

一种快速寻优的新型改进遗传算法^{*}

许义海¹, 李晓东²

(1. 广东女子职业技术学院现代教育技术中心, 广东 广州 511450
2. 中山大学电子与通信工程系, 广东 广州 510275)

摘 要: 在现有几种遗传算法 (GA) 的基础上, 对 GA 中的适应度函数、交叉策略和变异策略做了进一步的设计, 从而提出了一种新型改进 GA。新型改进 GA 以群体的多样性与算法的收敛速度、全局与局部搜索能力的综合均衡为设计重点, 较好地解决了一般 GA 收敛速度慢和局部搜索能力差的缺点。仿真结果表明: 该算法与常用的标准 GA 和采用算术交叉算子的实值编码改进 GA 相比, 有更快的收敛速度, 更高的收敛精度及全局收敛概率。

关键词: 遗传算法; 适应度函数; 交叉策略; 变异策略

中图分类号: TP301.6 **文献标识码:** A **文章编号:** 0529 6579 (2006) 02 0036 05

在对遗传算法 (Genetic Algorithm 简称 GA) 的实际研究与应用中, 人们发现 GA 存在许多不足和缺陷, 其中主要的有群体失去多样性而使算法产生早期收敛的问题、寻优时间长的问题以及局部搜索能力差的问题^[1-2]。多年来, 许多研究人员对 GA 做了大量的研究工作, 形成了多种改进的 GA^[2-4], 但现有改进的 GA 多从某一角度对算法进行了优化, 因此仍存在一定的局限性。本文在现有几种 GA 的基础上, 着重从适应度函数、交叉策略和变异策略的设计出发, 提出一种新型改进 GA。该算法以 GA 收敛速度慢和局部搜索能力差为主要改进目的, 将群体的多样性与算法的收敛速度、全局与局部搜索能力的均衡作为设计重点贯穿在新型改进 GA 的整个设计过程中。

1 快速寻优的新型改进遗传算法

1.1 适应度函数设计

合理的适应度函数, 有利于保持群体的多样性, 改善 GA 的整体性能^[5]。适应度函数的设计, 除了要满足不同个体能够利用其适应度进行优劣区分的基本功能外, 还应着重考虑以下两个问题: 一是在 GA 早期进化阶段, 要克服由于优秀个体的大量复制、群体失去多样性而引起的早期收敛问题, 就必须对优秀个体的适应度进行压缩, 对其他个体的适应度适当提升; 二是在 GA 进化的后期阶段, 多数个体都集中在最优解附近, 为突出进化的正确方向, 必须拉开最优个体与其它较优个体的适应

度, 以利于选择操作能做出敏感的选择。

基于上述适应度函数的设计思想, 参照文献 [5], 本文提出适应度函数设计方法:

$$f(y)=\begin{cases} 1-0.5\times(\frac{y-b}{a})^{\alpha} & \text{当 } y-b < a \text{ 时} \\ \frac{1}{1+(\frac{y-b}{a})^{\beta}} & \text{其它情况} \end{cases} \quad (1)$$

其中, y 为有待计算适应度的个体输出误差, b 为当前最优个体的输出误差, a 为当前群体所有个体的平均输出误差, $\alpha=1.5-t/T$, t 为当前进化代数, T 为最大进化代数。 β 取为 2 以贬低输出误差较大的个体适应度, 淘汰较差个体。适应度函数曲线如图 1 所示。

由图 1 (a) (b) (c) 可见, 群体适应度函数被限制在 [0 1] 的范围内, 且随着 GA 进化代数的增大, 适应度函数曲线能自动调节, 在进化早期能对优秀个体适应度进行压缩, 进化后期又能拉开最优个体与其他个体的适应度距离, 达到设计的目的。

1.2 交叉算子设计

在启发式交叉策略中^[6], 群体中所有个体都参与交叉, 即交叉概率 $p_c=1$, 并且利用两个交叉父代之间适应度的差异为交叉子代指示潜在的搜索方向, 大大地提高了 GA 的收敛速度, 其交叉策略

^{*} 收稿日期: 2005 08 26
基金项目: 国家自然科学基金资助项目 (60575006)
作者简介: 许义海 (1975 年生), 男, 助教; E-mail: xuyiha@21cn.com

(C)1994-2019 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

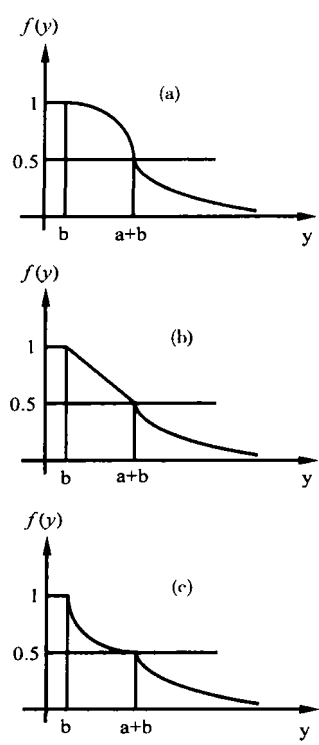


图 1 适应度函数图

Fig 1 Fitness function chart

(a) 进化早期 (b) 进化中期 (c) 进化后期

如下式所示：

$$\begin{cases} x_1^{n+1} = x_{kw}^n + a_1 (x_{high}^n - x_{kw}^n) d \\ x_2^{n+1} = x_{kw}^n + a_2 (x_{high}^n - x_{kw}^n) d \\ d = (f_{high} - f_{kw}) / f_{high} \end{cases} \quad (2)$$

其中 x_1^{n+1} 和 x_2^{n+1} 为交叉子代， x_{kw}^n 和 x_{high}^n 为适应度较低和较高的交叉父代， f_{kw} 和 f_{high} 分别为这两父代个体的适应度， d 为搜索方向， a_1 和 a_2 分别取 $[0 \ 1]$ 和 $[0 \ 2]$ 的随机数。由式 (2) 可见， x_1^{n+1} 在交叉父代空间范围内取值，属开发个体， x_2^{n+1} 在两倍于父代的空间内取值，在寻优中起开发和探测的作用，使算法具备较好的开发和探测能力。然而，该策略的缺点也是明显的：①交叉操作开发能力较强而探测能力不足；②当交叉父代两个体适应度接近时，交叉子代取值范围太窄；③两交叉父代的适应度相同时，交叉算子将失效。因此，本文对该交叉算子进行改进，提出如式 (3) 所示的交叉策略，其中 x_1^{n+1} ， x_2^{n+1} ， x_{kw}^n ， x_{high}^n ， f_{kw} ， f_{high} ， d 的参数意义与式 (2) 相同， x_3^{n+1} 和 x_4^{n+1} 为交叉操作产生的另外两个子代， a_1 ， a_2 ， a_3 ， a_4 分别取 $[0 \ 0.5]$ ， $[0.5 \ 1.5]$ ， $[1.5 \ 2.5]$ ， $[2.5 \ 3.5]$ 中的随机数。

$$\begin{cases} x_1^{n+1} = x_{kw}^n + a_1 (x_{high}^n - x_{kw}^n) d \\ x_2^{n+1} = x_{kw}^n + a_2 (x_{high}^n - x_{kw}^n) d \\ x_3^{n+1} = x_{kw}^n + a_3 (x_{high}^n - x_{kw}^n) d \\ x_4^{n+1} = x_{kw}^n + a_4 (x_{high}^n - x_{kw}^n) d \\ d = (f_{high} - f_{kw}) / f_{high} \end{cases} \quad f_{high} \neq f_{kw} \text{ 时}$$

$$a_1 \in [0 \ 0.5], \quad a_2 \in [0.5 \ 1.5], \quad a_3 \in [1.5 \ 2.5], \quad a_4 \in [2.5 \ 3.5]$$

新型改进 GA 采用近亲交叉回避策略，在两交叉父代适应度不同时才进行交叉，交叉产生四个新个体，从中选出较好的两个个体作为交叉子代。改进的启发式交叉算子能更好地兼顾交叉算子的开发和探测能力，相比式 (2) 交叉只产生两个个体来说，算法具有更大的局部搜索概率，由于 a_3 ， a_4 的取值较大，在父代适应度相近时，子代取值范围也相对较大，四个新个体具有连续的“爬峰能力”，对局部的单峰搜索有更快的收敛速度。当两交叉父代适应度相同时，算法不进行交叉而采用下述的一点非一致自适应变异，可以避免交叉失效。采用式 (3) 的交叉策略，产生的新个体有可能超出定义域，因此必须进行判断，当超出定义域时，可取相应的边界值。

1.3 一点非一致自适应变异策略

在实值非一致变异中^[7]，个体参数的变异量随进化代数的增大而逐渐变小，有利于在进化后期群体搜索到最优点附近时增加算法的局部搜索能力，它的主要缺点是变异操作是群体性的，即同一代内的个体都采用相同的变异范围。据文献 [8] 分析，当搜索点距全局最优点较远时，采用大的变异量更有好处，而搜索点在全局最优点附近时，大的变异量显然是有害的，即变异应针对不同的个体进行。由于全局最优点事先没法确知，而一般情况下，离全局最优点较远的个体，其输出误差也较大，相反，离全局最优点较近的个体，具有更小的输出误差，因此个体的输出误差可以作为变异取值范围的依据。基于这样的思想，本文提出一种新的变异策略：一点非一致自适应变异，即变异只在一个参数上进行，根据个体输出误差的大小自适应地选取变异取值范围，使较差个体有较好的探测能力，而较优个体能提高局部搜索能力，如下式所示：

$$C_k^{n+1} = \begin{cases} C_k^n + b_1 \cdot (UB - C_k^n) \cdot (1 - \frac{\text{error}}{2 \cdot \text{error}_k}) & z = 0 \\ C_k^n + b_2 \cdot (C_k^n - LB) \cdot (1 - \frac{\text{error}}{2 \cdot \text{error}_k}) & z = 1 \end{cases}$$

其中 n 为进化代数, C_k^n 为第 n 代中变异个体 x 的第 k 个参数, b_1, b_2 为 $[0, 1]$ 范围内的随机数, UB, LB 为该参数的定义域上界和下界, $error$ 为目标误差, $error_k$ 为第 n 代变异个体 x 的输出误差, z 为取值为 0 或 1 的随机数。

1.4 新型改进 GA 的算法流程及特点

1.4.1 算法流程 新型改进 GA 的算法流程图如图 2 所示, 其步骤如下:

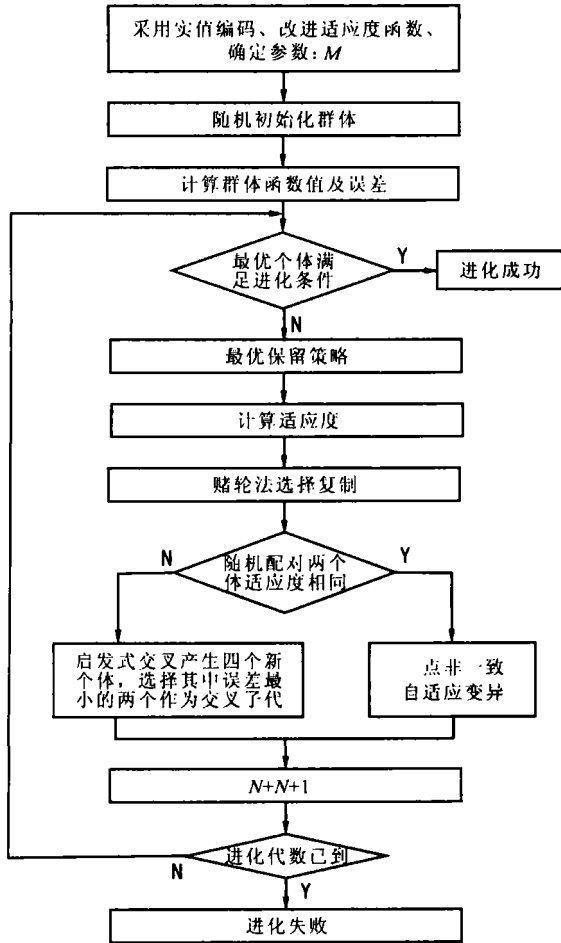


图 2 新型改进 GA 流程图

Fig. 2 New improved GA's flow chart

- (1) 确定群体规模 M 、最大进化代数 N 以及目标误差参数。
- (2) 采用实值编码策略, 在定义域内随机初始化群体。
- (3) 计算初始群体的个体函数值及误差。
- (4) 判断最优个体是否小于目标误差, 是的话则程序结束, 进化成功。
- (5) 采用最优保留策略保护最优个体。
- (6) 利用式 (1) 计算个体适应度。
- (7) 用赌轮法进行选择 and 复制。
- (8) 群体中的个体随机配对, 利用式 (3) 产生交叉子代, 利用式 (4) 产生变异子代, 并将子

代个体代替父代个体形成新群体。

(9) 进化代数加 1, 当代数达到 N 代时退出进化, 此时进化失败, 否则返回第 (4) 步进行新一轮的进化。

1.4.2 算法特点

(1) 采用改进的适应度函数, 能抑制超级个体的大量繁殖, 进化后期又能突出最优个体, 使 GA 能较好地保持群体的多样性, 避免发生早期收敛现象。

(2) 采用新型启发式交叉策略, 使 GA 有快速的收敛速度, 能兼顾全局寻优与局部搜索能力, 并且交叉子代具有与父代不同的新模式, 进一步保持了群体的多样性。

(3) 采用近亲回避策略, 交叉父代适应度相同时不交叉而进行一点非一致自适应变异。

(4) 交叉策略本身具备变异的效果, 可以省略基本的变异操作。

(5) 省略编码位数、交叉和变异概率, 算法具有更少的控制参数, 易于实现。

2 仿真实验

为验证本文提出的新型改进 GA 的有效性, 下面将其与另两种常用 GA 进行仿真比较, 通过仿真实验测试三种算法的全局收敛性, 即找到全局最优解的成功率及其速度。实验机器配置为: CPU 赛扬 IV 2.4G、内存 DDRAM 256M、操作系统 WinXP、工具软件 MATLAB 6.1。

2.1 测试函数的描述

本文选 Schaffer's F6 函数作为全局收敛性测试函数^[9], 其表达式为:

$$f(x, y) = 0.5 - \frac{\sin^2 \sqrt{x^2 + y^2} - 0.5}{(1 + 0.001(x^2 + y^2))^2} \quad (5)$$

Schaffer's F6 函数定义域为 $x, y \in [-100, 100]$, 不仅定义域大, 且寻优地形相当复杂, 除有大面积的平坦区域外, 还存在多处狭窄的峡谷和山峰, 整个定义域内有无数个局部极大点。函数的局部三维图如图 3 所示, 图 4 为全局最大点附近的剖面图, 可见全局最大点 $f(0, 0) = 1$ 周围有两圈脊, 取值分别为 0.990 284 和 0.962 776 与最大值 1 非常相近, 一般 GA 寻优极易陷入这两个极大值。因此, Schaffer's F6 函数是比较 GA 全局收敛性及局部搜索能力的理想函数。

2.2 参与仿真比较的 GA 算法描述

除本文提出的新型改进 GA 外, 参与比较的另外两种 GA 算法如下:

- (1) 标准 GA: 最优保留策略, 二进制编码,

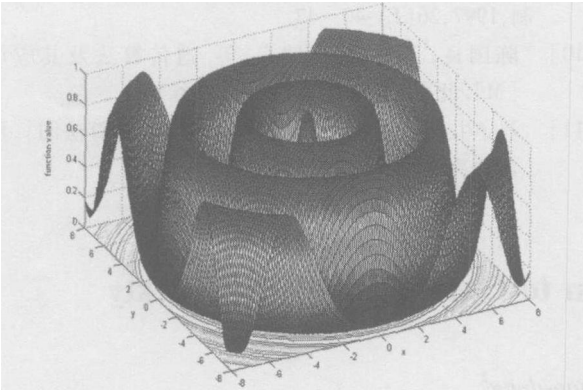


图 3 函数局部三维图

Fig 3 Local three - dimensional chart

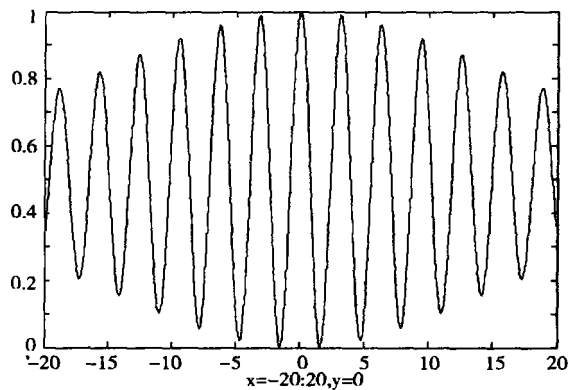


图 4 函数局部剖面图

Fig 4 Local section chart

下界法构建适应度函数，初始群体随机产生，赌轮选择复制，一点交叉，一点位取反变异^[10]。

(2) 采用算术交叉算子的实值编码改进 GA：最优保留策略，实值编码，下界法构建适应度函数，初始群体随机产生，赌轮选择复制，算术交叉，一点非一致变异^[3]。

2.3 仿真实验及结果

全局收敛标准规定：给定充分小的 ϵ ，在一定迭代次数内算法的收敛值 x 满足 $|x - x^*| \leq \epsilon$ ，则 x^* 为全局最优解^[11]。针对于 Schaffer' s F6 函数，全局最优解为 1，次优解为 0.990 284，因此，只要能寻找到误差 $\epsilon < 0.009$ 的个体就可认为找到全局最优解。用于比较全局收敛性的三种 GA 的实验参数设置为：群体规模为 300，目标误差为 0.009，最大进化代数为 50。三种 GA 运行 30 次后的结果如表 1 所示。

由表 1 可见，标准 GA 运行结果多次收敛于局部值 0.990 284，超出局部极大值找到全局最优值的成功率很低，说明标准 GA 局部搜索能力较差。采用算术交叉算子的实值改进 GA 寻优效果比标准

表 1 三种 GA 运行结果

Tab. 1 Result obtained by three kinds of GA

算 法	收敛概 率 %	平均进 化代数	最小进 化代数	最大进 化代数	典型 收敛值
标准 GA	6.6667	29	10	48	0.990 284
算术交叉 实值改进 GA	26.667	26.375	19	35	0.990 284
新型改进 GA	80	17.375	3	34	0.995

GA 方法好，性能有所改善，但全局收敛概率仍比较低。相比之下，本文提出的新型改进 GA 从全局收敛概率、平均进化代数及典型收敛值看，是成功的，说明新型改进 GA 既有很好的全局搜索能力，也有较强的局部搜索能力，并且收敛速度较快。

3 结 论

新型改进 GA 能较好地保持群体的多样性，防止发生早熟现象，较好地解决了 GA 收敛速度慢和局部搜索能力差的缺点，同时省略了二进制编码位数及交叉和变异概率，使算法更易于实现。仿真结果表明，新型改进 GA 和标准 GA 及采用算术交叉算子的实值改进 GA 相比^[3-10]，有更快的收敛速度，更高的收敛精度及全局收敛概率，是一种性能良好的改进 GA。

参考文献：

[1] LEUNG Yee, GAO Yong, XU Zongben. Degree of population diversity— a perspective on premature convergence in genetic algorithms and its markov chain analysis[J]. IEEE Transactions on Neural Networks, 1997, 8(5): 1165 - 1176.

[2] WEN Shaochu, LUO Fei, MO Hongqiang et al. The analysis of the local search efficiency of genetic neural networks and the improvement of algorithm[C]. Proceeding of the 4th World Congress on Intelligent Control and Automation, 2002, 1789 - 1793.

[3] YALCINOZ T, ALTUN H, UZAM M. Economic dispatch solution using a genetic algorithm based on arithmetic crossover[J]. IEEE Porto Power Tech Conference, 2001.

[4] CHIKAWA Y, ISHII Y. Retaining diversity of genetic algorithms for multivariable optimization and neural network learning[C]. IEEE Conference Proceedings, 1993, 1110 - 1114.

[5] 辛 甜, 朱 鳌 鑫. 遗传算法的适应度函数研究[J]. 系统工程与电子技术, 1998, 11: 58 - 62.

[6] 陈小平, 石 玉, 于 盛 林. 快速寻优的遗传交叉策略[J]. 控制理论与应用, 2002, 19(6): 981 - 984.

[7] DE CASTRO L N, IYODA E M, ZUBEN F J V, et al. Feed forward neural network initialization: an evolutionary

- approach[C]. IEEE Conference Proceedings, 1998
- [8] CHEN M usong LIAO Fonghang Neural networks training using genetic algorithms[C]. IEEE Conference Proceedings, 1998, 2436 – 2441.
- [9] 丁承民, 张传生, 刘辉. 遗传算法纵横谈[J]. 信息与控制, 1997, 26(1): 40 – 47.
- [10] 陈国良, 王煦法, 庄镇泉, 等. 遗传算法及其应用[M]. 北京: 人民邮电出版社, 1996
- [11] 刘守生, 于盛林, 丁勇, 等. 一种变焦遗传算法[J]. 控制与决策, 2002, 17(增刊): 731 – 734

A New Improved Genetic Algorithm to Obtain Solutions Quickly

XU Yihai¹, LIXiao-dong²

(1. Modern Educational Technical Center, Guangdong W omen's Polytechnic College, Guangzhou 511450, China)

(2. Department of Electronics and Communication Engineering, Sun Yat-Sen University, Guangzhou 510275, China)

Abstract Based on several existing Genetic Algorithm (GA), a improved GA is proposed through further facilitating the fitness function, crossover strategy and mutation strategy of GA. The new GA can effectively improve the poor performance of the existing one on convergence speed and local search ability because it emphasize the proportion of population diversity and convergence speed and also the proportion of globe search ability and local search ability. Simulation results show that the new GA has faster convergence speed, higher accuracy and higher globe convergence probability compared with the standard GA and the arithmetic crossover based real-coded GA.

Key words genetic algorithm; fitness function; crossover strategy; mutation strategy

· 简 讯 ·

本刊论文被 EI 收录数大幅增加

经多年努力, 中山大学学报 (自然科学版) 从 2005 年起被美国《工程索引》(EI Compendex) 列入收录的核心源期刊。据检索, 本刊近三年被 EI 收录的论文数不断增多, 2003 年被收录 95 篇, 2004 年为 106 篇, 2005 年增至 153 篇。

2005 年本刊被 EI 收录的论文中, 正刊 6 期被收录 118 篇, 下半年增刊被收录 35 篇, 收录论文数均超过 50%。

从被收录的学科分布看, 根据检索的结果, 2005 年第 5、6 期本刊物理学科 (含电子与通讯工程、力学、工学、信息科学) 发文 11 篇, 地学学科 (含地理、地质、大气、环境科学) 发文 13 篇, 全部被收录。其他学科, 如化学化工、数学与计算机应用等, 均有部分论文被收录。

2005 年下半年本刊增刊刊载“能源与环境材料国际研讨会专栏”论文 22 篇, 被 EI Compendex 收录 21 篇, 其他栏目论文收录 14 篇。

为提高被收录率, 根据 EI 数据库的要求, 论文摘要应包含研究的目的、方法、结果和结论, 而重点是结果和结论; 内容要求具体、充实, 以 200 – 300 字为宜, 切忌写成提要形式; 表达要简炼, 语义要确切, 结构要严谨, 表意要明白; 通常用第三人称表述, 英文与中文对应。请作者规范撰写论文摘要。

(本刊通讯员)