

# 全局混沌蝙蝠优化算法

崔雪婷<sup>1,2</sup>, 李 颖<sup>1,2</sup>, 范嘉豪<sup>1,2</sup>

(1. 吉林大学 计算机科学与技术学院, 吉林 长春 130012;  
2. 吉林大学 符号计算与知识工程教育部重点实验室, 吉林 长春 130012)

**摘 要:** 为提高蝙蝠算法进行特征选择的正确率, 提出全局混沌蝙蝠优化算法(GCBA)。首先, GCBA采用混沌映射方法使种群的初始化能够遍历整个解空间, 获取蝙蝠初始的最优位置, 使其具有更加丰富的种群, 解决了初始化种群随机性的问题。同时, GCBA引入当前粒子的最优解和当前种群的最优解跳出局部最优解, 可有效避免算法早熟, 有利于提高算法的全局搜索能力。蝙蝠算法(BA)、粒子群算法(PSO)与遗传算法(GA)在10个数据集上的测试结果表明, 所提算法具有更高的分类精度和更强的跳出局部最优的能力。

**关 键 词:** 蝙蝠算法; 混沌映射; 全局优化; 局部最优; 特征选择

**中图分类号:** TP 182 **文献标志码:** A **文章编号:** 1005-3026(2020)04-0488-05

## Global Chaotic Bat Optimization Algorithm

CUI Xue-ting<sup>1,2</sup>, LI Ying<sup>1,2</sup>, FAN Jia-hao<sup>1,2</sup>

(1. College of Computer Science and Technology, Jilin University, Changchun 130012, China; 2. Symbol Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun 130012, China. Corresponding author: FAN Jia-hao, E-mail: jihanfan@hotmail.com)

**Abstract:** In order to improve the accuracy of feature selection of bat algorithm, a global chaotic bat optimization algorithm(GCBA) was proposed. Firstly, GCBA adopts chaotic mapping method to enable the initialization of the population to traverse the entire solution space, obtain the optimal position of the bat, and make it more abundant. It solved the problem of initial population randomness. At the same time, GCBA introduces the optimal solution of the current particle and the optimal solution of the current population to jump out of local optimal solutions, which can effectively avoid the premature and improve the global search ability of the algorithm. The results of the bat algorithm(BA), particle swarm optimization(PSO) and genetic algorithm(GA) on 10 data sets showed that the proposed algorithm has higher classification accuracy and stronger ability to jump out of local optimum.

**Key words:** bat algorithm; chaotic mapping; global optimum; local optimum; feature selection

受蝙蝠回声定位行为的启发, 蝙蝠算法(bat algorithm, BA)于2010年由剑桥大学的Yang<sup>[1]</sup>提出, 随着研究的深入, 该算法逐渐引起人们的极大关注, 且在应用方面取得较大成功。Tsai等<sup>[2]</sup>将其改进后应用于解决数值优化问题; Yang<sup>[3]</sup>用其解决多目标优化问题; Bora等<sup>[4]</sup>将其应用在直流无刷马达问题上。蝙蝠算法是一种新的群智能算法, 在收敛速度和求解精度上还存在缺陷。该算法也存在易陷入局部最优、震荡、收敛速度慢和求解精度不高等不足。本文针对原始蝙蝠算法提出

了一种全局混沌蝙蝠优化算法(GCBA), 将混沌算法与最优解相结合用于解决特征选择问题。在种群初始化时GCBA算法覆盖整个解空间, 使用交叉算子做进一步优化; 此外, 将GCBA算法应用于特征选择领域<sup>[5]</sup>。为了验证GCBA算法的有效性, 将GCBA与其他3个优化算法在多个数据集上进行了测试。结果表明, GCBA算法能够获得更好的分类精度及更快的收敛速度。

## 1 蝙蝠算法

蝙蝠算法是一种采用回声定位原理模拟蝙蝠捕食过程的启发式群智能算法.不同于其他生物,蝙蝠采用回声定位来探测猎物、捕食猎物和寻找栖息地等.蝙蝠发出的脉冲频率通常在 25 ~ 150 kHz之间,波长范围为 2 ~ 14 mm,发出的声音响度能达到 110 dB.在探测猎物时,随着蝙蝠与猎物之间的距离逐渐缩短,蝙蝠发出声音的响度逐渐降低,脉冲发射率逐渐增加.

为了模拟蝙蝠回声定位的捕食行为,传统蝙蝠算法的3个理想化假设如下:

1) 假设蝙蝠利用回声定位感知自身与目标之间的距离,同时具有辨别目标与背景障碍物的能力.

2) 假设第  $i$  只蝙蝠在空间位置  $x_i$  以速度  $v_i$  随机飞行,以初始频率  $f_{\min}$ 、不同的波长  $\lambda$  和初始响度  $A^0$  进行目标搜索.蝙蝠结合自己和猎物之间的距离调整脉冲发射率  $r \in [0,1]$ .

3) 假设响度从最大值  $A^0$  变化到最小值  $A_{\min}$ .

基于以上3个理想化假设,在搜索空间中,蝙蝠的频率、速度与位置公式分别为

$$f_i = f_{\min} + (f_{\max} - f_{\min}) \times \beta, \quad (1)$$

$$v_i^{t+1} = v_i^t + (x_i^t - x^*) \times f_i, \quad (2)$$

$$x_i^{t+1} = x_i^t + v_i^{t+1}. \quad (3)$$

式中:  $f_i$  为第  $i$  只蝙蝠的脉冲频率;  $f_{\min}$  和  $f_{\max}$  分别为脉冲频率的最小值和最大值;  $\beta$  为  $[0,1]$  上服从均匀分布的随机数;  $v_i^{t+1}$  和  $v_i^t$  分别为第  $i$  只蝙蝠在第  $t+1$  和  $t$  代的飞行速度;  $x_i^t$  为第  $i$  只蝙蝠在第  $t$  代的位置;  $x_i^{t+1}$  为第  $i$  只蝙蝠在第  $t+1$  代的位置;  $x^*$  为当前群体中蝙蝠的最优位置.

在搜寻猎物的过程中,蝙蝠初始发出的超声波响度大,但发射速率低,有利于猎物在整个空间进行搜索.当蝙蝠发现猎物以后,蝙蝠逐渐减小响度、增加脉冲发射率,这样更有利于精确定位猎物的位置.公式为

$$r_i^{t+1} = r_i^0 [1 - \exp(-\gamma \times t)], \quad (4)$$

$$A_i^{t+1} = \alpha \times A_i^t. \quad (5)$$

式中:  $r_i^{t+1}$  为第  $i$  只蝙蝠在第  $t+1$  代的脉冲发射率;  $r_i^0$  为第  $i$  只蝙蝠的最大脉冲发射率;  $\gamma > 0$  为脉冲频度增加系数;  $A_i^{t+1}$  和  $A_i^t$  分别为第  $i$  只蝙蝠在第  $t+1$  和  $t$  代的音量响度;  $\alpha \in [0,1]$  为脉冲音量响度衰减系数.

BA 算法的伪代码如下:

目标方程  $f(x)$ ,  $x = (x_1, \dots, x_d)$

初始化蝙蝠种群  $x_i$  和  $v_i$  ( $i = 1, 2, \dots, n$ )

定义发射频率  $f_i$  在位置  $x_i$

初始化脉冲发射率  $r_i$  和响度  $A_i$

While ( $t < \text{Max number of iterations}$ )

    使用公式(1),(2)和(3)更新蝙蝠的频率、速度和位置

    形成新的解

    If ( $\text{rand} > r_i$ )

        选择出全局最优解

        形成一个局部解在最优解附近随机浮动

    End if

    通过随机飞行形成一个新的解

    If ( $\text{rand} < A_i \& f(x_i) < f(x_*)$ )

        接受新的解

        增大  $r_i$  并且减小  $A_i$  使用公式(4)和(5)

    End if

    对所有蝙蝠排序并且找到目前最好的解  $x_*$

End while

## 2 全局混沌蝙蝠优化算法

### 2.1 基于混沌映射的种群初始化

群智能搜索算法种群不断迭代,初始化种群的好坏对种群的收敛速度和寻优性能有着很大的影响<sup>[6]</sup>,多样性丰富的初始化种群更有利于找到种群的全局最优解.种群的全局最优解是一个未知量,没有任何先验知识能够知道全局最优解在什么位置,所以初始化种群的覆盖性很重要,蝙蝠尽量覆盖整个解空间.然而,BA 算法的种群初始化是随机产生的,不具有覆盖整个解空间的能力,这大大影响了 BA 算法的性能.

为了解决 BA 算法初始化种群随机性的问题,本文提出全局混沌蝙蝠优化算法(GCBA).该算法使种群的初始化能够遍历整个解空间.目前使用最广泛的混沌映射方法为 logistic 方法,该方法产生的初始化种群具有多样性并且能够遍历整个解空间.因此,本文采用改进的 logistic 映射方法来进行种群的初始化,其数学模型为<sup>[7]</sup>

$$y_i^{k+1} = |1 - 2 \times (y_i^k)^2|. \quad (6)$$

式中:  $y_i^k \in [0,1]$  为混沌变量,  $k = 1, 2, \dots, M$ ,  $M$  表示初始化种群维度,  $i = 1, 2, \dots, N$ ,  $N$  表示初始化种群数量.然后,针对  $y_i^k$  做逆映射得到相应解空间的初始化种群  $x_i^k$ :

$$x_i^k = l_i + (u_i - l_i) y_i^k. \quad (7)$$

式中  $l_i$  和  $u_i$  为变量取值范围的最小值和最大值.

2.2 局部最优和全局最优

在 BA 算法的后期,种群都趋向于最优区域,所以种群的多样性降低.如果当前种群个体找到的最优解并非全局最优解,则算法就会陷入局部最优.算法提前收敛,陷入早熟,这也是所有群智能算法的固有缺陷.为了解决这个问题,本文在每个蝙蝠进行位置更新的时候记录该蝙蝠局部最优位置和种群的全局最优位置.

$$x_i^{t+1} = x_i^t + v_i^{t+1} C_1 r_1 (P_i^l - X_i^t) + C_2 r_2 (P_g - X_i^t). \quad (8)$$

式中: $P_i^l$  为第  $i$  只蝙蝠的局部最优位置; $P_g$  为蝙蝠种群的全局最优位置; $C_1 = C_2 = 1.496\ 18$ ;  $r_1$  和  $r_2$  是两个在  $[0,1]$  之间的随机数.

2.3 GCBA 应用于特征选择

初始化蝙蝠种群时使用一个大小为  $n \times d$  的矩阵,其中,  $n$  表示蝙蝠样本数量,  $d$  表示特征个数.本文使用转换方程对蝙蝠个体的速度进行离散二进制操作,转换方程为

$$S(v_i^k(t)) = \frac{1}{1 + e^{-v_i^k(t)}}. \quad (9)$$

式中,  $v_i^k(t)$  为在第  $t$  次迭代过程中蝙蝠个体  $i$  在第  $k$  维的速度.

蝙蝠个体的位置更新方程为

$$x_i^k(t+1) = \begin{cases} 0, \text{rand} < S(v_i^k(t+1)); \\ 1, \text{rand} \geq S(v_i^k(t+1)). \end{cases} \quad (10)$$

式中,  $x_i^k(t)$  和  $v_i^k(t)$  分别为第  $t$  次迭代过程中蝙蝠个体  $i$  在第  $k$  维的位置和速度.

当蝙蝠个体  $i$  第  $t$  次迭代第  $k$  维的位置为 0 时,表示这个特征不会被选择;当蝙蝠个体  $i$  第  $t$  次迭代第  $k$  维的位置为 1 时,表示这个特征会被选择.

GCBA 的伪代码如下:

目标方程  $f(x)$ ,  $x = (x_1, \cdots, x_M)$   
使用混沌映射方法初始化蝙蝠种群  $x_i$  和  $v_i$ ,  
( $i = 1, 2, \cdots, n$ )  
定义发射频率  $f_i$  在位置  $x_i$   
初始化脉冲发射率  $r_i$  和响度  $A_i$   
While( $t < \text{Max number of iterations}$ )  
    使用公式(1),(2)和(8)更新蝙蝠的频率、速度和位置  
    形成新的解  
    使用式(9)把所有蝙蝠位置转换到二进制空间  
    If( $\text{rand} > r_i$ )  
        使用 SVM 分类器选择出全局最优解  $x_*$   
        形成一个局部解在最优解附近随机浮动  
End if

通过随机飞行形成一个新的解  
If( $\text{rand} < A_i \ \& \ f(x_i) < f(x_*)$ )  
    接受新的解  
    增大  $r_i$  并且减小  $A_i$  使用公式(4)和(5)  
End if  
End while

3 实验仿真及分析

本文选取 UCI 公共数据集测试 GCBA 算法性能,见表 1.样本数量从 76 到 1 473,特征数量从 4 到 24,数据维度的跨度大且特征数量分布广,数据集的选择具有实际意义.在实验设置上,把 GCBA 和 BBA<sup>[8]</sup>, PSO<sup>[9-10]</sup> 和 GA<sup>[11]</sup> 进行比较,结合 SVM 和十折交叉验证方法测试 10 个 UCI 数据集.这 10 个数据集分别是 Australian, Ballon, Breastcancer, Bupa liver, Contraceptive Method Choice, Cleveland heart, Diabetes, German, Glass 和 Heart.对于 BBA 算法,初始化参数设置:蝙蝠的数量 = 30, 响度 = 1.5, 脉冲发射率 = 0.5, 频率  $f_{\min} = 0$ ,  $f_{\max} = 1$ , 最大种群迭代次数 = 100.对于 PSO 算法,初始化参数设置:种群大小 = 30,  $\omega = 2$ , 系数  $c_1 = c_2 = 2$ , 最大种群迭代次数 = 100;对于 GA 算法,初始化参数设置:种群大小 = 30, 交换率 = 0.4, 变异率 = 0.2, 最大种群迭代次数 = 100.

表 1 数据集说明  
Table 1 Detail of data sets

数据集	样本数量	特征数量	类别数量
Australian	690	14	2
Ballon	76	4	2
Breastcancer	286	9	2
Bupa liver	345	6	2
Cleveland heart	303	13	2
Contraceptive Method Choice	1 473	9	3
Diabetes	768	8	2
German	1 000	24	2
Glass	214	10	7
Heart	270	13	2

鉴于篇幅限制,本文只给出 4 个算法. Breastcancer 和 German 数据集的分类正确率比较如图 1 和图 2 所示.从图 1 和图 2 可以看出, GCBA 算法分别在第 80 次到 90 次迭代期间获得

了最优解,而其他算法都提前收敛,陷入局部最优.通过与其他三个算法的比较,证明 GCBA 算法具有良好的全局搜索能力.

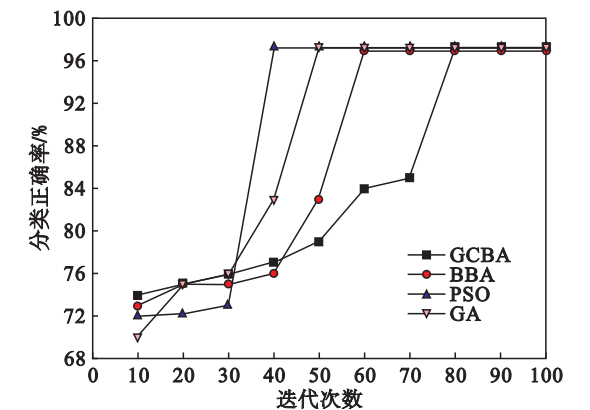


图 1 Breastcancer 数据集的分类正确率  
Fig. 1 Classification accuracy rate of Breastcancer dataset

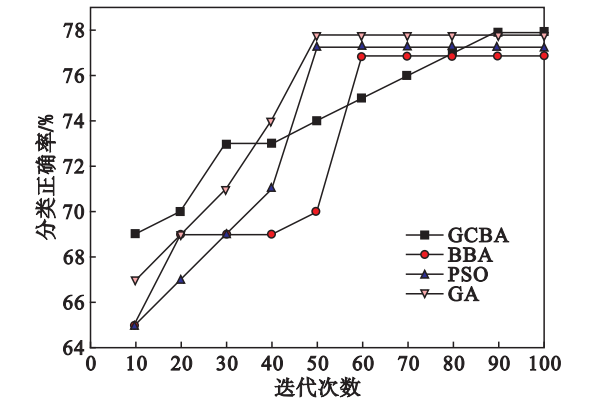


图 2 German 数据集的分类正确率  
Fig. 2 Classification accuracy rate of German dataset

从表 2 的实验结果可以看出:在所有的数据集上,GCBA 算法都达到最优正确率.全局最优和局部最优的引入,使算法能够有效跳出局部

表 2 实验结果  
Table 2 Experimental results

数据集	BBA		PSO		GA		GCBA	
	正确率/%	标准差	正确率/%	标准差	正确率/%	标准差	正确率/%	标准差
Australian	85.79	1.23	86.51	0.86	86.65	0.75	86.81	0.63
Ballon	77.75	0.5	80	0	80	0	80	0
Breastcancer	96.94	0.58	97.23	0	97.23	0.16	97.23	0
Bupa liver	70.71	0.13	70.71	0	70.71	0	70.71	0
Cleveland	82.85	1.56	83.55	0.89	83.18	0.48	84.26	0.26
Contraceptive	55.81	1.46	56.35	1.23	56.35	1.14	56.35	0.87
Diabetes	77.48	1.56	77.75	0.89	77.73	0.67	78.01	0.12
German	76.84	0.89	77.23	0.56	77.78	0.64	77.89	0.15
Glass	91.86	0.51	91.97	0.44	91.88	0.45	92.86	0.48
Heart	83.33	0.82	84.44	0.57	83.70	0.56	84.81	0.30

最优.此外,除了数据集 Glass,GCBA 在其余数据集上的稳定性表现良好.这是由于在初始化时,算法引入混动映射方法,使样本分布更广,提升了算法的稳定性.

从表 3 的实验结果可以看出:在所有的数据集上,BBA 算法结合局部最优( $P^l$ )和全局最优( $P_g$ )都达到更优的分类精度和较好的稳定性.由于全局最优和局部最优思想的引入,算法能够有效跳出局部最优解.从图 3 可以看出,BBA 算法在 Bupa liver 数据集上第 30 次迭代陷入了局部最优解,而引入局部最优和全局最优的 BBA 算法能够跳出局部最优,获得了全局最优解.

表 3 BBA 和  $BBA + P^l + P_g$  实验结果  
Table 3 Experimental results of BBA and  $BBA + P^l + P_g$

数据集	BBA		$BBA + P^l + P_g$	
	正确率/%	标准差	正确率/%	标准差
Australian	85.79	1.23	86.66	1.44
Ballon	77.75	0.5	80	0
Breastcancer	96.94	0.58	97.23	0
Bupa liver	70.71	0.13	70.71	0
Cleveland	82.85	1.56	83.22	0.53
Contraceptive	55.81	1.46	55.31	1.17
Diabetes	77.48	1.56	77.75	0.89
German	76.84	0.89	76.80	0.34
Glass	91.86	0.51	91.97	0.65
Heart	83.33	0.82	84.07	0.47