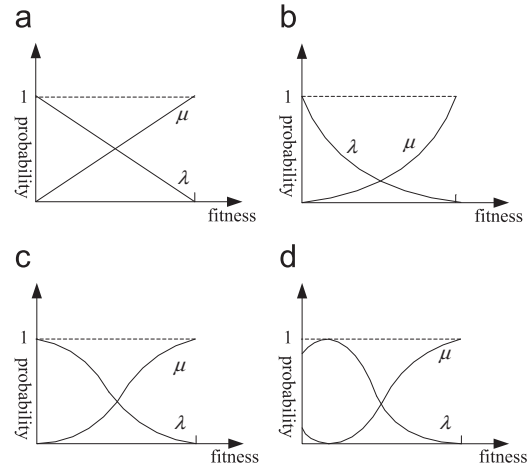


Fig. 1. Four migration model curves, where (a)–(d) respectively denote the linear migration curve, quadratic migration curve, sinusoidal migration curve, and generalized sinusoidal migration curve, respectively. λ is immigration rate and μ is emigration rate, and it is assumed that the maximum immigration rate and maximum emigration rate are both equal to 1.

curves can be obtained. In Ma (2010), the influences of six representative migration model curves on optimization performance are explored, based on empirical experiments, including three linear curves and three nonlinear curves. In this paper, the Markov model discussed above is used to verify the effect of migration curves on BBO. The three best migration models in Ma (2010) are used here, namely, the linear migration model, quadratic migration model, and sinusoidal migration model. In addition, a new model is introduced here, which is called the generalized sinusoidal migration model. The curves of these four models are shown in Fig. 1, where Fig. 1(a)–(d) respectively denote the linear migration curve, quadratic migration curve, sinusoidal migration curve, and generalized sinusoidal migration curve. In Fig. 1, λ denotes immigration rate and μ denotes emigration rate, and it is assumed that the maximum immigration rate and maximum emigration rate are both equal to 1. This subsection first reviews the previously proposed migration models, then discusses the generalized sinusoidal migration model.

According to Simon et al. (2010), the linear migration model is given as follows

$$\lambda = 1 - \text{fitness} \quad (11)$$

$$\mu = \text{fitness}$$

where $fitness$ denotes solution fitness, and is normalized to the range $[0,1]$. This model was first presented in the original BBO paper (Simon, 2008). It means that immigration rate λ and emigration rate μ are linear functions of solution fitness. This is illustrated by Fig. 1(a). The linear migration model does not exist in natural biogeography. Nevertheless this model exhibits features and properties of the process of migration that are much simpler than the general, nonlinear case.

The second model is the quadratic migration model, which is

$$\lambda = (1 - \text{fitness})^2 \quad (12)$$

$$\mu = (\text{fitness})^2$$

where migration rate λ and μ are concave quadratic functions of solution fitness, and where $fitness$ is again normalized to the range $[0,1]$. This is illustrated by Fig. 1(b). This model is inspired by island biogeography, which is developed to explain the species distribution of biological habitats. Based on an experimentally tested theory of island biogeography (Whittaker, 1998), we know that migration in a single habitat follows a quadratic function of the size of the habitat

and geographical proximity to other habitats. According to (12), when solution fitness is small, immigration rate rapidly decreases from its maximum while emigration rate slowly increases from zero. When solution fitness is large, immigration rate gradually decreases to zero and emigration rate rapidly increases to its maximum.

The sinusoidal migration model is given by

$$\begin{aligned} \lambda &= \frac{1}{2}(\cos(\text{fitness} \times \pi) + 1) \\ \mu &= \frac{1}{2}(-\cos(\text{fitness} \times \pi) + 1) \end{aligned} \tag{13}$$

where migration rate λ and μ are sinusoidal functions of solution fitness, and where *fitness* is again normalized to the range [0,1]. This model describes the migration curves to take into account species mobility, the evolution of particular species, and population size. These factors make the migration curves look like sinusoids. This is illustrated by Fig. 1(c). Based on (13), when solution fitness is small or large, immigration rate and emigration rate both change slowly from their extremes, and when solution fitness is medium, the migration rates change rapidly from their intersection.

Classical island biogeography theory indicates that the immigration rate decreases and emigration rate increases as the number of species increases in a habitat. In BBO this corresponds to a monotonic decrease in immigration rate and a monotonic increase in emigration rate as solution fitness increases, as shown in the previous three migration models, although their curve shapes are different. This means that as a solution becomes more fit, the probability of incorporating features from other solutions decreases. However, recent advances in biogeography indicate that for some pioneer species, at least for plants, an initial increase in species count results in an initial increase in immigration rate and an initial decrease in emigration rate (Lomolino et al., 2009; Whittaker, 1998). This is because the original unfavorable environmental conditions of the island are ameliorated by the first colonists, which make it more hospitable to additional species. That is, the positive effect of increased diversity due to initial immigration overcomes the negative effect of increased population size. In BBO this would correspond to an initial increase in immigration rate as a very poor candidate solution initially improves its fitness. This can be viewed as a temporary positive feedback mechanism in BBO. A very poor candidate solution accepts features from other solutions, increasing its fitness, which subsequently increases its likelihood of accepting even more features from other solutions. This idea can be incorporated into other EAs (Mühlenbein and Schlierkamp-Voosen, 1993) also, but its initial motivation comes from biogeography. This is depicted in Fig. 1(d), and is expressed as

$$\begin{aligned} \lambda &= \frac{1}{2}(\cos(\text{fitness} \times \pi + \beta) + 1) \\ \mu &= \frac{1}{2}(-\cos(\text{fitness} \times \pi + \beta) + 1) \end{aligned} \tag{14}$$

where β is a negative trigonometric offset angle (typically between $-\pi/2$ and 0), but in BBO, it denotes the degree of temporary positive immigration rate feedback. With this model, fitness is normalized to [0, $1 - \beta/\pi$]. This is called the generalized sinusoidal migration model. This proposed model shows that immigration initially increases with solution fitness. It gives improving solutions the momentum that they need to continue improving. As a solution continues to become fitter after the initial increase in immigration, immigration begins to decrease to give less fit solutions relatively greater opportunities to immigrate good solution features.

3. Result comparisons

This section first investigates the effect of the parameter β in the generalized sinusoidal migration model (Section 3.1), then compares the performance of the four migration models proposed in Section 2.3 using the BBO Markov chain model (Section 3.2), and finally compares the generalized sinusoidal migration model with the regular sinusoidal migration model using 23 benchmark testing functions (Section 3.3).

3.1. Generalized sinusoidal migration model: effect of the parameter β

In the first experiment the effect of the parameter β , which is the degree of temporary positive immigration rate feedback in the generalized sinusoidal migration model, is investigated. The limiting population distribution of the generalized sinusoidal migration model of BBO is given in Eq. (10). This is the probability, in the limit as the generation count approaches infinity, that the BBO population consists of any particular set of individuals. Test functions are limited to three-bit problems with a search space cardinality of eight and a population size of four, due to the exponential increase of Markov matrix sizes with problem size. Three fitness functions are investigated, which are given as

$$\begin{aligned} f_1 &= (1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8) \\ f_2 &= (1 \ 2 \ 3 \ 2 \ 1 \ 2 \ 3 \ 2) \\ f_3 &= (5 \ 2 \ 2 \ 3 \ 2 \ 3 \ 3 \ 4) \end{aligned} \tag{15}$$

where f_1 is a unimodal one-max problem, f_2 is a multimodal problem, and f_3 is a deceptive problem. Fitness values are listed in binary order, so the first element of each fitness function corresponds to the bit string 000, the second element corresponds to the bit string 001, and so on.

The parameter $\beta = 0, -\pi/4, -\pi/3,$ and $-\pi/2$ in Eq. (14) is used to investigate its influence on performance of the generalized sinusoidal migration model. In addition, simulation experiments are used to confirm the results. The other parameters of BBO are recommended as follows: population size of 50, maximum immigration rate and

Table 2

Optimization results of the generalized sinusoidal migration model for unimodal one-max problem f_1 when $\beta = 0, -\pi/4, -\pi/3,$ and $-\pi/2$. The table shows the probabilities of obtaining an all-optimal population and the probabilities of obtaining a no-optimal population using the BBO Markov model and simulations. The best performance is in **bold font** in each row.

Mutation rate	Population vector	Probability							
		$\beta = 0$		$\beta = -\pi/4$		$\beta = -\pi/3$		$\beta = -\pi/2$	
		Markov	Simulation	Markov	Simulation	Markov	Simulation	Markov	Simulation
0.1	All optimal	0.0456	0.0455	0.0475	0.0483	0.0526	0.0545	0.1001	0.1024
	No optimal	0.1994	0.1872	0.1975	0.1817	0.1756	0.1703	0.1542	0.1517
0.01	All optimal	0.6076	0.6008	0.6205	0.6260	0.6542	0.6618	0.7354	0.7361
	No optimal	0.0367	0.0351	0.0314	0.0306	0.0376	0.0371	0.0394	0.0317
0.001	All optimal	0.9062	0.9094	0.9335	0.9327	0.9363	0.9368	0.9456	0.9513
	No optimal	0.0151	0.0147	0.0112	0.0105	0.0102	0.0099	0.0103	0.0158

Table 3
Optimization results of the generalized sinusoidal migration model for multimodal problem f_2 when $\beta=0, -\pi/4, -\pi/3$, and $-\pi/2$. The table shows the probabilities of obtaining an all-optimal population and the probabilities of obtaining a no-optimal population using the BBO Markov model and simulations. The best performance is in **bold font** in each row.

Mutation rate	Population vector	Probability							
		$\beta=0$		$\beta=-\pi/4$		$\beta=-\pi/3$		$\beta=-\pi/2$	
		Markov	Simulation	Markov	Simulation	Markov	Simulation	Markov	Simulation
0.1	All optimal	0.2414	0.2419	0.2522	0.2592	0.2790	0.2716	0.2941	0.2824
	No optimal	0.1519	0.1433	0.1507	0.1578	0.1175	0.1214	0.1483	0.1311
0.01	All optimal	0.8584	0.8415	0.8774	0.8613	0.8806	0.8893	0.8972	0.9022
	No optimal	0.0241	0.0257	0.0213	0.0227	0.0199	0.0190	0.0287	0.0281
0.001	All optimal	0.9627	0.9617	0.9526	0.9604	0.9637	0.9619	0.9834	0.9712
	No optimal	0.0216	0.0244	0.0193	0.0187	0.0142	0.0153	0.0087	0.0090

Table 4
Optimization results of the generalized sinusoidal migration model for deceptive problem f_3 when $\beta=0, -\pi/4, -\pi/3$, and $-\pi/2$. The table shows the probabilities of obtaining an all-optimal population and the probabilities of obtaining a no-optimal population using the BBO Markov model and simulations. The best performance is in **bold font** in each row.

Mutation rate	Population vector	Probability							
		$\beta=0$		$\beta=-\pi/4$		$\beta=-\pi/3$		$\beta=-\pi/2$	
		Markov	Simulation	Markov	Simulation	Markov	Simulation	Markov	Simulation
0.1	All optimal	0.0714	0.0709	0.0866	0.0812	0.0890	0.0901	0.1127	0.1072
	No optimal	0.4821	0.4894	0.4648	0.4607	0.4321	0.4294	0.4078	0.4007
0.01	All optimal	0.7479	0.7422	0.7609	0.7745	0.7879	0.7822	0.8115	0.8110
	No optimal	0.1119	0.1214	0.1105	0.1184	0.1006	0.1034	0.0941	0.0935
0.001	All optimal	0.9224	0.9251	0.9164	0.9275	0.9312	0.9381	0.9689	0.9548
	No optimal	0.0322	0.0317	0.0267	0.0271	0.0204	0.0253	0.0186	0.0177

maximum emigration rate of 1, mutation rates of 0.1, 0.01, and 0.001 per bit per generation, generation count of 20,000, and 100 Monte Carlo runs for each function. Tables 2–4 show comparisons between theoretical (Markov) and simulation results of the generalized sinusoidal migration model with different values of β . The tables show the probability of obtaining a population in which all individuals are optimal, and the probability of obtaining a population in which no individuals are optimal. The mutation rates shown in Tables 2–4 are applied to each bit in each individual at each generation.

Several things are notable about the results in Tables 2–4. First, the mutation rate affects the performance of the generalized sinusoidal migration model. For three different test problems, the performance of the generalized sinusoidal migration model improves as the mutation rate decreases; that is, the probability of obtaining a population in which all individuals are optimal increases, and the probability of obtaining a population in which no individuals are optimal decreases. A high mutation rate of 0.1 per bit results in too much exploration, so the probability of obtaining the optimal population is low. However, as the mutation rate decreases to the values of 0.01 and 0.001, the probability of obtaining the optimal population significantly increases.

Second, when the parameter $\beta=-\pi/2$, the generalized sinusoidal migration model performs the best on all three test problems for most cases. For example, for the unimodal one-max problem (Table 2), the best performance is obtained by the generalized sinusoidal migration model with the parameter $\beta=-\pi/2$ and a mutation rate of 0.001 in its high probability of obtaining a population with all optimal individuals (94.6%), and in its low probability of obtaining a population with no optimal individuals (1.0%). When the parameter $\beta=0, -\pi/4$, and $-\pi/3$,

the probabilities are 90.6%, 93.3%, and 93.6%, respectively, for obtaining a population with all optimal individuals, and 1.5%, 1.1%, and 1.0%, respectively, for obtaining a population with no optimal individuals. This indicates that the value of parameter β can significantly affect the performance of the generalized sinusoidal migration model.

Third, from Tables 2–4, the Markov model results and the simulation results match well for all test problems, which confirms the Markov theory, and verifies the significance of parameter β for the generalized sinusoidal migration model.

Fig. 2 shows the probability of obtaining an all-optimal population for the generalized sinusoidal migration model for the unimodal one-max problem when the parameter $\beta=0, -\pi/4, -\pi/3$ and $-\pi/2$, and when the mutation rate is 1% per bit. It is seen that the results agree with Table 2, providing further confirmation of the Markov theory results.

3.2. Theoretical comparison of migration models

The next experiment investigates the effect of migration models on BBO performance using the BBO Markov chain model. The limiting population distribution of BBO with the four migration models proposed in Section 2.3 is compared using Eq. (10). Test functions and the parameters used in this experiment are the same as those described in the previous section. For the generalized sinusoidal migration model, $\beta=-\pi/2$, which provides the best performance based on the results in the previous section (when $\beta=0$, the generalized sinusoidal migration model reduces to the regular sinusoidal migration model). Tables 5–7 show comparisons between theoretical and simulated BBO with the four proposed migration models.

From Tables 5–7, first, the mutation rate affects the performance of BBO for all four migration models, which is similar to the results discussed in the previous section. Second, the generalized sinusoidal migration model performs better than the other three migration models for most cases. For example, for the deceptive problem (Table 7), the best performance is obtained by the generalized sinusoidal migration model with a mutation rate of 0.001 in its high probability of obtaining a

population with all optimal individuals (96.9%), and in its low probability of obtaining a population with no optimal individuals (1.9%). The probabilities of linear model, quadratic model, and sinusoidal model are 90.7%, 91.9%, and 92.2%, respectively, for obtaining a population with all optimal individuals, and 7.3%, 6.0%, and 3.2%, respectively, for obtaining a population with no optimal individuals. Third, the Markov theory results are confirmed by the simulation results.

Tables 5–7 indicate that changing the migration model curve can provide a valuable approach for enhancing BBO. From Tables 5–7, we further see that the nonlinear migration models (quadratic, sinusoidal, and generalized sinusoidal) are better than the linear migration model with different mutation rates for all test problems. Such results are similar to those reported in previous work (Ma, 2010). This confirms that when BBO migration model curves are closer to migration characteristics in nature, such as the generalized sinusoidal migration model, optimization performance is better.

3.3. Empirical comparison with sinusoidal migration model

To confirm the performance of the proposed generalized sinusoidal migration model, it is compared with the regular sinusoidal migration model using 23 benchmark functions, which are chosen from Hedar and Fukushima (2003). These functions are briefly summarized in Table 8. A more detailed description of these functions can be found in the literature (Yao et al., 1999), where functions f01–f07 are high-dimensional and unimodal, functions f08–f13 are high-dimensional and multimodal with many local minima, and functions f14–f23 are low-dimensional with only a few local minima. For both migration models, the

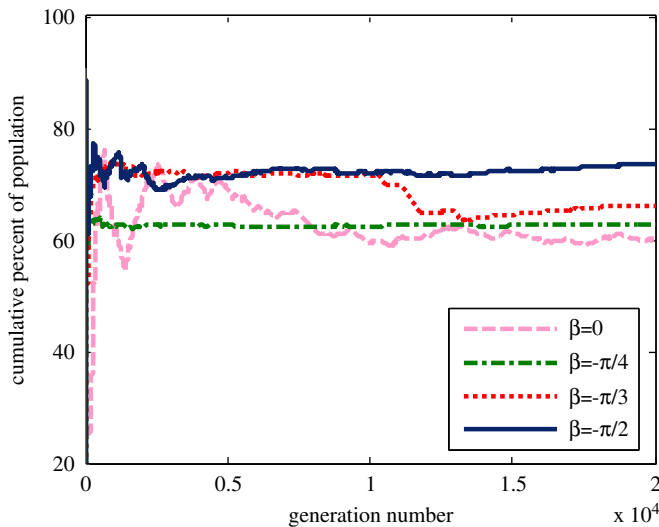


Fig. 2. Simulation results obtained by the parameter $\beta=0, -\pi/4, -\pi/3$ and $-\pi/2$ for the generalized sinusoidal migration model for a three-bit unimodal optimization problem with a mutation rate of 1% per bit. The figure shows the cumulative percent of obtaining an all-optimal population.

Table 5

Optimization results of four migration models for unimodal one-max problem f_1 . The table shows the probabilities of obtaining an all-optimal population and the probabilities of obtaining a no-optimal population using the BBO Markov model and simulations. The best performance is in **bold font** in each row.

Mutation rate	Population vector	Probability							
		Linear model		Quadratic model		Sinusoidal model		Generalized model	
		Markov	Simulation	Markov	Simulation	Markov	Simulation	Markov	Simulation
0.1	All optimal	0.0257	0.0214	0.0363	0.0360	0.0456	0.0455	0.1001	0.1024
	No optimal	0.2738	0.2709	0.2117	0.2108	0.1994	0.1872	0.1542	0.1517
0.01	All optimal	0.5049	0.5146	0.5529	0.5621	0.6076	0.6008	0.7354	0.7361
	No optimal	0.1004	0.1097	0.0416	0.0411	0.0367	0.0351	0.0394	0.0317
0.001	All optimal	0.8676	0.8504	0.9309	0.9274	0.9062	0.9094	0.9456	0.9513
	No optimal	0.0761	0.0755	0.0395	0.0356	0.0151	0.0147	0.0103	0.0158

Table 6

Optimization results of four migration models for multimodal problem f_2 . The table shows the probabilities of obtaining an all-optimal population and the probabilities of obtaining a no-optimal population using the BBO Markov model and simulations. The best performance is in **bold font** in each row.

Mutation rate	Population vector	Probability							
		Linear model		Quadratic model		Sinusoidal model		Generalized model	
		Markov	Simulation	Markov	Simulation	Markov	Simulation	Markov	Simulation
0.1	All optimal	0.1203	0.1231	0.2078	0.2166	0.2414	0.2419	0.2941	0.2824
	No optimal	0.1827	0.1759	0.2173	0.2078	0.1519	0.1433	0.1175	0.1214
0.01	All optimal	0.7760	0.7767	0.8311	0.8246	0.8584	0.8415	0.8972	0.9022
	No optimal	0.0515	0.0438	0.0354	0.0251	0.0241	0.0257	0.0287	0.0281
0.001	All optimal	0.9154	0.9212	0.9318	0.9412	0.9627	0.9617	0.9834	0.9712
	No optimal	0.0355	0.0541	0.0274	0.0264	0.0216	0.0244	0.0087	0.0090

Table 7
Optimization results of four migration models for deceptive problem f_3 . The table shows the probabilities of obtaining an all-optimal population and the probabilities of obtaining a no-optimal population using the BBO Markov model and simulations. The best performance is in **bold font** in each row.

Mutation rate	Population vector	Probability							
		Linear model		Quadratic model		Sinusoidal model		Generalized model	
		Markov	Simulation	Markov	Simulation	Markov	Simulation	Markov	Simulation
0.1	All optimal	0.0145	0.0144	0.0334	0.0332	0.0714	0.0709	0.1127	0.1072
	No optimal	0.7954	0.8012	0.5048	0.5007	0.4821	0.4894	0.4078	0.4007
0.01	All optimal	0.6407	0.6513	0.7009	0.7145	0.7479	0.7422	0.8115	0.8110
	No optimal	0.1921	0.1825	0.1193	0.0987	0.1119	0.1214	0.0941	0.0935
0.001	All optimal	0.9074	0.9017	0.9194	0.9152	0.9224	0.9251	0.9689	0.9548
	No optimal	0.0733	0.0691	0.0597	0.0590	0.0322	0.0317	0.0186	0.0177

Table 8
Benchmark functions. More details about these functions can be found in Yao. et al. (1999).

Function	Name	Dimension	Domain	Minimum
f01	Sphere model	30	$-100 \leq x_i \leq 100$	0
f02	Schwefel's problem 2.22	30	$-10 \leq x_i \leq 10$	0.
f03	Schwefel's problem 1.2	30	$-100 \leq x_i \leq 100$	0
f04	Schwefel's problem 2.21	30	$-100 \leq x_i \leq 100$	0
f05	Generalized Rosenbrock's function	30	$-30 \leq x_i \leq 30$	0
f06	Step function	30	$-100 \leq x_i \leq 100$	0
f07	Quartic function	30	$-1.28 \leq x_i \leq 1.28$	0
f08	Generalized Schwefel's problem 2.26	30	$-500 \leq x_i \leq 500$	-12569.5
f09	Generalized Rastrigin's function	30	$-5.12 \leq x_i \leq 5.12$	0
f10	Ackley's function	30	$-32 \leq x_i \leq 32$	0
f11	Generalized Griewank's function	30	$-600 \leq x_i \leq 600$	0
f12	Generalized Penalized function 1	30	$-50 \leq x_i \leq 50$	0
f13	Generalized Penalized function 2	30	$-50 \leq x_i \leq 50$	0
f14	Shekel's Foxholes function	2	$-65.536 \leq x_i \leq 65.536$	1
f15	Kowalik's function	4	$-5 \leq x_i \leq 5$	0.003075
f16	Six-Hump Camel-Back function	2	$-5 \leq x_i \leq 5$	-1.0316285
f17	Branin's function	2	$-5 \leq x_1 \leq 10, 0 \leq x_2 \leq 15,$	0.398
f18	Goldstein-Price's function	2	$-2 \leq x_i \leq 2$	3
f19	Hartman's function 1	3	$0 \leq x_i \leq 1$	-3.86
f20	Hartman's function 2	6	$0 \leq x_i \leq 1$	-3.32
f21	Shekel's function 1	1	$0 \leq x_i \leq 10$	-10.1532
f22	Shekel's function 2	1	$0 \leq x_i \leq 10$	-10.4029
f23	Shekel's function 3	1	$0 \leq x_i \leq 10$	-10.5364

following parameters of BBO have to be examined: population size, maximum migration rates, and mutation rate. In the literature (Ma, 2010) these parameters have been discussed in detailed. This paper uses a reasonable set of tuning parameters, but does not make any effort in finding the best parameter settings. For this experiment, the parameters of BBO used in the two migration models are the same: population size of 50, maximum immigration rate and maximum emigration rate of 1, mutation rate of 0.01 per bit per generation, generation limit of 20,000 for f01–f13, and 1000 for f14–f23, and 30 Monte Carlo runs. In addition, $\beta = -\pi/2$ for the generalized sinusoidal migration model.

Table 9 summarizes the performance on 23 benchmark functions for the generalized sinusoidal migration model and the regular sinusoidal migration model. It is apparent that the generalized sinusoidal migration model performs significantly better than the regular sinusoidal migration model in terms of the final results for the most functions. The generalized sinusoidal migration model performs the best on 16 functions, and the regular sinusoidal migration model performs the best on four functions (f06, f12, f17, f23). For functions f14, f19, f20, both models attain the global optimum.

Table 9 also indicates statistically significant differences of the two models based on the p value, which is the probability that the two sets of data come from the same distribution. From p value comparison between the generalized sinusoidal migration model and the regular sinusoidal migration model, there are 17 p values smaller than 0.05 (which is often used as the significance level or critical p value). Based on this result, the probability that two models are from the same distribution is low. It indicates that the parameter β is influential on BBO performance. Furthermore, the generalized sinusoidal migration model generally outperforms the regular sinusoidal migration model, which indicates that the parameter β , which is the degree of temporary positive immigration rate feedback in the generalized sinusoidal migration model, contributes to improve the optimization ability of BBO. Finally, note that the benchmark functions outperformed by the generalized sinusoidal migration model include high-dimensional and unimodal functions, high-dimensional and multimodal functions with many local minima, and low-dimensional functions with only a few local minima. Therefore, the type of benchmark function is not of importance for successful optimization using the generalized sinusoidal migration model.

Table 9

Comparison of experimental results over 30 Monte Carlo runs of the generalized sinusoidal migration model and the regular sinusoidal migration model. The table shows the best, mean and its standard deviation. The p value for each benchmark gives the probability that the two sets of results come from the same distribution. Best results for each benchmark function are shown in **bold font**.

Fun.	BBO						p value
	Sinusoidal model			Generalized model			
	Best	Mean	Stdev	Best	Mean	Stdev	
f01	2.17E-02	6.38E-02	5.46E-03	9.55E-04	8.24E-03	9.78E-04	8.24E-03
f02	1.84E-04	8.03E-04	3.84E-04	1.22E-04	1.47E-04	4.86E-05	0.47
f03	6.33E-02	8.31E-02	1.17E-02	2.71E-03	3.08E-03	5.71E-04	0.03
f04	5.68E-15	5.34E-14	2.54E-15	5.36E-15	4.64E-14	1.07E-15	0.16
f05	9.24E-01	3.47E+00	4.33E-01	6.54E-01	9.36E-01	2.44E-01	0.08
f06	0.00E+00	0.00E+00	0.00E+00	1.16E-15	8.21E-15	1.83E-16	7.19E-04
f07	1.37E-15	2.36E-15	1.96E-16	0.00E+00	0.00E+00	0.00E+00	6.58E-04
f08	2.63E-06	5.08E-06	2.74E-06	6.85E-09	9.38E-09	4.76E-10	1.28E-03
f09	1.55E-03	1.21E-02	3.78E-03	1.15E-04	4.38E-04	9.33E-05	0.01
f10	2.54E-01	9.71E-01	7.16E-02	1.14E-03	2.54E-03	8.25E-04	5.32E-03
f11	7.49E-01	1.97E+00	9.33E-01	2.55E-01	3.65E-01	1.72E-02	0.03
f12	2.26E-30	4.11E-30	8.45E-31	3.74E-25	7.81E-25	5.38E-26	9.21E-04
f13	1.28E-10	7.36E-10	6.32E-11	1.07E-11	9.05E-11	6.72E-12	0.02
f14	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.56
f15	3.19E-04	5.29E-04	6.27E-05	8.46E-06	7.18E-05	1.80E-06	0.02
f16	2.67E-09	1.51E-08	7.32E-09	7.11E-12	4.02E-11	1.22E-12	5.01E-04
f17	2.17E-15	1.44E-14	4.77E-15	5.34E-10	3.94E-09	3.89E-10	3.76E-04
f18	6.06E-15	7.05E-15	2.56E-16	0.00E+00	0.00E+00	0.00E+00	2.37E-04
f19	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.56
f20	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.56
f21	5.29E-08	6.14E-07	6.60E-08	2.69E-08	3.57E-08	6.96E-09	0.03
f22	9.60E-12	7.89E-10	2.31E-10	6.65E-15	3.47E-14	5.73E-15	4.88E-04
f23	3.55E-12	7.34E-11	5.78E-12	1.75E-10	4.54E-10	1.58E-10	0.03

4. Conclusions

This paper proposed a generalized sinusoidal migration model based on the natural migration relations in island biogeography, and explored optimization performance of BBO with different migration models based on Markov chain models of the BBO algorithm. New theoretical results for different migration models have been obtained, which are confirmed with simulation results. The experiments for a unimodal one-max problem, multimodal problem and deceptive problem: (1) show that the generalized sinusoidal migration model further improves BBO's performance. Namely, the parameter β , the degree of temporary positive immigration rate feedback in this model, can affect BBO's optimization ability; (2) verify that different migration models in BBO result in significant changes in performance based on Markov theory; (3) further show that BBO migration models which are closer to natural biogeography are significantly better than general models. Although the theoretical results are limited to small problem dimensions due to the factorial increase of the Markov transition matrix size with problem dimension, these results provide confidence that migration models based on island biogeography can improve BBO performance. In addition, to confirm the above conclusions, empirical performance comparison between the generalized sinusoidal migration model with the sinusoidal migration model was investigated through 23 benchmark functions. The results showed that generalized sinusoidal migration exhibits superior optimization performance.

Future work includes several important directions. The first is to explore additional migration model features as indicated by natural biogeography theory to obtain better BBO performance. It has been shown in this paper that the generalized migration model generally gives better BBO performance, but it remains to be seen how other migration model features will affect BBO performance. There are many other interesting possibilities for aligning BBO more closely with island biogeography. For example,

habitat similarity, species age criterion, resource competition, and migration time correlation, could inspire other variations to the BBO algorithm and to the shape of the migration curves. The second important direction for future research is the development of Markov theory results for BBO variations with the additional migration model features mentioned above. The third important direction for future research is the development of additional theoretical tools to study BBO performance. For example, the asymptotic convergence of BBO with different migration models and their convergence rates could be worth further study. Fourth, adaptive migration rates and their theoretical analysis could be considered.

Acknowledgment

This material is based upon work supported by the Zhejiang Provincial Natural Science Foundation of China under Grant no. Y1090866, and by Grant 0826124 in the CMMI Division of the Engineering Directorate of the National Science Foundation.

References

- Ahn, C., 2006. *Advances in Evolutionary Algorithms: Theory, Design and Practice*. Springer Publishing, New York.
- Davis, T., Principe, J., 1993. A Markov chain framework for the simple genetic algorithm. *Evolutionary Computation*. 1 (3), 269–288 fall 1993.
- Du, D., Simon, D., Ergezer, M., 2009. Biogeography-based optimization combined with evolutionary strategy and immigration refusal. In: *Proceedings of the IEEE Conference on Systems, Man, and Cybernetics*, San Antonio, Texas, October 2009, pp. 1023–1028.
- Ergezer, M., Simon, D., Du, D., 2009. Oppositional biogeography-based optimization. In: *Proceedings of the IEEE Conference on Systems, Man, and Cybernetics*, San Antonio, Texas, October 2009, pp. 1035–1040.
- Gardner, B., Simon, D., 2009. Evolutionary algorithm sandbox: A web-based graphical user interface for evolutionary algorithms. In: *Proceedings of the IEEE Conference on Systems, Man, and Cybernetics*, San Antonio, Texas, October 2009, pp. 583–588.

- Grinstead, C., Snell, J., 1997. *Introduction to Probability*. American Mathematical Society, Providence, Rhode Island.
- Hedar, A.R., Fukushima, M., 2003. Minimizing multimodal functions by simplex coding genetic algorithm. *Optimization Methods and Software* 18 (Apr. 2003), 265–282.
- Kundra, H., Kaur, A., Panchal, V., 2009. An integrated approach to biogeography based optimization with case based reasoning for retrieving groundwater possibility. In: *Proceedings of the 8th Annual Asian Conference and Exhibition on Geospatial Information, Technology and Applications*, Singapore, August 2009.
- Lomolino, M., Riddle, B., Brown, J., 2009. *Biogeography*. Sinauer Associates.
- Lundy, M., Mees, A., 1986. Convergence of an annealing algorithm. *Mathematical Programming*, 34 (1), 111–124 January 1986.
- Ma, H., 2010. An analysis of the equilibrium of migration models for biogeography-based optimization. *Information Sciences* 180 (18), 3444–3464 2010.
- Ma, H., Ni, S., Sun, M., 2009. Equilibrium species counts and migration model tradeoffs for biogeography-based optimization. 2009. In: *Proceedings of the IEEE Conference on Decision and Control*, Shanghai, December 2009, pp. 3306–3310.
- Ma, H., Simon D., 2010. Biogeography-based optimization with blended migration for constrained optimization problems. In: *Proceedings of the Genetic and Evolutionary Computation Conference*, Portland, OR, July 2010, pp. 417–418.
- Mühlenbein, H., Schlierkamp-Voosen, D., 1993. Predictive models for the breeder genetic algorithm I. Continuous parameter optimization. *Evolutionary Computation* 1 (1), 25–49 1993.
- Nix, A., Vose, M., 1992. Modeling genetic algorithms with Markov chains. *Annals of Mathematics and Artificial Intelligence* 5 (1), 79–88 1992.
- Panchal, V., Singh, P., Kaur, N., Kundra, H., 2009. Biogeography based satellite image classification. *International Journal of Computer Science and Information Security* 6 (2), 269–274 November 2009.
- Rarick, R., Simon, D., Villaseca, F.E., Vyakaranam, B., 2009. Biogeography-based optimization and the solution of the power flow problem. In: *Proceedings of the IEEE Conference on Systems, Man, and Cybernetics*, San Antonio, Texas, October 2009, pp. 1029–1034.
- Reeves, C., Rowe, J., 2003. *Genetic Algorithms: Principles and Perspectives*. Kluwer Academic Publisher., Norwell, Massachusetts.
- Schwefel, H.P., 1995. *Evolution and optimum seeking*. Wiley Press.
- Simon, D., 2008. Biogeography-based optimization. *IEEE Transactions on Evolutionary Computation*, 12 (6), 702–713 December 2008.
- Simon, D., Ergezer, M., Du, D., 2009. Population distributions in biogeography-based optimization algorithms with elitism. In: *Proceedings of the IEEE Conference on Systems, Man, and Cybernetics*, San Antonio, Texas, October 2009, pp. 1017–1022.
- Simon, D., Ergezer, M., Du, D., Rarick, R., 2010. Markov models for biogeography-based optimization. *IEEE Transactions on Systems Man, and Cybernetics—Part B: Cybernetics* 41 (no. 1), 299–306.
- Simon, D., Rarick, R., Ergezer, M., Du, D., April 2011. Analytical and numerical comparisons of biogeography-based optimization and genetic algorithms. *Information Sciences* 181 (no. 7), 1224–1248.
- Suzuki, J., 1995. A Markov chain analysis on simple genetic algorithms. *IEEE Transactions on Systems, Man, and Cybernetics, Part B* 25 (4), 655–659 April 1995.
- Suzuki, J., 1998. A further result on the Markov chain model of genetic algorithms and its application to a simulated annealing-like strategy. *IEEE Transactions on Systems, Man, and Cybernetics Part B*, 28 (1), 95–102 Feb. 1998.
- Whittaker, R., 1998. *Island Biogeography*. Oxford University Press.
- Yao, X., Liu, Y., Lin, G., 1999. Evolutionary programming made faster. *IEEE Transactions on Evolutionary Computation*, 3 (2), 82–102 June 1999.