

遗传算法求解 TSP 问题的研究进展

陈江华,林爱文,杨明,龚健

(武汉大学 资源环境学院,湖北 武汉 430079)

摘要: 文章介绍了 TSP 问题和遗传算法的基本原理以及特点;针对解决 TSP 问题,论述了遗传算法在编码表示和遗传操作算子等方面的应用情况,分别指出了顺序表示、路径表示和布尔矩阵表示的优缺点,阐述了三种基本的操作算子的应用现状;最后,简单说明了混合遗传算法在求解 TSP 问题中的应用并对遗传算法解决 TSP 问题的前景提出了展望。

关键词: TSP;遗传算法;编码;遗传操作;展望

中图分类号: O211.6 **文献标识码:** A **文章编号:** 1007-855X(2003)04-0009-05

Review of Genetic Algorithms for Traveling Salesman Problem

CHENG Jiang-hua, LIN Ai-wen, YANG Ming, GONG Jian

(School of Resources & Environment Science, Wuhan University, Wuhan 430079, China)

Abstract: TSP (Traveling Salesman Problem) is a typical NP-complete problem, and genetic algorithm (GA) is the perfect method for solving NP-complete problem. TSP and the basic theories and characteristics of GA are first introduced. Then the encoding model and genetic operation about GA in solving TSP are discussed. The advantages and disadvantages of ordinal representation, path representation and matrix representation are respectively indicated, and the application of the three basic genetic operators is elaborated. At last, the application of Hybrid genetic algorithm is briefly presented. It is pointed out that a better crossover or mutation routine can be found out which retains the structure from the parent chromosomes and still ends up with a legal tour in the child chromosomes, which leads to a better solution than ever before. And the prospect for the future of genetic algorithm in solving TSP is made.

Key words: TSP (Traveling Salesman Problem); genetic algorithm; encoding; genetic operation; prospect

0 引言

货担郎问题 (Traveling Salesman Problem, TSP), 也称为巡回旅行商问题, 是一个较古的问题。最早可以追溯到 1759 年 Euler 提出的骑士旅行问题。TSP 问题是一个典型的、容易描述但是难以处理的 NP 完全问题。同时 TSP 问题也是诸多领域内出现的多种复杂问题的集中概括和简化形式。所以, 有效解决 TSP 问题在计算理论上和实际应用上都有很高的价值。而且 TSP 问题由于其典型性已经成为各种启发式的搜索、优化算法的间接比较标准 (如遗传算法、神经网络优化法、列表寻优 (TABU) 法、模拟退火法等)。遗传算法就其本质来说, 主要是解决复杂问题的一种鲁棒性强的启发式随机搜索算法。因此遗传算法在 TSP 问题求解方面的应用研究, 对于构造合适的遗传算法框架、建立有效的遗传操作以及有效地解决 TSP 问题等有着多方面的重要意义。

1 TSP 问题

TSP 问题属于 NP 完全问题, 给定一组 n 个城市和他们两两之间地直达距离, 寻找一条闭合地旅程, 使得每个城市刚好经过一次而且总的旅行距离最短。TSP 问题的描述很简单, 简言之就是寻找一条最短的

收稿日期: 2002-11-21.

第一作者简介: 陈江华 (1978~), 男, 硕士研究生; 主要研究方向: 遗传算法以及地理信息系统的研究与开发. E-mail: chjhdj@163.com

遍历 n 个城市的路径,或者说搜索整数子集 $X = \{1, 2, \dots, n\}$ (X 的元素表示对 n 个城市的编号) 的一个排列 $(X) = \{v_1, v_2, \dots, v_n\}$, 使

$$T = d(v_i, v_i + 1) + d(v_i, v_n)$$

取最小值. 式中的 $d(v_i, v_i + 1)$ 表示城市 v_i 到城市 $v_i + 1$ 的距离.

2 遗传算法

2.1 基本原理

遗传算法是一种借鉴生物界自然选择和自然遗传机制的随机化搜索算法,由美国 J. Holland 教授提出,其主要特点是群体搜索策略和群体中个体之间的信息交换,搜索不依赖于梯度信息.它是一种全局化搜索算法,尤其适用于传统搜索算法难于解决的复杂和非线性问题.

选择(selection)算子、交叉(crossover)算子和变异(mutation)算子是遗传算法的3个主要操作算子^[1].

遗传算法中包含了如下5个基本要素^[1]:(1)对参数进行编码;(2)设定初始种群大小;(3)适应度函数的设计;(4)遗传操作设计;(5)控制参数设定(包括种群大小、最大进化代数、交叉率、变异率等).

2.2 特点

- 1) 遗传算法对优化问题没有太多的数学要求,而且只要知道目标函数的信息即可;
- 2) 遗传算法采用的是启发性的知识智能搜索算法,在搜索高度空间复杂问题上比以往有更好的效果;
- 3) 遗传算法是对问题参数或者变量编码群进行优化,而不是参数或变量本身;
- 4) 遗传算法使用的选择、交叉、变异算子都是随机的;

3 对 TSP 的遗传基因编码方法

在遗传算法的应用中,遗传基因的编码是一个很重要的问题.对遗传基因进行编码时,要考虑到是否适合或有利于交叉和变异操作.在遗传算法中有两种基于串的基因编码形式,一种是基于二进制的遗传算法(binary coded GA),一种是基于顺序的遗传算法(the order-based GA).在求解 TSP 问题时,采用基于二进制的遗传算法,不利于交叉和变异操作.

3.1 顺序表示

在求解 TSP 问题时,1985年,Grefenstette^[2]等提出了基于顺序表示(ordinal representation)的遗传基因编码方法.顺序表示是指所有城市依次排列构成一个顺序表(order list),对于一条旅程,可以依旅行经过顺序处理每个城市,每个城市在顺序表中的顺序就是一个遗传因子的表示,每处理完一个城市,从顺序表中去掉该城市.处理完所有城市以后,将每个城市的遗传因子表示连接起来,就是一条旅程的基因表示.例如,顺序表 $C = (1, 2, 3, 4, 5, 6, 7, 8, 9)$,一条旅程为 $5 - 1 - 7 - 8 - 9 - 4 - 6 - 3 - 2$.按照这种编码方法,这条旅程的编码为表 $l = (5\ 1\ 5\ 5\ 5\ 3\ 3\ 3\ 2\ 1)$.但是这种表示方法,在进行单点交叉的时候,交叉点右侧部分的旅程发生了随即变化,但是交叉点左侧部分的旅程未发生改变,由于存在这样的缺点,所以顺序表示的方法的适用性存在一定的问题.

3.2 路径表示

与顺序表示类似的还有路径表示(path representation)方法.路径表示是表示旅程对应的基因编码的最自然、最简单的表示方法.例如,旅程 $(5 - 1 - 7 - 8 - 9 - 4 - 6 - 2 - 3)$ 可以直接表示为 $(5\ 1\ 7\ 8\ 9\ 4\ 6\ 2\ 3)$,基于路径表示的编码方法,要求满足 TSP 问题的任一城市有且仅有一次访问的约束条件,即要求一个个体(即一条旅程)的染色体编码中不容许有重复的基因码.但是,在进行交叉操作之后所生成的个体一般不能满足这个约束条件.为了解决这个问题,人们提出了一些新的操作来解决这类问题.

3.3 布尔矩阵表示

1992年,Fox和McMahon等^[3]提出了旅程的矩阵表示方法.他们将一个旅程定义为一个优先权布尔矩阵 M ,当且仅当城市 i 排在城市 j 之前时矩阵元素 $m_{ij} = 1$.这种方法用 $n \times n$ 矩阵 M 代表一条旅程, M 具有如下三个性质:

- 1) 矩阵中 1 的数目为 $n(n-1)/2$;
- 2) $m_{ii} = 0, i = 1, 2, \dots, n$;
- 3) 若 $m_{ij} = 1$, 且 $m_{jk} = 1$, 那么 $m_{ik} = 1$.

基于 TSP 问题的染色体的表示,以及遗传操作的实现实质上可以看成是向量运算,储理才在 Matlab 环境下实现了遗传算法求解 TSP 问题的程序设计.

4 针对 TSP 的遗传操作算子

4.1 选择算子

对于求解 TSP 问题,常用的选择机制有轮盘赌选择(roulette wheel selection)机制,最佳个体保存(elitist model)选择机制,期望值模型(expected value model)机制,排序模型(rank-based model)选择机制,联赛选择模型(tournament selection model)机制,排挤模型(crowding model)等.

遗传算法中一个要求解决的问题是如何防止“早熟”收敛现象.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. Rudolph C^[4]提出采用精英选择策略即保持群体中最好的个体不丢失,以保证算法的收敛性. Konstantin Boukreev^[5]采用联赛选择方法和最优个体保存方法相结合的方法,通过在 VC++6.0 环境下编程实现了求解 TSP 问题的遗传算法,取得了很好的效果.谢胜利等提出浓度控制策略,当某种个体的浓度超过给定的浓度阈值时减少该种个体的数量,使之控制在给定的浓度阈值之内,并随机产生新的个体以补足种群的规模,他们通过实验证明该策略很好的解决了遗传算法中群体的多样性问题.

4.2 交叉算子

1985年,Goldberg等^[6]提出针对 TSP 问题的基于路径表示的部分匹配交叉(Partially Matched Crossover, PMX)操作,PMX 操作要求随机选取两个交叉点,以便确定一个匹配段,根据两个父个体中两交叉点之间的中间段给出的映射关系生成两个子个体.

1985年,Davis等^[7]针对 TSP 问题也提出了一种交叉操作方法——基于路径表示的顺序交叉(Ordered Crossover, OX)操作.这种 OX 操作能保留排列并融合不同排列的有序结构单元.当两个父个体交叉时,通过选择父个体 1 的一部分,保留父个体 2 中城市码的相对顺序生成子个体.例如, $p_1: (1\ 2\ 3\ | \ 4\ 5\ 6\ 7\ | \ 8\ 9)$, $p_2: (3\ 5\ 7\ | \ 8\ 2\ 1\ 9\ | \ 6\ 4)$ 两个交叉点之间的中间段保持不变,然后把 p_2 从第二个交叉点开始排序,到达表尾时返回表头记录,得到 6-4-3-5-7-8-2-1-9,在此基础上去掉 p_1 中保留的 4 5 6 7,得到 3-8-2-1-9 并把它复制到 p_1 ,从第二个交叉点开始,到达表尾时再返回表头继续.这样,我们可以得到子个体 1 为: $s_1: (2\ 1\ 9\ 4\ 5\ 6\ 7\ 3\ 8)$,同理可生成子个体 2 为: $s_2: (5\ 6\ 7\ 8\ 2\ 1\ 9\ 3\ 4)$.

在使用遗传算法求解问题时的一个重要问题是收敛速度问题.1985年,Grefenstette^[8]提出贪心交叉法(Greedy Crossover),基本原理是:在一个父个体中选择第一个城市作为子个体的第一个城市,然后在两个父个体中进行比较并找到与已经选择的那个城市相邻且距离较近的城市作为下一个城市扩展到旅程中;如果与该城市相邻的两个城市有一个已经在旅程中,那么选择另外一个,如果两个都在旅程中,那么就选择其它没有被选择的的城市.许多人(如谢胜利^[9]、Konstantin Boukreev^[5]等)采用这一交叉算子都取得了很好的效果.使用贪心交叉法,不仅提高了寻优的速度,而且也考虑到城市之间的连接关系.

1987年,Oliver等^[10]针对 TSP 问题提出了循环交叉(Cycle Crossover, CX)操作.

上述操作基本上考虑的是城市的位置和顺序,没有考虑城市间的联系.1989年,Whitle^[11]等提出了一种边重组(Edge Recombination, ER)交叉操作,使个体能够从父个体继承 95%~99%的边信息.ER 操作是根据继承两个父个体定义的旅程中城市间的相邻关系生成子个体.但是 ER 操作可能存在边失败(edge failure)的问题.可喜的是,在 1991年,Starkweather^[12]等提出了一种改进的方法,在 ER 操作中不再保留父个体中共同部分的序列.实验结果表明这种处理方法比含随机选择的处理的性能有相当的改善.

4.3 变异算子

遗传算法强调的是交叉的功能.从遗传算法的观点来看,解的进化主要靠选择机制和交叉策略来完

成,变异只是为选择、交叉过程中可能丢失的某些遗传基因进行修复和补充,变异在遗传算法的全局意义上只是一个背景操作.针对 TSP 问题,陈国良等^[1]介绍了四种主要的变异技术.(1)点位变异,变异仅以一定的概率(通常很小)对串的某些位做值的变异;(2)逆转变异,在串中,随机选择两点,再将这两点内的子串按反序插入到原来的位置中;(3)对换变异,随机选择串中的两点,交换其值(码);(4)插入变异,从串中随机选择 1 个码,将此码插入随机选择的插入点中间.

此外,对于变异操作还有一些变体形式,如 Sushil J. Louis^[13]提出的贪心对换变异(greedy - swap mutation),其基本思想是从一个染色体中随机的选择两个城市(即两个码值),然后交换它们,得到新的染色体,以旅程长度为依据比较交换后的染色体与原来的染色体的大小,保留旅程长度值小的染色体.谢胜利^[10]等提出倒位变异算子,该算子是指在个体编码串中随机选择两个城市,是第一个城市的右城市与第二个城市之间的编码倒序排列,从而产生一个新个体.如,若有父个体 $P(145236)$,假设随机选择的城市是 4,3,那么产生的新个体为 Offspring(143256).

Hiroaki SENGO KU^[14]等采用 2opt^[15]方法的思想,提出了 Greedy Subtour Crossover (GSC) 变异算子.这种方法与贪心对换变异有相同的思想,但是更易扩张更有效率. Konstantin Boukreev^[5]通过实验发现:当城市大小在 200 之内时,该变异算子可以大大改善程序的运行速度,随着城市数目的增加,尤其是当城市数目达到 1 000 以上时,程序运行速度非常慢.

5 混合遗传算法

随着遗传算法的发展,遗传算法广泛的与其它算法相结合,产生许多混合遗传算法(Hybrid genetic algorithm).王俊海^[16]将 Memetic 算法与遗传算法结合起来,给出了一种基于 Memetic 算法的高效遗传算法.将这种算法应用于 TSP 问题取得了满意的结果,把局部优化的高效性与遗传算法的鲁棒性很好的结合起来了. Ranieri Baraglia^[17]等提出一种解决 TSP 问题的并行混合启发式(Parallel Hybrid Heuristic)算法. Jog 等^[18]提出启发式算法(Heuristic Algorithm)也取得了很好的效果.

6 结论与展望

综上所述,我们可以看到,针对求解 TSP 问题,各种遗传算法总是依赖于问题的编码以及遗传操作算子.我们要发展遗传算法也要以这几个方面作为突破口.各种求解 TSP 问题的算法里面都包含启发式的信息或者是在编码过程中包含有边的信息(如矩阵表示和矩阵交叉等),我认为这是一种趋势.尽管对遗传算法有很多改进和发展,但是总是可以找到更好的算法.对于遗传算法求解 TSP 问题,我认为最大的一个问题是:很难保证染色体的完整结构,从父染色体开始一直到生成一个合法的旅程而得到的子个体的整个过程中,染色体的结构总是会发生变化.也许以后会找到更好的交叉和变异规则而使染色体的结果保持不变.

参考文献:

- [1] 陈国良,等.遗传算法及其应用[M].北京:人民邮电出版社,1996.5~10.
- [2] Grefenstette J J. Genetic Algorithms for the Salesman Problem[C]. In: Proceedings of the First International Conference on Genetic Algorithms, Lawrence Erlbaum Associates, Publishers, 1985. 160~165.
- [3] Fox B R, McMahon M B. Genetic Operators for the Sequencing Problems. Foundations of Genetic Algorithms[M]. In: Rawlins G J E. Morgan Kaufmann Publishers, 1991. 284~300.
- [4] Rudolph C. Convergence Properties of Canonical Genetic Algorithms[J]. IEEE Trans. Neural Networks, 1994;5(1):96~101.
- [5] Konstantin Boukreev. Genetic Algorithm and Traveling Salesman Problem[OL]. <http://www.codeguru.com/misc/index.shtml>
- [6] Goldberg D E, & Lingle R. Alleles, loci, and the Traveling Salesman Problem[C]. Proceedings of an International Conference on Genetic Algorithms and Their Applications, 1985. 154~159.
- [7] Davis L. Job Shop Scheduling with Genetic Algorithms[C]. In: Proceedings of an International Conference on Genetic Algorithms and Their Applications, 1985. 136~140.
- [8] Grefenstette J J, Gopal R, Rosmaita B J, Gucht D V. Genetic Algorithms for the Traveling Salesman Problem[C]. In J. J. Grefenstette, Editor, Proceedings of the First International Conference on Genetic Algorithms, Lawrence Erlbaum, 1985. 160~

168.

- [9] 谢胜利,等. 求解 TSP 问题的一种改进的遗传算法[J]. 计算机工程与应用,2002(8):58~245.
- [10] Oliver L M, Smith D J, Holland J R C. A study of Permutation Crossover Operators on the Traveling Salesman Problem[C]. In: Proceedings of the Second International Conference on Genetic Algorithms, Lawrence Erlbaum Associates, 1987. 224~230.
- [11] Whitley L D, Starkweather T, Fuquay D A. Scheduling Problems and Traveling Salesman: The Genetic Edge Combination Operator[C]. In: Proceedings of the 3rd International Conference on Genetic Algorithm, Morgan Kaufmann Publishers, 1989. 133~140.
- [12] Starkweather T, McDaniel S, Mathias K, Whitley C, Whitley D. A Comparison of Genetic Sequencing Operators[C]. Proceedings of the 4th International Conference on Genetic Algorithm. Morgan Kaufmann, Los Altos, 1991. 69~76.
- [13] Sushil J. Louis. Genetic Algorithm for TSP[OL]. <http://www.cs.unr.edu/~sushil/papers/conference/newpapers/99/gecco99/iga/GECCO/gecco/node3.html>.
- [14] Hiroaki SENGOKU, Ikuo YOSHIHARA. A Fast TSP Solver Using GA on JAVA[OL]. <http://www.gcd.org/sengoku/docs/arob98.pdf>.
- [15] Croes G A. A Method for Solving Traveling Salesman Problems[J]. Operations Research, 1958, 6:791~812.
- [16] 王俊海. TSP 问题的一种高效 Memetic 算法[J]. 交通与计算机, 2002, 20(1):14~17.
- [17] Ranieri Baraglia, Jose Ignacio Hidalgo, Raffaele Perego. A Parallel Hybrid Heuristic for the TSP[OL]. http://malvasia.cnuce.cnr.it/~raffaele/papers_ps/EvoWorkshop.pdf.
- [18] Prasanna Jog, Jung Y. Suh, and Dirk Van Gucht. Parallel Genetic Algorithms Applied to the Traveling Salesman Problem[J]. SIAM Journal of Optimization, 1991, 1(4):515~529.

(上接第8页)

在 MapInfo 平台下把数据进行处理并输入数据库,主要是把 1:1 万的土地利用现状图、地形图、行政区划图、用地适宜性评价图等为基本工作底图,利用手扶数字化仪或扫描仪,在 GIS 软件支持下按特定编码分别录入计算机.对于图形数据,要进行分层输入,以便图形数据的分层管理、图形输出.

导入数据之后就可以实现规划管理的许多功能,其中最重要的一项功能是建设用地的动态选址功能.利用 MapBasic 语言开发 MapInfo 的自动选址.其核心语句是:Select * from 表名 where + 条件.选址之后,让系统进行自动统计用地类别、用地权属以及用地数量.并以报表的形式给出结果.从而让决策者能根据上述报表快速作出决策.

土地利用总体规划中有一项很重要的内容,即专题规划.利用 GIS 技术,可以对空间数据进行分析,也就是对规划区内的用地类别进行分析,依此得出用地量的预测.在此基础上,用 MapInfo 平台分类、总结,并采用一系列的专题图例自动生成规划区用地专题图.本文在制作基本农田保护区规划中,依据预测结果,把基本农田分成叁大类即一级农田保护区、二级基本农田保护区和一般农田保护区.然后,分别用叁种不同的图例(叁种不同的颜色和图案)予以标识.从而把规划期内农田的数量和质量单独列出,为保护耕地、实现耕地的动态平衡以及给决策者的决策,提供科学依据.

6 结束语

本文对土地利用总体规划的具体技术方法、技术路线作了较为详细的阐述,在实践中结合 GIS 技术,取得了较好的成果.但各业用地量的预测及方案编制还需进一步的探讨.

参考文献:

- [1] 宋 键,等.人口预测和人口控制[M].北京:人民出版社,1982.51~62.
- [2] 袁占良,黄俊华.基于网络的土地利用总体规划管理信息系统[J].桂林工学院学报,2000,4(3):2~3.
- [3] 尹 君,张天增,赵世强,林 培,卢树昌.土地利用规划的多目标规划模式探讨[J].地理学与国土研究,1997,13(4):1~2.