

自适应策略在实数编码遗传算法中的应用研究

安建成, 靳海娟

(太原理工大学 计算机与软件学院, 山西 太原 030024)

摘 要: 为解决简单遗传算法收敛速度慢以及局部收敛问题, 在研究自适应策略的基础上, 提出了一种基于实数编码, 综合精英保留策略、2/4 竞争选择策略和自适应策略的改进遗传算法. 该改进算法将自适应策略及其在遗传算法中的应用方法做了改进, 仿真实验证明, 该算法可以提高收敛速度, 有效实现全局最优化.

关键词: 实数编码; 精英保留; 自适应策略; 2/4 竞争选择

中图分类号: TP301

文献标识码: A

文章编号: 1000-7180(2011)04-0140-03

Research on Application of Self-adaptive Strategy in Real-coding Genetic Algorithm

AN Jian-cheng, JIN Hai-juan

(College of Computer and Software, Taiyuan University of Technology, Taiyuan 030024, China)

Abstract: To solve the slow convergence rate and local convergence of Simple Genetic Algorithm, on the basis of researching self-adaptive strategy, an improved genetic algorithm (IGA) with real-coding, elite reservation, 2/4 competitive choosing and self-adaptive strategy is proposed. Self-adaptive strategy and its application in genetic algorithm are improved in the paper. The experiment shows that the improved algorithm is more effective in realizing the global optimization and promoting evolution efficiency.

Key words: real-coding; elite reservation; self-adaptive strategy; 2/4 competitive choosing

1 引言

遗传算法 (Genetic Algorithm, GA) 由美国 Michigan 大学的 Holland 于 1975 年首先提出^[1,3,4], 是模拟达尔文的遗传选择和自然淘汰的生物进化论的计算模型. 它作为一种有效的全局并行优化搜索工具, 具有简单、通用、鲁棒性强和适于并行分布处理的特点.

常用的简单遗传算法采用的是二进制编码、适应度比例选择、单点交叉和单点变异的遗传操作方式, 该算法在应用中有一些明显的不足, 主要是不成熟收敛和收敛速度慢的问题. 另外, 采用二进制编码也有不足之处, 如在变异时, 二进制各基因位所表现出的位权重不一样, 导致变异后解增量的突变现象, 如二进制串 0110, 若第一位变异 (为 0111), 串值改变量是 1, 而第四位变异后 (为 1110), 串值改变量是

8, 变异越往高位, 串值的改变量越大, 结果可能导致接近最优点的个体遗漏^[3,7].

基于实数编码的 GA 是对连续参数优化问题的自然描述, 不存在编码和解码过程, 能大大提高解的精度和收敛速度; 便于进行大空间搜索; 既可以克服二进制编码引起的海明悬崖 (Hamming cliffs) 问题, 又具有微调功能; 同时具有便于和其他搜索技术相结合等特点^[5,6]. 鉴于以上原因, 文中采取实数编码方式. 采用实数编码时, 个体的表示形式为:

$$X_i^t = [x_{i1}^t, \dots, x_{ik}^t, \dots, x_{in}^t] \quad (1)$$

式中, x_{ik}^t ($k = 1, \dots, n$) 为实数变量, t 为进化代数, 初始代 $t = 0$, 个体编号 $i = 1, 2, \dots, N$, n 为变量个数, N 为种群的个体数目.

2 一种改进的遗传算法

针对基本遗传算法的缺陷, 很多人提出了不同

的改进方法. 郑生荣、赖家美等^[3]提出了网格分布法来产生初始种群,采用动态自适应交叉率、近亲杂交回避机制及 2/4 竞争生存法相结合的混合杂交方法,并提出了一种新的动态交叉概率计算方法和变异概率计算方法. 陈明杰^[2]和高学金^[8]均针对传统自适应遗传算法各自提出了新的自适应交叉概率和变异概率计算方法.

文中针对基本遗传算法的缺陷,对自适应交叉遗传算法的应用以及自适应交叉率和变异率的计算方法做出了改进,通过仿真实验证明,改进的遗传算法在提高收敛速度和改善局部收敛问题上均有所改进.

2.1 初始种群生成

在大部分遗传算法的设计中,初始群体中的个体是随机产生的. 根据问题的先验知识,分析最优解子空间在整个问题空间中的分布情况,将最优解所在子空间设定为初始群体生成区域. 这种方法没有考虑初始群体中个体在解空间的分布情况,因而可能导致多个个体集中在某一局部区域内,不利于扩大搜索空间和搜索到全局最优点. 初始群体分布性差会严重影响整个算法的收敛性能甚至导致算法不收敛. 在此本文采用文献^[8]使用的小区间生成法,即先把各待优化参数的取值范围分成群体总数个小小区间,在各小区间中分别随机生成初始个体. 这样生成的初始个体将会均匀地分布在整个解空间上,并能保证随机产生的个体间有明显的差距,保证了初始群体含有较丰富的模式,增强了搜索收敛于全局最优解的可能.

2.2 选择

基本遗传算法使用比例选择算子,利用与每个个体适应度成比例的概率决定其子孙的遗传可能性. 若设种群数为 M ,个体 i 的适应度为 f_i ,则个体 i 被选取的概率为 $P_i = f_i / \sum_{k=1}^M f_k$,当个体的概率给定后,由 $[0, 1]$ 间的随机数来决定参加交配的个体,正是由于这样选择的随机性,使得适应度大的个体也可能落选^[2,3,9]. 鉴于此,本文按照基于适应度排序的方法选择个体进入下一步操作.

2.3 交叉和变异策略

文中将进化过程分为进化初期和进化后期. 进化初期,种群中个体的适应度相差比较大,容易产生新的个体,因而使用固定的交叉率和变异率,交叉率较大,变异率较小. 在进化后期由于种群日趋成熟,个体间的适应度差异逐渐变小,不容易产生新的个

体,进化停滞,因而使用自适应的交叉率和变异率,以期提高算法的收敛速度^[2,3].

Srinivas^[8,9]等提出了传统的自适应遗传算法,变异率和交叉率能随适应度自动改变,当种群各个体适应度趋于一致或者趋于局部最优时,是交叉率和变异率增大,而当种群适应度比较分散时,是交叉率和变异率减少. 其中,交叉率 P_c 和变异率 P_m 按如下公式进行自适应调整:

$$P_c = \begin{cases} k_1(f_{\max} - f') / (f_{\max} - f_{\text{avg}}), & f' \geq f_{\text{avg}} \\ k_2, & f' < f_{\text{avg}} \end{cases} \quad (2)$$

$$P_m = \begin{cases} k_3(f_{\max} - f) / (f_{\max} - f_{\text{avg}}), & f \geq f_{\text{avg}} \\ k_4, & f < f_{\text{avg}} \end{cases} \quad (3)$$

式中, $f_{\max}$ 是种群最大的适应度值, f_{avg} 是每代种群的平均适应度, f' 是要交叉的两个个体中较大的适应度值, f 是要变异的个体的适应度. k_1, k_2, k_3, k_4 均是 $(0, 1)$ 区间的值.

当个体适应值低于平均适应值时,说明该个体性能不好,应对它采用较大的 P_c 和 P_m . 如果适应度值高于平均适应度值,说明该个体性能优良,应赋予其相应的 P_c 和 P_m . 可以看出,当适应值越接近最大适应值时, P_c 和 P_m 越小,当适应度值等于最大适应值时 P_c 和 P_m 的值为 0. 这种方法对于群体处于进化后期比较合适,但对于进化初期不利,因为进化初期群体中的较优个体几乎处于一种不发生变化的状态,而此时的优良个体不一定是优化的全局最优解,容易使进化趋向局部最优解. 因此需要做进一步改进,使群体中最大适应度值的个体的 P_c 和 P_m 不为 0.

考虑到动态自适应交叉率与当代种群的最大适应度、平均适应度及参与杂交的两个个体中适应度较大个体的适应度有关,根据推荐交叉概率为 0.4 ~ 0.9^[9] 的情况,提出如下改进的自适应交叉率计算公式:

$$p_c = \begin{cases} 0.4 + 0.5 \cdot k_1 \cdot (f_{\max} - f_{\text{avg}}) / f_{\text{big}} & (f_{\max} - f_{\text{avg}}) \leq f_{\text{big}} \\ 0.4 + 0.5 \cdot k_2 & (f_{\max} - f_{\text{avg}}) > f_{\text{big}} \end{cases} \quad (4)$$

根据推荐变异概率为 0.0001 ~ 0.1 的情况,改进的自适应变异率计算公式如下:

$$p_m = \begin{cases} 0.0001 + 0.0999 \cdot k_1 \cdot (f_{\max} - f_{\text{avg}}) / f & (f_{\max} - f_{\text{avg}}) \leq f \\ 0.0001 + 0.0999 \cdot k_2 & (f_{\max} - f_{\text{avg}}) > f \end{cases} \quad (5)$$

其中, $f_{\max}$ 是当代种群最大适应度, f_{avg} 是当代种群

平均适应度, f_{big} 是交叉的两个个体中较大的适应度, f 是变异的个体的适应度. k_1, k_2 是 $[0, 1]$ 之间的随机数. 这样, 当 $(f_{\text{max}} - f_{\text{avg}}) \leq f_{\text{big}}$ 或 $(f_{\text{max}} - f_{\text{avg}}) \leq f$ 时, 个体的适应度较高, 个体比较优良, 因而减小交叉率和变异率; 反之, 则相应增大交叉率和变异率.

为了防止子代适应度退化, 可在进化初期将交叉对与交叉得到的两个子个体实行 2/4 生存竞争选择^[3,9], 四个个体中两个优秀者进入下一步操作, 另两个抛弃. 在进化后期, 为了防止自适应进化阶段优良个体被破坏, 使用精英保留策略来保留优良个体. 精英保留策略就是把每一代遗传操作后产生的新一代种群的最高适应度与上一代种群的最高适应度作比较, 如果小于上一代的最高适应度, 就随机淘汰新一代中的一个个体, 把上一代中具有最高适应度的个体加入到新一代中^[2,6].

3 仿真实验

文中将改进的遗传算法应用到两个典型的函数的测试, 并将其与简单遗传算法 (SGA) 进行了比较.

测试函数 1: 单变量函数优化

$$f_1(x) = x \sin(10\pi \cdot x) + 2.0, x \in [-1, 2] \quad (6)$$

测试函数 2: 多元单峰函数优化

$$f_2(x) = \sum_{i=1}^n x_i^2, -512 \leq x_i \leq 512 \quad (7)$$

第一个测试函数的仿真实验中, 种群个体数目为 40, 每个个体长度为 20, 最大遗传代数为 25, 交叉率为 0.7, 变异率为 0.001. 基本遗传算法某次执行后的结果如图 1 所示, 改进的遗传算法某次执行后的结果如图 2 所示.

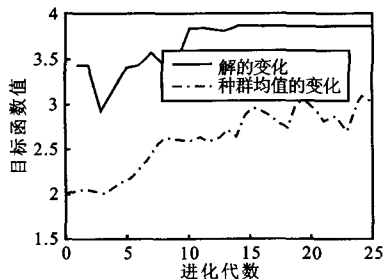


图 1 SGA 对函数 1 的结果

由图 1 和图 2 可以看出, 基本遗传算法进化速度慢, 进化有停滞现象, 收敛速度慢, 在第 17 代时方可收敛到 3.8499. 而改进的遗传算法进化速度很快, 种群优良性稳步提高, 算法能跳出局部最优, 在

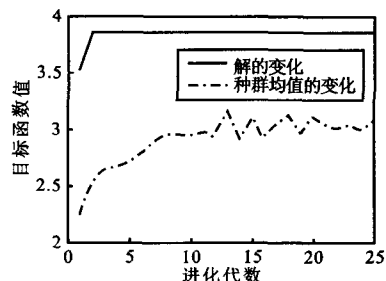


图 2 IGA 对函数 1 的结果

第 2 代即收敛到 3.8503. 另外, 将两种算法均执行 250 次, 基本遗传算法收敛了 142 次, 平均收敛值为 3.8497. 而改进的遗传算法收敛了 147 次, 平均收敛值为 3.8501.

第二个测试函数的仿真实验中, 种群个数为 40, 最大遗传代数为 250, 种群个体长度为 20, 交叉率为 0.7, 变异率为 0.001. 基本遗传算法某次执行后的结果如图 3 所示. 改进后的遗传算法某次执行后的结果如图 4 所示.

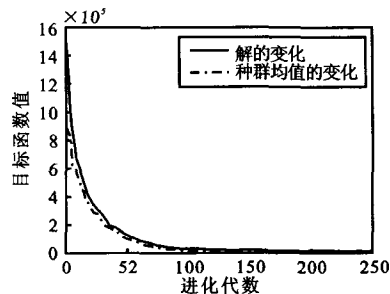


图 3 SGA 对函数 2 的结果

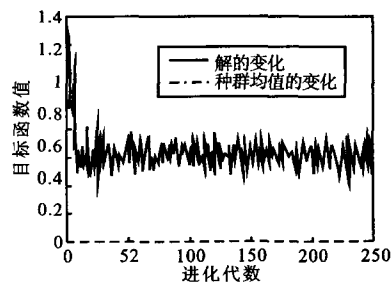


图 4 IGA 对函数 2 的结果

由图 3 和图 4 可以看出, 进化后期算法的解均在一定范围内震荡, 但是基本遗传算法得到的最优解是用 10^5 数量级来衡量的. 改进的遗传算法得到的解均小于 1, 而且改进的遗传算法的种群均值从初始化开始就接近于 0. 将两种算法分别执行 100 次, 基本遗传算法未收敛, 得到的最优解是 273.09, 而改进的遗传算法收敛了 100 次, 得到的最优解为 0.0040246.

(下转第 146 页)

算法协调机制去避免冲突,有效的避免多个节点在同一时刻发送信息丢失信息包的情况^[5].

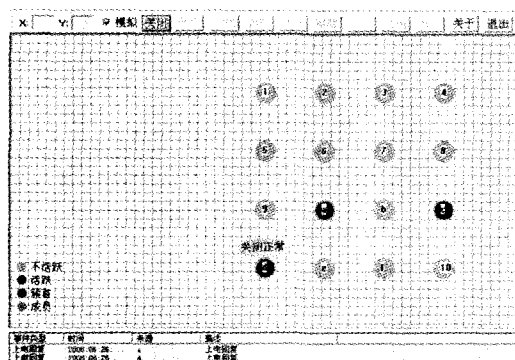


图5 节点关闭测试图

5 结束语

文中基于网域化微系统遥控方案的研究,解决了传统遥控设备中功能单一、控制距离有限、只能对单一设备控制等问题,通过软件编程可以任意设定指令功能,增加了控制的灵活性.同时,利用组合定位方式,提高了系统对网域内节点的定位时间和定

位精度,从而快速建立起对网域内的动态监测.实例表明,该系统对网域监测效果良好,同时还具有通用性,对于不同网域,可以通过修改相应的电路及相关的程序即可实现同样的功能.

参考文献:

- [1] 黄宜,黄如庆. 网络智能化视频监控系统的研究[J]. 探测与控制学报,2007,8(29):42-46.
- [2] 蒋挺,赵成林. 紫蜂技术及其应用[M]. 北京:北京邮电大学出版社,2006.
- [3] 熊远生,钱苏翔,吴伟雄. 网域控制系统的研究现状综述[J]. 工业仪表与自动化装置,2006(1):11-14.
- [4] 李文国,张凤伟. 超低功耗单片机实现智能遥控器设计[J]. 嵌入式系统,2005,2:95-97.
- [5] Otanez Paul G, Moyne James R, Tilbury Dawn M. Using deadbands to reduce communication in networked control systems [C]// American Control Conference, Alaska, USA,2002,(4):3015-3020.

作者简介:

杨卫男,(1957-),高工. 研究方向为动态测试技术.

(上接第142页)

由以上实验分析可知,改进的遗传算法在收敛速度和最优解精度上都比基本遗传算法有所提高,能够在短时间内使得最优解向全局最优解逼近.

4 结束语

文中改进了自适应遗传算法的应用策略以及自适应交叉和变异概率的计算方法,对实数编码的遗传算法进行了更进一步的综合改进,仿真结果证明该算法缩短了收敛时间,提高了收敛速度,比基本遗传算法的性能有很大提高.

参考文献:

- [1] 吴涛,毕笃彦. 一种基于改进遗传算法的自适应区域生长法研究[J]. 微电子学与计算机,2008,25(4):90-92.
- [2] 陈明杰,刘盛. 改进自适应遗传算法在函数优化中的应用[J]. 哈尔滨工程大学学报,2007,28(8):875-879.
- [3] 郑生荣,赖家美. 一种改进的实数编码混合遗传算法[J]. 计算机应用,2006,26(8):1959-1962.
- [4] 张海建,陈向东,王世海. 一种遗传算法优化的图像模糊边缘检测方法[J]. 微电子学与计算机,2008,25(4):105-107.

- [5] 李鹏,董聪. 基于实数编码的广义遗传算法及其在优化问题中的应用[J]. 控制与决策,2002,17(4):487-490.
- [6] Vicente Valls, Francisco Ballestin, Sacramento Quintanilla. A hybrid genetic algorithm for the resource-constrained project scheduling problem[J]. European Journal of Operational Research,2008,18(5):495-508.
- [7] Wu ChihHung, Tzeng GwoHshung, Goo YeongJia, et al. A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy[J]. Expert Systems with Applications,2007(32):397-408.
- [8] 高学金,王普,孙崇正. 基于实数编码的自适应遗传算法及应用[J]. 北京工业大学学报,2007,33(2):144-148.
- [9] 王小平,曹立明. 遗传算法—理论、应用与软件实现[M]. 西安:西安交通大学出版社,2002.

作者简介:

安建成男,(1963-),副教授. 研究方向为数据库、数据挖掘.

靳海娟女,(1982-),硕士. 研究方向为数据挖掘、人工智能.