

DOI: 10.11992/tis.201706061

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20180423.0946.002.html>

遗传算法求解多旅行商问题的相对解空间分析

赵新超, 郭赛

(北京邮电大学理学院, 北京 100876)

摘要: 首先介绍了多旅行商问题的模型, 并指出遗传算法解决多旅行商问题的关键是染色体编码方案的设计, 为了减少冗余解带来的代价, 本文给出了传统的两种染色体编码方案 (单染色体和双染色体), 以及最新的两段式染色体编码方案; 接着引入相对解空间概念, 以此定量地给出不同染色体方案对应解空间的相对大小关系; 基于相对解空间概念, 本文分析了 3 种染色体编码方案对应的解空间在极限意义下的相对大小关系, 并分析了旅行商数与城市数在不同情形下解空间的近似相对大小关系。本文对搜索空间定量分析的理论结果对工程问题的求解提供了科学的指导意义。

关键词: 多旅行商问题; 遗传算法; 染色体编码; 相对解空间; Stirling 公式

中图分类号: O224; TP18 **文献标志码:** A **文章编号:** 1673-4785(2018)05-0760-09

中文引用格式: 赵新超, 郭赛. 遗传算法求解多旅行商问题的相对解空间分析[J]. 智能系统学报, 2018, 13(5): 760-768.

英文引用格式: ZHAO Xinchao, GUO Sai. Analysis on the relative solution space for MTSP with genetic algorithm[J]. CAAI transactions on intelligent systems, 2018, 13(5): 760-768.

Analysis on the relative solution space for MTSP with genetic algorithm

ZHAO Xinchao, GUO Sai

(School of Science, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: This paper introduces the concept of multiple traveling salespersons problem (MTSP) and a chromosome encoding design method for solving the MTSP using genetic algorithm. In order to reduce the cost of redundant solution, two traditional chromosome design methods (single and double chromosome designs) are proposed, as well as the current two-part chromosome encoding design. Then the concept of relative solution space is introduced to quantitatively compare the relative size order of spaces for different solutions. Then on based on the relative solution space, the relative size order of the solution spaces of three chromosome-encoding designs are analyzed under an infinite limit. Subsequently, the relative size order of the solution spaces is analyzed individually when the relationship presents different situations between the number of travelers and the number of cities. The theoretical results of the quantitative analysis on the search space given in this paper are of great significance in guiding engineering applications and solving practical problems.

Keywords: multiple traveling salespersons problem; genetic algorithm; chromosome encoding; relative solution space; Stirling formula

多旅行商问题 (multiple traveling salesman problem, MTSP) 可直观描述为一个旅行商团队要分头遍历若干个城市, 每个城市至少被一个旅行

商访问一次且只访问一次, 如何安排 m (大于 1) 位旅行商遍历 n (大于 m) 个城市, 使得总的访问行程最小^[1]。该问题最常见的应用领域是车间调度领域, 在生产线上的作业调度通常被建模为一个旅行商问题 (traveling salesman problem, TSP)。如果生产经营扩展到有多条平行线, 工作可以分配,

收稿日期: 2017-06-17. 网络出版日期: 2018-04-23.

基金项目: 国家自然科学基金项目 (61375066, 11671052, 71772060).

通信作者: 郭赛.E-mail: saisaismily@163.com.

这个问题可以建模为一个多旅行商 MTSP^[2]; 另一个经常被建模为多旅行商问题的就是车辆调度问题 (VSP)^[3]。车辆调度问题是指对一些车辆进行调度, 所有的车辆离开并回到原地点, 使得每个地方只能而且必须被访问一次。

由于多旅行商问题的复杂性, 必须根据实际问题的大小采用启发式解决方案^[4]。遗传算法 (genetic algorithms, GA)^[5]是一种基于生物自然选择和基因遗传学原理的高效并行、随机全局优化搜索算法。Keshavarz 等^[6]学者发现遗传算法对调度问题有很强的适应性, MTSP 有效的遗传算子也激发了研究者的极大兴趣; 张永库等^[7]研究了改进的遗传算法在模糊聚类中的表现; Katayama 等^[8]提出的混合变异遗传算法在旅行商问题中的应用提高了遗传算法解决旅行商问题的性能。多旅行商问题在实际问题中有很广泛的应用。Huang 等^[9]研究了多旅行推销员问题在热轧规划中的应用; 江贺等^[10]研究了近年来出现的新 NP-难解问题, 即黑白旅行商问题 (BWTSP), 给出了遗传算法解决问题的一种新思想; Triguí 等^[11]提出一种解决多机器人系统多目标多旅行推销员问题的模糊逻辑方法。最近出现了多种解决 TSP 问题的方法, 王宇平等^[12]提出的求解 TSP 的量子遗传算法使用了有关量子方面的知识。研究者们也在其他经典优化算法^[13]中找到了许多新思想。

遗传算法求解 MTSP 的关键是要根据问题设计一种好的染色体编码方法, 而好的遗传算法染色体编码方案应该能够从候选解中减少或消除冗余的解。冗余解是指以一种以上的方式表示同一染色体, 并多次出现在种群中的染色体编码方式。相同解的多次再现不仅增加了查找空间, 而且降低了查找效率。

本文首先列出了传统的两个染色体方案 (单染色体设计方案和双染色体方案), 以及最近提出的较新的两段式染色体设计方案; 本文引入相对解空间概念, 以此定量地衡量不同染色体方案给出的相对解空间大小关系, 先给出对旅行商数目和城市数目在趋于无穷时的极限相对大小关系, 接下来近似分析了旅行商数目与城市数目在不同情形下解空间的相对大小关系。

1 3 种染色体方案

MTSP 染色体的设计主要有以下 3 种^[1]。

1.1 单染色体设计

图 1 描述了 MTSP 问题城市数 $n=15$ 、旅行商

数 $m=4$ 时染色体编码的一种设计方法的一个实例。这种设计使用了一个染色体编码长度为 $n+m-1$ 的染色体, 因此被称为“单染色体”(one chromosome) 设计。该设计方案中, n 个城市用从 $1\sim n$ 正整数编码表示, 并进行无规则的排列, 代表着 n 个城市的整数编码列被 $-1\sim-(m-1)$ 这 $m-1$ 个负整数分为 m 段, 分别依次对应着有序的 m 个旅行商所访问的城市, 因此这个长度为 $n+m-1$ 的整数列的任何一种排列组合都可能是这个 MTSP 问题的解。以图 1 为例, 第一个旅行商将访问编号为 2、5、14、6 的这几座城市, 第二个旅行商将访问编号为 1、11、8、13 的这几座城市, 依次类推。在这种染色体设计中, MTSP 所有问题的解将可能是由 $(n+m-1)!$ 个解构成的一个空间。这些解中有很多是冗余的, 例如, 旅行商 1 与旅行商 2 的访问城市全部按顺序相互调换位置后的解和现有解。

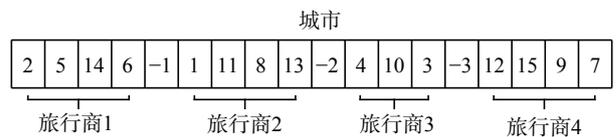


图 1 单染色体方案

Fig. 1 One chromosome

1.2 双染色体设计

图 2 描述了同样一个 MTSP 问题由遗传算法中染色体编码方案的另一种设计方法。一般称它为“双染色体 (two chromosomes)”设计。这种设计使用两个长度为 n 的染色体表示一个解。图 2 中, 第一个染色体表示 n 个城市的一组排列, 第二个染色体的每一位数对应上面相应的城市, 由对应编码的旅行商来访问。例如, 旅行商 1 访问的城市为 5、14、10、15, 旅行商 2 访问的城市为 2、8、12、9。使用这种编码设计, 对应的空间是由 $n!m^n$ 个可能的解构成的一个集合。同样, 这种编码方案也有许多冗余的解。例如, 上述事例解中头两个基因位交换位置后得到的解与现有解就是相同的。



图 2 双染色体方案

Fig. 2 Two chromosomes

1.3 两段式染色体设计

图 3 描述了一个新的染色体设计方案, 它由前后两个部分组成, 因此称为“两段式染色体 (two-part chromosome)”设计。前段是整数 $1\sim n$

的一个排列,代表了 n 个城市的一个顺序排列;后段长度为 m ,按顺序分别表示 m 个旅行商在前段编码中对应访问的城市数,并且这 m 个数之和等于 n 。图 3 所示旅行商 1 访问前段中的头 4 个城市顺序分别为 2、5、14、6;旅行商 3 访问从第 4+4+1 个开始的连续 3 个城市,分别为 4、10、3。使用新的两段式染色体设计,不仅减少了查找空间,而且由于固定了旅行商的顺序而消除了部分(但不是全部的)冗余的解。使用这种染色体编码设计,对于前段可能会有 $n!$ 种排列,后段由于是一个和为 n 的正整数向量,因此需要 C_{n-1}^{m-1} 种可能的 m 维正整数表示,才能满足要求。从而,可以得到解空间大小为 C_{n-1}^{m-1} 。

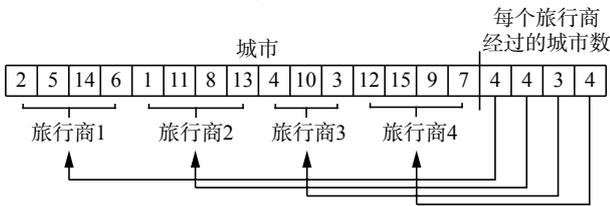


图 3 两段式染色体方案
Fig. 3 Two-part chromosome

2 3 种染色体编码方案下的相对解空间分析

上一节介绍了遗传算法求解多旅行商问题的 3 种染色体编码方案。文献[1]给出了 3 种染色体编码方案对应的搜索空间的大小,仅是定性分析了 3 种方案的优劣,即两段式染色体方案对应的解空间优于前两种编码方案对应的解空间,而第一种单染色体设计方案优于双染色体编码方案。如果只是定性地了解 3 种编码方案对应解空间大小关系,对染色体编码方案和算法设计以及实际工程应用并没有太大的指导意义。因此本文首先引入相对解空间概念,然后详细地定量分析了 3 种染色体方案对应的解空间相对大小,这对多旅行商问题的研究和求解具有现实的指导意义。

2.1 相对解空间

3 种染色体编码方案对应的解空间大小^[1]分别为

$$C_{\text{one}} = (m+n-1)! \quad (1)$$

$$C_{\text{two}} = n!m^n \quad (2)$$

$$C_{2\text{-part}} = n!C_{n-1}^{m-1} \quad (3)$$

式中: C_{one} 表示单染色体编码对应的解空间规模, C_{two} 表示双染色体编码对应的解空间规模, $C_{2\text{-part}}$ 表示两段式染色体解空间的规模。从 3 种编码方案解空间的代数表达式可以看出,它们都是城市数 n 和旅行商数目 m 的二元函数,自变量 n 和

m 的取值范围是正整数,且一般 $m < n$ (旅行商人数小于城市数)。

式 (1) ~ (3) 按大小排序有: $C_{2\text{-part}} < C_{\text{one}} < C_{\text{two}}$ ^[1] ($n > m$),这几个代数式给出的解空间大小从数量上来讲是对解空间的绝对衡量,假若将其两两作比值运算(较大的作分母),得到的结果是个相对值,当 m 和 n 取较大的正整数时,比值结果是一个小于 1 的真分数,我们称之为“相对解空间”,分别用 C_{31} 、 C_{32} 和 C_{12} 表示,即

$$C_{31} = \frac{C_{2\text{-part}}}{C_{\text{one}}} = \frac{n!C_{n-1}^{m-1}}{(m+n-1)!} \quad (4)$$

$$C_{32} = \frac{C_{2\text{-part}}}{C_{\text{two}}} = \frac{n!C_{n-1}^{m-1}}{n!m^n} \quad (5)$$

$$C_{12} = \frac{C_{\text{one}}}{C_{\text{two}}} = \frac{(m+n-1)!}{n!m^n} \quad (6)$$

2.2 相对解空间的极限分析

在现实的工程应用中,比如车辆调度问题, m 、 n 一般取值较大,因此随着自变量的增大或问题规模的扩大,我们需要了解相对解空间对应的比值结果的变化趋势,从而从数学表达式中讨论 m 和 n 取较大正整数时,相对解空间的极限行为,也可以认为是讨论 m 、 n 趋于无穷时的极限。

2.2.1 C_{31} 的极限分析

1) m 适当大时

在多旅行商问题中,城市数目与旅行商数目之间的差距一般都比较大,因此当讨论 C_{31} 的极限性质时,不妨假设 $n \gg m$,当 n 充分大、 m 适当大时,对 C_{31} 取对数,分析可得:

$$\begin{aligned} \ln C_{31} &= \sum_{k=1}^n \ln k + \sum_{k=1}^{n-1} \ln k - \sum_{k=1}^{m+n-1} \ln k - \sum_{k=1}^{n-m} \ln k - \\ &\sum_{k=1}^{m-1} \ln k = \sum_{k=1}^n \ln k + \sum_{k=1}^{n-1} \ln k - \left(\sum_{k=1}^{n-1} \ln k + \sum_{k=n}^{n+m-1} \ln k \right) - \\ &\left(\sum_{k=1}^n \ln k - \sum_{k=n-m+1}^n \ln k \right) - \sum_{k=1}^{m-1} \ln k = - \sum_{k=n}^{n+m-1} \ln k + \\ &\sum_{k=n-m+1}^n \ln k - \sum_{k=1}^{m-1} \ln k < \sum_{k=1}^{m-1} \ln k < 0 \\ C_{31} &= e^{-\ln(m-1)!} = \frac{1}{(m-1)!} \end{aligned}$$

所以 $C_{31} < 1$ 。

$C_{31} < 1$ 意味着第三种编码方案对应的解空间严格小于第一种方案对应的解空间,这严格证明了 Carter^[1] 给出的解空间的定性比较结果。当 n 足够大时,所需的旅行商数也有相应增大,因此当 n 充分大、 m 适当大时, $C_{31} \ll 1$ 。例如: $m=10$, $C_{31} < e^{-\sum_{k=1}^9 \ln k} \approx 2.76 \times 10^{-6}$ 。

2) m 是不大的常数时

$(n-1)!$ 可与 $(n+m-1)!$ 相抵消, $n!$ 可与 $(n-m)!$ 相抵消, 所以当 n 趋于无穷时得出:

$$C_{31} = \frac{C_{t-part}}{C_{one}} = \frac{n!n-1!}{(n+m-1)!(n-m)!(m-1)!} \rightarrow \frac{1}{(m-1)!}$$

即当旅行商人数 m 取不大的常数时, 随着城市数目的增加, 单染色体编码方案对应的空间约是两段式染色体方案解空间的 $(m-1)!$ 倍, 此结果与两种解空间的极限分析是一致的。

2.2.2 C_{32} 的极限分析

1) m 适当大时

依然假设 $n \gg m$, 当 n 充分大、 m 适当大时, 取对数分析可得:

$$\ln C_{32} = \sum_{k=1}^{n-1} \ln k - \sum_{k=1}^{n-m} \ln k - n \ln m - \sum_{k=1}^{m-1} \ln k = \sum_{k=n-m+1}^{n-1} \ln k - n \ln m - \sum_{k=1}^{m-1} \ln k$$

$\sum_{k=n-m+1}^{n-1} \ln k$ 中有 $m-1$ 项, 通项是城市数 n 的对数, 而 $n \ln m$ 中有 n 项, $\ln m$ 是旅行商数目的对数, 并且当 n 充分大时, $\lim_{n \rightarrow \infty} \frac{n}{\ln n} = +\infty$ ^[14], 所以

当 $n \rightarrow \infty$ 时, $\sum_{k=n-m+1}^{n-1} \ln k - n \ln m < 0$ 。因为

$$\sum_{k=n-m+1}^{n-1} \ln k - n \ln m < (m-1) \cdot \ln(n-1) - n \ln m < m \ln n - n \ln m < 0$$

所以, $\ln C_{32} < -\sum_{k=1}^{m-1} \ln k < 0$, 即 $C_{32} < 1$ 。

$C_{32} < 1$ 同样意味着第三种编码方案对应的解空间小于第二种方案对应的解空间。同 C_{31} 的分析类似, 当 n 足够大、 m 适当大时, $C_{32} \ll 1$ 。

2) m 是不大的常数时

$$C_{32} = \frac{C_{t-part}}{C_{two}} = \frac{n!C_{n-1}^{m-1}}{n!m^n} = \frac{(n-1)!}{(m-1)!(n-m)!m^n} = \frac{(n-1)(n-2)\cdots(n-m+1)}{(m-1)!m^n} < \frac{n^m}{(m-1)!m^n}$$

由此可知, $C_{32} < \frac{1}{(m-1)!}$ 。

由对 C_{32} 的分析可知, 当 m 取大于 2 的常数时, $C_{32} < 1$ 。此结果显示, 第三种编码方案对应的解空间小于第二种编码方案对应的解空间, 此结果与两种解空间的极限分析是一致的。

2.2.3 C_{12} 的极限分析

假设 $n \gg m$, 当 n 充分大、 m 适当大时, 对 C_{12} 取对数, 即

$$\ln C_{12} = \sum_{k=1}^{n+m-1} \ln k - \sum_{k=1}^n \ln k - n \ln m = \sum_{k=n+1}^{n+m-1} \ln k - n \ln m < (m-1) \ln(n+m-1) - n \ln m$$

当 n 充分大时, $\lim_{n \rightarrow \infty} \frac{n}{\ln n} = +\infty$, 因此 $\ln C_{12} < 0$,

$C_{12} < 1$ 。 C_{12} 的大小说明第一种编码方案对应的解空间小于第二种编码方案对应的解空间。

至此, 在合理适当的条件下: n 适当大, 对 3 种解空间的相对大小给出详细严密的论证, 很好地补充了文献[1]中定性的结论, 即 $C_{two} > C_{one} > C_{2-part}$, 而且由分析可知, 当 n 充分大、 m 适当大时, $C_{two} \gg C_{one} \gg C_{2-part}$ 。

以上分析中, 对一般的 m 和 n 给出的是相对解空间理论上的对比, 但是如果只知道相对大小, 对实际问题应用几乎没有实质性的帮助, 例如求 C_{31} 的极限, 实质上 m 为常数, n 逐渐增大, 也就是说, 在某一个多旅行商问题中, 我们保持商人的数目不变, 增大工程的规模, 让城市数目逐渐增大, 这也就意味着单个商人访问的城市数目随着城市数目的增多而无限增加下去, 这显然不符合实际情况, 因为每个旅行商访问的城市数目是有上限的, 即最大负荷能力。因此单一用这个极限结果去估计解空间的大小, 虽然能得出严密的定量分析结果, 但是其代表的实际意义是不乏片面性的。

我们不仅需要知道染色体编码方案 C_{2-part} 对应的解空间小, 更重要的是需要知道到底小多少, 这也是本文最初的研究目的。通过研究发现, 对一般性的 m 和 n , 很难给出解空间的具体差距, 因此本文在旅行商数 m 与城市数 n 之间成几种典型函数关系的条件下给出相对解空间的差距, 包括城市数 n 与旅行商数 m 的线性关系、幂函数关系和指数关系。

2.3 相对搜索空间的粗略估计

图 4 描述了 15 个城市, 是旅行商数量从 1 增加到 14 时, 3 种染色体编码设计的搜索空间的变化示意图。图 4 的纵轴表示搜索空间的数量级。由图 4 可以看出, 在旅行商逐步增加的情况下, 单染色体编码对应的搜索空间数量的增加由快到慢; 双染色体编码对应的搜索空间数量的增加比较平稳; 而两段式染色体编码对应的搜索空间数量呈现两端相当、中间略微增加的大体对称的情况。因此图 4 与 2.2 节的解空间分析结果显示出了两段式染色体编码设计的初步优越性。

众所周知, 当城市数目逐步增加时, 旅行商人数也应该相应增加, 而 m 一般会随 n 的变化而变化。以下详细讨论 m 与 n 有特定关系时的相对搜索空间。例如: $n=10m$, $n=m^2$, $n=e^m$ 三种情形下, 相对搜索空间的变化趋势。

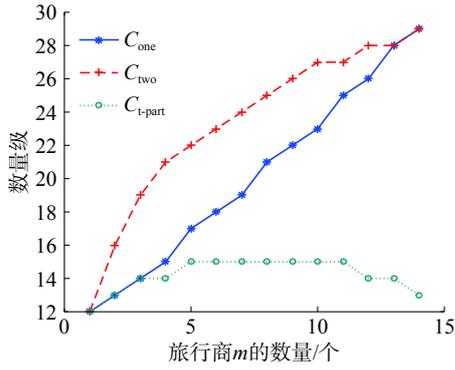


图 4 解空间变化示例

Fig. 4 Example of solution space change

从图 5~7 所示的 3 组函数图像可以看出, 3 种情形下的相对搜索空间变化, 总体来说, 从相对搜索空间上比较, 在 $n=10m$ 和 $n=m^2$ 情形下,

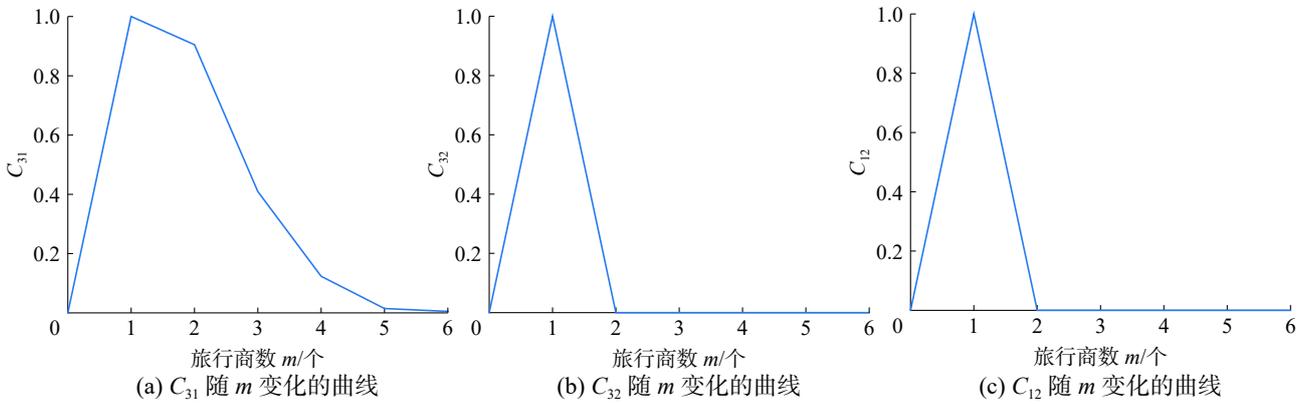


图 5 $n=10m$ 情形下的相对搜索空间变化曲线图

Fig. 5 Solution space change in the case of $n=10m$

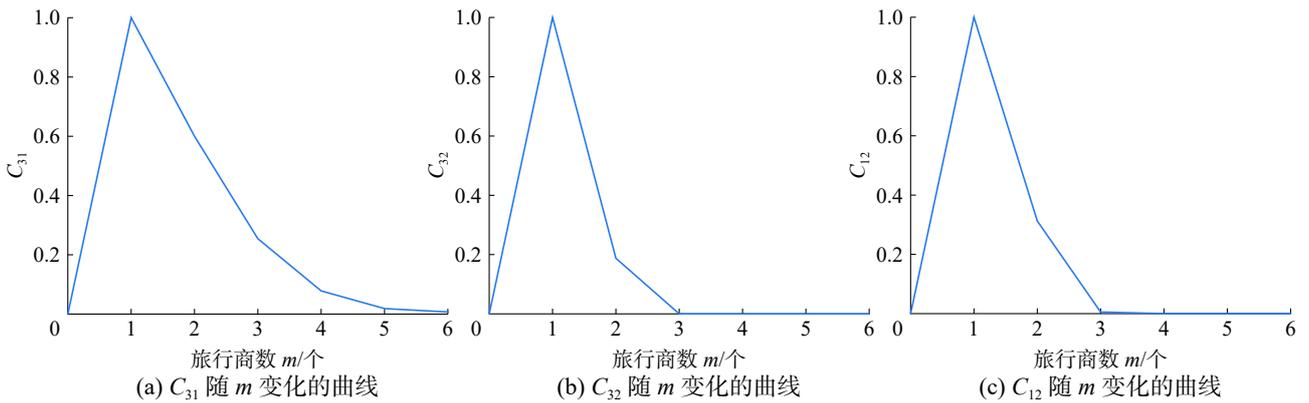


图 6 $n=m^2$ 情形下的相对搜索空间变化曲线图

Fig. 6 Solution space change in the case of $n=m^2$

2.4 分情形下的相对解空间大小分析

2.4.1 Stirling 公式

相对解空间表达式中出现阶乘项, 增加了相应的分析和对比运算, 而我们在此要讨论的是 m, n 取较大正整数时相对解空间的变化属性, 因此本文用 Stirling 公式来近似化简分式^[14](当 n 取充分大的正整数时, 用多项式代替阶乘运算), 图 8

C_{31} 和 C_{32} 在前期的差别不大, 纵轴随着 m 的增大很快收敛, 且 C_{12} 在 $n=m^2$ 情形下比在 $n=10m$ 情形下收敛得更快些; 在 $n=e^m$ 情形下 C_{32} 相对搜索空间和 C_{12} 相对搜索空间比其他两种情形收敛更快, 但 C_{31} 在前期收敛得有些缓慢, 而在 $n=e^m$ 情形下明显比其他两种情形收敛更快。从同一情形下看 3 种相对搜索空间的变化时, 可以看出 C_{31} 、 C_{32} 和 C_{12} 都是纵轴随着 m 的增大很快收敛, C_{32} 和 C_{12} 收敛较快些, 在 $m \leq 6$ 时就能定性地看出收敛; C_{31} 收敛最慢, 在 m 取较大值时能判断出也是收敛的。从两种判断方法中我们能够看出收敛, 但是很难直观地从函数图像上直接看出 m 取较大值时的收敛状况, 也即只能定性得出收敛的结论, 这与前面的结论亦是吻合的。

为阶乘项和等价的多项式项的变化示意图。Stirling 公式为

$$\lim_{n \rightarrow \infty} \frac{n!}{\sqrt{2\pi n} \left(\frac{n}{e}\right)^n} = 1 \quad (7)$$

2.4.2 $n=10m$ 时相对解空间分析

当 $n=10m$ 时:

$$C_{31} = \frac{C_{2\text{-part}}}{C_{\text{one}}} = \frac{n!C_{n-1}^{m-1}}{(m+n-1)!} \sim \frac{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n \sqrt{2\pi(n-1)}\left(\frac{n-1}{e}\right)^{n-1}}{\sqrt{2\pi(m+n-1)}\left(\frac{m+n-1}{e}\right)^{m+n-1} \sqrt{2\pi(m-1)}\left(\frac{m-1}{e}\right)^{m-1} \sqrt{2\pi(n-m)}\left(\frac{n-m}{e}\right)^{n-m}} \sim$$

$$\sqrt{\frac{5(10m-1)}{9\pi(11m-1)(m-1)}} \cdot \frac{m^m\left(m-\frac{1}{10}\right)^{10m-1}}{\left(m-\frac{1}{11}\right)^{11m-1} (m-1)^{m-1}} \cdot \frac{10^{20m-1} \cdot e^{m-1}}{11^{11m-1} \cdot 9^m} \quad (8)$$

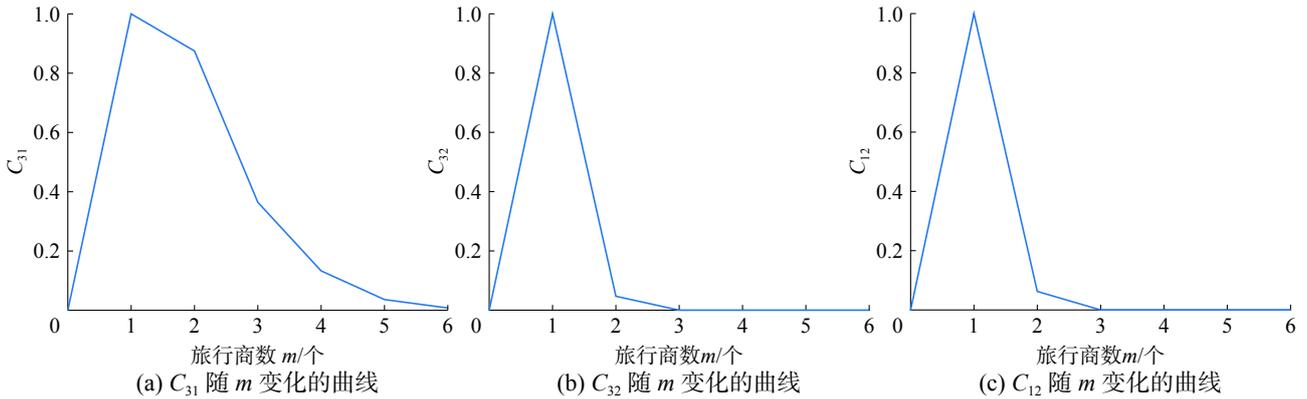


图 7 $n=e^m$ 情形下的相对搜索空间变化曲线图

Fig. 7 Solution space change in the case of $n=e^m$

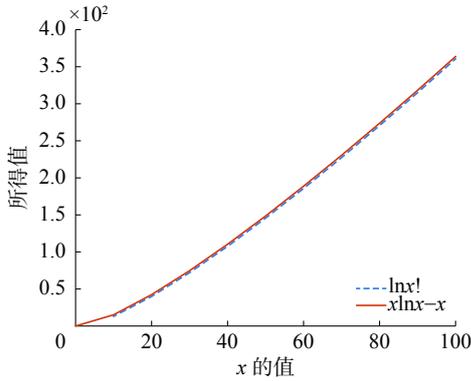


图 8 $n!$ 与 $\sqrt{2\pi n}\left(\frac{n}{e}\right)^n$ 变化示意图

Fig. 8 Diagram of changes in $n!$ and $\sqrt{2\pi n}\left(\frac{n}{e}\right)^n$

式(8)是一个关于 m 的分式, 可分为 3 个部分。当 m 取较大值时, 第一部分 $\sqrt{\frac{5(10m-1)}{9\pi(11m-1)(m-1)}}$

可以用 $\sqrt{\frac{1}{2\pi m}}$ 近似代替, 第二部分 $\frac{m^m\left(m-\frac{1}{10}\right)^{10m-1}}{\left(m-\frac{1}{11}\right)^{11m-1} (m-1)^{m-1}}$

是幂指数次项, 在 m 较大时可以用 $\frac{1}{m^{m-1}}$ 代替, 第三部分 $\frac{10^{20m-1} \cdot e^{m-1}}{11^{11m-1} \cdot 9^m}$ 可以化简为 e^{m-1} 。在这三项中, 增长速度主要体现在第二项和第三项幂指数项, 也就是说, 我们评价整个代数表达式随着 m 的增大其增长速度可以用第二项和第三项代替。

在计算机程序设计中, 衡量一个算法的好坏,

通常会用到时间复杂度和空间复杂度。性按照某一速度增长。它只表示增长的速度 (Order), 这个时间或空间复杂度并不是与程序复杂性严格相等的数学表达式, 它只抽象表示增长速度的类型, 通常用大 O 表示法。本文中的解空间的大小也称为空间复杂度, 也同样用算法设计中的空间复杂度表示法 (大 O 表示法) 来衡量本文中讨论的 3 种染色体设计方案的相对解空间。

因增长速度主要体现在第二项和第三项幂指数项, 也就是说, 我们评价整个代数表达式随着 m 的增大其增长速度可以用第二项和第三项代替。于是, 式(8)化简结果为

$$C_{31} = \frac{C_{2\text{-part}}}{C_{\text{one}}} \sim \sqrt{\frac{1}{2\pi m}}\left(\frac{e}{m}\right)^m \quad (9)$$

式(9)的结果可以用 $\left(\frac{e}{m}\right)^m$ 来代表其收敛速度, 则该等式可表示为

$$C_{\text{one}} = C_{2\text{-part}} \cdot O\left(\left(\frac{m}{e}\right)^m\right) \quad (10)$$

同样

$$C_{32} = \frac{C_{2\text{-part}}}{C_{\text{two}}} = \frac{n!C_{n-1}^{m-1}}{n!m^n} \quad (11)$$

当 m 取较大正整数时, 指数项可以用二项展开式的前两项近似替代, 具体过程不在此赘述, 以下相似替代过程将不特殊说明, 式(11)化简为

$$C_{32} \sim \sqrt{\frac{1}{2\pi m}} \cdot \frac{1}{m^{10m}} \cdot \frac{10^{10m-1}}{9^m} \quad (12)$$

则该等式可表示为

$$C_{\text{two}} = C_{2\text{-part}} \cdot O\left(m^{10m}\right) \quad (13)$$

对于 C_{12} 同样可得到:

$$C_{12} = \frac{C_{\text{one}}}{C_{\text{two}}} = \frac{m+n-1!}{nm^n} = \frac{(11m-1)!}{(10m)!m^{10m}} \sim \frac{\sqrt{2\pi(11m-1)}\left(\frac{11m-1}{e}\right)^{11m-1}}{\sqrt{2\pi(10m)}\left(\frac{10m}{e}\right)^{10m} m^{10m}} \sim \sqrt{\frac{11}{10}} \cdot \frac{1}{m^{9m}} \cdot \frac{1}{e^{m-1}} \quad (14)$$

则该等式可表示为

$$C_{\text{two}} = C_{\text{one}} \cdot O(m^{9m} \cdot e^m) \quad (15)$$

由前两个关系式:

$$C_{\text{one}} = C_{2\text{-part}} \cdot O\left(\left(\frac{m}{e}\right)^m\right) \quad (16)$$

$$C_{\text{two}} = C_{2\text{-part}} \cdot O(m^{10m}) \quad (17)$$

用等式 (17) 除以等式 (16) 可得

$$\frac{C_{\text{two}}}{C_{\text{one}}} = \frac{C_{2\text{-part}} \cdot O(m^{10m})}{C_{2\text{-part}} \cdot O\left(\left(\frac{m}{e}\right)^m\right)} = O(m^{9m} \cdot e^m) \quad (18)$$

式 (18) 等价于 $C_{\text{two}} = C_{\text{one}} \cdot O(m^{9m} \cdot e^m)$, 这与上述推导中的等式 (15) 是吻合的, 由此该方法计算三者解空间大小在逻辑上是严密的, 从式 (15) ~ (17) 可知, $C_{\text{two}} = C_{\text{one}} \cdot O(m^{9m} \cdot e^m) = C_{2\text{-part}} \cdot O(m^{10m})$, 可以看出, 随着问题规模的逐渐增大, 3 种编码方案对应解空间的差距呈幂指函数的关系。由此可以看出, 两段式染色体编码方案的优越性。

$$C_{31} = \frac{C_{2\text{-part}}}{C_{\text{one}}} = \frac{n!C_{n-1}^{m-1}}{(m+n-1)!} = \frac{m^2!(m^2-1)!}{(m^2+m-1)!(m-1)!(m^2-m)!} \sim \frac{\sqrt{2\pi m^2}\left(\frac{m^2}{e}\right)^{m^2}}{\sqrt{2\pi(m^2-1)}\left(\frac{m^2-1}{e}\right)^{m^2-1}} \quad (19)$$

$$\frac{\sqrt{2\pi(m^2+m-1)}\left(\frac{m^2+m-1}{e}\right)^{m^2+m-1}}{\sqrt{2\pi(m-1)}\left(\frac{m-1}{e}\right)^{m-1} \sqrt{2\pi(m^2-m)}\left(\frac{m^2-m}{e}\right)^{m^2-m}}$$

同样用二项展开式的前两项去近似代替指数项部分, 式 (19) 可简化为

$$C_{31} \sim \frac{1}{\sqrt{2\pi m}} \cdot \frac{1}{m^m} \cdot e^{m-1} \quad (20)$$

式 (20) 的收敛速度主要取决于第二项幂指数因式 $\frac{1}{m^m}$ 和第三项 e^m , 所以式 (20) 关系可表示为

$$C_{\text{one}} = C_{2\text{-part}} \cdot O\left(\left(\frac{m}{e}\right)^m\right) \quad (21)$$

同样将 $n=m^2$ 代入式 (22):

$$C_{32} = \frac{C_{2\text{-part}}}{C_{\text{two}}} = \frac{n!C_{n-1}^{m-1}}{n!m^n} = \frac{(n-1)!}{(m-1)!(n-m)!m^n} = \frac{(m^2-1)!}{(m-1)!(m^2-m)!m^{m^2}} \sim \frac{\sqrt{2\pi(m^2-1)}\left(\frac{m^2-1}{e}\right)^{m^2-1}}{\sqrt{2\pi(m-1)}\left(\frac{m-1}{e}\right)^{m-1} \sqrt{2\pi(m^2-m)}\left(\frac{m^2-m}{e}\right)^{m^2-m} m^{m^2}} \sim \frac{1}{\sqrt{2\pi(m-1)}} \cdot \frac{1}{m^{m^2-m}} \cdot \frac{1}{m} \quad (22)$$

2.4.3 平方关系下的相对解空间分析

上一小节讨论了 3 种染色体方案在城市数目与旅行商人数为线性关系时, 解空间的相对大小关系, 实际上, 在线性关系 $n=km$ 下, 系数 k 代表的是平均每个商人访问的城市数目。因此讨论相对解空间时, 让 m 逐渐递增时, k 是作为常数处理的, 也就是说, 平均每个商人访问的城市数目是大体不变的。在现实的工程问题中, 随着工程规模的增加, 即待访问城市数目的急剧增多, 旅行商的人数也应该随之增加。考虑这个实际问题时, 不应该单纯依靠旅行商人数的增加来完成所有城市的访问, 也应该随着城市的增加, 相应地增大每个商人的工作负载, 即访问城市的数目。于是, 我们应该找到一个函数来表示城市数目和旅行商人数的关系, 满足随着城市数目的增多, 旅行商人数也随之增加, 同时平均每个商人访问的城市数目也随之相应增加。考虑到城市数与旅行商人数的关系和商人访问城市效率的不同, 本文选择两种常见的不同增速关系, 即二次幂函数与指数增长关系。

本节讨论 n 与 m 在二次函数关系下相对解空间的收敛速度。

将 $n=m^2$ 代入式 (19):

式 (22) 的增长速度仍取决于第二项的幂指数项, 所以式 (22) 关系可表示为

$$C_{\text{two}} = C_{2\text{-part}} \cdot O(m^{m^2-m}) \quad (23)$$

继而将 $n=m^2$ 代入 C_{12} 化简得到

$$C_{\text{two}} = C_{\text{one}} \cdot O\left(\frac{m^{m^2-2m}}{e^m}\right) \quad (24)$$

由式 (21)、(23)、(24) 可得 $C_{\text{two}} = C_{\text{one}} \cdot O\left(\frac{m^{m^2-2m}}{e^m}\right) = C_{2\text{-part}} \cdot O(m^{m^2-m})$, 看出在旅行商人数和城市数为二次关系时, 双染色体编码对应的解空间是两段式染色体编码对应的解空间的 m^{m^2-m} 倍, 因此不同染色体编码方案的选择对问题的求解经常起主要的作用, 可以指导实际问题的求解。

2.4.4 指数关系下的相对解空间分析

本节讨论 n 与 m 在指数函数关系下相对解空间的收敛速度。

将 $n=e^m$ 代入式 (4):

$$C_{31} = \frac{C_{1-part}}{C_{one}} = \frac{n!C_{n-1}^{m-1}}{(m+n-1)!} \sim \frac{\sqrt{2\pi e^m} \left(\frac{e^m}{e}\right)^{e^m} \sqrt{2\pi(e^m-1)} \left(\frac{e^m-1}{e}\right)^{e^m-1}}{\sqrt{2\pi(m+e^m-1)} \left(\frac{m+e^m-1}{e}\right)^{m+e^m-1} \sqrt{2\pi(m-1)} \left(\frac{m-1}{e}\right)^{m-1} \sqrt{2\pi(e^m-m)} \left(\frac{e^m-m}{e}\right)^{e^m-m}} \sim \frac{1}{\sqrt{2\pi m}} \cdot \frac{1}{m^m} \cdot e^{m-1} \quad (25)$$

同样用二项展开式的前两项取近似代替指数项部分, 式(25)可化简为

$$C_{31} \sim \frac{1}{\sqrt{2\pi m}} \cdot \frac{1}{m^m} \cdot e^{m-1} \quad (26)$$

式(26)的收敛速度主要取决于第二项幂指数因式 $\frac{1}{m^m}$, 所以式(26)关系可表示为

$$C_{one} = C_{2-part} \cdot O\left(\left(\frac{m}{e}\right)^m\right) \quad (27)$$

同样将 $n=e^m$ 代入式(5)得

$$C_{32} = \frac{C_{2-part}}{C_{two}} = \frac{n!C_{n-1}^{m-1}}{n!m^n} = \frac{(n-1)!}{(m-1)!(n-m)!m^n} \quad (28)$$

式(28)的增长速度仍是取决于第二项的幂指数项, 所以式(28)关系可表示为

$$C_{two} = C_{2-part} \cdot O(m^{e^m}) \quad (29)$$

继而将 $n=e^m$ 代入 C_{12} 化简得到:

$$C_{two} = C_{one} \cdot O\left(\frac{m^{e^m-m}}{e^m}\right) \quad (30)$$

由式(27)、(29)、(30)可得 $C_{two} = C_{one} \cdot O\left(\frac{m^{e^m-m}}{e^m}\right) = C_{2-part} \cdot O(m^{e^m})$, 可以看出在旅行商人数和城市数为指数关系时, 双染色体编码对应的解空间是两段式染色体编码对应的解空间的 m^{e^m} 倍, 不同染色体编码方案的选择对问题的求解效率有非常重要的影响, 因此本文得到的对 3 种编码方案对应解空间的大小分析对实际问题的求解提供了定量的指导意义。

3 结束语

本文首先介绍了多旅行商问题的概念和 3 种染色体编码方案, 已有文献只是定性地给出 3 种编码方案对应解空间的大小关系, 并没有给出严格证明。本文首先给出相对解空间的概念, 然后在城市数充分大和城市数大于旅行商人数的条件下, 在极限意义下严格证明了 3 种解空间的相对大小关系, 最后利用 Stirling 公式严密分析了几种染色体编码方案中旅行商人数和城市数目在不同关系下相对解空间的大小, 得出了相对解空间的理论增长速度, 给实际工程问题求解中染色体编码方案的选择提供了科学的参考依据和指导意义。

基于 3 种染色体方案求解多旅行商问题的相对解空间大小关系分析, 就目前方案给出的解空间仍然有冗余解存在。两段式染色体方案在单染色体方案和双染色体方案基础之上消除了相当大的一部分冗余, 虽然没有完全消除, 但是给了我们很大启发, 去寻找更加合理的新的染色体编码方案以进一步减小解空间的冗余。

参考文献:

- [1] CARTER A E, RAGSDALE C T. A new approach to solving the multiple traveling salesperson problem using genetic algorithms[J]. European journal of operational research, 2006, 175(1): 246–257.
- [2] YUAN Shuai, SKINNER B, HUANG Shoudong, et al. A new crossover approach for solving the multiple travelling salesmen problem using genetic algorithms[J]. European journal of operational research, 2013, 228(1): 72–82.
- [3] WANG Xuping, XU Chuanlei, HU Xiangpei. Genetic algorithm for vehicle routing problem with time windows and a limited number of vehicles[C]//Proceedings of 2008 International Conference on Management Science and Engineering 15th Annual Conference. Long Beach, CA, USA, 2010: 128–133.
- [4] BRANKE J, NGUYEN S, PICKARDT C, et al. Automated design of production scheduling heuristics: a review[J]. IEEE transactions on evolutionary computation, 2016, 20(1): 110–124.
- [5] GOLDBERG D E. Genetic algorithms in search, optimization, and machine learning[M]. Reading, Mass: Addison-Wesley Publishing Company, 1989.
- [6] KESHAVARZ T, SALMASI N, VARMAZYAR M. Minimizing total completion time in the flexible flowshop sequence-dependent group scheduling problem[J]. Annals of operations research, 2015, 226(1): 351–377.
- [7] 张永库, 尹灵雪, 孙劲光. 基于改进的遗传算法的模糊聚类算法[J]. 智能系统学报, 2015, 10(4): 627–635.
ZHANG Yongku, YIN Lingxue, SUN Jinguang. Fuzzy clustering algorithm based on the improved genetic algorithm[J]. CAAI transactions on intelligent systems, 2015, 10(4): 627–635.
- [8] KATAYAMA K, SAKAMOTO H, NARIHISA H. The efficiency of hybrid mutation genetic algorithm for the trav-

- elling salesman problem[M]. The Netherlands: Elsevier Science Publishers B. V., 2000.
- [9] 黄可为, 汪定伟. 热轧计划中的多旅行商问题及其计算方法[J]. 计算机应用研究, 2007, 24(7): 43–45, 57.
HUANG Kewei, WANG Dingwei. Multiple traveling salesman problem and its application to hot rolling planning[J]. Application research of computers, 2007, 24(7): 43–45, 57.
- [10] 江贺, 张宪超, 车皓阳, 等. 带多项式量级约束条件的多商品流 BWTSP 线性规划[J]. 计算机研究与发展, 2007, 44(10): 1796–1800.
JIANG He, ZHANG Xianchao, CHE Haoyang, et al. A multi-commodity flow linear programming with polynomial constraints for black and white traveling salesman problem[J]. Journal of computer research and development, 2007, 44(10): 1796–1800.
- [11] TRIGUI S, CHEIKHROUHOU O, KOUBAA A, et al. FL-MTSP: a fuzzy logic approach to solve the multi-objective multiple traveling salesman problem for multi-robot systems[J]. Soft computing, 2017, 21(24): 7351–7362.
- [12] 王宇平, 李英华. 求解 TSP 的量子遗传算法[J]. 计算机学报, 2007, 30(5): 748–755.
WANG Yuping, LI Yinghua. A novel quantum genetic algorithm for TSP[J]. Chinese journal of computers, 2007, 30(5): 748–755.
- [13] 赵新超, 刘国莅, 刘虎球, 等. 基于非均匀变异和多阶段扰动的粒子群优化算法[J]. 计算机学报, 2014, 37(9): 2058–2070.
ZHAO Xinchao, LIU Guoli, LIU Huqiu, et al. Particle swarm optimization algorithm based on non-uniform mutation and multiple stages perturbation[J]. Chinese journal of computers, 2014, 37(9): 2058–2070.
- [14] 张筑生. 数学分析新讲 (第 2 册)[M]. 北京: 北京大学出版社, 1990.

作者简介:



赵新超, 男, 1976 年生, 教授, 博士生导师, 主要研究方向为进化计算、群体智能及其运筹优化。



郭赛, 女, 1993 年生, 硕士研究生, 主要研究方向为群体智能理论。