

基于有导向变异算子的 GM-EA 算法*

闭应洲^{1,2}, 陆建波¹, 丁立新², 元昌安¹

(1. 广西师范学院 计算机与信息工程学院, 南宁 530004; 2. 武汉大学 软件工程国家重点实验室, 武汉 430072)

摘要: 为了进一步提高演化算法的效率, 提出基于有导向变异算子的 GM-EA 算法 (guided mutation evolutionary algorithm)。通过结合粒子群优化的方法改进郭涛算法, 更好地利用当前最优解指导变异, 并将算法分为探索与开采两个阶段; 在开采阶段基于模拟退火方法决定是否用新个体取代旧个体, 在巩固所获取的建筑物成分的同时, 尽可能克服早熟收敛问题。实验结果证明了新算法的有效性。

关键词: 有导向的变异; 郭涛算法; 粒子群优化; 模拟退火

中图分类号: TP301.6 **文献标志码:** A **文章编号:** 1001-3695(2010)04-1249-03

doi:10.3969/j.issn.1001-3695.2010.04.012

GM-EA: guided mutation evolutionary algorithm

BI Ying-zhou^{1,2}, LU Jian-bo¹, DING Li-xin², YUAN Chang-an¹

(1. College of Computer & Information Engineer, Guangxi Teachers Education University, Nanning 530004, China; 2. State Key Laboratory of Software Engineering, Wuhan University, Wuhan 430072, China)

Abstract: To design a more effective evolutionary algorithm, this paper introduced a new guided mutation evolutionary algorithm by combining Guotao algorithm with the idea from particle swarm optimization, which focused on exploiting the global best solution in population to direct the mutation. In order to preserve the components of building-blocks and avoid the premature problem, separated the search process as the exploration phase and exploitation phase, and in exploitation phase simulated annealing was applied as the replace policy. The experimental results show that the proposed algorithm is significantly superior to Guotao algorithm.

Key words: guided mutation; Guotao algorithm; particle swarm optimization (PSO); simulated annealing

0 引言

变异算子作用在一个个体上, 然后产生一个新的与父代不一致的子代个体, 它实现了演化搜索的基本步骤: 生成一个后代就相当于在搜索空间中从一点搜索到另一点。由于变异算子的作用一般是使得某些位以小的概率产生变化, 或者是个体的一小部分产生变化, 在变异算子的作用下, 父代个体与子代个体变化不大。从这个角度看, 变异算子有其理论意义, 它可以保证搜索空间是连接的, 即搜索空间中任意点在变异算子的作用下是可达的。在给定充足时间的前提下, 演化算法找到最优解的关键在于表示潜在解的基因型在变化算子的作用下是可达的^[1]。

满足这个条件的最简单方法就是, 允许变异算子将任意位上的等位基因以非零概率转换为任意的等位基因, 但实际应用中很少采用这种策略。因为面对“爆炸性”的组合空间, 这种完全随机的变异算子会严重影响算法的性能。为了提高算法的性能, 一系列有导向的变异算子 (guided mutation operator)^[2] 被提出; 文献[3] 提出一个带导向的变异算子, 它通过从当前种群中获取线索来指导倒序; 文献[4] 通过应用免疫原理设计有导向的变异算子取代经典演

化算法中盲目的变异算子, 有效抑制优化过程中的退化现象; 文献[5] 提出将分布估计算法与传统的变异算子结合, 在概率模型的指导下, 变异算子修改父代个体产生新个体。由于概率模型刻画了有潜质的区间, 所以新解很有可能属于或靠近有前途的区域, 所有这些有导向的变异算子都取得了很好的效果。

粒子群算法 (PSO) 是模拟鸟群觅食过程的一种仿生算法。在 PSO 中, 只有当前全局最优解或局部最优解影响粒子的搜索行为。与遗传算法比较, 在大多数的情况下, PSO 搜索更强调利用已有的最优解^[6]。近似最优原理^[7] 认为: 好的解是相似的, 因此围绕已知最优解来产生新的解是合理的。但需要注意的是, 强调对已知最优解的利用在加快搜索进程向较优解靠拢的同时, 往往导致早熟收敛。在探索 (exploration) 与开采 (exploitation) 之间找到平衡是各种启发式算法需要面对的难题, 为了解决这个难题, 文献[8] 提出将搜索过程分为探索阶段与开采阶段。

基于这些认识, 本文提出应用 PSO 改进郭涛算法, 设计具有更强导向性的变异算子, 并将搜索过程分为探索与开采阶段。但在早期探索阶段只允许新个体有改进时才能取代旧个体, 通过高的选择压力来加速探索过程; 在后期开采阶段, 基于

收稿日期: 2009-09-02; **修回日期:** 2009-10-29 **基金项目:** 国家自然科学基金资助项目 (60763012, 40761027); 广西自然科学基金资助项目 (0991104)

作者简介: 闭应洲 (1967-), 男, 副教授, 博士, 主要研究方向为智能计算、智能信息处理 (byzhou@163.com); 陆建波, 男, 讲师, 硕士, 主要研究方向为智能信息处理; 丁立新, 教授, 博士, 主要研究方向为智能计算; 元昌安, 男, 教授, 博士, 主要研究方向为智能计算。

模拟退火算法决定是否用新个体取代旧个体,允许以一定的小概率接受差解,尽可能克服早熟收敛问题。

1 交叉算子与变异算子的作用比较

由模式定理可知,具有低阶、短定义距以及平均适应度高于种群平均适应度的模式在子代中将以指数增长,Holland 将这类模式称为建筑块(building blocks),Goldberg 提出了建筑块假设,认为建筑块在遗传算子的作用下相互结合产生适应度更高的个体,从而找到更优的可行解。Harrik 也指出^[9],遗传算法能够超越随机搜索(random search)的根本原因在于新解的生成不是随机的,在产生新种群时,它有效地利用当代种群包含的建筑块信息。

根据上述理论,遗传算法的工作过程就是,从选择相互竞争的低阶、短定义距的模式开始,逐步将它们组合形成高阶的模式,然后不断重复这个过程,直到一个长度为 $l-1$,阶为 l 的模式(即全局最优解)被创建和发现。下面从模式处理的角度来分析交叉算子和变异算子的作用。交叉算子或变异算子的破坏性行为是指一个属于阶为 k 的模式 H_k 的个体,交叉或变异后它的后代不再属于模式 H_k ,而建设性行为是指一个属于阶为 m 的模式 H_m 的个体,交叉或变异后得到一个属于阶为 n 的模式 H_n 的个体, $m < n$ 。

对于演化算法中交叉和变异算子的破坏性和建设性行为,文献[10]作了系统的分析,并得到以下结果:a)模式的阶越高,模式的长度越长,交叉算子的破坏性越大,但随着种群的个体趋于相似,交叉算子的破坏性趋于下降;b)交叉算子的破坏性越大,它的建设性越强,通过比较交叉算子的破坏性与建设性,进一步证实了“没有免费的午餐”理论(no-free-lunch theorem);c)模式的阶越高,变异算子的破坏性越大,而且变异的水平(概率)越大破坏性越大;d)高水平的变异并没有带来高的建设性;e)由于交叉能够将低阶的建筑块(模式)组合成高阶的建筑块,从建设性的角度来说,交叉算子具有更高的建设性,但笔者认为,得出这个结论的原因在于变异算子是完全随机的。

在不同的演化算法中,变异算子所起的作用也不一样。在遗传算法中,变异算子被看做是一个背景算子(background operator),它的作用仅仅是给种群添加新鲜血液,而在遗传程序设计,则根本不用变异算子;但在演化规划中,变异算子是惟一的变化算子,通过对个体中的每个变量应用适应性的变异,使得 EP 在不用交叉算子的情况下一样取得很好的效果。

基于对交叉算子和变异算子的建设性和破坏性作用的分析,笔者认为,在演化搜索的早期阶段,个体所包含的模式阶较低、定义距较短,交叉算子和变异算子的建设性大于破坏性,因此,算法的适应值提高很快;而在演化搜索的后期,个体所包含的模式阶较高、定义距较长,因此交叉算子和变异算子的破坏性远大于建设性。所以,算法的适应值很难提高,甚至“早熟”收敛。为了解决这个问题,有导向的变异算子逐步成为研究热点。

2 基于有导向变异算子的演化算法

为了设计更有效的变异算子加速搜索进程,本文提出结合

粒子群优化算法的思想,改进郭涛算法。但强调对历史最优解的利用,加速向局部最优靠拢往往会导致“早熟”收敛,为此本文提出:在算法的后期阶段结合模拟退火算法的特点,以较小的概率允许新个体作小的后退。

记 $P_{m,t}$ 为第 t 代种群中的个体 m ,它的目标函数值为 $f(P_{m,t})$, $O_{m,t}$ 为对 $P_{m,t}$ 施加以有导向的变异后得到的新个体,相应的目标函数值为 $f(O_{m,t})$ 。对于求最小值优化问题来说,若 $f(O_{m,t}) < f(P_{m,t})$,则变异后的新个体更好,用新个体取代旧个体, $P_{m,t+1} = O_{m,t}$;否则计算概率

$$p = \exp[f(P_{m,t}) - f(O_{m,t})/T] \quad (1)$$

若 $p > \text{random}[0, 1)$,用新个体取代旧个体, $P_{m,t+1} = O_{m,t}$;否则 $P_{m,t+1} = P_{m,t}$ 。其中 $\text{random}[0, 1)$ 表示区间 $[0, 1)$ 上的一个均匀随机数, T 为模拟退火中的温度,随着演化的推进逐步减小。

改进后的算法与郭涛算法在以下两个方面不一样:a)郭涛算子中以较小的概率(一般 $p = 0.02$)进行随机倒序,而以较大的概率(一般 $p = 0.98$)从种群中随机选择另一个体指导倒序;而有导向的变异算子以概率($p = 0.2 \sim 0.5$)选择当前最优解指导倒序,以概率 $1 - p$ 从种群中随机选择另一个体指导倒序;b)对个体施加倒序运算后立即评估,在早期探索阶段,只有新个体比旧个体更好时才能取代旧个体;而在后期开采阶段,基于模拟退火算法决定是否用新个体取代旧个体。模拟退火允许以一定的小概率接受差解,使得搜索可以跳出局部最优,它是解决早熟收敛问题的重要方法。下面为带导向的变异算子的描述:

```

初始化种群  $P_0, t = 0$ ;
计算  $P_0$  中个体的适应值;
while(不满足终止条件) do
{
    对种群  $P_t$  中的每个个体  $S_i$  do
    {  $S' = S_i$ 
      从  $S'$  中随机选择一个城市作为  $C$ 
      while (true) do
      { if (rand() ≤  $p$ )
        从当前发现的最优解中选择另一城市  $c'$ 
        else
        | 从种群  $P_t$  中随机选择另一个个体  $S'$ ;
          在  $S'$  中指定  $C$  的下一个城市为  $C'$ ;
        }
      若  $S$  中  $C'$  是与  $C$  相邻的城市,则不进行倒序操作退出循环。
      对  $S$  中  $C$  的下一个城市与  $C'$  间的城市实施倒序操作
       $C = C'$ 
      评估新个体
      if ( $t < n\text{Expoitation}$ ) // 探索阶段
        如果新个体更好就用新个体取代旧个体,退出循环;
      else // 开采阶段
        基于模拟退火方法决定是否用新个体取代旧个体,退出循环;
    }
}
t = t + 1
}
    
```

3 实验结果

为了比较新算法的性能,选用了随机产生的旅行商问题:TSP200、TSP500,它们分别包含 200 和 500 个城市。这些城市

的位置都是通过随机函数在二维空间生成的: $x \in [0, 1000]$, $y \in [0, 1000]$,城市之间的距离为欧氏距离。计算机实验平台是:CPU 1.73 GHz(双核), RAM 1 GB, 操作系统 Windows XP, 开发工具 VC 6.0。

算法1为郭涛算法,算法2是本文提出的新算法,两个算法中种群规模都是100。算法1中 $p = 0.02$,表示算法1以0.02的概率执行随机倒序,以0.98的概率执行有导向的倒序,即从种群中随机选择另一个体指导倒序。算法2中 $p = 0.5$,表示算法2以0.5的概率应用当前最优解来指导倒序,以0.5的概率从种群中随机选择另一个体指导倒序。为了比较两个算法的性能,设计以迭代的次数为停机条件,表1、2的代数就是指算法运行到指定的代数后停机。

表1是针对TSP200问题,算法1和算法2分别独立运行20次的运行结果。表中的数值是指周游所有城市的距离,最小值是指算法独立运行20次所得到的最短距离,平均值是指算法独立运行20次所得到的平均距离,时间是指算法独立运行20次平均每次花费的时间。算法2的搜索过程分为两个阶段,早期的5000代为探索阶段,经变异产生的新个体必须有改进才能取代旧个体;从5001代开始进入开采阶段,基于模拟退火算法决定是否用新个体取代旧个体。在模拟退火中,初始温度设为15,每代下降0.001。

表1 不同算法基于TSP200的运行结果比较

代数	算法1(郭涛算法)			算法2(有导向的变异算子)		
	最小值	平均值	时间/s	最小值	平均值	时间/s
5 000	11 524	11 758	26.8	10 820	10 890	14.3
10 000	10 688	10 841	32.8	10 673	10 716	18.5
15 000	10 733	10 775	38.7	10 632	10 680	23.7
20 000	10 655	10 693	45.5	10 620	10 641	27.5

表2是针对TSP500问题,算法1和算法2分别独立运行20次的运行结果。图1是算法2找到的最优解,每个“*”表示城市的位置。算法2的搜索过程同样分为两个阶段,早期的10000代为探索阶段,经变异产生的新个体必须有改进才能取代旧个体;从10001代开始进入开采阶段,基于模拟退火算法决定是否用新个体取代旧个体。其中,初始温度设为20,每代下降0.0005。

表2 不同算法基于TSP500的运行结果比较

代数	算法1(郭涛算法)			算法2(有导向的变异算子)		
	最小值	平均值	时间/s	最小值	平均值	时间/s
10 000	29 118	29 240	228	17 714	17 830	102.5
20 000	17 653	17 970	270.2	16 705	16 780	131.2
30 000	16 912	17 020	305.3	16 533	16 594	162.5
50 000	16 647	16 772	364.6	16 364	16 428	195.7

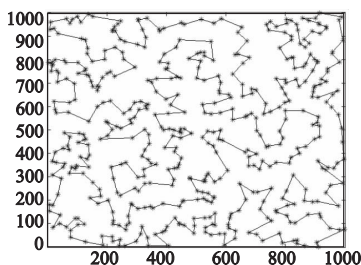


图1 基于有导向的变异算子找到的TSP500最优解

从表1、2的结果可以看出,在运行同样代数的条件下,算

法2比算法1花费更少的时间而取得更好的结果。需要注意的是,应用模拟退火算法时,初温的选择和温度下降的幅度对算法的性能影响很大。笔者在实验过程中发现,初温设置较大时,如设为100,即使应用较大的温度下降幅度来保证后期温度很低,基于模拟退火选择的效果也不好。究其原因,笔者认为,经过早期的探索阶段,算法已找到较好的解,如果初温过高,就会导致算法作较大的后退,破坏已有的建筑块,从而影响算法的性能。

4 结束语

设计演化算法的关键问题之一就是如何产生后代。在演化算法中,传统的变异算子对父代作小的随机扰动,但由于变异算子没有有效利用种群中的全局信息,种群中的个体经变异后也可能变得较差。研究表明,郭涛算法通过从当前种群中获取线索来指导变异,有利于建筑块的组合,降低了随机倒序对已有建筑块的破坏,因此取得了比随机倒序算法更好的性能。本文将粒子群优化的方法与郭涛算法结合,更好地利用当前最优解指导变异,并通过将算法分为探索与开采两个阶段,在开采阶段基于模拟退火方法决定是否用新个体取代旧个体,从而在探索与开采之间实现了较好的平衡,取得了比郭涛算法更好的效果。实验结果证明了带导向的变异算子是一种有效的识别和组合建筑块的算子。

参考文献:

- [1] EIBEN A E, SMITH J E. Introduction to evolutionary computing [M]. Berlin: Springer-Verlag,2003.
- [2] STORN R, PRICE K. Differential evolution: a simple and efficient adaptive scheme for global optimization over continuous spaces, TR-95-012 [R]. Berkely: International Computer Science Institute, 1995.
- [3] GUO Tao, MICHALEWICZ Z. Inver-over operator for the TSP[C]// Proc of the 5th International Conference on Parallel Problem Solving from Nature. Berlin: Springer,1998:803-812.
- [4] 闭应洲,丁立新,杨小雄.基于免疫原理降低交叉算子破坏性的研究[J].计算机工程与应用,2007,43(18):42-44.
- [5] ZHANG Qing-fu, SUN Jian-yong, TSAN G E. An evolutionary algorithm with guided mutation for the maximum clique problem [J]. IEEE Trans on Evolutionary Computation, 2005, 9(2): 192-200.
- [6] 张利彪.基于粒子群和微分进化的优化算法研究[D].长春:吉林大学,2007.
- [7] GLOVER F W, LAGUNA M. Tabu search[M]. Norwell, MA: Kluwer Academic Publishers,1998.
- [8] KEMENADE C H M. Building block filtering and mixing, SEN-R9837[R]. Amsterdam: Centre for Mathematics and Computer Science,1998.
- [9] HARIK G R. Learning gene linkage to efficiently solve problems of bounded difficulty using genetic algorithms[D]. Ann Arbor: University of Michigan,1997.
- [10] SPEARS W M. The role of mutation and recombination in evolutionary algorithms[D]. Fairfax: George Mason University,1998.