

一种求解函数优化问题的单亲遗传算法^{*}

王斌¹ 李元香¹ 王治²

(武汉大学软件工程国家重点实验室 武汉430072)¹ (阿尔卡特上海贝尔股份有限公司 上海200083)²

A Partheno Genetic Algorithm for Traveling Salesman Problem

WANG Bin¹ LI Yuan-Xiang¹ WANG Zhi²

(State Key Laboratory of Software Engineering, Wuhan University, Wuhan 430072)¹

(Alcatel Shanghai Bell Inc., Shanghai 200083)²

(wzfq@yeah.net)

Abstract In this paper, a kind of Partheno Genetic Algorithm (PGA) based on Path Representation scheme is proposed for solving Traveling Salesman Problem (TSP). This algorithm employs only mutation and selection operators to produce the offspring, instead of traditional crossover operator. A specific mutation operator is designed combining the insertion operator with inversion operator, which ensures its strong searching capability. This algorithm simulates the recurrence of nature evolution process, while providing fewer control parameters. Experiments based on Chinese 144 cities (CHN144) and 7 instances selected from TSPLIB are used to test the performance of this algorithm. They prove that it can reach the satisfying optimization at a faster speed. Especially, for the CHN144, the best path it finds is better than any other available one.

Keywords TSP, Combinatory operator, Evolution cycle, Partheno-genetic algorithm

1 引言

求解函数优化问题,特别是多峰函数的优化问题,由于传统的遗传算法将杂交算子作为主要算子,在选择压力的作用下,群体的多样性在进化过程中存在逐步降低的现象,最后以致趋同,进而过早收敛。文[1]也指出“杂交算子在搜索过程中存在着严重的成熟化效应,它在起搜索作用的同时,不可避免地使群体的多样性趋于0,从而逐渐减小自己的搜索范围,引起过早收敛”。

为防止遗传算法过早收敛,目前已提出了许多改进的算法。文[2]提出了“多种群自聚集演化算法”,其思想是设法使群体在演化过程中自行分割成种群,不同的种群向不同的局部最优方向收敛,而种群间不再交换信息。虽然通过种群隔离来保护全局最优附近的种群,该算法可以防止过早收敛,但是它存在着参数难以控制的问题,具体为算法中个体的领域的确定方式难以控制,取值过大或过小都将影响算法的性能。文[3]提出了一种改进的基于二进制编码的自适应变异遗传算法,其思想是:对个体的每一位编码位有一个相应的变异概率,初值设为0.5,在每一代运行中对个体的每一编码位,依据各自的变异概率进行变异,再依据生成的子个体、父体的适应值以及父体的编码位的变异概率计算子个体的相应的编码位的变异概率,使变异概率逐渐减小,最后收敛到全局最优解。该算法的额外开销大,每一代进化都需要重新计算个体的每一位的变异概率,而且计算中要用到前后两代个体的适应值。从文[3]给出的实验计算结果来看计算时间长,而且对于复杂的函数求解收敛率不高。

本文提出的单亲遗传算法(PGA),是基于实数编码的改进的遗传算法。它只使用变异算子,并且一子个体只由一个父

体变异产生,从而保证个体间的隔离,彼此不交换信息,有效地保持群体的多样性。针对变异算子全局勘探能力强,但局部开发能力差的特点,PGA算法采用了文[4]提出的非均匀变异算子,它的行为表现为:在算法的前期能均匀地搜索解空间,在后期能对局部进行越来越细微的搜索。PGA算法通过最优个体的保留,和每代的用保留的最优个体替换当前群体的最差的个体,使群体最终收敛到全局最优点。PGA算法简单,参数少且易于控制,具有较高的求解精度和较快的收敛速度,在全局寻优和防止过早收敛方面有明显的优势,实验结果证明了这一点。

2 算法描述及分析

2.1 非均匀变异算子(nonuniform mutation)

非均匀变异算子是由 Janikow 和 Michalewicz^[4]给出的,下面给出它的描述:对一个体 x , 如果选择其中的元素 x_k 变异,得到新个体为 $x' = \langle x_1, \dots, x'_k, \dots, x_n \rangle$, 这里

$$x'_k = \begin{cases} x_k + \Delta(t, \text{right}(k) - x_k) & \text{如果随机二进制数为0} \\ x_k - \Delta(t, x_k - \text{left}(k)) & \text{如果随机二进制数为1} \end{cases}$$

其中 $\Delta(t, y) = y \cdot r \cdot (1 - \frac{t}{T})^b$, 这里 r 为区间 $[0, 1]$ 里的一个随机数, t 为当前代数, T 为最大代数, b 为确定非均匀度的系统参数。

2.2 单亲遗传算法(PGA)描述

```

procedure PGA ;
begin
  随机生成 N 个个体形成群体 P;
  计算 P 中个体的适应值;
  M-OPT := P 中最好的个体; t := 1;
  while t <= T do
  begin
    r := P 中最好的个体;
    If r 的适应值 > M-OPT 的适应值 then M-OPT := r;
  end
end

```

^{*} 国家自然科学基金(69703011)资助项目,教育部骨干教师资助计划,王斌 硕士研究生,研究方向为演化计算和并行计算,李元香 教授,博士生导师,研究方向为演化计算和并行计算,王治 硕士,研究方向为演化计算,智能网管。

```

用 M-OPT 替换掉 P 中最差的个体;
对 P 中的每个个体进行变异操作;
t:=t+1;
end;
end;

```

2.3 算法思想及说明

PGA 算法保留历史最优个体,在每一代淘汰最差的个体,用保留的历史最优个体替换掉当前代最差的个体,并对每代个体使用非均匀变异算子进行变异,产生下一代个体。其控制参数有3个,即:群体规模 N、控制算法的终止的进化代数 T、非均匀变异算子中的确定非均匀度的系统参数 b。这里没有复杂的参数控制。

要克服遗传算法的过早收敛,应该是在算法的前期,使个体保持分散,均匀地搜索解空间,此时遗传算子具有全局的均匀搜索的能力,使搜索能到达尽可能多的局部区域,而又不被局部最优所吸引。在前期进行充分的全局搜索的基础上,算法后期群体会收缩到全局最优个体所在的局部区域进行局部的精确搜索,此时的遗传算子应该具有局部的细微搜索能力。PGA 算法的总体思想就是利用专门设计的变异算子实现这种前期的充分全局勘探,发现最有希望找到最优解的区域,进而在后期实现群体逐步集中到该区域时的精细局部开发。

非均匀变异算子的行为是随着演化代数的变化而变化的,在算法的前期演化代数 t 较小,Δ(t, y) 的可变化范围较大,这样能均匀地进行全局搜索。在算法的后期 t 较大,Δ(t, y) 的可变化范围随着 t 的增大越来越小,这样能集中在局部区域搜索。如果控制参数 T 越大,随着演化代数增加,在搜索的前期就越能进行充分的勘探,使得有个体可以进入最优点的吸引域,并且在后期在一定的选择压力下逐渐使群体集中到最优点的吸引域内,从而防止了算法的过早收敛。

2.4 结合实例的 PGA 算法分析

下面结合一个多峰函数优化问题,分析 PGA 算法的行为。下面的 F1 函数是 Schaffer 函数,它的全局极大点是(0, 0)。而在距全局极大点大约 3.14 范围内的隆起部有无限多的次全局极大点。由于 F3 的强烈振荡性质及其全局最优个体被次最优个体所包围的特性使得一般算法很难找到它的全局最优解。许多改进的遗传算法都把 F1 作为测试函数。如文[2]中的 f1,文[3]中的 f2。

$$F1: 0.5 - \frac{(\sin^2 \sqrt{x_1^2 + x_2^2} - 0.5)}{[1 + 0.001(x_1^2 + x_2^2)]^2}, -5 \leq x_1, x_2 \leq 5$$

下面的图1—图4是 PGA 算法对函数 F1 进行优化求解时,运行过程中的不同进化代群体的个体分布图。图中‘.’表示全局最优个体,‘□’表示个体。

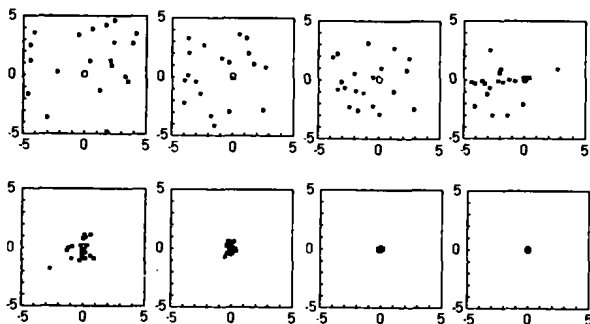


图1 参数 N=20, T=600, b=2 时,PGA 运行的第75、150、225、300、375、450、525、600代的群体的个体分布图

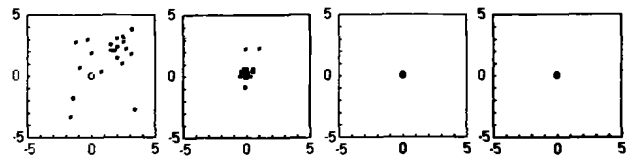


图2 参数 N=20, T=300, b=10 时,PGA 运行的第75、150、225、300代的群体的个体分布图

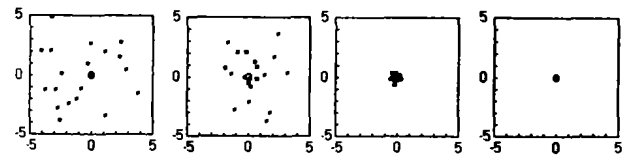


图3 参数 N=20, T=1200, b=2 时,PGA 运行的第300、600、900、1200代的群体的个体分布图

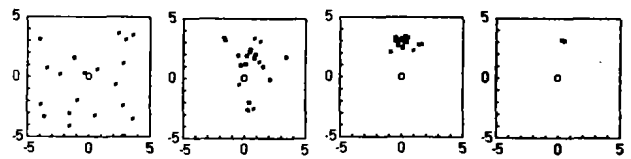


图4 参数 N=20, T=300, b=2 时,PGA 运行的第75、150、225、300代的群体的个体分布图

从图1看出 PGA 在300代前个体是分散的,在300代后个体开始向全局最优个体聚集,到600代群体收缩到全局最优个体。比较图2和图1:当 b 取值增大,其他参数相同的情况下,算法的收敛速度明显加快。比较图3和图1:当 T 取值增大,其他参数相同的情况下,算法的收敛速度明显减慢。比较图4和图1:当 T 取值过小,其他参数相同的情况下,算法的收敛速度大大加快,但最后收敛到了局部最优个体。

从图1到图4可看出,群体在前期分散勘探,以发现全局最优个体所在的区域,在后期群体向全局最优个体区域聚集,进行集中开采的过程。参数 T 对算法的性能有明显的影, T 过大,虽然能收敛到全局最优个体,但收敛速度减慢,而 T 过小,由于群体在前期的勘探时间过短,最后收敛到局部最优个体。但只要 T 给得足够大,就肯定能收敛到全局最优个体。参数 b 能影响算法的收敛速度, b 值增大,提高了算法的收敛速度,但 b 值增大,也使非均匀变异算子的搜索步长减小,所以 b 值不能取得太大。

图1所得最优解 $x = (0.0000101351, 0.0000290526)$, $f(x) = 0.9999999905228$, 迭代600次,图2所得最优解 $x = (0.0000000059, 0.0000000006)$, $f(x) = 1.00000000000000$, 迭代300次。文[2]的多种群自聚集演化算法所得的最优解为 $x = (0.0003, -0.0002)$, $f(x) = 0.9996$, 迭代1000次。所以 PGA 算法在求解 Schaffer 函数时的收敛速度和求解精度都要高于文[2]的多种群自聚集演化算法给出的结果。文[3]的自适应变异遗传算法是在群体规模为100,5000次迭代下,100次实验66次算得最优解,收敛时间太长,成功率不高。

3 实验测试

为了进一步地测试 PGA 算法收敛到全局最优个体的能力和算法的收敛速度,选用了容易使遗传算法过早收敛的二维多峰函数、二维单峰函数和高维函数对算法进行测试,以下是实验用到的测试函数:

$$F2: \sum_{i=1}^5 i \cos[(i+1)x_1 + i] \cdot \sum_{i=1}^5 i \cos[(i+1)x_2 + i]$$

其中 $-10 \leq x_i \leq 10$ 当 $i=1,2$

这是一个二维多峰函数,函数有760个局部最小,其中18个是全局最小,其值为-186.73

$$F3: 100(x_1^2 - x_2)^2 + (1 - x_1)^2, -2.048 \leq x_i \leq 2.048$$

这是一个非凸函数,虽然它是一个二维单峰函数,但它是病态的,难以进行全局极小化。它的全局极小值为0,极小点为(1,1)

$$F4: - \sum_{i=1}^n x_i \sin(\sqrt{|x_i|}) (-500 \leq x_i \leq 500)$$

函数的次全局极小点远离全局极小点,函数的全局极小点 $(x_i = 420.96875, i=1, \dots, n)$

对函数 F2, F3 各运行10次,函数值的计算保留小数点后14位有效数值。算法的参数的取值为 $N=20, b=10$, 因为 T 的取值越大,求得最优解的概率越大,求得最优解的精度也越高,所以实验测试连续运行算法10次,其解的 T 值较为精确。结果如下:

表1 二维函数计算的结果

函数	算法求得的最优函数值	迭代次数 T
F2	0.00000000000000	15000
F3	186.73090883102394	500

对函数 F4, 取维数 $n=20, 50, 100, 200, 500$ 各做1次实验,函数值的计算保留小数点后14位有效数值,算法的参数的取值为 $N=20, b=2$, 所得最优解向量的最小分量值,最大分量值,以及 T 的取值如下(T 的取值更大的话,求解的精度也将更高):

表2 高维函数计算的结果

维数	求得的最优点的最小分量值	求得的最优点的最大分量值	迭代次数 T
20	420.96840375635936	420.96883072006694	35000
50	420.96691741297690	420.96942438125824	85000
100	420.96574323812429	420.96925839726208	170000
200	420.96458019978458	420.96946954489171	360000
500	420.96670311248499	420.96923486539667	1260000

(上接第130页)

组件中心负责扫描整个消息队列,发现未处理的消息并将其提取出来,解析其中内容,然后交给 ModelFactory 处理。

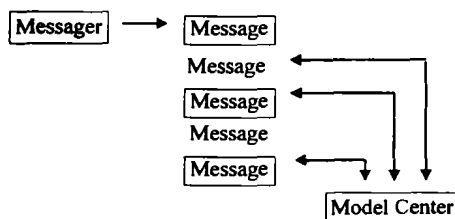


图3 采用消息队列进行缓冲处理

引入这种队列缓冲机制后,我们就可以根据需要对整个应用系统进行扩展以提高其性能。例如,为了提高系统的处理能力,可以通过编程使模型/组件中心支持多线程的工作方式,这样,系统对请求的处理能力就可以大幅度地提高;另外,我们看到,引入缓冲队列这一中间处理过程之后,我们只需要再附加少量的模块就可以把组件中心这个庞大的模块部署到

对函数 F4, 若取参数 $N=20, b=5$, 函数的维数为20维, 迭代次数 $T=35000$, 算得的最优解向量的最小分量值为 -302.52493475233606 , 最大分量值为 420.96874730329939 , 算法收敛到局部最优点。若将 T 的值增大, $T=70000$, 最优解向量的最小分量值为 420.96874459692141 , 最大分量值为 420.96874785510272 , 算法收敛到全局最优点。

结论 由以上运算实例可见,采用最优个体保留,每代用保留的最优个体替换当前群体的最差个体,只使用非均匀变异算子的单亲遗传算法,由于在前期能有效勘探出最优区域,在后期能对最优区域进行精细的搜索,从而能有效地防止遗传算法的过早收敛,又有较高的求解质量,算法的收敛速度也很快。

同其他的改进的遗传算法相比,PGA 算法的参数较少,而且易于控制。算法的收敛速度和求解精度受终止代数及最大迭代次数 T 的控制, T 越大解的质量越高,当然收敛速度也减慢, T 的取值可由对解的质量要求来定。算法的参数 b 对算法收敛速度有影响, b 增大,算法收敛速度会加快,但也容易出现算法的过早收敛,其原因是较大的 b 值,限制着非均匀算子的搜索步长,使前期的勘探不够充分,使个体进入最优点的吸引域的概率降低,从而造成算法最终收敛到局部最优值,当然如果加大迭代次数,也是可以收敛到全局最优点的。根据经验值, b 的取值范围在1到10之间,对于简单低微的函数,可以取较大的 b 值,对于复杂的高函数,可以取较小的 b 值。

参考文献

- 徐宗本,高勇.遗传算法过早收敛现象的特征分析及其预防[J].中国科学(E辑),1996,26(4):364~375
- 黄樟灿,陈思多,李亮.演化计算中的种群隔离与自聚集.软件学报,2002,13(4):827~832
- 李海民,吴成柯.自适应变异遗传算法及其性能分析.电子学报,1999,27(5):90~93
- Janikow C, Michalewicz Z. An experimental comparison of binary and floating point representations in genetic algorithms. In Belew and Booker [21], pages 151~157

多个服务器上去,在多个服务器之间实现分布式处理和负载均衡。

总结 综合上面对该系统模型的分析,我们可以看到,这个改进后的以 MVC 设计模式构建的 Web 应用系统框架,其结构较传统的框架清晰,系统各个模块之间的耦合相对宽松,特别是它在很大程度上解决了长期困扰 Web 开发人员如何将系统的事务逻辑和表现形式分离的难题,实现了开发工作的并行开展。并且,在引入了模型组件库和逻辑页面模板库之后,整个系统变得更为灵活,其可维护性和可扩展性得到了极大的提高。

参考文献

- Gamma E. Design Patterns. U.S. Addison-Wesley Pub Co.
- Martin D. Professional XML. U.K. Wrox Press
- Mohr S T. Designing Distributed Applications with XML, ASP, IE5, LDAP & MSMQ. U.K. Wrox Press
- Sun Microsystems Inc. Java™ Servlet 2.3 and JavaServer Pages™ 1.2 Specifications
- Sun Microsystems Inc. JavaServer Pages™ Standard Tag Library Specification