

基于差分进化生物地理学优化的 多层感知器训练方法*

王娟^{1,2}, 吴宪祥², 曹艳玲²

(1. 西安翻译学院基础课部, 西安 710105; 2. 西安电子科技大学空间科学与技术学院 ICIE 研究所, 西安 710071)

摘要: 针对生物地理学优化训练多层感知器存在的早熟收敛以及初始化灵敏等问题, 提出一种基于差分进化生物地理学优化的多层感知器训练方法。将生物地理学优化(biogeography-based optimization, BBO)与差分进化(differential evolution, DE)算法相结合, 形成改进的混合 DE_BBO 算法; 采用改进的 DE_BBO 来训练多层感知器(multi-layer perceptron, MLP), 并应用于虹膜、乳腺癌、输血、钞票验证四类数据分类。与 BBO、PSO、GA、ACO、ES、PBIL 六种主流启发式算法的实验结果进行比较表明, DE_BBO_MLP 算法在分类精度和收敛速度等方面优于已有方法。

关键词: 生物地理学优化; 差分进化; 多层感知器; 数据分类

中图分类号: TP18 **文献标志码:** A **文章编号:** 1001-3695(2017)03-0693-04

doi:10.3969/j.issn.1001-3695.2017.03.012

Multi-layer perceptron using hybrid differential evolution and biogeography-based optimization

Wang Juan^{1,2}, Wu Xianxiang², Cao Yanling²

(1. Dept. of Basic Courses, Xi'an Fanyi University, Xi'an 710105, China; 2. ICIE Institute, School of Aerospace Science & Technology, Xidian University, Xi'an 710071, China)

Abstract: The problems of premature convergence and initialization-sensitive are often experiencing when train the multi-layer perceptron using the biogeography-based optimization. This paper proposed a novel multi-layer perceptron training method using hybrid differential evolution and biogeography-based optimization. This paper introduced the differential evolution to the biogeography-based optimization to construct the hybrid DE_BBO algorithm and then used the hybrid DE_BBO algorithm for training MLPs. In order to investigate the efficiencies of DE_BBO in training MLPs, this paper employed four classification datasets, including the Iris dataset, the breast cancer dataset, the blood transfusion datasets and the banknote authentication dataset. Comparing with six well-known heuristic algorithms, including BBO, PSO, GA, ACO, ES, and PBIL in a statistically significant way, the experimental results show that training MLPs using hybrid DE_BBO is significantly better than the current heuristic learning algorithms in terms of convergence speed and convergence accuracy.

Key words: biogeography-based optimization; differential evolution; multi-layer perceptron; data classification

0 引言

生物地理学优化(biogeography based optimization, BBO)算法是由 Simon 于 2008 年提出的一种进化算法, 该算法通过研究生物物种迁移数学模型, 借鉴其他仿生智能优化算法的框架而形成^[1]。文献[2,3]表明 BBO 算法具有良好的优化能力, 因此被应用到一些实际工程优化问题, 如图像分割^[4,5]、天线设计^[6,7]、数据分类^[8]、电力系统优化^[9]等。

前馈神经网络(feedforward neural network, FNN)是神经网络中最简单和使用最广泛的形式。FNN 包括单层感知器(single-layer perceptron, MLP)和多层感知器(multi-layer perceptron, MLP)两类。多层感知器由不同层的多个感知器构成, 使得其适合解决非线性问题, 因而被广泛应用于数据分类等领域^[10]。

文献[11]证明, 具有一个隐藏层的多层感知器可以近似任何的连续或者非连续函数。与其他神经网络一样, MLPs 的学习能力对于算法的收敛性能至关重要。学习过程的最终目标是找到在神经网络中连接权重和偏差的最佳组合来实现用于训练和测试样本的最小误差。传统的学习算法普遍存在局部极优和对初值敏感等问题。智能启发式优化方法为这类问题提供了新的解决思路。基因算法^[12]、粒子群算法^[13]、蚁群算法^[14]、进化策略^[15]和声搜索算法^[16]等智能优化算法被成功应用于训练 MLPs, 改善了 MLPs 的全局寻优能力和收敛速度。文献[17]将生物地理学优化应用于训练 MLPs 进行虹膜、乳腺癌、输血、钞票验证^[18]等数据分类, 提高了 MLPs 的分类准确率。

本文将差分进化与生物地理学优化相结合, 利用差分变异算子的开发能力, 构造 DE_BBO 混合算法, 有效改善 BBO 算法

收稿日期: 2015-09-29; **修回日期:** 2016-01-21 **基金项目:** 国家自然科学基金资助项目(61105066); 中央高校基本科研业务费专项资金资助项目(JB141305)

作者简介: 王娟(1980-), 女, 讲师, 主要研究方向为进化计算、图像工程; 吴宪祥(1980-), 男(通信作者), 副教授, 硕士, 博士, 主要研究方向为智能控制与图像工程(wuxianxiang@163.com); 曹艳玲(1989-), 女, 硕士研究生, 主要研究方向为进化计算、图像工程。

变异能力差的情况,避免陷入局部极优,增强算法的全局搜索能力,并利用 DE_BBO 优化 MLPs 输入数据中的特征向量,进一步提高了数据分类的速度和准确性。

1 生物地理学优化算法

1.1 迁移模型

自然界中,各生物种群分布的地方不同,称之为栖息地。每个栖息地的环境均不一样,并且各物种对每个栖息地的适应程度也不同,这造就了物种分布的多样性、迁移和灭绝等现象。每个栖息地具有不同的适宜度指数(habitat suitability index, HSI)。若某栖息地 HSI 较高,表明该栖息地适宜物种生存,其物种较多;HSI 较低,表明该栖息地不适合物种生存,其物种较少。

生物地理学中常用图 1(a)所示的线性迁移率数学模型来描述生物种群的迁移规律^[19]。每个栖息地迁入率 λ_k 、迁出率 μ_k 由式(1)决定:

$$\begin{cases} \lambda_k = I(1 - \frac{k}{n}) \\ \mu_k = E \frac{k}{n} \end{cases} \quad (1)$$

其中: k = 物种数, $n = S_{\max}$, I 和 E 分别为迁入率、迁出率的最大值。栖息地种群数量达到平衡点 S_0 时,该点处迁入率与迁出率相等。

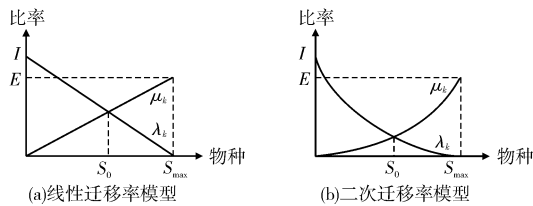


图 1 生物地理学种群迁移率模型

现实中,当栖息地的生物种群数量为零,迁入率取最大值 $\lambda = I$,迁出率 $\mu = 0$ 。随着迁入该栖息地的生物种群增加,栖息地越来越拥挤,越来越少的迁入种群能够存活下来,越来越多的生物种群选择离开该栖息地,迁入率呈现递减趋势,迁出率呈现递增趋势。当种群数量达到饱和时, $\lambda = 0, \mu = E$ 。所以,利用图 1(b)所示的二次迁移率模型描述生物种群的迁移规律更加符合实际情况,迁入率 λ_k 、迁出率 μ_k 由式(2)决定:

$$\begin{cases} \lambda_k = I(1 - \frac{k}{n})^2 \\ \mu_k = E(\frac{k}{n})^2 \end{cases} \quad (2)$$

1.2 BBO 算法流程

2008 年,Simon 在对生物物种迁移数学模型研究的基础上,提出了生物地理学优化算法,其算法流程如下:

a)生成初始栖息地。设定栖息地数量、栖息地内移民的最大容量、最大迁入率 I 、最大迁出 E 和最大变异率 $m_{\max}$ 。为了保证栖息地分散性提高 BBO 搜索性能,每个栖息地定义如下变异率:

$$m_i = m_{\max}(1 - \frac{p_i}{p_{\max}}) \quad (3)$$

其中: $m_{\max}$ 为最大变异率; p_i 为第 i 个栖息地的变异概率; $p_{\max} = \arg \max(p_i), i = 1, 2, \dots, N$ 。

b)计算所有栖息地的 HSI。计算过后按 HSI 从高到低对栖息地排序。

c)根据迁移率模型更新每个栖息地的迁入率 λ_k 、迁出率 μ_k 、变异率 m_i 。

d)修改栖息地。根据每个栖息地的迁入率和迁出率,移民按照迁移策略修改栖息地,并且重新计算 HSI。

e)根据变异算子更新物种,重新计算适宜度。

f)保留每次迭代最佳解。因突变会破坏较优方案的栖息地特征,所以在算法的迭代过程中,保留部分群体精英个体,可以在 HSI 发生突变后能恢复到最佳解决方案。

g)结果是否满意。此步骤主要看是否满足终止的条件,如不满足,则重回到步骤 b),满足条件则输出迭代最优解^[1]。

2 DE_BBO 算法

差分进化(differential evolution, DE)由 Storn 和 Price 于 1995 年提出^[20],是一种基于群体差异的进化算法。在选择作用的影响下,个体间的差异随着进化代数的增加会逐渐降低。个体差异性的减少又影响变异所带来的多样性,导致算法早熟收敛。为此,本文采用一种基于 DE 与 BBO 算法的混合迁移算子,利用差分进化过程产生满足种群需求的步长移动,再通过迁移思想对群体资源进行有效协调与分配,充分发挥原算法的特点且弥补各自的不足之处^[18,21,22]。

2.1 混合迁移算子

混合算子的主要思想:在 BBO 算法的迁移算子中以概率形式引入 DE 的交叉策略,构造混合迁移交叉算子,在保留原算法快速收敛的前提下,扩大了算法的搜索范围,提高了算法的局部搜索能力^[23]。

2.2 变异操作

DE 算法的变异操作如下:

$$U_i(j) = X_{r1}(j) + F[X_{r2}(j) - X_{r3}(j)] \quad (4)$$

其中: i 表示第 i 个栖息地; j 表示适应度向量的第 j 个变量; X 表示当前一代变量值; U 表示重新生成的新一代的变量值。从 $1 \sim N$ 中随机选取 r_1, r_2, r_3 ,并使 r_1, r_2, r_3 和 i 为 4 个互不相等的整数。

若只满足 BBO 算法的条件,则根据迁出率选择迁出的栖息地:

$$U_i(j) = X_k(j) \quad (5)$$

其中: k 表示根据迁出率选择的迁出栖息地。

BBO 与 DE 结合算法,变异操作采用 DE 的差分公式,这样就相当于用 DE 算法取代 BBO 算法中的变异操作,不仅保存了 BBO 算法中迁移操作的更新速度,还大大提高了变异的能力,使其在可行域内的遍历能力增强。

2.3 DE_BBO 算法步骤

基于差分进化的生物地理优化算法步骤如下:

a)设置种群规模 NP ,阈值 I, E ,精英保留个数 $keep$,最大变异概率 p_m ,及 DE 算法的相关初始参数,随机生成初始种群并评估适应值。

b)按照适应值排列群体 P 。

c)若满足停止条件,则输出最优解;否则,转步骤 d)。

d)计算 P 中个体的物种数及移民移出概率。

e)根据混合迁移算子改变群体 P 形成目标群体 P_1 ,并评估适应值。

f)根据选择方案挑选 P 与 P_1 中优秀个体形成 P_2 。

g)对 P_2 进行变异得到下一代群体 P ,评估适应值,转步骤b)。

3 DE_BBO 优化 MLPs

3.1 多层感知器

如图2,给出了三层 MLPs 的结构,其中的输入节点的数目是 n ,隐藏节点的数目是 h ,输出节点的数目是 m 。

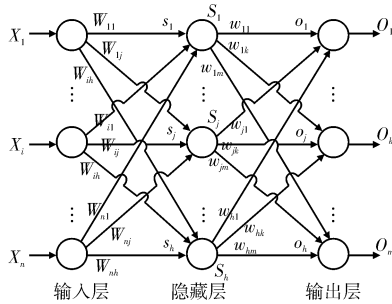


图2 $n-h-m$ 结构的多层感知器

从图2中可以看出,由于MLP属于FNN簇,所以结构中节点间单向连接。首先,计算得到输入的加权和:

$$s_j = \sum_{i=1}^n (W_{ij} \cdot X_i) - \theta_j \quad j=1,2,\dots,h \quad (6)$$

其中: n 是输入节点的数目; W_{ij} 是从输入层第 i 节点到隐藏层第 j 节点的连接权重; θ_j 是第 j 个隐含节点的偏置; X_i 表示第 i 个输入。每一隐藏节点输出可由式(7)计算:

$$S_j = \text{sigmoid}(s_j) = \frac{1}{1 + \exp(-s_j)} \quad j=1,2,\dots,h \quad (7)$$

利用式(8)和(9)分别计算隐节点加权和以及MLP的最终输出:

$$o_k = \sum_{j=1}^h (w_{jk} \cdot S_j) - \theta'_k \quad k=1,2,\dots,m \quad (8)$$

$$O_k = \text{sigmoid}(o_k) = \frac{1}{1 + \exp(-o_k)} \quad k=1,2,\dots,m \quad (9)$$

其中: w_{jk} 是从第 j 个隐藏节点到第 k 个输出节点的连接权重, θ'_k 是第 k 个输出节点的偏置。

3.2 DE_BBO 训练 MLPs

多层感知器最重要的是连接权重和偏置。从上述方程中看出,权重和偏置影响最终输出值。为了使实际输出更接近期望输出,需寻找最佳的连接权重和偏置来训练多层感知器^[17,24]。这需要通过训练集完成多层感知器的学习过程。所有训练样本的均方误差(MSE)为

$$E = \sum_{k=1}^q \frac{\sum_{i=1}^m (o_i^k - d_i^k)^2}{q} \quad (10)$$

其中: q 是训练样本的数量; m 为输出的数量; d_i^k 是在 k 个训练样本下的第 i 个输入单元的期望输出; o_i^k 是 k 个训练样本下的第 i 个输入单元的实际输出。第 i 个栖息地的HSI通过式(11)计算:

$$\text{HSI}(\text{Habitat}_i) = E(\text{Habitat}_i) \quad (11)$$

算法的流程如图3所示。首先,生成一组随机MLP集作为栖息地,每个MLP对应一个栖息地,每个权重和偏置对应栖息地上的居民。经过初始化后,每个MLP的均方误差通过式(10)来计算。然后,通过式(1)~(3)来更新迁入率、迁出率和突变率;结合迁出率和迁入率重新选择MLP;在栖息地突变率的基础上对MLP进行突变。最后,进行精英的选择,其中较好的MLP被保存起来以防止在下一代中遭到进化和变异算子的破坏;重复操作,直到满足终止条件。

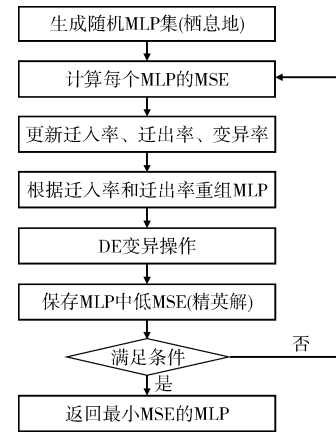


图3 DE_BBO 训练 MLP 流程图

4 实验结果

4.1 样本数据集分类准确性

为了验证本文提出的DE_BBO训练多层感知器的有效性,采用DE_BBO、BBO、PSO、GA、ACO、ES、PBIL算法分别训练多层感知器并应用于虹膜、乳腺癌、输血、钞票验证数据集(来自(UCI)的机器学习资料库^[25])进行对比实验。上述算法每次运行时产生的随机种子数不同,为了消除实验结果的随机性,每个算法运行200次,求取多次运行结果的平均值,各数据集的分类准确率比较如表1所示。

表1 虹膜、乳腺癌、输血、钞票验证数据集分类准确性比较

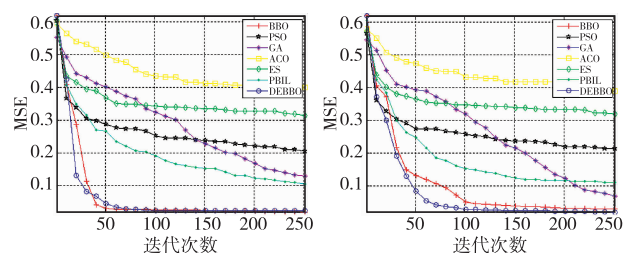
优化方法	虹膜分类准		乳腺癌分类		输血分类		钞票验证分类	
	线性	二次	线性	二次	线性	二次	线性	二次
DE_BBO	90.67	91.33	98.00	95.00	79.44	79.44	100.00	100.00
BBO	89.33	80.67	95.00	98.00	75.81	75.81	99.58	100.00
PSO	35.89	39.60	13.05	21.25	75.93	76.79	90.95	91.60
GA	72.44	68.43	95.60	95.85	77.14	76.55	99.79	99.95
ACO	20.33	34.63	38.30	39.85	76.57	76.51	66.53	67.91
ES	42.11	46.37	3.10	5.95	75.81	76.07	94.22	94.29
PBIL	50.89	61.57	17.40	11.30	76.77	75.69	95.81	96.56

从表1可知,无论对线性还是二次迁移率模型,本文提出的DE_BBO_MLP分类精确率均明显高于其他优化算法。而且,二次迁移率模型数据分类准确率优于线性迁移率模型。

4.2 不同迁移模型下收敛曲线

为了验证DE_BBO_MLP的收敛速度,图4和5分别给出了各种方法进行虹膜分类和输血分类的收敛曲线。

1)虹膜分类 虹膜数据集包含150个样本,分成Setosa、Versicolor和Virginia三类。所有样本都有四个特征:sepal长度、sepal宽度、petal长度和petal宽度。虹膜数据集采用结构4-9-3的多层感知器(MLPs)。收敛曲线如图4所示。



(a) 线性迁移率模型

(b) 二次迁移率模型

图4 虹膜数据集分类收敛曲线

2) 输血分类。随机选择 748 捐赠者的捐赠数据库和 4 个属性(自去年到现在捐献时间(月), 捐献总次数, 总的捐血量, 从第一次捐献到现在的时间(月))。输出采用二进制变量来表示, 1 代表捐赠血, 0 代表不献血。本文用 500 个样本数据训练 4-9-1 结构的多层感知器(MLP), 其余的 248 个样本作为测试集。收敛曲线如图 5 所示。

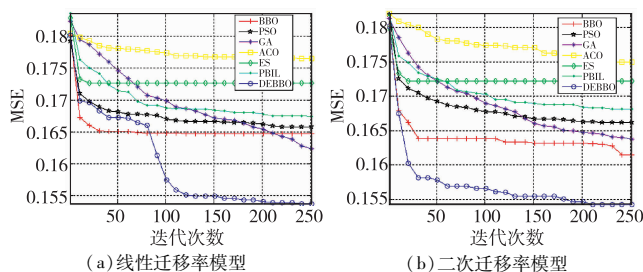


图5 输血数据集分类收敛曲线

根据收敛曲线可以看出, 本文提出的 DE_BBO_MLP 算法在线性迁移率模型和二次迁移率模型下收敛速度均优于其他算法。

5 结束语

本文将差分进化与生物地理学优化相结合, 构造混合 DE_BBO 算法, 利用 DE_BBO 算法进行 MLP 的连接权重和偏置的优化问题。实验数据对比表明, DE_BBO 与传统的 BBO、PSO、GA、ACO、ES 和 PBIL 相比, DE_BBO 训练 MLPs 可以较好地避免局部极小值, 收敛速度快, 收敛鲁棒性高, 有效提高了分类的准确率。

参考文献

- [1] Simon D. Biogeography-based optimization[J]. *IEEE Trans on Evolutionary Computation*, 2008, 12(6): 702-713.
- [2] Ma Haiping, Simon D. Biogeography-based optimization with blended migration for constrained optimization problems[C]//Proc of the 12th annual conference on Genetic and evolutionary computation, ACM, 2010: 417-418.
- [3] Simon D, Rarick R, Ergezer M, et al. Analytical and numerical comparisons of biogeography-based optimization and genetic algorithms[J]. *Information Sciences*, 2011, 181(7): 1224-1248.
- [4] 郑肇葆. 生物地理学优化(BBO)在图像分割中的应用[J]. *武汉大学学报:信息科学版*, 2011, 36(8): 932-935.
- [5] 张新明, 尹欣欣, 冯梦清, 等. 改进的生物地理学算法及其在图像分割中的应用[J]. *电光与控制*, 2015, 22(12): 24-28, 58.
- [6] Singh U, Singh D, Singh P. Concentric circular antenna array design using hybrid differential evolution with biogeography based optimization[C]//Proc of IEEE International Conference on Computational Intelligence and Computing Research. [S. l.]: IEEE Press, 2013: 1-6.
- [7] Singh U, Kumar H, Kamal T S. Design of Yagi-Uda antenna using biogeography based optimization[J]. *IEEE Trans on Antennas and Propagation*, 2010, 58(10): 3375-3379.
- [8] 朱黎辉, 李晓宁. 面向图像分类的多层感知机 BBO 优化方法[J]. *四川师范大学学报:自然科学版*, 2015, 38(6): 930-936.
- [9] Boussaid I, Chatterjee A, Siarry P, et al. Hybridizing biogeography-based optimization with differential evolution for optimal power allocation in wireless sensor networks[J]. *IEEE Trans on Vehicular Technology*, 2011, 60(5): 2347-2353.
- [10] Kruse R, Borgelt C, Klawonn F, et al. Multi-layer perceptrons[M]//Computational Intelligence. London: Springer, 2013: 47-81.
- [11] Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators[J]. *Neural Networks*, 1989, 2(5): 359-366.
- [12] Leung F H F, Lam H, Ling S, et al. Tuning of the structure and parameters of a neural network using an improved genetic algorithm[J]. *IEEE Trans on Neural Networks*, 2003, 14(1): 79-88.
- [13] Mendes R, Cortez P, Rocha M, et al. Particle swarms for feedforward neural network training[C]//Proc of the 2002 International Joint Conference on Neural Networks. [S. l.]: IEEE Press, 2002: 1895-1899.
- [14] Blum C, Socha K. Training feed-forward neural networks with ant colony optimization: an application to pattern classification[C]//Proc of the 5th International Conference on Hybrid Intelligent Systems. [S. l.]: IEEE Press, 2005: 233-238.
- [15] Wienholt W. Minimizing the system error in feedforward neural networks with evolution strategy[C]//Proc of ICANN'93. London: Springer, 1993: 490-493.
- [16] Kulluka S, Ozbakira L, Baykasoglu A. Training neural networks with harmony search algorithms for classification problems[J]. *Engineering Applications of Artificial Intelligence*, 2012, 25(1): 11-19.
- [17] Mirjalili S, Mirjalili S M, Lewis A. Let a biogeography-based optimizer train your multi-layer perceptron[J]. *Information Sciences*, 2014, 269(8): 188-209.
- [18] 陈基涛. 基于高斯变异的生物地理学优化模型[J]. *计算机仿真*, 2013, 30(7): 292-295.
- [19] 马海平, 李雪. 生物地理学优化算法的迁移率模型分析[J]. *东南大学学报:自然科学版*, 2009, 39(1): 16-21.
- [20] Stornand R, Price K. Differential evolution: a simple and efficient heuristic for global optimization over continuous spaces[J]. *Journal of Global Optimization*, 1997, 11(4): 341-359.
- [21] Boussaid I, Chatterjee A, Siarry P, et al. Two-stage update biogeography-based optimization using differential evolution algorithm[J]. *Computer and Operations Research*, 2011, 38(8): 1188-1198.
- [22] Xiong Guojang, Shi Dongyuan, Duan Xianzhong. Enhancing the performance of biogeography-based optimization using polyphyletic migration operator and orthogonal learning[J]. *Computers & Operations Research*, 2014, 41(1): 125-139.
- [23] 叶开文, 刘三阳, 高卫峰. 基于差分进化的生物地理学优化算法[J]. *计算机应用*, 2012, 32(11): 2981-2984.
- [24] Mirjalili S, Hashim S Z M, Sardroudi H M. Training feedforward neural networks using hybrid particle swarm optimization and gravitational search algorithm[J]. *Applied Mathematics Computation*, 2012, 218(22): 11125-11137.
- [25] Blake C, Merz C J. UCI repository of machine learning databases[EB/OL]. 1998. <http://www.ics.uci.edu/~mllearn/LRepository.html>.