

用于求解 TSP 问题的改进遗传算法

文 艺 潘大志

(西华师范大学数学与信息学院 南充 637009)

摘 要 TSP 问题是一个典型的组合优化问题,也是一个 NP 难题,一般很难精确地求出其最优解,因而找出有效的近似解算法具有重要意义。针对基本遗传算法在解决 TSP 问题时所存在的收敛速度慢、容易“早熟”的问题,在选择算子中引入选择因子,同时提出一种改进的交叉算子和基于种群相似度的更新策略。改进的交叉算子是先比较两个城市间距离再进行交换城市序号,因此加快了收敛的速度,而基于种群的相似度更新策略则在算法的后期可以有效地防止早熟。通过对实例 144 进行测试,证明该算法在解决该类问题上取得了较好的效果。

关键词 TSP, 遗传算法, 改进交叉算子, 相似度

中图法分类号 TP301 文献标识码 A

Improved Genetic Algorithm for Traveling Salesman Problem

WEN Yi PAN Da-zhi

(College of Mathematics and Information, China West Normal University, Nanchong 637009, China)

Abstract Traveling salesman problem(TSP) is a typical combination optimization problem, which is also a NP hard problem. It's hard to find a precision result, and it is very important to search for the near result. Based on basic genetic algorithm in solving TSP problems of slow convergence speed and easy to premature, an improved crossover operator and the updating strategy based on similarity of a population was put forward in this paper. The improved crossover operator exchanges the two cities' number by comparing the distance between the two cities, thus accelerates the convergence speed. And the updating strategy based on population can effectively prevent the population from premature in the later stages of the algorithm. Especially, for the CHN144, the best path it finds is better than the basic one.

Keywords TSP, Genetic algorithm, Improved crossover operator, Similarity

1 引言

遗传算法(GA)是模拟自然界生物进化的一种概率搜索算法,它最早是由 Michigan 大学的 Holland 教授提出的,20 世纪 70 年代,De Jong 基于遗传算法思想在计算机上进行大量纯数值函数优化计算实验,80 年代由 Goldberg 归纳形成基本框架[1]。目前,遗传算法主要被用于优化问题[2]、图像处理[3]、通信工程[4]等领域。

TSP 问题是典型的组合优化问题,对 TSP 问题的一般描述为:一个旅行商从某个城市出发,遍历 n 个城市后返回到出发的城市,且每个城市只能访问一次,如何规划旅游路线才能使旅行商的旅游路线总距离最短。对于 n 个城市而言,可能存在的路径有 $n!/2n$ 条,而且随着 n 的增加,路径的条数呈指数增长。所以用传统的方法求解该问题存在很大的难度,甚至当城市较大时,基本不可能实现。

对于求解小规模 TSP 问题的传统算法有穷举法、分支限界法、动态规划法[5,6]等,但这些都是基于搜索的确定性算法,当 n 较大时,它的时间和空间复杂度都会呈指数增长。所以,对于大规模的 TSP 问题,目前较好的算法有遗传算

法[7,8]、蚁群算法、模拟退火算法、启发式搜索算法以及一些混合的智能算法[9]。目前,已经有很多学者从不同方面对遗传算法进行改进,取得了良好的结果,如刘全等借鉴了精英策略和协同进化的思想,选择两个相异的、高适应度的个体(精英个体)作为进化操作的核心,加快了算法的收敛速度[10];姚明海等将遗传算法和模拟退火算法混合,遗传算法的全局搜索能力弥补了模拟退火算法容易陷入局部最优的问题[11];谢艳丽等根据交叉算子和变异算子的特点,在现有的 GA 基础上,引入拉普拉斯算子改进交叉算子以及结合黄金分割法对变异算子做了进一步改进,提高了算法收敛速度并有效防止了“早熟”。针对遗传算法在搜索过程中的收敛过慢和容易早熟问题[22],且交叉被认为是贡献于遗传算法的全局搜索能力最重要的因素,本文提出一种改进交叉算子和基于种群相似度更新策略,改进交叉算子加快了算法的收敛速度,基于群体相似度的更新策略又可以有效防止算法早熟。从这两方面改进遗传算法,通过对求解中国 144 个城市(CHN144)TSP 问题进行了测试,该改进的遗传算法优于基本遗传算法。

2 基本遗传算法

遗传算法是建立在自然选择和群体遗传学原理基础上的

本文受四川省教育厅自然科学基金(14ZA0127),西华师范大学博士启动基金(12B022),校级创新团队(CXTD2015-4)资助。

文艺(1994-),女,硕士生,主要研究方向为智能计算、数值计算,E-mail:243933701@qq.com;潘大志(1974-),男,教授,硕士生导师,主要研究方向为智能计算、算法设计。

随机迭代、进化、具有广泛适应性的搜索算法。它从一组随机产生的初始群体开始搜索操作，群体中的每一个个体代表问题的一个解，这些个体通过选择、交叉、变异算子进行反复迭代，最终得到优良个体。

基本遗传算法的主要步骤如下：

- Step1 编码；
- Step2 初始化种群；
- Step3 计算种群中每个个体的适应度值；
- Step4 选择；
- Step5 交叉；
- Step6 变异；
- Step7 终止条件判断，若满足终止条件，则转至下一步，否则，转至 Step2；
- Step8 输出结果。

3 用于求解 TSP 的改进遗传算法

3.1 编码

采用顺序表示的编码方法，每一个数字代表一个城市。例如，一条染色体(1 2 3 4 5 6 7)代表对 7 个城市进行排序，从城市 1 依次经由 2、3、4、5、6、7 最后回到城市 1 的一条路径。

3.2 适应度函数

适应度是对个体好坏的评价，适应度越大，说明个体越好。在 TSP 问题适应度函数用路径总距离的倒数来表示。所以使用如下所示的适应度函数：

$$f(x) = \frac{1}{\sum_{i=1}^n d(C_i, C_{i+1}) + d(C_1, C_n)}$$

其中， $d(C_i, C_{i+1})$ 表示第 i 个城市到第 $i+1$ 个城市的距离。

3.3 选择算子

选择操作即从旧群体中以一定的概率选择个体到新种群中。为了使群体中的优良个体得以保留，本文设计一种新的选择算子，在子代种群的选择生成时，先采用比例选择算子（即每个个体被选择的概率为 $\frac{f(x_i)}{\sum f(x_i)}$ ）选择上代 $\alpha \times M$ 个体，种群中剩余的 $(1-\alpha) \times M$ 个体按适应度从高到低从父代选取，最终生成子代的 M 个个体，其中 α 为选择因子。

3.4 交叉算子

遗传算法解决 TSP 问题常用的交叉算子有顺序交叉算子、循环交叉算子、部分交叉算子等。本文针对 TSP 问题提出一种改进的交叉算子，其基本思想是：进行交叉操作的两个父体，一个称作原父体，另一个称作参照父体。先随机选择一个基因位作为标准位，再比较两父体中标准位与后继基因位所表示的城市之间的距离，如果原父体的两基因位的距离大于参照父体中两基因位的距离，则交换原父体中两个父体中此基因位相邻的下个基因位的位置，否则以原父体中的当前基因位的后继基因位为标准位，继续如上操作（如果标准位为这条染色体的最后一个基因位，则它的后继基因位为第一个基因位）。具体操作如下：

(1) 一个子代的产生

按轮盘赌选择两个父体

父体 1: 1 2 3 4 5 6 7

父体 2: 7 6 3 4 2 1 5

1) 选择父体 1 作为原父体，父体 2 作为参照父体；
2) 随机产生一个城市序号如 5，称 5 号城市为标准城市；
3) 比较两个父体中标准城市和下一个城市的距离，则比较 $d(5,6)$ 和 $d(5,7)$ ，如果 $d(5,6) > d(5,7)$ ，则交换原父体中 6 和 7 的位置，产生子代(1 2 3 4 5 7 6)；否则，当前位置加 1，以之为标准城市，继续如上操作。

(2) 另一个子代的产生

以父体 2 作为原父体，父体 1 作为参照父体，继续(1)中步骤 2)、步骤(3)的操作。

3.5 变异算子

这里采用基本遗传算法中的交换变异算子。先随机产生两个城市序号，然后将其对换位置，如产生 $[1, 7]$ 范围内的随机数 2 和 3，确定两个位置：

1 | 2 | 4 5 6 7 | 3 |
变异后：

1 | 3 | 4 5 6 7 | 2 |

3.6 基于群体相似度的种群更新策略

遗传算法中较难解决的一个问题是如何较快地找到最优解并防止“早熟”收敛问题。为了保证遗传算法的全局收敛性，就要维持群体中个体的多样性，避免有效基因的丢失。另一方面，为了加快收敛速度，就要使群体较快地向最优状态转移，这样又会减少群体的多样性，容易陷入局部最优解。许多学者提出了各种改进方案来提高算法的性能，如精英选择策略、对参数进行自适应调整策略^[13]等。

这里运用基于群体相似度的种群更新机制，使种群保持较好的多样性，以提高算法的性能。

(1) 求出 M 个个体的平均距离：

$$T(M) = \frac{1}{M} \sum T_i(M)$$

其中， $T_i(M)$ 为第 i 个个体的路径长度。

(2) 若 $\sum |T_i(M) - T(M)| \leq A$ ，产生 P 个新个体，使个体总数为 $M+P$ ，否则进行下一代操作。

(3) 选择适应度较高的 M 个个体组成的新群体，产生规模为 M 的下一代群体。

4 算法测试分析

4.1 选择因子 α 取值分析

在改进的选择算子中引入了选择因子 α ，为了确定参数 α 的最佳取值，这里以 CHN144 作为测试用例，算法中其他参数设置如下：交叉概率 $p_c = 0.88$ ，变异概率 $p_m = 0.07$ ，种群容量 $M = 200$ ，最大遗传代数为 $T = 2000$ ， $P = 10$ ， $A = 1000$ 。测试中 α 的取值范围为 $0.65 \leq \alpha \leq 0.95$ ，并以 0.05 为步长进行测试。为了保证测试的有效性，算法在每组参数下运行 20 次，并比较算法在每组参数下的最短距离和平均距离，其测试结果如表 1 所列。

表 1 参数 α 在不同取值下的运行结果

α	0.65	0.70	0.75	0.80	0.85	0.90	0.95
Avg	46102	44522	44522	43648	43995	46154	48797
Best	40114	39578	39584	38944	38690	40662	41406

由表 1 可知，当选择因子 α 在 $[0.8, 0.85]$ 之间时，测试用例计算的结果最好（表中 Avg 表示平均距离，Best 表示最短距离）。

4.2 算法测试

为了测试改进遗传算法(IGA)在求解 TSP 问题上的性能,本文以 CHN144 为例,将其运算结果与基本遗传算法(GA)进行对比。算法基本参数取值如下:交叉概率 $p_c = 0.88$,变异概率 $p_m = 0.07$ 。IGA 中选择因子 $\alpha = 0.85$ 。GA 和 IGA 分别在不同的种群容量 M 和遗传代数 T 下各自进行运行 20 次,各自获得最优结果如表 2 所列。为了体现 GA 和 IGA 各自的迭代收敛过程,给出了在表 2 中 1 组和 2 组参数、3 组和 4 组参数、5 组和 6 组参数下,GA 和 IGA 各自最优解的迭代过程,其迭代曲线如图 1—图 3 所示。

表 2 基本遗传算法和改进遗传算法结果比较

数据	算法	M	P	A	T	最优结果
1	GA	100			2000	65734
2	IGA	100	10	10000	2000	41711
3	GA	500			5000	58342
4	IGA	100	50	500000	5000	38761
5	GA	500			10000	50781
6	IGA	500	50	500000	10000	33361

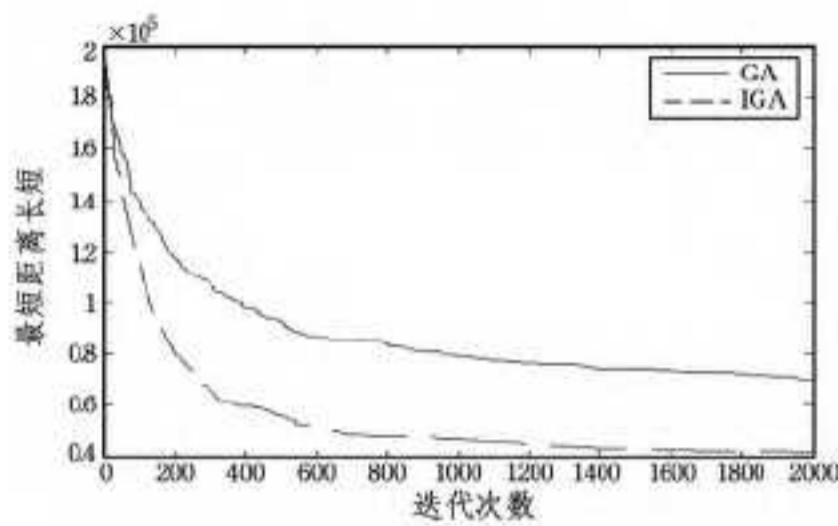


图 1 测试 CHN144 实例中 1 组和 2 组参数迭代过程的比较

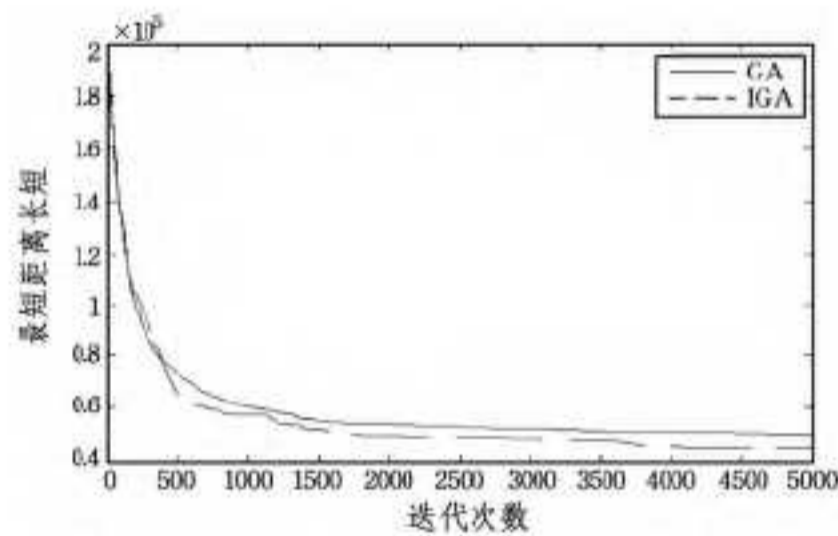


图 2 测试 CHN144 实例中 3 组和 4 组参数迭代过程的比较

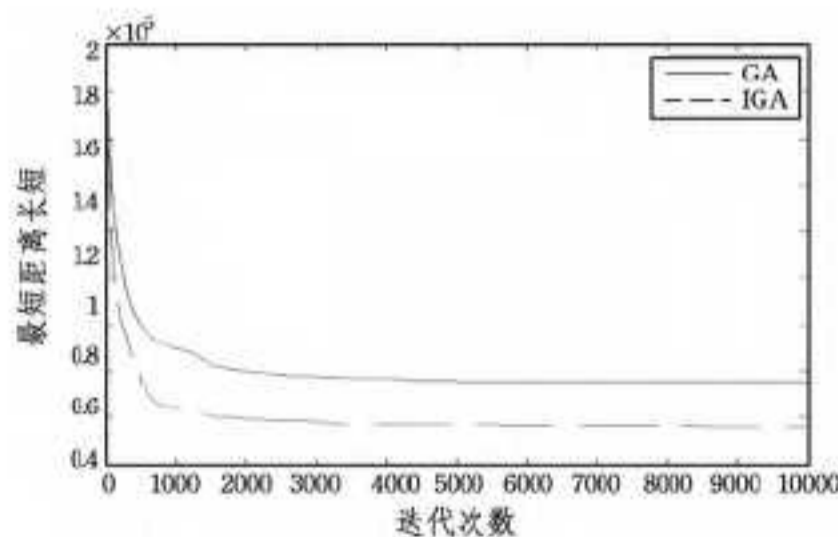


图 3 测试 CHN144 实例中 5 组和 6 组参数迭代过程的比较

4.3 结果分析

由表 2 可知,当交叉概率为 0.88,变异概率为 0.07 时,通过分别比较数据 1 和数据 2、数据 3 和数据 4、数据 5 和数据 6 的最优结果发现,当种群容量和遗传代数一定时,改进遗传算

法的结果要优于基本遗传算法,通过比较数据 2 和数据 3,当种群容量为 100 时,虽然改进遗传算法的遗传代数要少于基本遗传算法,但它的结果仍然优于基本遗传算法,通过比较数据 3 和数据 4,当遗传代数为 5000 时,改进遗传算法的种群容量大于基本遗传算法时,它的最优解比基本遗传算法要好得多。总的来说,要使算法得到较好解,需要足够的遗传代数,为了保持种群的多样性,需要较大的种群规模。从图 1—图 3 可以看出,在算法的初期,最短路径长度下降得比原算法快,说明改进交叉算子比原交叉算子作用明显,而后期算法趋于平缓,但仍有波动,说明基于相似度的种群更新起了作用。

结束语 在基本遗传算法的基础上,在选择算子中引入选择因子,同时提出一种改进的交叉算子以加快算法的收敛速度。在加快收敛的同时为了防止种群多样性的缺失,在算法中又加入了基于种群相似度的浓度更新策略,有效地防止了“早熟”。通过 CHN144 城市的实验说明,改进遗传算法比基本遗传算法具有更好的稳定性,收敛速度更快,表明对基本遗传算法的相关改进操作是有效可行的。对于更大规模的城市来说,改进遗传算法比基本遗传算法更有效。

参考文献

- [1] 周明,孙树栋. 遗传算法原理及应用[M]. 北京:国防出版社,1999:60-80
- [2] 袁满,刘耀林. 基于多智体遗传算法的土地利用优化配置[J]. 农业工程学报,2014,30(1):191-199
- [3] 门慧勇. 基于遗传算法的图像分割优化研究[D]. 长春:东北师范大学,2012
- [4] 陈侠. 基于改进的遗传算法的网络编码优化方法研究[D]. 武汉:华中科技大学,2012
- [5] 周康,强小利,同小军,等. 求解 TSP 算法[J]. 计算机工程与应用,2007,43(29):43-47
- [6] 赵颂华. 城市公共资源监管设计新思维[J]. 科技资讯,2015,15:52-55
- [7] Cheang B, Gao Xiang, Lim A, et al. Multiple pickup and delivery traveling salesman problem with last-in-first-out loading and distance constraints[J]. European Journal of Operational Research, 2012, 223(1): 60-75
- [8] Nagata Y, Soler D. A new genetic algorithm for the asymmetric traveling salesman problem[J]. Expert Systems with Applications, 2012, 39(10): 8947-8953
- [9] 周永权,黄正新,刘洪霞. 求解 TSP 问题的离散型萤火虫群优化算法[J]. 电子学报,2012,40(6):1164-1170
- [10] 刘全,王晓燕,傅启明,等. 双精英协同进化遗传算法[J]. 软件学报,2012,23(4):765-775
- [11] 姚明海,王娜,赵连朋. 改进的模拟退火算法和遗传算法解决 TSP 问题[J]. 计算机工程与应用,2013,49(14):60-65
- [12] 谢燕丽,许春林,姜文超. 一种基于交叉和变异算子改进的遗传算法研究[J]. 计算机应用与发展,2014,24(14):80-83
- [13] 汪金刚,罗辞勇. 求解 TSP 问题的自适应领域遗传算法[J]. 2010,46(27):20-24