

基于染色体差异与变异衰减原则的进化算法

房茂燕¹, 汪民乐¹, 毕义明¹, 郭 慧²

(1. 第二炮兵工程大学理学院, 西安 710025; 2. 第二炮兵驻 159 厂军事代表室, 北京 100074)

摘 要: 传统进化算法将种群中的每个染色体看作独立的个体, 没有充分挖掘其内部关系。为解决该问题, 提出一种基于染色体差异与变异衰减原则的进化算法。利用染色体的指标值和染色体的结构进行交叉运算, 给出经验交叉概率公式, 对每个染色体设置相亲次数, 按照生物进化规律, 将变异衰减原则引入到进化算法的变异算子中。实验结果表明, 该算法 3 种函数的平均进化代数最高为 78, 搜索成功率最低为 99%, 优于基本遗传算法和引入分裂算子的遗传算法。

关键词: 个体差异; 遗传算法; 进化; 种群; 交叉; 变异衰减

Evolution Algorithm Based on Chromosome Difference and Variation Attenuation Principle

FANG Mao-yan¹, WANG Min-le¹, BI Yi-ming¹, GUO Hui²

(1. College of Science, The Second Artillery Engineering University, Xi'an 710025, China;

2. The Second Artillery Military Representative Office in 159 Plants, Beijing 100074, China)

【Abstract】 Traditional evolutionary algorithm sees each chromosome of the population as independent individuals, and doesn't make fully minning between its internal relations. In order to solve this problem, this paper proposes an evolution algorithm based on chromosome difference and variation attenuation principle. It uses the chromosome index and chromosome structure to do cross operation, gives experience crossover probability formula, set blind date frequency for each chromosome, according to the evolution law, variation attenuation principle is introduced into the mutation operator of evolutionary algorithm. Experimental results show that under 3 functions, the highest average evolution of this algorithm is 78, minimum search success rate is 99%, and is better than basic Genetic Algorithm(GA) and GA introduced division operator.

【Key words】 individual difference; Genetic Algorithm(GA); evolution; population; crossover; variation attenuation

DOI: 10.3969/j.issn.1000-3428.2013.05.049

1 概述

传统的进化算法如遗传算法(Genetic Algorithm, GA), 将每一个候选解看作一个独立的个体, 即候选解之间没有相互关系^[1-2]。而目前对遗传算法的改进主要集中在改进遗传操作和优化控制参数上^[3], 文献[4-5]改进的遗传算法重点也是在这些方面。然而, 通过一定选择机制筛选出来的候选解集包含整体遗传信息, 代表着进化算法的进化方向和强度, 影响着算法的有效性。因此, 在进化过程中, 忽视这一部分优良解集的整体属性, 以及算法对搜索空间的探索能力势必引起以下情况:

- (1) 种群进化过程中的进化遗传信息不完整。
- (2) 进化的同时不能探索待优化问题的内部结构。
- (3) 容易造成算法收敛到局部解, 引起早熟现象。

文献[6]利用贝叶斯网络对种群进行进化, 解决了上面的问题, 但是贝叶斯网络的结构学习是一个 NP 难题, 计算量非常大^[7]。众所周知, 差异越大的染色体交叉后产生新型染色体的概率越高, 差异越小的染色体交叉产生新型染色体的概率相对越小, 这个原则为充分挖掘优良种群中各染色体之间的相互关系, 对最优解可能存在的方向进行搜索, 避免了标准进化算法的缺陷和贝叶斯优化算法的不足。本文以上述内容为基础, 根据染色体差异设计交叉概率计算模型, 依据变异衰减原则构造变异概率计算公式, 从而提出一种基于染色体差异与变异衰减原则的进化算法。

2 算法设计

对于 GA 种群中的染色体进化来说, 在进化过程中怎样能沿着最优的方向发展, 又能跳出局部最优而寻找全

基金项目: 陕西省自然科学基金资助项目(2011JM8013)

作者简介: 房茂燕(1979—), 男, 讲师、博士研究生, 主研方向: 人工智能, 神经网络; 汪民乐、毕义明, 教授、博士生导师; 郭 慧, 助理工程师

收稿日期: 2012-05-14 **修回日期:** 2012-07-24 **E-mail:** fmy19791019@163.com

局最优优点是遗传算法面临的主要问题。解决此问题的关键是寻找新的进化方向, 即产生的新一代染色体更加多样。

2.1 选择算子介绍与基于染色体差异的交叉算子

选择算子采用基于适应度函数的赌轮选择算子, 同时对每一代中的最优个体进行保护, 即将上一代中最优的一个染色体不经过进化操作直接进入下一代, 从而避免退化现象^[8-9]。

在种群染色体的交叉进化过程中, 染色体的差异主要体现在 2 个方面: (1) 由染色体不同引起的指标值差异。(2) 染色体自身的差异, 特别是 0-1 编码的染色体, 便于比较染色体本身的差异。对于染色体的进化来说, 2 个染色体的差异越大, 其交叉产生新染色体的多样性越大, 跳出局部最优的概率越大。因此, 本文通过比较 2 个染色体指标值的差异来确定 2 个染色体的交叉概率。

2.1.1 由染色体对应指标值差异引起的交叉概率

根据上面的差异原则, 在交叉算子中, 当 2 个染色体的指标值差异较大时, 交叉概率也应该较大; 差异较小时, 交叉概率应该较小。设定一个基础交叉概率为 p_0 ($0 < p_0 < 1$), 即指标值相差最小的 2 个染色体的交叉概率接近 p_0 , 指标值相差最大的 2 个染色体的交叉概率为 p_1 。根据这个原则, 本文给定 2 个染色体的交叉概率公式:

$$p = p_0 + (p_1 - p_0) \frac{|t(i) - t(j)|}{M_{\max} - M_{\min}} \quad (1)$$

其中, p 表示第 i 个、第 j 个染色体交叉的概率; M_{\max} 表示种群中所有染色体对应的指标值中最大的指标值; M_{\min} 表示种群中所有染色体对应的指标值中最小的指标值; $t(i)$ 表示第 i 个染色体的指标值; $t(j)$ 表示第 j 个染色体的指标值; $0 \leq 1 - \frac{|t(i) - t(j)|}{M_{\max} - M_{\min}} < 1$, 根据指数函数的单调性,

由于 $0 < p_1 < 1$, $p_0 < p \leq p_1$, 而且 2 个染色体指标值的差异性越大, 交叉概率越大, 当第 i 个、第 j 个染色体的指标值分别为最大最小值时, 交叉概率 $p = p_1$ 。

2.1.2 由染色体本身差异引起的交叉概率

染色体本身差异是指由染色体编码不同引起的差异, 本文利用差异度概念来衡量染色体的差异, 差异度是用 2 个染色体基因编码不相同的位数与染色体的基因位数的比值确定。设 2 个染色体的差异度为 p_s , 则:

$$p_s = \frac{n'}{n} \quad (2)$$

其中, n' 表示基因不相同的位数; n 表示染色体的基因位数。

根据 2 个染色体差异越大、交叉概率越大的原则, 下面确定 2 个染色体的交叉概率:

$$p = p_0^{1-p_s} \quad (3)$$

其中, p_0 为基础交叉概率。

2.1.3 相亲原则

没有参与交叉的染色体称为单身染色体。在 2 个染色体被选中, 称为一次相亲, 然后按照其对应的交叉概率进行运算, 如果两者依概率不能交叉, 那么这 2 个染色体仍为单身, 如果它们参与的相亲次数不超过设定的阈值 k , 即最大相亲次数, 它们可以继续和单身染色体进行相亲, 直到所有染色体或者参与交叉运算, 或者相亲次数达到最大相亲次数。

2.2 基于进化衰减原则的变异算子

在物种的进化过程中, 在新物种产生时, 需要对环境进行适应, 变异率相对较高, 随着物种不断发展繁衍, 物种逐渐适应了环境, 从而趋于相对稳定, 变异率逐步减小, 这便是变异衰减原则。体现在遗传算法上便是随着进化代数的增加, 变异率逐步减小。但是, 虽然变异率逐步减少, 变异的可能仍然存在, 因此, 变异率不能够为 0, 设定变异最小值 p_{m0} 。在遗传算法的进化过程中, 变异概率采用下面的变化公式:

$$p_m = p_{m1} a^{N-1} \quad (4)$$

$$p_{mN} = \begin{cases} p_m & p_m > p_{m0} \\ p_{m0} & p_m \leq p_{m0} \end{cases} \quad (5)$$

其中, p_{m1} 表示第 1 代的变异率; p_{mN} 表示第 N 代的变异率; a 表示衰减速率, $0 < a < 1$; N 表示进化代数。

基于神经网络学习规则的进化算法操作流程如下:

Step1 随机产生一组规模为 m 的种群。

Step2 计算每一个体的适应度。

Step3 对当前群体作用选择算子, 选择出 M 个体, 构成繁殖下一代的中间群体。

Step4 按照染色体差异与相亲原则对种群进行交叉运算。

Step5 根据变异率衰减原则进行变异算子运算。

Step6 判断是否满足终止条件, 则输出当前最优解, 否则, 转 Step2。

3 仿真实验

本文针对参考文献[10-12]提出的 3 个函数, 利用本文算法通过 Matlab 软件编程得到了各自的结果, 与文献[5-6]的结果进行比较。本文将基于指标值差异的基础交叉概率设为 0.65, 最大交叉概率为 0.9, 基于染色体本身的基础概率为 0.7, 相亲次数为 1, 初始变异率为 0.3, 变异率按照每代 0.95 的比率降低, 最低不小于 0.01。设定初始种群规模为 100, 终止规则为进化达到 100 代, 或连续 10 代出现相同的结果, 对如下函数分别进行 100 次仿真实验, 3 种算法得到结果如表 1 所示, 其中, 种群规模指染色体个数。

函数 1 $f_1(x) = |\sin(30x)| \times (1 - \frac{x}{2})$, $x \in [0, 1]$, 该函数有多个极大值点, 最大值点 $x = 0.0517900$, $\max f_1(x) =$

0.973 962 6，精度为 10^{-6} 。

函数 2 函数 $f_2(x) = x \sin(10\pi x) + 2.0, -1 \leq x \leq 2$ 存在多个极值点，最大值比 3.85 略大。

函数 3
$$f_3(x_1, x_2) = 0.5 - \frac{\sin^2 \sqrt{x_1^2 + x_2^2} - 0.5}{[1 + 0.001(x_1^2 + x_2^2)]^2}$$
$$-100 < x_1, x_2 < 100$$

此函数的全局极大点是(0,0)，极大值为 1，而在距全局极大点大约 3.14 范围内的隆起有无限多的局部极大点，设定收敛精度为 0.01。

由表 1 可知，本文算法求解的函数平均值、进化收敛的速度、搜索成功率优于其他 2 种算法，即本文算法具有预防 GA 早熟和加速进化的功效，收敛效果较好。

表 1 3 种算法对函数进行 100 次仿真实验的结果

函数	种群规模	算法	函数平均值	平均进化代数	搜索成功率/(%)
$f_1(x)$	100	基本遗传算法	0.971 438 70	139	82.3
	100	引入分裂算子的遗传算法	0.972 246 20	106	87.8
	100	本文算法	0.973 962 60	78	100.0
$f_2(x)$	100	基本遗传算法	3.850 000 00	16	99.5
	100	引入分裂算子的遗传算法	3.850 170 00	134	99.7
	100	本文算法	3.850 270 00	50	100.0
$f_3(x_1, x_2)$	200	基本遗传算法	0.987 837 76	43	90.3
	500	引入分裂算子的遗传算法	0.988 702 77	34	93.5
	100	本文算法	0.996 596 65	28	99.0

4 结束语

本文提出一种基于染色体差异与变异衰减原则的进化算法。在种群染色体的交叉进化过程中，根据 2 个染色体的差异越大，其交叉产生新染色体的多样性越大，跳出局部最优的概率也越大的原则，给出不同情况下的交叉概率公式。在进化过程中，考虑物种变异率的衰减原则，使变异率随着进化代数的增加逐步减少，直到某一值而不再减少。实验结果表明，该算法的收敛效果较好。今后将添加新的参数，研究其对算法效率的影响，并改进各参数之间的关系。

参考文献

[1] 杨有龙. 基于图形模型的智能优化[D]. 西安: 西北工业大学, 2003.

[2] Davis L. Handbook of Genetic Algorithm[M]. [S. l.]: Van Nostrand Reinhold, 1991.

[3] 帅训波, 马书南, 邵艳伟, 等. 基于两种新型遗传算子的优化组合遗传算法[J]. 计算机系统应用, 2010, 19(7): 98-101.

[4] 刘 刚, 曹 勇, 李华德. 几种改进遗传算法的性能比较[J]. 微计算机信息, 2007, 23(30): 190-192.

[5] 丛 飏, 陈卓然. 浅析遗传量子算法与遗传算法在函数极值问题中的比较法[J]. 吉林师范大学学报: 自然科学版, 2008, 29(1): 34-37, 56.

[6] 钟小平, 李为吉, 赵 燕. 改进的贝叶斯优化算法及应用[J]. 机械科学与技术, 2006, 25(4): 497-500.

[7] 谢娟英, 张 琰, 王春霞, 等. 基于分裂算子的遗传算法和自适应遗传算法[J]. 计算机工程与应用, 2010, 46(33): 28-31.

[8] 袁丽华, 黎 明, 李军华. 基于一因多效的遗传算法研究[J]. 系统工程与电子技术, 2007, 29(11): 1969-1972.

[9] 汪民乐, 高晓光, 刘 刚. 遗传算法早熟问题的定量分析及其预防策略[J]. 系统工程与电子技术, 2006, 28(8): 75-80.

[10] 钟流清. BN 结构和参数学习算法改进研究与设计[J]. 微型电脑应用, 2001, 17(5): 8-10.

[11] 潘正君, 康立山, 陈毓屏. 演化计算[M]. 北京: 清华大学出版社, 1998.

[12] Michalewicz Z. Genetic Algorithm+Data Structure=Evolution Programmings[M]. 3rd ed. [S. l.]: Spring-Verlag, 1996.

编辑 刘 冰