

基于辅助种群分类的遗传算法

涂井先, 刘 伟

(广东工业大学 应用数学学院, 广东 广州 510520)

摘要: 提出了基于辅助种群分类的遗传算法, 该算法克服了辅助种群多样性不好的缺点, 利用先验知识将辅助种群分为若干类, 分类后辅助种群与主种群杂交更有利于后代的进化, 同时也更好保证了种群的多样性. 数值试验表明, 改进的算法优于当前一些较好的遗传算法, 并能跳出局部最优解从而求解出全局最优解.

关键词: 遗传算法; 辅助种群; 主种群; 分类

中图分类号: TP18

文献标志码: A

文章编号: 1007-7162(2012)01-0039-04

A Genetic Algorithm Based on the Classification of the Auxiliary Group

Tu Jing-xian, Liu Wei

(Faculty of Applied Mathematics, Guangdong University of Technology, Guangzhou 510520, China)

Abstract: A genetic algorithm, based on the classification of the auxiliary group, was proposed. The algorithm overcame the weakness that diversity of the auxiliary group was not perfect, and with prior knowledge it classified the auxiliary group into several subgroups. After the classification, it is more favourable for the auxiliary group and the main group to evolve. Meanwhile, the diversity of the auxiliary group is maintained. The results show that the improved algorithm is more effective than some other existing genetic algorithms. The improved algorithm avoids being trapped into the local optimum, thus deriving the global optimum.

Key words: genetic algorithm; auxiliary population; main population; classification

遗传算法是一种模拟生物进化的智能优化算法. 由于遗传算法具有全局搜索能力强、容易实现且对目标函数的性质要求低的优点, 遗传算法被广泛应用于实际优化问题中, 成为最有效的优化算法之一^[1-3].

传统的遗传算法主要存在两个问题, 第一个问题是传统的遗传算法局部搜索能力差, 不少学者将一些局部搜索算子嵌入到遗传算法中, 增强了遗传算法的局部搜索能力^[4-6]. 第二个问题是传统的遗传算法容易出现早熟现象, 早熟现象产生的原因是在进化过程中, 种群的多样性过早地丧失. 为了克服遗传算法容易出现早熟现象的缺点, 不少学者相继提出了基于适应值共享的策略^[7-9]和基于双种群的策略, 这些策略在一定程度上维持了种群多样性. 基于适应值策略的缺点是要准确地设定小生境半径很

难. 基于双种群的策略是近几年提出来的, 成为维持种群多样性策略的研究热点之一.

方必和等^[10]人提出了基于淘汰机制的双种群遗传算法, 该算法对两个种群引入了战争, 通过战争淘汰适应值小的个体, 然后随机生成一定数量的个体填入种群, 虽然该算法通过随机生成个体来维持种群多样性, 但是该算法减弱了遗传算法的局部搜索能力. Taejin Park 等^[11]人提出了基于自适应多样性控制的双种群遗传算法, 该算法引进了主种群和辅助种群的概念, 主种群类似于通常意义下的种群, 辅助种群的个体的适应值是由该个体与主种群的中心距离和给定的半径参数决定的, 在进化过程中, 辅助种群的个体与主种群的中心的距离由大到小再由小到大交替变化, 这种变化规律使主种群跳出局部最优解和局部搜索的两个功能交替进行. 虽然该算

法既能有效地防止了早熟现象的产生,又能增强遗传算法的局部搜索能力,但是辅助种群的多样性不好,不利于扩展搜索空间。

为了克服辅助种群的多样性不好的缺点,本文提出了基于辅助种群分类的遗传算法。利用先验知识求出能成功繁衍后代^[11]的父代之间的若干个距离,根据这些距离把辅助种群分为若干类,这些类与主种群杂交产生优良的后代是有希望的。该算法既能增强主种群跳出局部最优解的能力,又能提高遗传算法的局部搜索能力。数值实验表明,该算法优于基于自适应多样性控制的双种群遗传算法和基于适应值共享的遗传算法。

1 基于辅助种群分类的遗传算法

优化问题模型描述如下:

$$\max f(x),$$

其中 $x \in \mathbf{R}^n$, $\mathbf{R}$ 代表实数域, $x = (x_1, x_2, \dots, x_n)$, $x_i \in [a_i, b_i]$ ($i = 1, 2, \dots, n$)。

1.1 辅助种群的适应值函数

主种群的个体用目标函数值作为适应值,辅助种群的个体的适应值与该个体和主种群的中心距离有关。辅助种群的规模为 N , 在选择操作中, 分别用 K 个不同的适应值函数评价辅助种群及其后代, 从这些个体中根据不同的适应值函数评价法分别选择 $\frac{N}{K}$ 个个体到下一代的辅助种群, 选择后的辅助种群规模仍为 N 。辅助种群的适应值函数定义为

$$f_r^i(x) = 1 - |\delta^i - d(\bar{M}, x)|, \quad (1)$$

其中 $i = 1, 2, \dots, K$, δ^i 是第 i 个适应值函数的半径参数, $x = (x_1, x_2, \dots, x_n)$ 为辅助种群的个体, $d(\bar{M}, x)$ 是主种群的中心与 x 的欧氏距离, 主种群的中心可以是所有个体的平均值或者最好的个体, 本文采用主种群所有个体的平均值作为主种群的中心, 假设 $c = (c_1, c_2, \dots, c_n)$ 为主种群的中心, 那么

$$d(\bar{M}, x) = \sqrt{\frac{1}{n} \sum_{j=1}^n \left(\frac{c_j - x_j}{b_j - a_j} \right)^2}. \quad (2)$$

当 $d(\bar{M}, x)$ 越靠近 δ^i 时, $f_r^i(x)$ 越大, 反之 $f_r^i(x)$ 越小。

1.2 半径参数的估计和变化规律

选择合适的半径参数是至关重要的。如果半径参数太小, 辅助种群主要起着对主种群进行局部搜索的功能, 不能维持主种群多样性, 如果半径参数太大, 辅助种群与主种群杂交产生优良的后代则希望

很小。

δ^i 用下面的公式(3)更新:

$$\delta_{t+1}^i = (1 - \alpha)\delta_t^i + \alpha\Delta_t^i, \quad (3)$$

其中, $i = 1, 2, \dots, K$; α 是更新速度, $0 < \alpha < 1$; Δ_t^i 是目标半径, 计算公式为

$$\Delta_t^i = \lambda_1 + (i - 1) \frac{\lambda_2 - \lambda_1}{K - 1}. \quad (4)$$

$$\lambda_1 = \begin{cases} \min\{d(p, p') \mid p, p' \in G_t\} & \text{如果 } |G_t| \geq 2, \\ \delta^1 & \text{其他;} \end{cases} \quad (5)$$

$$\lambda_2 = \begin{cases} \max\{d(p, p') \mid p, p' \in G_t\} & \text{如果 } |G_t| \geq 2, \\ \delta^K & \text{其他;} \end{cases} \quad (6)$$

其中 G_t 是第 t 代成功繁衍^[11]集合。由式(3)~式(6)容易得出下面两个结论:

$$1) \Delta_t^1 \leq \Delta_t^2 \leq \dots \leq \Delta_t^K;$$

$$2) \Delta_t^2 - \Delta_t^1 = \Delta_t^3 - \Delta_t^2 = \dots = \Delta_t^K - \Delta_t^{K-1}.$$

本文利用先验知识把辅助种群分为 K 类, 这些类与主种群杂交产生优良的后代是有希望的。这些类位于不同的层上, 第一类中的个体位于最内层, 第二类的个体位于第二层, 最后一类的个体位于最外层, 这些层在搜索空间比较均匀地排列。首先初始化半径参数: $\delta^1 = \frac{1}{K}$, $\delta^2 = \frac{2}{K}$, $\dots$, $\delta^K = 1$ 。因为 λ_1 和 λ_2 的总体趋势会越来越小, 所以半径参数也会越来越小, 当半径参数很小的时候, 辅助种群的个体与主种群的个体会很相似, 此时有利于主种群进行局部搜索, 但不利于主种群跳出局部最优解。为了增强算法的全局搜索能力, 当 δ^1 连续 τ 代没有变小时, 按照上面的初始化方法重新设定半径参数。总的来讲, 半径参数的变化规律是由大变小和由小变大交替进行, 使主种群跳出局部最优解和主种群局部搜索的两个功能交替进行。该方法既能增强遗传算法的局部搜索能力, 又能有效地防止早熟现象的产生。假设 $\delta_{\min}$ 为进化过程中 δ^1 的最小值, LastUpdated 为 $\delta_{\min}$ 最近一次更新时的进化代数。

调整半径参数的算法步骤如下。

步骤 1: $t = 1$ 。

步骤 2: 若 $t < \text{LastUpdated} + \tau$, 用公式(3)更新 $\delta_t^i, \delta_t^i = \max(0, \min(1, \delta_t^i))$ 。

若 $\delta_t^1 < \delta_{\min}$, $\delta_{\min} = \delta_t^1$, LastUpdated = t 。否则 $\delta_t^1 = 1$, $\delta_{\min} = 1$, LastUpdated = t 。

步骤 3:如果 $t \leq T, t = t + 1$, 返回步骤 2, 否则结束.

1.3 基于辅助种群分类的遗传算法的主要步骤

步骤 1:参数的初始化. 设定主种群规模为 M , 辅助种群规模为 N , 进化代数 t , 最大进化的代数 T , 杂交概率 p_c , 变异概率 p_m , 辅助种群的分类数为 K , 半径参数 $\delta^i, i = 1, 2, \dots, K$.

步骤 2:初始化主种群 $\overline{M}^0$ 和辅助种群 $\overline{N}^0$.

步骤 3:对主种群和辅助种群进行杂交变异操作. 主种群进行杂交变异操作产生的后代的集合记为 $I_{\overline{M}}$, 辅助种群进行杂交变异操作产生的后代的集合记为 $I_{\overline{N}}$, 主种群和辅助种群之间进行杂交变异操作产生的后代的集合记为 C .

步骤 4:对父代种群及其后代种群的个体进行适应值评价.

$$\begin{aligned}\overline{M}^t &= \overline{M}^t \cup I_{\overline{M}} \cup C, \\ \overline{N}^t &= \overline{N}^t \cup I_{\overline{N}} \cup C.\end{aligned}$$

用 $f_{\overline{M}}$ 评价 $\overline{M}^t$, 分别用 $f_r^1, f_r^2, \dots, f_r^K$ 评价 $\overline{N}^t$. 其中, $f_{\overline{M}}$ 为目标函数.

步骤 5:按优胜劣汰的原则从种群中选择一定数量的个体组成下一代的种群. 用锦标赛法从 $\overline{M}^t$ 中选择 M 个个体到主种群 $\overline{M}^{t+1}$ 中, 依次以 $f_r^1, f_r^2, \dots, f_r^K$ 为评价个体的标准, 用锦标赛法从 $\overline{N}^t$ 中选择 $\frac{N}{K}$ 个个体到辅助种群 $\overline{N}^{t+1}$ 中, 在进行锦标赛选择前, 将最好的个体选入下一代种群中.

步骤 6:如果 $t \leq T, t = t + 1$, 返回步骤 3, 否则运算结束.

2 数值实验

本文选择了 4 个经典复杂的多峰函数优化问题, 测试了 3 种算法的求解精度和收敛速度.

f_1 是 Rohachevsky 测试函数, 最小值约为 -0.24 .

$$\min f_1(x, y) = x^2 + y^2 - 0.3 \cos(3\pi x) + 0.3 \cos(4\pi y) + 0.3, \text{ 其中 } x, y \in [-1, 1].$$

f_2 是 Shubert 测试函数, 有 760 个点达到局部最小值, 其中只有 18 个点达到全局最小值, 全局最小值约为 -186.73 .

$$\min f_2(x, y) = \left\{ \sum_{i=1}^5 i \cos[(i+1)x + i] \right\} \times \left\{ \sum_{i=1}^5 i \cos[(i+1)y + i] \right\}, \quad x, y \in [-10, 10].$$

f_3 的全局最小值为 0, 有无穷多个点达到局部最小值.

$$\min f_3(x, y) = (x^2 + y^2)^{0.25} (\sin^2 50(x^2 + y^2)^{0.1} + 1.0), \quad x, y \in [-5.12, 5.12].$$

f_4 是 Schaffer 函数, 全局最小值为 0.

$$\min f_4(x, y) = 0.5 + \frac{\sin^2 \sqrt{x^2 + y^2} - 0.5}{[1 + 0.001 \times (x^2 + y^2)^2]^2}, \quad x, y \in [-10, 10].$$

数值实验在 matlab7.0 中完成, 本文将基于辅助种群分类的遗传算法(记为 CAGA)与另外两个遗传算法进行比较, 分别是文献[11]中的基于自适应性控制的双种群的遗传算法(记为 DPGA)和文献[9]中的基于适应值共享的遗传算法(记为 FSGA). 3 种算法的参数均取: $p_c = 0.9, p_m = 0.1, M = 50, T = 500$. 在 CAGA 和 DPGA 中, $N = 100, \tau = 10, \alpha = 0.2$. 在 CAGA 中, $K = 5$.

3 种算法对每一个测试函数都独立重复运行 20 次. 为了比较求解精度, 本节比较了 20 次实验解的平均值和标准差(见表 1). 为了比较收敛速度, 本文给出了一次实验的进化曲线图(见图 1).

表 1 3 种算法收敛精度的比较

Tab.1 The comparision of the convergence rate between the three algorithms

测试函数	算法	μ	σ
f_1	CAGA	-2.400 32 E-01	2.883 25 E-05
	DPGA	-2.400 24 E-01	1.198 78 E-05
	FSGA	-2.400 27 E-01	9.992 70 E-06
f_2	CAGA	-1.863 25 E+02	1.432 94 E-02
	DPGA	-1.860 56 E+02	2.132 91 E-02
	FSGA	-1.867 25 E+02	1.208 47 E-02
f_3	CAGA	6.283 40 E-02	9.450 47 E-02
	DPGA	8.345 97 E-02	1.048 36 E-01
	FSGA	3.867 80 E-01	2.327 79 E-01
f_4	CAGA	5.673 29 E-03	1.895 92 E-02
	DPGA	9.077 07 E-03	2.056 42 E-02
	FSGA	7.355 83 E-02	1.656 84 E-02

由表 1 可以看出, 在优化 f_1, f_3 和 f_4 时, CAGA 的求解精度优于 DPGA 和 FSGA, 在优化 f_2 时, CAGA 的求解精度优于 DPGA, FSGA 要优于 CAGA. 由图 1 可以看出, 在优化 f_1, f_2, f_3 和 f_4 时, CAGA 的收敛速度比 DPGA 和 FSGA 要快. 总体上说, CAGA 优于 DPGA 和 FSGA.

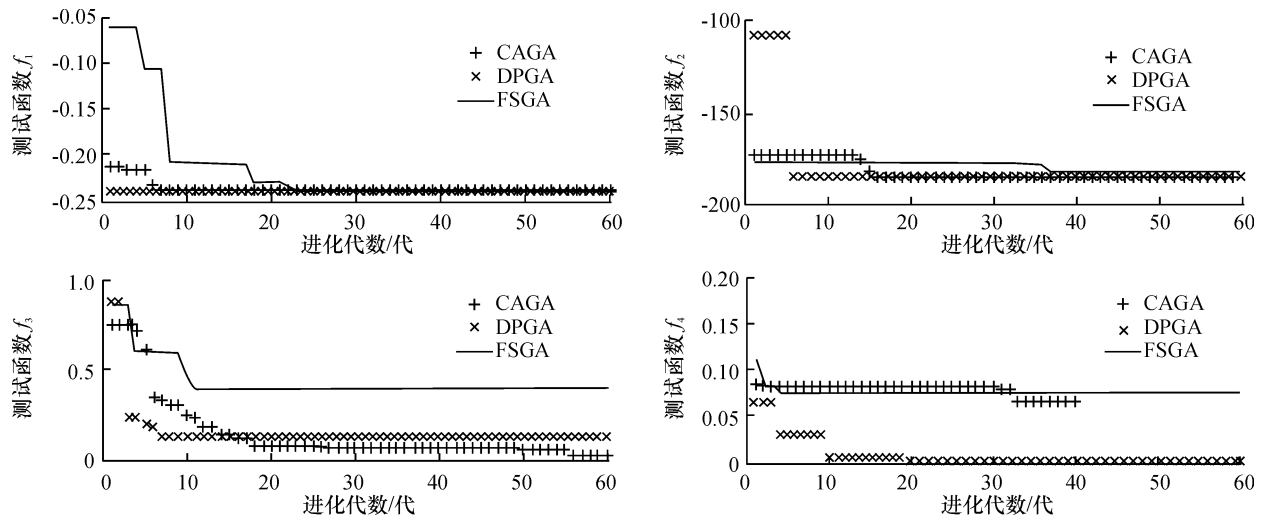


图 1 f_1 、 f_2 、 f_3 和 f_4 的进化曲线图

Fig.1 Evolutionary curves in f_1 , f_2 , f_3 and f_4

3 小结

本文提出了基于辅助种群分类的遗传算法,该算法克服了双种群遗传算法中辅助种群多样性不够好的缺点,利用先验知识对辅助种群进行分类,将辅助种群分为有希望与主种群成功繁衍的若干类.数值实验表明,新算法既能求解到精度比较高的解,又有较快的收敛速度.

参考文献:

[1] Tu Cheng-yuan, Zeng Yan-jun. A new genetic algorithm based upon globally-optimal choosing and its practices [J]. Engineering Science, 2003, 5(02): 28-29.

[2] Ming Zhou, Sun Shu-dong. Principle of Genetic Algorithmn and Application [M]. Beijing: National Defence Industry Press, 1999.

[3] Chuanxin Zhao, Yimu Ji. Particle swarm optimization for 0/1 knaps problem [J]. Microcomputer Develepment, 2005 (10): 23-25.

[4] 唐焕文,秦学志. 实用最优化方法[M]. 3 版. 大连:大连理工大学出版,1999.

[5] 周明,孙树栋. 遗传算法原理及其应用[M]. 北京:国防工业出版社,1999.

[6] 刘伟,刘海林. 基于外点法的混合遗传算法求解约束优化问题[J]. 计算机应用, 2007, 27(1): 238-240.

[7] Goldberg D E, Richardson J. Genetic Algorithms with Sharing for Multimodal Function Optimization[M]. Hillsdale: Lawrence Erlbaum, 1987.

[8] 于歆杰,王赞基. 对适应值共享遗传算法的分类及评价[J]. 模式识别与人工智能, 2001, 14(1): 42-47.

[9] Antonio D C, Claudio D S. On the Role of Population Size and Niche Radius in Fitness Sharing[J]. IEEE Transactions on Evolutionary Computation, 2004, 8(6): 580-592.

[10] 方必和,于蕾蕾. 基于淘汰机制的双种群遗传算法[J]. 计算机技术与发展, 2009, 19(9): 101-104.

[11] Taejin Park, Kwang R R. A Dual-Population Genetic Algorithm for Adaptive Diversity Control[J]. IEEE Transactions on Evolutionary Computation, 2010, 14(6): 580-592.