

TSP 的邻域搜索算法的分析和改进^①

赵 赫 杜端甫

(北京航空航天大学管理学院, 100083)

摘要 本文分析了 TSP 的邻域搜索算法 (K-OPT) 的主要特点和一些不足之处, 并在此基础上提出了一个较为有效的并且具有较低的算法复杂度的启发式算法。

关键词: TSP, K-OPT, 启发式算法

1. 引言:

TSP 是近年来被广泛研究的众多组合最优化难题之一, 在众多的领域当中有着广泛的应用 (比如说: 流水车间的零件排序, 公共汽车的巡回路线, 航班机组人员的安排, 机床的布局 [1], NP 难度问题和 NPC 问题的探讨 [3] 等等), 长期以来一直被应用数学工作者、计算机科学界所广泛的关注。关于 TSP 的定义可以参见 [1] [3]。目前 TSP 已经有很多算法, 这些算法大致分成两大类, 一种是启发式算法, 另一种是最优化算法。一般而言, 最优化算法所解决问题的规模被限制在有限的结点数范围之内, 使得在实际应用中问题的最优解常常无法得到 (譬如说: 机床上的零件排序问题一般都是几十个至上百个); 而对于启发式算法来说, 问题的规模一般没有什么限制, 但是该算法在大多数的情况下得不到最优解, 只能是近似最优解, 随着问题规模的增大, 由于最优解得不出, 因而, 也很难判断所得到的近似解的好坏。但是就目前的研究来看, 学者们主要还是将精力放在启发式算法的评估上面, 也就是说对 TSP 进行分类研究, 并且设定一些评判的规则 [4] [5] [6]。从研究成果来看 K-OPT 方法在大多数场合下还是比较有效的算法, 本文主要也讨论的是这种邻域搜索算法。

邻域搜索算法最早在 60 年代提出 [3], 在许多文献当中又被称为 K-OPT。有关该算法的详细迭代步骤, 请参见 [1] [3], 本文主要是分析较常用的几种邻域搜索算法, 并在此基础上提出一个较为有效的改进算法。

2. 对邻域搜索算法的分析:

2.1 分析的意义:

邻域搜索算法作为一种比较“老”的算法, 虽然在算法的思想上不如近年来提出的一些算法的思想新颖 (如: 与生物学理论相结合的遗传算法, 与热学相结合的模拟退火算法等等)。但是

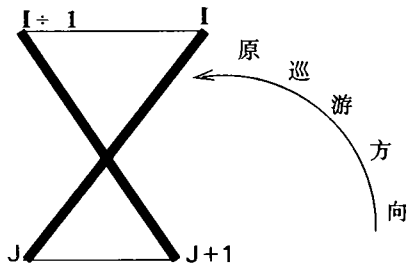
^① 本文 1996 年 10 月 29 日收到。

这些算法常常要与邻域搜索算法配合起来一道来寻优, 否则的话就无法起到相应的寻优效果(2)。从上述意义上讲, 对该算法的研究就显的很有意义了。

在该算法当中常用的是 2-OPT 算法。2-OPT 算法的算法渐进时间复杂度是 $O(N^2)$ (3), 因其具有较低的复杂度, 因而本文主要对该算法进行分析和改进。

2.2 对 2-OPT 算法的分析:

2-OPT 算法的基本操作思想如下图所示 (K-OPT 方法与此同理):



如果 $W(I, I+1) + W(J, J+1) > W(I, J) + W(I+1, J+1)$

则用两条邻域边 $I-J$ 和 $I+1-J+1$ 替换 $I-I+1, J-J+1$

得到新的巡游方向: $I-J-*****-I+1-J+1-***I$

2-OPT 算法由于是每两条边的考虑来实施回路中的替代, 考虑的邻域中的边数较少 (对于一个无向完全图而言, 任一条边的邻边数有 $N-1$ 条), 因而往往容易落入局部最优。

如: 假设有一个对称的 TSP, 其路径权值矩阵为:

$$\begin{pmatrix} *, 7648, 1054, 6134, 9371, 1073 \\ 7648, *, 1084, 7627, 2643, 535 \\ 1054, 1084, *, 7419, 6964, 2106 \\ 6134, 7627, 7419, *, 2227, 4601 \\ 9371, 2643, 6964, 2227, *, 1145 \\ 1073, 535, 2106, 4601, 1145, * \end{pmatrix}$$

该问题用 LITTLE 算法 (1) 求解得到最优解为: $LMIN=12179$, 路: 1, 3, 2, 6, 5, 4, 1。用 2-OPT 算法求解, 初始解为:

1, 4, 6, 5, 2, 3, 1。迭代的最终结果为: $L=12682$, 路:

1, 6, 4, 5, 2, 3, 1。在开始执行 2-OPT 算法的时候, 因为 W_{1-6} 与 W_{4-5} 的和为 $3300 < W_{1-4} + W_{5-6} = 7279$, 根据 2-OPT 替换法则, 1-4 边与 5-6 边替入回路, 1-6 边与 4-5 边退出回路, 回路此时的长度为 12682, 继续进行, 没有符合规则的边可以替入回路, 算法终止。但是假如我们不急于替换, 而是继续扫描初始回路, 我们发现边 4-5 与边 2-6 的和为 $2227 + 535 = 2762 <$ 边 4-6 与边 2-5 之和 (7244), 用边 4-6 与边 2-5 替换边 4-5 与边 2-6 则得回路为: 1, 3, 2, 6, 5, 4, 1, 此时路长为 12179, 得到了最优解; 又假定我们选择的初始回路为:

4, 6, 5, 2, 1, 3, 4, 则用 2-OPT 方法可以得到最优解。

从而我们可以看出 2-OPT 方法的特点: 该方法易于落入局部最优; 且与初始条件有关。正是由于这些特点, 所以本文在后面提出了一个改进的算法。

类似的分析可以得到结论: 一般而言 K-OPT 方法的最大缺陷就是由于对邻域边搜索的不

全面, 从而造成局部的收敛。那么自然想到, 能否使用较大的 K 值来增强搜索的效果呢? 实际上, 从大量的实验可以发现 3-OPT 方法确实比 2-OPT 方法有了明显的改进, 而且它的算法时间复杂度是 $O(N^3)$, 亦是接受的; 但进一步增大 K 值, 时间上增大了许多 (对于一个 30 个结点的中型 TSP, 若令 $K=10$, 则对于一个每秒可运算一千万次的计算机而言, 算法要执行的时间约为 1.8 年), 但是解的改进不大 (3)。经分析我们认为是如下原因造成的: 虽然增大了 K 值使得考查了更多的邻域边, 然而 K -OPT 方法的规则是要求非邻域中的 K 条边之和必须要小于邻域中的相应的 K 条边之和, 同时可以将邻域中的 K 条边替换出来并且形成相应的回路, 换句话说, 假如对于邻域中的三条边, 若其满足 3-OPT 的替换条件的话, 也可能引入了第四条邻域中的边时, 用 4-OPT 方法来衡量时不合要求了 (可能这四条边的和大于相应的四条边的和; 或虽然这四条边的和小于相应的四条边的和, 但是有可能该四边无法于邻域中的其余边形成相应的回路), 所以如果邻域中的边很多 (相当于增大 N), 那么就可能找到相应的满足 4-OPT 条件的第四条边。可见: K -OPT 方法有效的条件是 (当 K 较大时譬如 K 大于 4), N 应足够大。至于对应于每一种 K , 究竟多大的 N 值更加有效, 还须大量的实验。

3. 对邻域搜索算法的改进

本文提出一个基于 2-OPT 算法的改进算法: 即将 2-OPT 方法第一次的求解结果记录下来 (以字符串的形式), 随机的对串中的路线顺序进行调换 (如: 1,2,5,4,3,1 \rightarrow 5,2,1,4,3,5), 即类似于遗传算法的变异算子作用机理 (2), 得到新的字符串 (对应与新的路径), 在对此路很用 2-OPT 算法来进行改进, 将得到的结果再次记录下来重复上面的“变异”和寻优过程, 每次都将会用 2-OPT 寻优的结果记录, 最后将几次的结果比较得到一个满意解。算法的过程可以如下描述:

```

PROCEDURE NEW-OPT
  PROCEDURE INITIALIZE(初始化)
    FOR I=1 TO LP
      PROCEDURE 2-OPT(寻优)
      PROCEDURE RECORD(记录结果)
      PROCEDURE MUTATION(变异)
    NEXT I
  PROCEDURE COMPARE(比较输出结果)
END PROCEDURE

```

之所以这样做得理由是: 因为前文分析了 2-OPT 方法的特点, 因而此处用变异的方法打破局部最优, 同时因为 2-OPT 方法与初始条件关系较大, 因而每次从不同的初始条件出发, 必定能得到差异较明显的解, 我们又知道, 2-OPT 方法的算法时间复杂度为 $O(N^3)$, 假定我们一共用 2-OPT 方法求解 C 次 ($C < N$), 而每求解一次包括两步: 即是 2-OPT + 变异, 变异一次可以在 $O(N)$ 完成, 那么每次可在 $O(N^3)$ 完成, 总时间为 $C \cdot O(N^3)$, 又 $C < N$, 所以最终算法的时间复杂度为 $O(N^3)$, 即不增加算法的复杂度而对算法进行了改进。

下面看一下该方法与 2-OPT 方法和 3-OPT 算法对于中小型的 TSP ($N=20$) 和中型 TSP ($N=30$) 一共 20 个算例进行比较的结果:

编 号	规 模	最 优	2-OPT	3-OPT	改进方法
1	30	3883	5046	4079	3902
2	30	2706	3612	2858	2914
3	30	3607	4515	4234	3984
4	30	3304	5535	4032	4041
5	30	3341	4817	4326	3736
6	30	2792	4281	3081	3225
7	30	3375	4800	3484	3410
8	30	3192	4646	3646	3304
9	30	3963	5236	4634	4150
10	30	3188	3299	3390	3299
11	20	2281	2434	2666	2281
12	20	1819	1858	1942	1858
13	20	1914	2179	2002	2179
14	20	2212	2516	2375	2212
15	20	1193	2002	1193	1228
16	20	2153	2302	2435	2244
17	20	2188	2653	2352	2199
18	20	1526	1704	2133	1548
19	20	1493	1734	1558	1493
20	20	1259	1901	1459	1334

从以上的结果可以看出:

1. 改进之后的算法每次的运算结果都不劣与 2-OPT 算法所得到的结果, 而且改进的程度还比较的显著, 相对于最优解而言, 也是比较的接近, (平均误差为 5.2%);

2. 改进之后的算法复杂度低于 3-OPT 算法的复杂度 ($O(N*N*N)$), 但是 3-OPT 算法的运算平均误差为 (12.78%), 大于改进算法的误差;

3. 改进算法在处理 30 个结点的问题时相对于 3OPT 的结果不如在处理 20 个结点时所得到的结果好, 因为它毕竟是以 2-OPT 算法为基础的, 因而它也有 2-OPT 算法的缺点 (每次收敛的过早), 当然我们也可以类似地构造 K-OPT 算法的改进算法, 但是由于 K-OPT ($K \geq 2$) 对解的初始情况反应不敏感, 也就是说, 不同的初始解所得到的结果差别很小, 因而用本文中提到的方法来改进意义就不是很大。

4. 总结:

本文首先分析了传统的邻域搜索算法中较常用的 2-OPT 算法和 3-OPT 算法, 并且分析了该算法的不足之处, 即容易过早的收敛到局部最优。传统的算法需要不断的和新思想相结合, 比如说: 将其与遗传算法相结合, 事实证明, 将邻域搜索算法与其他算法相结合还是比较成功的, 因而对该算法进行适当的改进, 在不增加算法复杂度的基础上, 提高算法的寻优率还是一件比较有意义的工作, 可以预测将来好的启发式算法必然是多种算法的混合体。

参考文献

- (1) 杜端甫, 运筹图论, 北航出版社, 1990
- (2) 刘勇, 康立山等, 非数值并行算法 (第二册), 遗传算法, 武汉大学出版社, 1995
- (3) [美] PAPADIMITRIOU, K. STEIGLITZ 著, 刘振宏, 蔡茂诚译, 组合最优化算法和复杂性分析, 清华大学出版社, 1985
- (4) BUDDHADEV ROYCHOUDHURY, JOHN F. MUTH, THE SOLUTION OF TRAVELLING SALESMAN PROBLEMS BASED ON INDUSTRIAL DATA, JOURNAL OF OPERATIONANL RESEARCH SOCITY, 37-353, 1995
- (5) K. NURMI, TRAVELLING SALESMAN PROBLEM TOOLS FOR MICROCOMPUTERS, COMPUTERS OPS RES VOL 18. PP 741-749, 1991
- (6) H. L. ONG AND H. C. HUANG ASYMPTOTIC EXPECTED PERFORMANCE OF SOME TSP HEURISTICS: AN EMPITICAL EVALUATION, EUROPEAN JOURNAL OF OPERATIONAL RESEARCH 43 (1989) 231-238

Analysis and Improvement of the Neighbouring Search Algorithm for FSP

Zhao He Du Duanfu

(Management School of Beijing University of Aero. & Astr)

Abstract This paper analyses the features of neighbor area search algorithms (K-OPT), and its shortcoming. Based on these, We proposed an effcent and lower algorithms complexity herustic method.

Key words: TSP, K-OPT, herustic