

基于个体差异的遗传选择算子设计

乔家庆, 付 平, 孟升卫

(哈尔滨工业大学自动化测试与控制系, 黑龙江哈尔滨 150001)

摘 要: 本文针对遗传算法采用基于局部竞争机制的选择算子不易跳出局部最优的问题, 设计了一种基于个体差异的局部竞争选择算子, 局部竞争作用的区域由种群内个体间的差异决定, 有利于保持种群的多样性, 增强了算法从局部最优中跳出的能力. 仿真结果表明, 遗传算法采用基于个体差异的选择算子, 在全局极大值附近存在剧烈震荡的情况下, 有效地保证了算法优化成功率和效率.

关键词: 遗传算法; 选择算子; 局部竞争

中图分类号: TP18 **文献标识码:** A **文章编号:** 0372-2112 (2006) 12A-2414-03

A Genetic Selection Operator Based on Difference Among Individuals

QIAO Jia-qing, FU Ping, MENG Sheng-wei

(Department of Automatic Test and Control, Harbin Institute of Technology, Harbin, Heilongjiang 150001, China)

Abstract: When the local competition based selection is adopted in the genetic algorithm, the algorithm is difficult to escape from the local optimum. In this paper, a selection operator based on the difference among individuals is proposed. Competition occurs within an area defined by the difference, which helps to maintain the population's diversity and enhance the ability to escape from the local optimum. The experiments show that, the genetic algorithm adopting the improved selection operator can effectively ensure the success and effectiveness of the optimization, even when there is severe vibration around the global optimum.

Key words: genetic algorithm; selection operator; local competition

1 引言

遗传算法是模仿自然界“优胜劣汰”进化过程的一种随机搜索算法, 它将问题的可能解作为遗传操作的个体, 由个体组成的种群在选择、杂交和变异三种遗传算子的作用下搜索问题的最优解. 由于其对搜索空间的连续性和分散性没有限制而且可以并行实现^[1,2], 因而得到了广泛地应用, 如控制领域^[3]、图像处理^[4]等.

遗传算法的主要问题是早熟收敛与收敛速度^[5]. 传统的遗传算法在解决复杂问题时, 由于种群在遗传过程中失去了多样性, 往往陷入局部最优解, 不能收敛到全局最优解. 解决这个问题关键在于保持算法收敛速度的同时, 保持种群的多样性^[6].

选择算子在遗传算法中直接决定个体被复制的数目, 对种群多样性有着重要影响^[7]. 因此, 改进选择算子, 是提高遗传算法全局寻优能力的重要手段. 本文设计了一种基于个体差异的局部竞争选择算子, 竞争作用的区域由种群内个体间的差异决定, 削弱了局部最优解在选择过程中的优势, 有利于保持种群的多样性, 增强了算法从局部最优中跳出的能力.

本文第二部分介绍一般选择算子及其特点, 第三部分提出了基于个体差异的局部竞争选择算子, 分析了算子对局部最优解在选择优势的抑制, 并在第四部分通过仿真实验验证了其有效性.

2 选择算子及特点

在遗传算法中, 选择算子使搜索的重点放在适应值较高

的个体上, 使其获得较高的复制数目, 充分利用了群体内当前所具有的有效信息, 但同时选择算子的作用也使群体内个体失去多样性. 不同的选择算子导致不同的选择压力, 较大的选择压力使最优个体具有较高的复制数目, 从而使算法收敛速度快, 但容易陷入局部最优解; 而较小的压力, 不易陷入局部最优解, 但收敛速度慢. 因此, 在收敛速度与跳出局部最优的能力之间进行有效的权衡是设计选择算子时面临的最重要的问题. 选择算子可分为基于适应值比例、基于排名和基于局部竞争三类.

基于适应值比例的选择是以个体在种群中的相对适应值为评价指标的选择策略. 这类选择算子有繁殖池和轮盘式选择, 其中繁殖池选择是根据相对适应值确定个体复制的个数, 适应值小的个体将被剥夺生存机会; 轮盘式选择根据相对适应值确定个体被选择的概率, 群体中所有成员都有被选择的机会, 因而选择压力与繁殖池选择相比较小.

基于排名的选择通过个体适应值在种群中的排名来计算其被选择的概率, 然后根据这个概率使用轮盘选择.

与以上两种选择策略不同, 基于局部竞争机制的选择过程中, 只涉及种群局部的个体间适应值的大小关系, 而不涉及适应值信息, 锦标赛选择算子是一种基于这种策略的选择算子. 选择过程中先随机地在群体中选择 k 个个体进行比较, 适应值大的个体被选择. 参数 k 被称为竞赛规模, 这种选择机制也使得适应值大的个体有较大的概率被选择, 但选择的标准与适应值大小不直接成比例, 因而也能避免群体中出现“超级个体”, 可以一定程度上避免出现早熟收敛现象和停滞现象.

3 局部竞争机制的改进

常用的选择算子通过利用个体的相对适应值或者局部竞争机制避免算法的早收敛和停滞,但对于比较复杂的问题,如在全局最优解附近剧烈震荡的情况下,算法仍易陷入局部最优。

在当前种群中,个体 x 的适应值表示为 $f(x)$, 个体 x^* 不是全局最优解, 但存在 $\delta > 0$, 使得若 $\|x - x^*\| \leq \delta$, 有 $f(x^*) \geq f(x)$, 则称 x^* 为一个局部最优解。

在基于局部竞争机制的锦标赛选择过程中, 进行比较的个体随机的在群体中选择, 设群体规模为 N , 竞赛规模为 2, 群体中适应值最高的局部最优解 x^* 的数目为 $m (2 \leq m \leq N)$, 因为 x^* 是当前种群中最优的, 因此只要被选中的两个个体中有 x^* , 则 x^* 就可以在竞争中胜出, 即被复制到子代种群中。局部最优解被选中, 从而在竞争中胜出的概率 p^* 可表示为:

$$p^* = \frac{C_m^1 C_{N-m}^1 + C_m^2}{C_N^2} \quad (1)$$

p^* 与 m 的关系如图 1 ($N = 30$)。

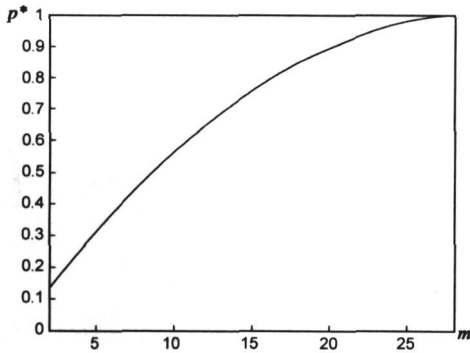


图 1 p^* 与 m 关系图

可见 m 增大, 局部最优解被选中从而在比较中胜出的概率也随之增大, 算法跳出局部最优的概率相应减小。这种基于局部竞争的选择中, 算法陷入局部最优的概率由局部最优解在种群中所占比例决定。因此对选择算子改进, 减小局部最优解在种群中所占比例的影响, 可以增大算法跳出局部最优的概率。

自然界的遗传竞争一般发生在相似或者临近的个体中, 即自然选择在生物圈的局部作用, 模仿这一现象, 可以对基于局部竞争机制的选择算子进行改进。

选择过程中, 某一个体 x_i 与种群中各个个体差异表示为:

$$L_{ij} = \sqrt{\sum_k (x_j(k) - x_i(k))^2} \quad (2)$$

其中 $j = 1, 2, \dots, N$ 。规定局部竞争作用发生在以 x_i 为中心的区域, 这一局部区域内适应值最大的个体被选择, 并规定区域半径

$$R_i = r \times \max\{L_{ij}\}_{j=1,2,\dots,N} \quad (3)$$

其中 r 是 $[0, 1]$ 上的一个平均分布的随机数。

选择作用的局部区域表示为:

$$S_i = \{x_j | L_{ij} < R_i\} \quad (4)$$

显然, x_i 属于 S_i , 参与局部竞争的个体与 x_i 的差异小于 R_i , 即竞争发生在类似个体间。由于区域半径受随机数 r 影响, 这就使得局部竞争的胜出者既可能是目前种群中的最优个体(即局部最优个体), 也可能是与差异 x_i 较小的相对优秀的个体。

对局部竞争机制的改进, 将随机挑选局部竞争候选个体, 改为在以每个父代个体为中心的类似个体间选择适应值较高的个体, 削弱了局部最优个体在选择过程中的优势, 有利于搜索从局部最优中跳出。

设局部最优个体 x^* 与 x_i 的差异为:

$$L^* = \lambda \times \max\{L_{ij}\}_{j=1,2,\dots,N} \quad (5)$$

显然 $0 \leq \lambda \leq 1$ 。若 $\lambda = 1$, x^* 胜出的概率 $p^* = 0$ 。 $\lambda < 1$ 时,

$$p^* = p(L^* < R_i) = p(\lambda < r) \quad (6)$$

由于 r 是 $[0, 1]$ 上平均分布的随机数, 易得:

$$p^* = 1 - \lambda \quad (7)$$

局部最优解被选中从而在比较中胜出的概率与局部最优解在种群中所占比例无关, 而且局部最优个体 x^* 与 x_i 的差异较大时, 参与竞争的概率较小, 有利于保持种群的多样性, 即搜索从局部最优中跳出的概率比改进前大。而且, 这种基于个体差异的选择算子, 父代中每个个体都有一次机会作为局部竞争区域的中心, 因此不会淘汰局部最优解(若父代局部最优解数目为 m , 子代中局部最优解至少为 m), 不会使搜索发散。

4 仿真实验

为了验证基于个体差异选择算子的有效性, 本文选择了优化算法研究中常用的三个 Benchmark 函数进行了仿真试验, 实验结果与采用锦标赛选择的遗传算法进行对比, 仿真中使用的三个函数 f_1, f_2, f_3 分别为^[8]:

为了使仿真的结果更具对比性, 其他参数都进行完全相同的设置, 具体设置为: 种群大小为 30, 最大循环次数为 2000, 采用算术杂交算子和自适应变异算子。分别对三个 benchmark 函数都进行了 50 次仿真, 运算结果分别如表 2 至表 4 所示, 比较的指标分别为优化的成功率、 G_{\min} (最小迭代次数)、 G_{\max} (最大迭代次数)和 G_{average} (平均迭代次数)。

表 1 Benchmark 函数

函数	解空间
$f_1(x_1, x_2) = f_a(x_1, x_2) * f_b(x_1, x_2)$ $f_a(x_1, x_2) = \frac{1}{1 + (x_1 + x_2 + 1)^2(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)}$ $f_b(x_1, x_2) = \frac{1}{30 + (2x_1 - 3x_2)^2(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)}$	$x_1, x_2 \in [-2, 2]$
$f_2(x_1, x_2) = x_1 * \sin(4\pi x_1) - x_2 * \sin(4\pi x_2 + \pi) + 1$	$x_1, x_2 \in [-1, 1]$
$f_3(x_1, x_2) = \frac{1}{1 + \frac{\sin^2(\sqrt{x_1^2 + x_2^2} - 0.5)}{(1 + 0.001(x_1^2 + x_2^2))^2}}$	$x_1, x_2 \in [-100, 100]$

表 2 f_1 函数仿真结果

选择算子	成功率	G_{\min}	G_{\max}	G_{average}
基于个体差异选择	100%	9	178	62
锦标赛选择	100%	120	1922	928

表 3 f_2 函数仿真结果

选择算子	成功率	G_{\min}	G_{\max}	G_{average}
基于个体差异选择	100%	13	1086	139
锦标赛选择	92%	135	2000	1585

表 4 f_3 函数仿真结果

选择算子	成功率	G_{\min}	G_{\max}	G_{average}
基于个体差异选择	100%	36	803	288
锦标赛选择	48%	313	1989	1156

从实验数据可以看出,对于 f_1 和 f_2 两个函数,采用两种选择算子的遗传算法找到全局最优解的成功率都很高;对于比较复杂的函数 f_3 ,由于其在全局极大值附近剧烈震荡,而其它位置相对平坦,因此算法很容易陷入局部极值。对 f_3 的实验中,采用锦标赛选择算子的遗传算法的成功率较低,而采用基于个体差异选择算子的遗传算法仍保持了比较高的寻优成功率。

在对三个函数 f_1, f_2, f_3 的仿真实验中,成功找到全局最优解的情况下,采用基于个体差异选择算子的遗传算法的平均迭代次数远小于采用锦标赛选择算子的遗传算法。但是,由于基于个体差异选择算子需要计算种群内个体间差异,确定属于选择作用区域的个体,并且要找出选择作用区域内的最优个体,因此计算量大于锦标赛选择算子。但考虑到采用基于个体差异选择算子时较高的优化成功率和较小的迭代次数,基于个体差异选择算子的性能仍优于锦标赛选择算子。

5 结论

本文设计了一种基于个体差异的局部竞争选择算子,通过种群内个体间的差异规定选择作用的区域,减小了局部最优解被选择的概率,有利于保持种群的多样性,增强了遗传算法从局部最优中跳出的能力。对三个 Benchmark 函数的实验结果表明,采用基于个体差异选择算子的遗传算法相对于采用标准锦标赛选择的遗传算法的性能有了较大改善,在全局极大值附近存在震荡的问题可以取得比较理想的优化结果。

参考文献:

- [1] Shu Fa Hwang, Rong-Song He. A hybrid real-parameter genetic algorithm for function optimization[J]. *Advanced Engineering Informatics*, 2006, (20): 7-21.
- [2] Chen Hai-ying, Guo Qiao, Xu Li. Grid-based pseudo-parallel genetic algorithm and its application[J]. *Journal of Beijing Institute of Technology*, 2006, 15(1): 48-52.
- [3] Kristinsson K. System identification and control using genetic algorithm[J]. *IEEE Trans Syst Man Cybern*, 1992, 22(5): 1033-1046.

- [4] Huang HC, Pan JS, Lu ZM, Sun SH, Hang HM. Vector quantization based on genetic simulated annealing[J]. *Signal Process*, 2001, 81(7): 1513-1523.
- [5] Andre J, Siarry P, Dognon T. An improvement of the standard genetic algorithm fighting premature convergence in continuous optimization[J]. *Advances in Engineering Software*, 2001, (1): 49-60.
- [6] Chang-Yong Lee. Entropy-Boltzmann selection in the genetic algorithms[J]. *IEEE Trans Syst Man Cybern-Part B: Cybernetics*, 2003, 33(1): 138-142.
- [7] 潘正君, 康立山, 陈毓屏. 演化计算[M]. 北京: 清华大学出版社, 1998.
Pan Zheng-jun, Kang Li-shan, Chen Yu-ping. *Evolutionary Computation*[M]. Beijing: Tsinghua University Press, 1998. (in Chinese)
- [8] 孙宁, 彭喜元, 乔立岩. 引导型免疫算法研究[J]. *电子学报*. 2005, 12(A): 2401-2405.
Sun Ning, Peng Xi-yuan, Qiao Li-yan. Study on guiding immune algorithm[J]. *Acta Electronica Sinica*, 2005, 12(A): 2401-2405. (in Chinese)

作者简介:



乔家庆 男, 1978 年出生于哈尔滨, 在读博士研究生, 哈尔滨工业大学自动化测试与控制系讲师。主要研究方向为测试优化与压缩、演化计算。E-mail: qiaojiaqing@hit.edu.cn



付平 男, 1965 年生于黑龙江哈尔滨, 博士学位。现为哈尔滨工业大学电气工程及自动化学院自动化测试与控制系研究员, 博士生导师, 中国电子学会及中国计量测试学会高级会员, 主要研究方向为计算机自动测试与控制、图像处理。E-mail: fuping@hit.edu.cn



孟升卫 男, 1970 年出生于河北省藁城县, 在读博士研究生, 哈尔滨工业大学自动化测试与控制系副教授。主要研究方向为信号高速测量、精密时间间隔测量和数字通讯技术等。E-mail: mengsw@hit.edu.cn