

文章编号:1001-9081(2005)08-1903-03

一种可实现精确优化的小生境遗传算法

张兆中¹,张问银²

(1. 山东临沂师范学院 数学系, 山东 临沂 276005;

2. 中国科学院 成都计算机应用研究所, 四川 成都 610041)

(touchang@163.com)

摘要:提出了一种基于排挤机制的小生境遗传算法,其特点是能够在进化过程中动态地记录个体的进化方向。仿真实验表明,该算法不但可以有效地保持种群的多样性,找到较多的解,还可提高解的精度。

关键词:遗传算法;小生境;进化方向

中图分类号: TP18 **文献标识码:** A

A novel niche genetic algorithm of accurate optimization performance

ZHANG Zhao-zhong¹, ZHANG Wen-yin²

(1. Department of Mathematics, Linyi Normal University, Linyi Shandong 276005, China;

2. Chengdu Institute of Computer Application, Chinese Academy of Sciences, Chengdu Sichuan 610041, China)

Abstract: Based on crowding mechanism, a novel niche genetic algorithm was proposed which can record evolutionary direction dynamically during evolution. Simulation shows that the algorithm can not only keep population diversity but also find accurate solutions.

Key words: genetic algorithm; niche; evolutionary direction

遗传算法(Genetic Algorithm, GA)是模拟生物在自然环境中的遗传和进化过程的一种全局优化概率搜索算法。通过对当前群体施加选择、交叉、变异等一系列操作,产生新一代种群,逐步使种群进化到包含或接近最优解的状态^[1]。它具有良好的通用性、鲁棒性。但也有明显缺点:1)由于其强大的全局搜索能力,当目标函数有多个解时,往往只能搜索到一个局部最优解;2)由于搜索的随机性,使结果精度不高。

本文提出了一种可标记进化方向的小生境遗传算法DRN-GA(Direction Recorded Niche Genetic Algorithm)。该算法的特点是:1)基于“排挤机制”^[2],在进化过程中,能不断排挤掉相似的个体,有效地保持种群的多样性,使种群的个体分布于不同解的领域,从而搜索到较多的解;2)利用进化过程中有用信息,为每个个体标记进化方向。所以,在DRN-GA算法执行后,如果以每个个体为初始点,按标记的进化方向继续局部寻优,会进一步提高解的精度。

1 DRN-GA 算法

1.1 进化方向

定义 设目标函数为 $J = \max f(X) = \max_{x_i \in A_i} f(x_1, x_2, \dots,$

$$\underbrace{d_1 d_2 \dots d_n}_D \underbrace{b_{11} b_{12} \dots b_{1m_1}}_{x_1} \underbrace{b_{21} b_{22} \dots b_{2m_2}}_{x_2} \dots \underbrace{b_{i1} b_{i2} \dots b_{im_i}}_{x_i} \dots \underbrace{b_{n1} b_{n2} \dots b_{nm_n}}_{x_n}$$

图1 个体编码结构

其中, $b_{i1} b_{i2} \dots b_{im_i}$ 为 x_i 的二进制编码, d_i 记录 x_i 的进化方向。遗传进化过程中, D 区段不参与交叉和变异操作。

1.3 小生境实现原理

小生境遗传算法常用于求解多峰函数优化问题,利用它

$x_n)$, $X_0 = (x_{01}, x_{02}, \dots, x_{0n})$, $x_{0i} \in A_i$ 。称 $D = (d_1, d_2, \dots, d_n)$ 为 J 在 X_0 点的进化方向,若 $d_i (i = 1, 2, \dots, n)$ 满足:

(1) $d_i \in \{-1, 0, 1\}$;

(2) $d_i = 1$, 若存在小正数 ε , 使 $f(x_{01}, x_{02}, \dots, x_{0i} + \varepsilon, \dots, x_{0n}) > f(x_{01}, x_{02}, \dots, x_{0i}, \dots, x_{0n})$;

(3) $d_i = -1$, 若存在小正数 δ , 使 $f(x_{01}, x_{02}, \dots, x_{0i} - \delta, \dots, x_{0n}) > f(x_{01}, x_{02}, \dots, x_{0i}, \dots, x_{0n})$;

(4) $d_i = 0$, 其他。

若 D 为目标函数 J 在 X_0 点的进化方向, $step$ 为一小正数, 当 d_i 均不为 0 时, 显然有:

$$f(x_{01} + step \cdot d_1, \dots, x_{0n} + step \cdot d_n) > f(x_{01}, x_{02}, \dots, x_{0n})$$

即: $f(X_0 + step \cdot D) > f(X_0)$ (1)

根据进化过程中的信息对每个个体 (X_0) 标记进化方向, 进化后再从每个个体出发, 按标记方向进行局部搜索, 可以提高解的精度。由此必须设计合理的个体编码结构。

1.2 个体编码结构

如上所述, 个体编码中除应包含决策变量的编码, 还要有记忆进化方向的部分。为适应本算法, 设计个体编码方案如图 1 所示。

能求出较多(全局或局部)最优解。其实现思想是确保种群在遗传进化中保持多样性。已有多种小生境实现思想^[2-5]。本文给出的实现方法描述如下:

记第 i 个个体为 P_i , 第 j 个个体为 P_j , $\|P_i - P_j\|$ 为 P_i, P_j

收稿日期:2005-02-28

作者简介:张兆中(1972-),男,山东临沂人,硕士研究生,主要研究方向:数值方法及软件、软件工程;张问银(1972-),男,山东临沂人,博士研究生,主要研究方向:自动推理、图像检索、信息隐藏。

间的欧氏距离, L 为一较小正数, $f(\cdot)$ 为适应度函数。小生境的实现是在普通遗传算法的进化过程中的每一代的选择、交叉、变异操作之前, 加入如下操作: 比较种群中任意两个个体的相似性, 若 $\|P_i - P_j\| < L$, 表明 P_i 与 P_j 相似程度较大, 则排挤掉其中适应度较小者, 方法是对适应度较小者施加一较强的惩罚函数, 使之适应度变为极小。例如, 如果 $f(P_i) > f(P_j)$, 则令: $f(P_j) = \text{penalty} \cdot f(P_j)$ (penalty 是一很小正数)。在以后的进化中 P_j 会以极大的概率被淘汰掉。根据 De Jong 的排挤原理, 这样可以实现小生境的进化环境。

1.4 进化方向的标记

个体进化方向的标记可以在上面的小生境实现中同步完成。设单变量函数 $y = g(x)$, $x \in A$ 的图像如图 2, $x_1, x_2 \in A$ 且 $|x_1 - x_2| < \rho, \rho$ 为一较小正数。

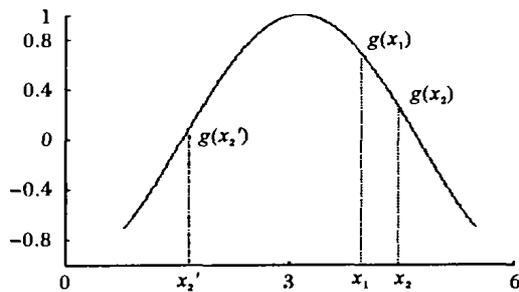


图 2 进化方向原理

设目标函数为 $J = \max_{x \in A} g(x)$, 由图可知, x_1 比 x_2 更优。

可以根据 x_1 与 $x_2, g(x_1)$ 与 $g(x_2)$ 的大小关系, 定出 x_1 (更优者) 的进化方向: 由 $x_1 < x_2, g(x_1) > g(x_2)$ 及 $|x_1 - x_2| < \rho$, 所以在 x_1 的一个邻域内 $g(x)$ 是下降的, 可推出, 存在一个很小正数 ε , 使得 $g(x_1 - \varepsilon) > g(x_1)$, 即 $(x_1 - \varepsilon)$ 是比 x_1 更优的点, 所以 x_1 的进化方向为 -1 。

上面由 x_1 与 $x_2, g(x_1)$ 与 $g(x_2)$ 的大小关系确定较优者进化方向的前提是 x_1 与 x_2 距离小 (即 $|x_1 - x_2| < \rho$)。若距离较大, 如图 2 中的 x_1 与 x_2' , 无法由 x_1 与 $x_2', g(x_1)$ 与 $g(x_2')$ 大小关系推出 x_1 或 x_2' 的进化方向。类似地, 有下面的结论:

结论: 若记 x_{ik} 为个体 P_i 的第 k 个决策变量, x_{jk} 为个体 P_j 的第 k 个决策变量。当 $\|P_i - P_j\| < L$ 时, 可以通过比较 $f(P_i)$ 与 $f(P_j), x_{ik}$ 与 x_{jk} 的大小关系确定出 P_i 与 P_j 中更优者的进化方向分量 $d_k (k = 1, 2, \dots, n)$ 。

由上面例子的分析, 可以得出:

当 $f(P_i) > f(P_j)$ 时 (此时 P_i 为更优者), P_i 的进化方向为:

$$d_k = \begin{cases} -1 & \text{如果 } x_{ik} < x_{jk} \\ 0 & \text{如果 } x_{ik} = x_{jk} \\ 1 & \text{如果 } x_{ik} > x_{jk} \end{cases} \quad k = 1, 2, \dots, n$$

当 $f(P_i) < f(P_j)$ 时 (此时 P_j 为更优者), P_j 的进化方向为:

$$d_k = \begin{cases} -1 & \text{如果 } x_{ik} > x_{jk} \\ 0 & \text{如果 } x_{ik} = x_{jk} \\ 1 & \text{如果 } x_{ik} < x_{jk} \end{cases} \quad k = 1, 2, \dots, n$$

将上述标记方法融入 1.3 节介绍的小生境实现中, 可得能标记进化方向的小生境遗传算法 DRN-GA。

1.5 DRN-GA 实现步骤

(1) 确定种群规模, 采用 1.2 节的个体编码方案, 选择、

交叉和变异的方法分别采用按比例选择、单点交叉和均匀变异, 算法终止条件是进化代数等于进化终止代数。

(2) 取 $t = 0$, 以一定的方式产生初始种群 $\vec{P}(0)$ 。

(3) 计算种群 $\vec{P}(t)$ 中个体 $P_i(t)$ 的适应度 $f(P_i(t))$, $i = 1, 2, \dots, N$, 计算其中任两个个体 $P_i(t), P_j(t)$ 的欧氏距离 $\|P_i(t) - P_j(t)\|$, 若 $\|P_i(t) - P_j(t)\| < L$, 用 1.3 节中的方法处理其中适应度较小者, 并用 1.4 节中的方法为适应度较大者标记进化方向。根据个体的适应度按比例选择 N 个个体。

(4) 对选中的 N 个个体进行交叉操作和变异操作。

(5) 若 $t = T$, 则进化终止。否则, 令 $t = t + 1$ 转到步骤 (3)。

2 局部寻优算法和精确优化

执行完 DRN-GA 算法后得到种群 $\vec{P}(T)$ 。对 $\vec{P}(T)$ 中的任意个体 P_i , 取出所保存的进化方向 $D = (d_1, d_2, \dots, d_n)$ 及决策变量值 $X_0 = (x_1, x_2, \dots, x_n)$ 。局部寻优即从 X_0 开始, 沿方向 D 以步长 step 寻找比 X_0 更优的解。在这里要用到式 (1)。流程如图 3 所示。

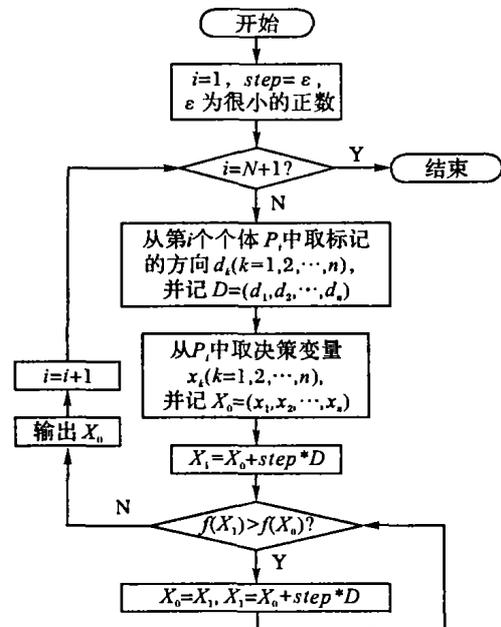


图 3 局部寻优流程

执行完 DRN-GA 算法后, 种群的每个个体中都保存了进化方向。用局部寻优算法, 对每个个体沿进化方向继续搜索, 可进一步提高解的精度。两算法串行执行, 如图 4 所示。

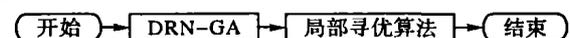


图 4 精确优化结构

3 仿真

为验证算法的搜索能力和优化精度方面的改进, 考虑如下两个函数:

(1) Camel 函数:

$$f(x, y) = (4 - 2.1x^2 + \frac{x^4}{3})x^2 + xy + (-4 + 4y^2)y^2$$

$$-2 < x < 2, -1.3 < y < 1.3$$

此函数有 6 个局部极小点 (其中有两个病态极小点, GA 搜索不到), 其中有两个全局最小点。在这两点的函数值为

-1.0316。函数变量的精度要求为 0.01,则每个变量所占二进制编码长度为 9,个体编码为 20 位(加两个进化方向标记位)。取目标函数为 $J = \max[20 - f(x,y)]$,种群规模为 80。本例要验证算法的两个方面:1)验证提出的 DRN-GA 的寻找到较多的解的能力,并和标准遗传算法 SGA(Standard GA)比较。为便于比较,两算法均采用单点交叉,交叉概率 $p_c = 0.6$,均匀变异,单点变异概率为 $p_m = 0.002$ 。具有相同的初始种群。终止代数 100。由于算法的随机性,所以对两个算法均反复运行 30 次。两算法搜索到的解的个数对比见表 1,种群的变化如图 5 中(a)~(c);2)验证优化精度的提高。即给出 DRN-GA 执行完后的解及继续局部寻优后的解,给出后者相对于前者精度的变化情况。局部优化时, $step = 0.0001$ 。运行的实验结果数据如表 2 所示,种群的变化情况如图 5 中(c)~(d)。

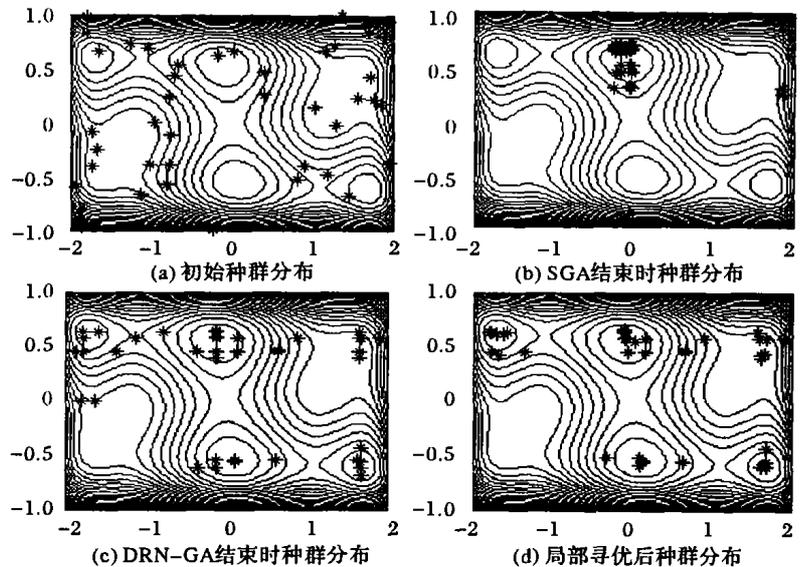


图 5 Camel 函数下三种算法的种群分布变化

表 1 DRN-GA 与 SGA 寻解能力对照

算法	收敛次数	不收敛次数	找一个解次数	找二个解次数	找三个解次数	找四个解次数	找五个解次数	找六个解次数	平均执行时间
SGA	22	8	19	3	0	0	0	0	7.8s
DRN-GA	30	0	0	2	4	24	0	0	26.4s

表 2 局部寻优后各解的精度变化(数据为运行 30 次的最优值)

算法	第一个解			第二个解			第三个解			第四个解		
	x	y	f(x,y)	x	y	f(x,y)	x	y	f(x,y)	x	y	f(x,y)
DRN-GA	-0.086	0.698	-1.0299	0.084	-0.701	-1.0305	-1.694	0.803	-0.2140	1.706	-0.784	-0.2138
局部寻优	-0.0898	0.7127	-1.0316	0.0896	-0.7108	-1.0316	-1.7020	0.7960	-0.2154	1.7020	-0.7969	-0.2154

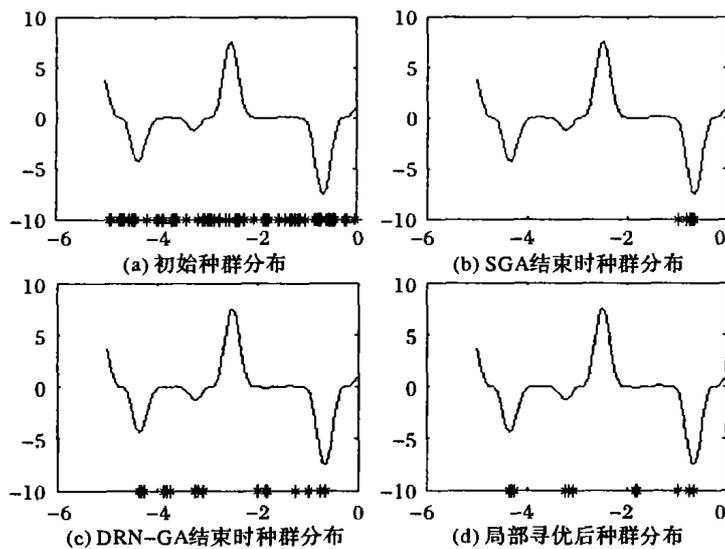


图 6 三种算法下种群分布变化

结果分析:1)由图 5 (a)~(c)可看出,DRN-GA 执行后的个体分布于四个不同的解周围,而 SGA 执行后的个体集中于一个解领域;从表 1 可知,在 SGA 的 22 次收敛中,有 19 次收敛于一个解,收敛于两个以上解的情况只有 3 次。而 DRN-GA 在 30 次收敛中,有 24 次收敛于 4 个解。实验结果表明,提出的 DRN-GA 算法有良好的寻解能力,且 DRN-GA 比 SGA 有较大的时间消耗;2)由图 5 中(c)~(d)可以直观地看出,进一步局部寻优后,大部分个体都向最优解附近靠拢,分布更集中,有些个体重叠。从表 2 中可看出,局部寻优后,各解的精度都有提高。

(2) 函数 $f(x) = (\sin 2x + \cos 5x)^3, -5 < x < 0$

此函数有四个局部极小值,其中在 $x = -0.6464$ 取得全

局最小值 -7.5010。取种群规模为 80,编码为 18 位长的二进制串,交叉概率为 0.9,变异概率为 0.01,采用单点交叉,终止代数 100,运行次数为 20 次。种群的变化如图 6 所示。从图 6 可以直观地看出,与 SGA 相比,DRN-GA 可以找到更多的解,同时,局部寻优可以进一步提高解的精度(种群个体分布于 X 轴上)。

4 结语

本文提出的 DRN-GA 算法有两方面的良好性能:1)与普通遗传算法相比能搜索到更多的解。2)沿着标记的方向继续局部搜索可以进一步提高解的精度。仿真结果证明了这两点。在一些对时间要求不高而对精度要求较高的场合,这种算法有较高应用价值。

参考文献:

[1] HOLLAND JH. Adaptation in Natural and Artificial Systems [M]. Michigan: The University of Michigan Press, Ann Arbor, 1975.

[2] DE JONG KA. An Analysis of the Behavior of a Class of Genetic Adaptive System[D]. Ph. D Dissertation, University of Michigan, No. 76-9381, 1975.

[3] CAVICCHIO DJ. Reproductive Adaptive Plans[A]. Proceedings Of the ACM. 1972 Annual Conference[C]. 1972. 1-11.

[4] CAVICCHIO DJ. Adaptive Search Using Simulated Evolution[R]. Report 03296-4-T, Ph. D Dissertation, University of Michigan, 1970.

[5] GOLDBERG DE, RICHARDSON J. Genetic Algorithm with Sharing for Multimodal Function Optimization[A]. Proceedings of 2nd International Conference on Genetic Algorithm, Lawrence Erlbaum Associates[C]. 1987. 41-49.