

文章编号: 1001-0920(2006)07-0792-04

一种改进的实数自适应遗传算法

潘伟^{1,2}, 刁华宗¹, 井元伟²

(1. 沈阳炮兵学院 电子侦察系, 沈阳 110162; 2 东北大学 信息科学与工程学院, 沈阳 110004)

摘要: 研究了基于实数编码的遗传算法的改进问题. 针对实数编码在搜索后期存在搜索效率低、易早熟收敛等现象, 讨论了遗传算法的参数调节问题. 提出一种自适应交叉概率和变异概率, 既考虑了进化代数对算法的影响, 又考虑到每代不同个体适应度的作用, 给出一种改进的实数自适应遗传算法. 最后利用 3 个测试函数对算法进行验证, 在函数的最终值、平均运行代数、收敛概率几方面都取得了较好的结果.

关键词: 自适应遗传算法; 实数编码; 交叉概率; 变异概率

中图分类号: TP18 **文献标识码:** A

An Improved Real-value Adaptive Genetic Algorithm

PAN Wei^{1,2}, DIAO Hua-zong¹, JIN G Yuan-wei²

(1. Department of Electric Detection, Shenyang Artillery College, Shenyang 110162, China; 2 College of Information Science and Engineering, Northeastern University, Shenyang 110004, China Correspondent: PAN Wei, Email: pan.w@126.com)

Abstract: The improvement problem of genetic algorithm is studied based on real coding. To the shortcomings that the search is inefficient and it is easy to premature convergence, the parameter adjusting problem of genetic algorithm is discussed. The adaptive crossover probability and adaptive mutation probability are proposed, considering the influence of every generation to algorithm and the effect of different individual fitness in every generation. Three testing functions are used to validate the algorithm. The results show that the ultimate value, the average algebraic sum and the convergence probability all obtain the preferable values.

Key words: Adaptive genetic algorithm; Real code; Crossover probability; Mutation probability

1 引言

遗传算法是模拟生物进化论的计算模型, 是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. 最初的遗传算法是基于二进制串的, 类似于生物染色体结构, 可用生物遗传理论来解释, 各种遗传操作易于实现, 算法处理的模式多. 但是二进制编码不能直接反映问题固有的结构特征, 个体长度大, 占用计算机内存多, 数值优化时精度不高, 且稳定性不如实数编码^[1]. 在具体问题中, 直接采用解空间的形式进行编码, 可在解的表现型上进行遗传操作, 从而引入特定领域的启发式信息, 取得比二进制编码更高的效率

这在优化理论和神经网络应用中具有重要意义^[2~6].

尽管实数编码精度高, 适合于复杂大空间搜索, 但易使遗传算法在搜索后期效率低下和未成熟收敛. 因此, 遗传算法的各参数特别是交叉概率和变异概率的选择显得更加重要^[7,8]. 交叉概率和变异概率通常根据理论分析中参数的大致范围来选择, 或根据经验来确定某个实际问题的参数组合, 具有一定的盲目性.

本文在群体规模确定的情况下, 设计出一种改进的实数自适应遗传算法. 算法中提出的自适应交叉概率和变异概率, 既考虑到进化代数对算法的影

收稿日期: 2005-10-19; 修回日期: 2006-04-10

基金项目: 国家自然科学基金项目(69974007); 教育部高等学校博士学科点专项科研项目(20020145007).

作者简介: 潘伟(1975—), 男, 哈尔滨人, 讲师, 博士, 从事非线性系统的鲁棒控制、遗传算法应用等研究; 井元伟(1956—), 男, 辽宁西丰人, 教授, 博士生导师, 从事复杂系统的对称性、相似性及稳定性等研究.

响,又考虑到每代中不同个体适应度对算法的作用,这两个重要参数可随算法进化的需要自动调整

2 实数自适应遗传操作

遗传算法中的编码策略和遗传操作对算法起到至关重要的作用,也是实数遗传算法与二进制遗传算法区别最大之处。下面对编码策略和遗传操作分别加以介绍

2.1 编码策略

实数编码与二进制编码的方法相似^[9],只是每个基因有 0~9 共 10 种可能取值。若每个变量用 L 位十进制数表示,变量个数为 m ,则染色体长度为 $m * L$ 。实值编码策略不对变量进行编码,而将每个变量当作一位基因直接处理

2.2 遗传操作

(1) 排序选择法 + 最佳个体保存法

选择是从群体中选择优胜个体,淘汰劣质个体的操作,目的是把优化的个体遗传给下一代

排序选择法是在计算每个个体的适应度后,根据适应度大小对群体中个体排序,并把事先设计好的概率表按序分配给个体,作为各自的选择概率。这样,个体的适应度不直接影响后代的选择,保持了一定的选择压力,从而在一定程度上抑制了标准遗传算法的早熟和局部收敛

最佳个体保存法是对群体中适应度最高的个体,不进行遗传操作而直接复制到下一代。该方法可以保证某一代的最优解不被破坏

(2) 非一致交叉

设进行交叉的两父代个体为 X_1 和 X_2 ,交叉后得到的子代个体为

$$\begin{cases} X_1 = rX_1 + (1-r)X_2, \\ X_2 = rX_2 + (1-r)X_1. \end{cases} \quad (1)$$

其中 $r \in [0, 1]$,可取定值也可变化

当 r 取固定值时为一致交叉,两对相同的交叉父代只能得到相同的两对新个体,不利于增加群体多样性,并且 r 的不同取值对交叉算子也有直接影响。当 r 取 $[0, 1]$ 中的随机值时为非一致交叉,它有利于将一致交叉中 r 取不同值产生的效果进行平均,以改善搜索特性

(3) 实值变异

变异的本质是挖掘群体中个体的多样性,同时提高算法的局部随机搜索能力

变异步长的选择比较困难,最优步长视具体情况而定。本文变异步长在优化过程中根据群体的进化进程自动改变,变异步长为

$$X = X \pm 0.5L\Delta \quad (2)$$

其中: X 和 X' 分别为变异前和变异后的个体; $\Delta = \sum_{i=0}^m \frac{a(i)}{2^i}$, $a(i)$ 以 $1/m$ 的概率取值 1,以 $1 - 1/m$ 的概率取值 0,本文设 $m = 20$; L 为各代群体中个体最大值与最小值之差

3 交叉概率和变异概率的自适应确定

首先说明以下各量的意义: P_c 为交叉概率, P_m 为变异概率, f 为当前个体适应度值, f_{\max} 为当前群体中最大个体适应度值, \bar{f} 为当前群体平均适应度值

3.1 交叉概率的自适应确定

交叉算子主要用来产生新个体,实现算法的全局搜索,它在遗传操作中起到关键作用。从群体整体进化过程看,交叉概率应随进化过程逐渐变小,最后趋于某一稳定值,以免算法不能收敛或收敛过程加长。从产生新个体的角度看,群体中的所有个体在交叉操作上应有同等地位和相同概率,从而使 GA 在搜索空间具有各个方向的均匀性。为此,本文设计与进化代数相关的交叉概率

$$P_c = \frac{1}{1 + e^{\alpha G}} + \beta \quad (3)$$

其中: G 为进化代数, α 代表交叉概率的变化速率, β 代表交叉概率的收敛极限。交叉概率的变化曲线如图 1 所示

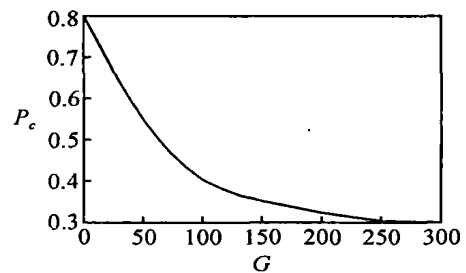


图 1 交叉概率曲线

3.2 变异概率的自适应确定

变异算子主要用来维持群体多样性,防止出现未成熟收敛。在算法早期,群体中个体多样性丰富,此时变异概率应小些,以提高算法的运行速度。随着进化的进行,个体越来越向适应度高的个体靠近,致使个体越来越单一,此时变异概率应大些,以维持群体的多样性。同理,同一代群体中个体的变异概率应随个体的优劣而变化,即加大优质个体的变异概率。为此,设计与遗传进化代数和个体适应度相关的自适应变异概率

$$P_m = \begin{cases} \frac{k_1}{1 + e^{-\alpha G}} \frac{f - \bar{f}}{f_{\max} - \bar{f}}, & f > \bar{f}; \\ k_2, & f \leq \bar{f}. \end{cases} \quad (4)$$

其中: G 为进化代数; α 代表变异概率的变化速度; k_1

是为保证遗传算法不退化为随机搜索 P_m 所取的最大值; k_2 为较小的变异概率, 一般取 0.001.

关于 k_1 的取值, 有如下结论:

定理 1 若父代群体的全部 N 个个体均分布在超平面 π_{t_i} 上, 变异概率为 P_m , 则群体经变异操作后, 子代个体逃离超平面 π_{t_i} 的概率为 $1 - (1 - P_m)^{N \cdot O(H)}$. 当此概率接近于 1 时, 意味着子代的生成与父代没有关系, 算法退化为随机搜索方法. 此时的变异概率 P_m 即为定常系数 k_1 的值, 其表达式为

$$k_1 = 1 - e^{\ln \epsilon^{N \cdot O(H)}} \quad (5)$$

其中: ϵ 为极小的正数, N 为群体规模, $O(H)$ 为模式阶.

证明 设

$$\epsilon = (1 - P_m)^{N \cdot O(H)},$$

经过指数变换, 有

$$P_m = 1 - e^{\ln \epsilon^{N \cdot O(H)}},$$

即有式 (5) 成立.

4 改进算法的性能测试

本节将改进的自适应遗传算法应用于一类测试函数, 并与一般遗传算法进行比较. 选取 3 个具有相当复杂度的测试函数^[10], 这些测试函数都包含多个极值. 所选函数如下:

$$f_1(x, y) = 100(x^2 - y)^2 + (1 - x^2),$$

- 2.048 < x, y < 2.408;

$$f_2(x, y) = 0.5 - \frac{\sin^2 \sqrt{x^2 + y^2} - 0.5}{(1 + 0.001(x^2 + y^2))^2},$$

- 100 < x, y < 100;

$$f_3(x, y) = - (x^2 + 2y^2 - 0.4 \cos(3\pi x) - 0.6 \cos(4\pi y)),$$

- 10 < x, y < 10

函数 f_1 是一个二参数单极值的非二次函数, 属于单峰值函数, 但该函数却是病态的, 在 $y = x^2$ 处有一条狭长深谷, 极易陷入局部解, 难以进行全局最小化. 该函数的全局最小点在 (1, 1), 最小值为 0. 函数图形如图 2 所示.

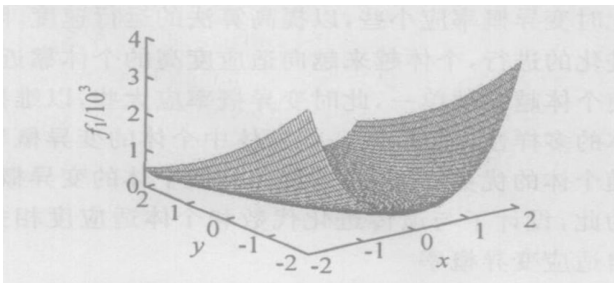


图 2 测试函数 f_1 图形

函数 f_2 有无数个局部极大点, 但只有一个 (0, 0) 点为全局最大点, 最大值为 1. 此函数的最大峰周

围有两圈脊, 它们的取值分别为 0.990 284 和 0.962 776, 因此优化过程中很容易停滞于这些局部极大点. 函数图形如图 3 所示.

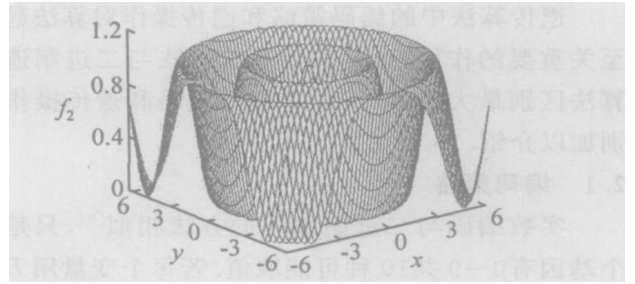


图 3 测试函数 f_2 图形

函数 f_3 是一个多峰函数, 每个峰顶都是一个局部最大点, 但只有一个全局最大点 (0, 0), 此时函数最大值为 1. 函数图形如图 4 所示.

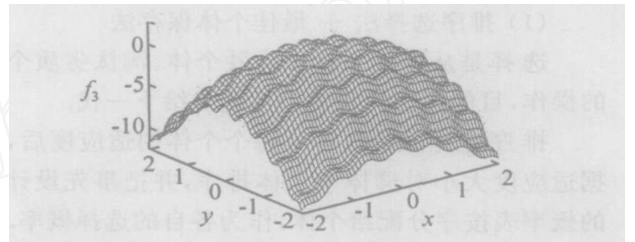


图 4 测试函数 f_3 图形

对于函数 f_1, f_2 和 f_3 , 均采用 3 种实验方法进行测试. 为叙述简便, 将简单遗传算法记为 T_1 , 二进制改进自适应遗传算法记为 T_2 , 实数改进自适应遗传算法记为 T_3 .

T_1 编码方式为二进制编码, 染色体长度为 10, 采用基本遗传操作 (适应度比例选择, 单点交叉, 单点变异), 交叉概率 $P_c = 0.6$, 变异概率 $P_m = 0.01$.

T_2 编码方式为二进制编码, 也采用基本遗传操作, 交叉概率和变异概率分别为改进的自适应交叉概率 P_c 和自适应变异概率 P_m .

T_3 编码方式为实数编码, 遗传操作采用排序选择 + 最佳个体保留的选择算子, 非一致的交叉算子, 变步长的变异算子, 交叉概率和变异概率分别为改进的自适应交叉概率 P_c 和自适应变异概率 P_m .

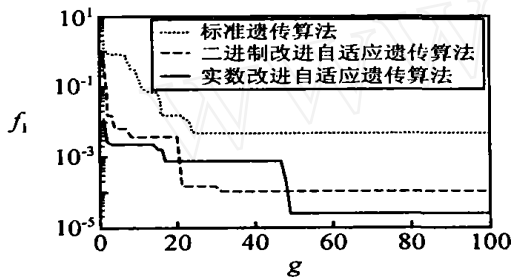
3 种实验方法中群体规模均为 100, 总进化代数设为 100. 分别用两种方法实验 50 次, 所得结果如表 1 所示. 其中 x, y 和 $f(x, y)$ 为算法结束时的典型值, g 表示平均运行代数, p 为收敛到全局最优的概率.

函数 f_1 较为简单, 是一种单极值函数, 所以 T_1, T_2 和 T_3 方法的效果也基本相同. 但仍能看出, 后两种方法在函数的最终值和收敛概率上占有优势. 至于平均运行代数的增加, 与变异概率随代数增加而增大有关.

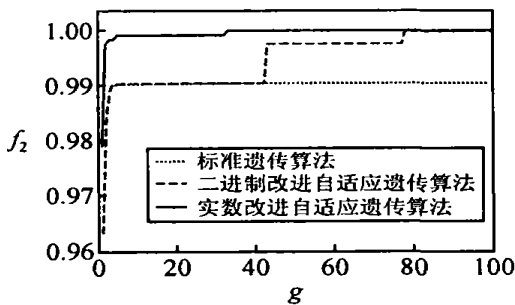
表 1 测试函数 3 种算法结果比较

f	T	x	y	$f(x, y)$	g	$p/\%$
f_1	T_1	0.993 157	0.993 157	0.004 666	25	20
	T_2	0.998 976	0.998 976	0.000 106	31	91
	T_3	0.999 581	0.999 581	0.000 018	48	100
f_2	T_1	1.906 158	- 2.492 669	0.990 284	42	10
	T_2	- 0.009 775	0.009 775	0.999 809	78	44
	T_3	0.000 982	0.000 982	0.999 998	37	93
f_3	T_1	0.009 775	0.009 775	0.993 496	84	56
	T_2	0.000 196	0.000 196	0.999 997	42	82
	T_3	0.000 098	0.000 098	0.999 999	31	98

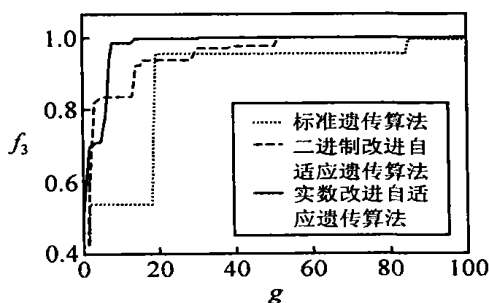
函数 f_2 存在无穷多个局部最小值, 采用 T_1 方法运行多次, 收敛于未成熟解的概率很大 T_2 方法有了明显的改善, 但由于 x 和 y 的取值区间较大, 平均运行代数较多 T_3 方法函数的最终值、平均运行代数和收敛概率都是成功的, 特别是平均运行代数有了大幅降低



(a) 函数 f_1



(b) 函数 f_2



(c) 函数 f_3

图 5 3 种方法的性能比较

函数 f_3 局部极大点较少, 采用 T_1 方法收敛概率尚可, 但平均运行代数明显大于 T_2 和 T_3 方法 本文提出的方法适于求解函数 f_3 , T_2 方法得到了良好的结果, T_3 方法使得 x, y 和 $f(x, y)$ 最终值更加精确

由表 1 可以看出, 本文改进的自适应交叉和变异概率, 在寻优精度、运行代数和收敛概率几方面, 都优于标准遗传算法, 而实数编码比二进制编码的优势更加明显 图 5 给出了 3 个测试函数在 3 种不同实验方法 T_1, T_2 和 T_3 下的性能比较 其中横坐标为算法进化代数, 纵坐标为最优个体对应的适应度值, (a) 的纵坐标为对数坐标

从图 5 可以清楚地看出, 本文改进的自适应遗传算法的性能明显优于标准遗传算法的性能, T_2 和 T_3 方法收敛于全局最优的速度明显快于 T_1 方法 这是由于改进方法具有指导搜索方向的能力, 后代可向适应度改善区域不断进化 换言之, 将进化代数和个体适应度差异引入交叉概率和变异概率, 使后代沿着成功的搜索方向进化, 而此方向是从父代中识别出的 通过比较表明, 标准遗传算法不能集中到潜在的搜索方向, 使得最好个体以及整个群体适应度改善的速度受到限制

5 结 论

本文提出一种新型实数自适应遗传算法, 不但把实数引入自适应遗传算法, 而且对遗传算法中的重要参数交叉概率和变异概率作了新的设计 设计中即考虑到进化代数对算法的影响, 又考虑到每代中不同个体适应度对算法的作用 最后利用 3 个测试函数对算法进行验证 从测试函数看出, 对于单峰值函数, 算法的作用并不明显; 对于复杂的多峰值函数, 算法的作用相当明显, 在函数的最终值、平均运行代数和收敛概率几方面都取得了较好的结果

参考文献 (References)

- [1] Bessau M, Slarry P. A Genetic Algorithm with Real-value Coding to Optimize Multimodal Continuous Functions[J]. *Structure Multitask Optimization*, 2001, 23(1): 63-74
- [2] Blanco A, Delgado M, Pegalajar M C. A Real Coded Genetic Algorithm for Training Recurrent Neural Networks[J]. *Neural Networks*, 2001, 14(1): 93-105
- [3] Baskar S, Subberaj P, Rao M V C. Hybrid Real Coded Genetic Algorithm Solution to Economic Dispatch Problem [J]. *Computers and Electrical Engineering*, 2003, 29(3): 407-419
- [4] Tsutsui S, Goldberg D E. Search Space Boundary Extension Method in Real Coded Genetic Algorithm s [J]. *Information Sciences*, 2001, 133(3): 229-247

(下转第 800 页)

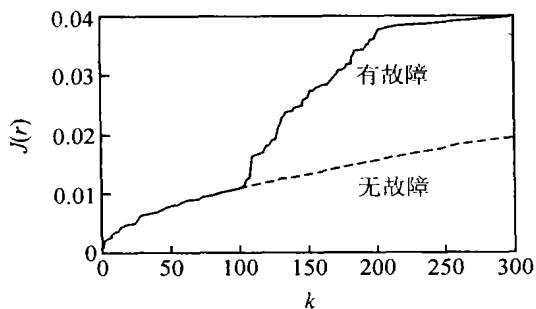


图1 残差评价函数

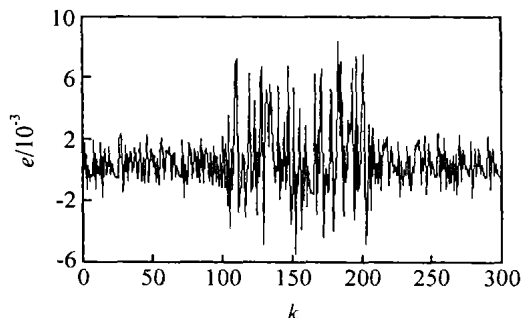


图2 残差信号

5 结论

本文研究基于滤波器的时滞离散马尔可夫跳跃系统的鲁棒故障检测问题,给出了相应滤波器的设计方法。该滤波器设计方法采用较为成熟的线性矩阵不等式方法进行求解,所得的故障检测滤波器在系统存在未知输入和控制输入时,能准确检测到故障信号。

参考文献(References)

- [1] Frank P M, Ding S X. Survey of Robust Residual Generation and Evaluation Methods in Observer-based Fault Detection Systems [J]. *J of Process Control*, 1997, 7(6): 403-424
- [2] Ding S X, Jeansch T, Frank P M. A Unified Approach to the Optimization of Fault Detection Systems [J]. *Int J of Adaptive Control Signal Processing*, 2000, 14(7): 725-745
- [3] Niemann H, Saberi A, Stoorvogel A, et al. Exact, Almost and Delayed Fault Detection: An Observer-based Approach [J]. *Int J of Robust and Nonlinear Control*, 1999, 9(4): 215-238
- [4] Patton R J, Chen J. On Eigenstructure Assignment for Robust Fault Diagnosis [J]. *Int J of Robust and Nonlinear Control*, 2000, 10(14): 1193-1208
- [5] Nobrega E G, Abdalla M O, Grigoriadis K M. LM I-based Filter Design for Fault Detection and Isolation [A]. *Proc of the 39th Conf on Decision and Control [C]*. Sydney, 2000, 4329-4334
- [6] Zhong M Y, Ding S X, Lam J, et al. An LM I Approach to Design Robust Fault Detection Filter for Uncertain LTI Systems [J]. *Automatica*, 2003, 39(3): 543-550
- [7] Cao Y Y, Lam J. Stochastic Stabilizability and H Control for Discrete-time Jump Linear Systems with Time Delay [J]. *J of Franklin Institute*, 1999, 336(8): 1263-1281
- [8] Gao H J, Lam J, Xu S Y, et al. Stabilization and H Control of Two-dimensional Markovian Jump Systems [J]. *IMA J of Mathematical Control and Information*, 2004, 21(4): 377-392
- [9] Xu S Y, Chen T W, Lam J. Robust H Filtering for Uncertain Markovian Jump Systems with Mode-dependent Time Delays [J]. *IEEE Trans on Automatic Control*, 2003, 48(5): 900-907
- [10] Oliveira M C, Gerome J C, Bernussou J. Extend H_2 and H_∞ Norm Characterization and Controller Parameterizations for Discrete-time Systems [J]. *Int J of Control*, 2002, 75(9): 666-679
- [5] Goldberg D E. Real Coded Genetic Algorithms, Virtual Alphabets and Blocking [J]. *Complex Systems*, 1991, 5(2): 139-167
- [6] Srinivas M, Patnaik L M. Adaptive Probabilities of Crossover and Mutation in Genetic Algorithms [J]. *IEEE Trans on Systems, Man and Cybernetics*, 1994, 24(4): 656-667
- [7] 陈长征, 王楠. 遗传算法中交叉和变异概率选择的自适应方法及作用机理 [J]. *控制理论与应用*, 2002, 19(1): 41-43
(Chen C Z, Wang N. Adaptive Selection of Crossover and Mutation Probability of Genetic Algorithm and Its Mechanism [J]. *Control Theory and Applications*, 2002, 19(1): 41-43)
- [8] 张晓绩, 方浩, 戴冠中. 遗传算法的编码机制研究 [J]. *信息与控制*, 1997, 26(2): 134-139
(Zhang X J, Fang H, Dai G Z. Studies on Encoding Mechanism of Genetic Algorithm [J]. *Information and Control*, 1997, 26(2): 134-139)
- [9] 侯格贤, 吴成柯. 遗传算法的性能分析 [J]. *控制与决策*, 1999, 14(3): 257-260
(Hou G X, Wu C K. Performance Analysis for Genetic Algorithm [J]. *Control and Decision*, 1999, 14(3): 257-260)
- [10] Whitley D, Beveridge R, Graves C, et al. Test Driving Three Genetic Algorithms: New Test Functions and Geometric Matching [J]. *J of Heuristics*, 1995, 1(1): 77-104

(上接第 795 页)