

旅行商问题最优路径的改进免疫遗传算法

孔令夷

(西安邮电大学经济与管理学院, 陕西 西安 710061)

摘要: 本文研究了一种改进的求解旅行商问题最优路径的免疫遗传算法. 结合随机法与贪心法生成初始种群, 利用亲和度排序而选取抗体以得到复制群体, 引入轮盘赌及克隆选择获取高亲和度抗体, 并实施疫苗接种及免疫记忆更新抗体. 运用免疫记忆机理的闭环逻辑, 证明了该算法生成的城市序列是全局收敛的. 数值实验证明该算法是有效的.

关键词: 旅行商问题; 免疫遗传算法; 克隆选择; 疫苗接种

MR(2010) 主题分类号: 90C39

中图分类号: O242.28

文献标识码: A

文章编号: 0255-7797(2015)02-0361-07

1 引言

现实中的旅行商问题 (Travelling salesman problem, TSP) 是指旅行商要经过所有 N 个城市, 各 1 次, 最终回到起点. 本研究的目的就是寻找最短路径. 设 N 个城市的集合

$$C = \{c_1, c_2, \dots, c_N\}, \quad (1.1)$$

每两个城市之间的距离为 $d(c_i, c_j)$, 属于正实数集合, 其中

$$c_i, c_j \in C (1 \leq i, j \leq N), \quad (1.2)$$

求使目标函数

$$T_d = \sum_{i=1}^{N-1} d(c_{\Pi(i)}, c_{\Pi(i+1)}) + d(c_{\Pi(N)}, c_{\Pi(1)}) \quad (1.3)$$

最小的路径序列 $c_{\Pi(1)}, c_{\Pi(2)}, \dots, c_{\Pi(N)}$, 其中 $\Pi(1), \Pi(2), \dots, \Pi(N)$ 是 $1, 2, \dots, N$ 的全排列. 遗传算法 (Genetic Algorithm, GA) 遵循“适者生存”的规律, 通过模仿自然选择而不断探索最优解, 求解过程包括选择 (Selection)、交叉 (Crossover)、变异 (Mutation) 操作. 免疫算法 (Immune Algorithm, IA) 通过高度模拟自然免疫系统, 兼具确定性及随机性, 在启发式算法中的勘探及挖掘能力较为显著, 不仅优选亲和度较高的抗体, 而且也抑制并适度保留高浓度且亲和度低的抗体, 实现抗体的多类型, 也就确保了 IA 的开放性及其全局寻优能力. 由于 GA 和 IA 都对目标问题的限制不多, 对目标函数的数学特性要求也不严格, 因而相比传统的运筹规划方法, 在处理复杂问题时显得游刃有余, 几乎都能找到最优解. GA 和 IA 的应用也非常广泛, 适用于处理大多数科学与工程复杂问题, 应用价值很高, 其中, 旅行商问题就是一个典

*收稿日期: 2012-09-29 接收日期: 2014-04-20

基金项目: 国家自然科学基金资助 (71102149).

作者简介: 孔令夷 (1977-), 男, 山东烟台, 副教授, 主要研究方向: 智能计算.

型, 该问题在图论中具有代表性, 已被证明是高维非线性完全问题 (NPC, Non-deterministic Polynomial Complete Problem)^[1], 具有高度的计算复杂性, 多年来都是理论界面临的难题, 这两种主流的智能算法应用于 TSP 求解已成为众多学者研究的对象^[2].

2 改进免疫遗传算法设计

求解旅行商问题的传统方法存在有明显的缺陷: 设计变量少, 与现实不相符; 假设较多, 对目标函数有可微或连续的限制, 会产生较多非可行解, 难以得到全局最优解^[3]. 而 GA 的优势较明显, 不直接处理显性变量, 而是对变量编码, 编码之后可以任意组合成串, 即染色体, 这才是其处理的对象, 因此对 TSP 没有限制; 其次, 在求最优解的过程中, 能够接受各种类型的目标函数, 线性或非线性, 离散或连续, 可微或不可微, 这体现了 GA 的普适性; 第三, GA 是基于面上的一代种群求解, 从上一代 (父代) 的很多染色体 (个体) 开始评估选择, 择优后再经过交叉变异等方式形成下一代 (子代), 保留更多优良个体, 淘汰较劣个体, 几代遗传之后就能基本保证得到全局最优解; 再者, GA 的择优去劣过程, 只以适应度值作为判断, 不需要更多信息, 操作简便易行; 最后, GA 基于概率论的知识而进行遗传操作, 求出较高可信度的最优解, 不排除进一步的改善, 确保其灵活性和可改进性. 以上优点决定了 GA 适合于求解高维、大样本、非线性、非结构性的复杂问题^[4]. 然而, 传统遗传算法 (Traditional Genetic Algorithm, TGA) 严重依赖于参数设置和交叉变异算子, 存在早熟收敛和冗余迭代的缺陷. 近年来, 国内外学者纷纷提出了改进办法以克服 TGA 的缺陷^[5,6]. 受免疫算法中的克隆选择原理、亲和度评判、免疫识别以及免疫系统的学习和记忆能力等生物免疫系统信息处理机理的启发^[7], 本文拟在相关文献基础上, 提出改进免疫遗传算法 (Improved Immune Genetic Algorithm, IIGA). 新算法融合模糊论、改进遗传算子、局部邻域搜索和免疫系统, 运用克隆选择原理获取亲和度较高的抗体, 从而使抗体逼近于抗原 (最优解), 以避免局部优化, 而实现旅行商路径搜寻全局最优. 最后, 以我国 31 个中心城市的地理数据为例测试对比了新算法及 TSP 的寻优性.

2.1 编码方式

本研究的编码采用基于旅行商需要遍历的所有城市的次序, 这也是最常用的编码方式, 以有限的城市数量作为搜索范围, 有助于提高搜索效率. 设

$$S = (1, 5, 4, 3, 2, 6, 7), \quad (2.1)$$

这就可以看成是从城市 1 出发, 依次经过城市 5、4、3、2、6、7, 最终回到起点的一个路径.

2.2 初始种群生成

初始种群的质量对后续求解具有关键作用. 若按随机方式产生初始种群, 难以保证其优良性, 易产生很多非可行解, 从而降低算法效率. 本研究拟综合随机法与贪心法来生成初始染色体种群. 贪心法是指每一个步骤都求得局部最优的方法.

2.3 适应度及亲和度函数

适应度函数是评价个体优劣的指标, 由于本文研究的路径问题是最小化路径长度, 因此,

本研究适应度函数取线性定标

$$f = \frac{\alpha\sqrt{NM}}{T_d}, \quad (2.2)$$

式中 α 为预先设定的常数, N 为城市的数目, M 为包含所有城市的最小正方形的边长, T_d 就是实际行进路径长度. 生物免疫系统利用亲和度评价抗体的好坏. 与抗原结合强的抗体, 即具有较高适应度的抗体具有较高亲和度, 抗体与抗体之间结合程度较弱的抗体, 即浓度较低的抗体具有较高亲和度. 抗体的适应度计算分为两步: 首先对抗体解码, 得到一个旅行商问题的完整解, 然后计算该解的路径长度, 具有较短路径长度的抗体具有较高适应度. 在免疫识别原理中, 免疫识别包括识别“自我”与“非我”, 其中, 识别“自我”的方法通常是扫描整个字符串, 具有相同或相近模式的抗体为“自我”, 而模式差别较大的抗体为“非我”. 然而, 这种“自我”识别方法会极大增加 IA 的计算量. 不难发现, 具有相同适应度的抗体通常具有相同或相似的编码结构, 因此, 可以简单认为具有相同适应度的抗体属于“自我”, 反之适应度不同者属于“非我”. 综上, 提出抗体亲和度计算公式

$$A = \frac{f}{1 + \beta * \ln(1 + D)}, \quad (2.3)$$

式中 β 为一个设定的正常数, D 是和该抗体具有相同适应度的抗体在整个群体里的比例, 即浓度. 可见, 亲和度与适应度成正比, 与浓度呈负相关. 适应度越高, 或者浓度越低, 则亲和度越大; 反之反亦.

2.4 交叉变异操作

基于旅行商遍历城市次序的编码方式, 个体内部基因存在先后关系, 若在交叉变异操作中破坏了这种自然关系, 就有可能产生大量不可行子代个体, 造成算法早熟收敛或冗余迭代. 本研究拟选用优先保留交叉, 操作过程是: 随机产生两个父代个体, 并产生一个等长的1,2随机串; 扫描随机串, 如果第 k 位是 1, 则提取第一个父代染色体最左边的城市号作为子代第 k 位, 如果第 k 位是 2, 则提取第二个父代染色体最左边的城市号, 然后删除两个父代中的该城市号, 重复以上操作, 直到随机串被扫描完. 可见, 该法与映射交叉、次序交叉或循环交叉相比, 能够更好地继承父代优良基因. 当每个抗体按照亲和度比例进行复制后, 子群体中的每个抗体都要经历高概率变异过程. 新算法拟选择平移变异, 是指随机选择插入码和插入点, 进行平移操作. 比如

$$S = (1, 5, 4, 3, 2, 6, 7). \quad (2.4)$$

若随机选取插入码为 6, 插入点为 5 与 4 之间, 则有

$$S' = (1, 5, 6, 4, 3, 2, 7). \quad (2.5)$$

该法与对换变异、目标导向变异等相比, 更好地保持基因间的先后次序, 继承父代的优良性.

2.5 局部邻域搜索

IIGA 拟引入局部邻域搜索, 这是旅行商问题中常用的一种子代择优方法, 有助于进一步加快算法的收敛, 缩短求最优解的运行时间. 其操作内容是: 以交叉变异操作产生的子代个

体为基体, 对每个基因采用右邻基因交换的方法产生新的局部邻域子代个体. 例如

$$S' = (1, 5, 6, 4, 3, 2, 7). \quad (2.6)$$

将基因 2 与其右邻的基因 7 交换, 就能生成一个新个体

$$S'' = (1, 5, 6, 4, 3, 7, 2). \quad (2.7)$$

因此, 局部邻域搜索能产生 $N - 1$ 个局部邻域子代个体, 从中选择具有最大适应度的邻域个体与基体再做比较, 以适应度大者为更新后的子代基体.

2.6 双重选择操作

选择操作是对生物进化论的“适者生存”思想的体现, 越优良的个体具有越大概率进入下一代, 种群性能就会不断优化. 本研究首先采用 GA 的轮盘赌法 (roulette wheel selection), 保证种群中最优个体随机替换掉下一代中的某个体, 这对于算法不断寻优具有关键作用. 接着, 再引入 IA 的克隆选择操作, 根据 2.4 及 2.5 已经执行过的交叉变异操作及局部邻域搜索之后的复制子群体中的所有抗体, 对其亲和度 A 作出降序排列, 随机性选取复制子群体中的若干个高亲和度抗体, 替换父群体中低亲和度抗体.

2.7 疫苗接种

对抗体接种疫苗是指基于先验知识修改个体基因, 使所得个体以较高概率保有高适应度. 疫苗是从先验知识中凝练而成, 它包含的信息量及准确性对免疫算法的功效性影响很大. 对于 TSP, 假设在某一时刻, 旅行商从一个城市出发, 优先考虑的下一站应该是距离当前城市最近且未被走访的城市. 基于此, 在最优的解决方案中必然包含大量相邻城市间距离最短的路径. 因此, 具体设计 TSP 抗体疫苗接种操作为: 假定所有要经过的城市中与城市 c_i 最近的城市是 c_j , 且二者并无直连路径, 即存在下面的路径

$$\begin{cases} l1 = \{\dots, c_{i-1}, c_i, c_{i+1}, \dots\}, \\ l2 = \{\dots, c_{j-1}, c_j, c_{j+1}, \dots\}. \end{cases} \quad (2.8)$$

那么现有的旅行商行进的完整路径为

$$\pi = \{\dots, c_{i-1}, c_i, c_{i+1}, \dots, c_{j-1}, c_j, c_{j+1}, \dots\}. \quad (2.9)$$

免疫算子会将 c_i 的最邻近城市 c_j 调至其紧邻位, 由此旅行商路径变为

$$\pi' = \{\dots, c_{i-1}, c_i, c_j, c_{i+1}, \dots, c_{j-1}, c_{j+1}, \dots\}. \quad (2.10)$$

这是针对某一特定疫苗的免疫接种过程. 当然, 这次调整只是利用先验知识在局部对某个路径作出变更, 这种变更到底对旅行商路径有无优化, 还需要选择机制的进一步判别.

2.8 免疫记忆

运用免疫系统优化抗体群, 表现为整个抗体群历经一个渐变式进化过程. 鉴于交叉、变异及局部邻域搜索的随机性, 有可能在多次迭代后, 难免会丢失先前抗体群中对进化有贡献

的若干优良基因. 在此情况下, 抗原相对于迭代后的群体就可以认为是一个“新抗原”. 免疫记忆机理能够帮助新算法记忆迭代之前的抗体群中的某些优良个体, 若经过若干次迭代后还无法得到全局最优解或最优解未能被改进, 则调用记忆单元的抗体进入到目前抗体群中, 同时启发免疫算子向不同方向搜索, 这使 IIGA 能够不断向最优解方向搜寻, 见图 1.

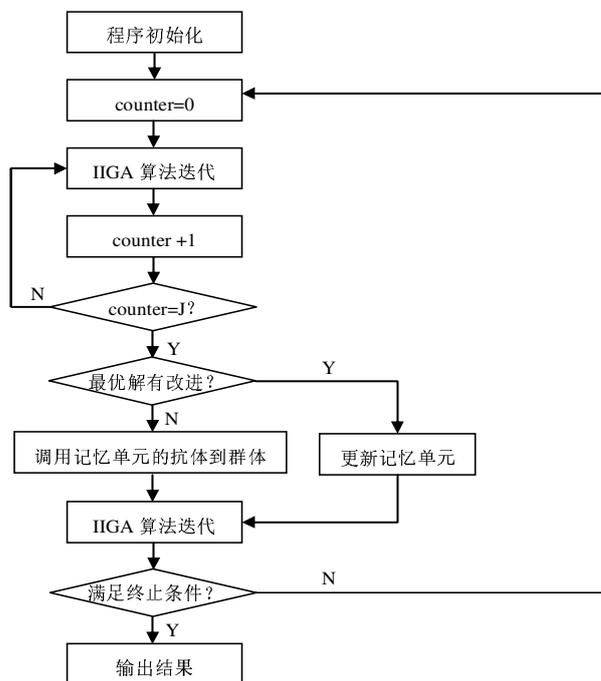


图 1: 免疫记忆机理的逻辑架构图

2.9 算法步骤

步骤 1: 初始化. 设置预定常数、最大迭代次数、交叉概率 P_c 、变异概率 P_m 等参数.

步骤 2: 采用遍历城市排序的编码方法, 综合随机法与贪心法生成初始种群.

步骤 3: 计算初始种群中所有个体 (抗体) 的适应度及亲和度, 基于亲和度降序排列, 按比例选取 m 个抗体复制到复制子群体之中.

步骤 4: 按 P_c 概率执行交叉操作, 按 P_m 概率执行变异操作, 再做局部邻域搜索.

步骤 5: 二次选择操作. 首先根据适应度 $f(S)$, 基于轮盘赌法执行 GA 的选择操作, 确定子代个体, 保证优良染色体能够保留下来; 接着基于亲和度 A 执行 IA 的克隆选择操作.

步骤 6: 实施抗体疫苗接种操作.

步骤 7: 评价最优解的适应度、亲和度及浓度等, 判断当前最优解是否相比前几代父抗体群被改进, 如果有改进, 则将当前抗体群中适应度较高的前 n 个抗体作为记忆细胞, 以更新记忆单元; 反之, 引入免疫记忆算子, 调用记忆单元的 n 个抗体随机替换当前群体的子抗体, 恢复到父代抗体群中 n 个更优良的抗体.

步骤 8: 判别 IIGA 是否达到最大迭代次数, 若是, 则求得最大适应度及亲和度的个体抗体作为最优解. 反之, 则回到步骤 3 继续迭代运行.

3 算法检验与分析

选取我国 31 个中心城市的地理数据用于算法检验. 设定 P_c 为 0.2, P_m 为 0.5, MaxItr 为 1000 次, 鉴于启发式智能算法求解往往依赖于先验知识, 不能严格验证其最优解存在, 因此用最满意解的提法更妥当. 新算法的最满意值为 14268km, 明显优于 TGA, 见表 1、图 2、图 3. 说明改进算法取得了最优解. 这是由于 TGA 的初始种群生成产生了大量不可行解, 在交叉变异过程中缺失全局寻优能力. 而 IIGA 引入局部邻域搜索、疫苗接种以及基于亲和度评判的克隆选择和免疫记忆操作, 确保了子代个体持续不断地向最优解方向收敛及发展.

对比两种算法的极差也能看出, IIGA 的种群离散程度较小, 向最优解的收敛性更好. 不难发现这主要源于以下两点: 其一, IIGA 的初始种群质量优于 TGA, 其可行染色体比例较高, 避免了初期大量不满意染色体的生成. 其二, IIGA 的局部搜索操作、疫苗接种以及基于亲和度评判的克隆选择和免疫记忆操作, 提高了子代抗体向最优解的收敛性. 总之, IIGA 的求最优解能力更强, 比 TGA 更高效.

表 1: 两种算法的检验结果

运行结果	TGA	IIGA	改进程度
平均值 (km)	16917	14657	13.4%
最满意值 (km)	15387	14268	7.3%
最不满意值 (km)	17293	14903	13.8%
极差 (km)	1906	635	66.7%

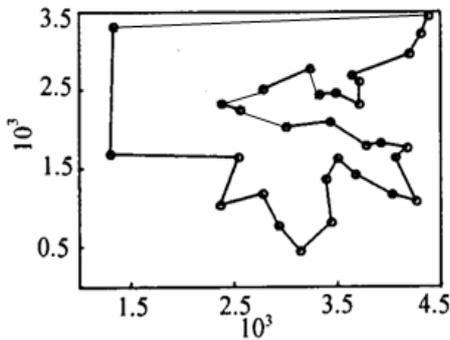


图 2: IIGA 的最满意值

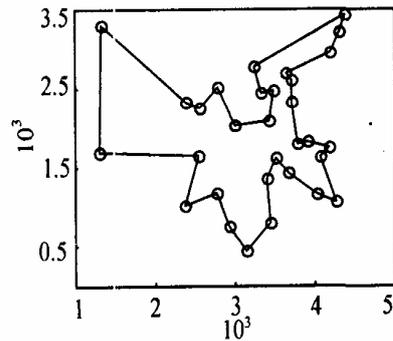


图 3: TGA 的最满意值

4 结束语

本研究将免疫遗传算法应用于 TSP 路径寻优, 更逼真地模拟了现实中各种复杂因素, 研究价值较高. 首先, 利用随机选取与贪心法相结合的方式生成初始种群, 确保了初始种群的优良性能, 克服了 TGA 在随机方式下生成大量非可行解的缺陷; 接着通过设计面向 TSP 的适值函数及亲和度函数, 计算初始种群中所有抗体的适应度及亲和度, 基于亲和度降序排列, 按比例选取 m 个抗体复制到复制子群体之中; 再执行交叉变异以及局部邻域搜索操作, 以确保优良基因, 加速染色体向最优解收敛; 接着引入经典的轮盘赌选择法, 同时基于亲和度

执行 IA 的克隆选择操作; 进一步地, 实施 IA 的抗体疫苗接种操作, 并引入免疫记忆算子; 最后, 给出了最优解判据. 今后的研究可着眼于最优解非可行时初始种群的调整, 或者多目标条件下 TSP 的自适应免疫遗传算法设计, 同时应开发更快收敛的免疫遗传算子.

参 考 文 献

- [1] YANG J H, Wu C G, Lee H P, et al. Solving travelling salesman problems using generalized chromosome genetic algorithm[J]. *Progress in Natural Science*, 2008, 18(7): 887–892.
- [2] YU Hong, YANG Dachun. An incremental rule acquisition algorithm based on rough set[J]. *The Journal of China Universities of Posts and Telecommunications*, 2005, 12(1): 47–52.
- [3] BHATIA A K, BASU S K. Tackling 0/1 knapsack problem with gene induction[J]. *Soft Computing*, 2003, 8(1): 1–9.
- [4] ZHAO X C, Gao X S. Affinity genetic algorithm[J]. *Journal of Heuristics*, 2007, 13(2): 133–150.
- [5] CHENG R, GEN M, TSUJIMURA Y. A tutorial survey of Job-shop scheduling problems using genetic algorithms part ii: hybrid genetic search strategies[J]. *Computers & Industrial Engineering*, 1999, 36(2): 343–364.
- [6] 张国辉, 高亮, 李培根. 基于遗传规划的作业车间调度算法研究 [J]. *控制与决策*, 2008, 23(8): 924–928.
- [7] Liu R C, Jiao L C, Du H F. Clonal strategy algorithm based on the immune memory[J]. *Journal of Computer Science and Technology*, 2005, 20(5): 728–734.

IMPROVED IMMUNE GENETIC ALGORITHM FOR OPTIMAL PATH OF TRAVELLING SALESMAN PROBLEMS

KONG Ling-yi

(School of Economics and Management, Xi'an University of Post and Telecommunications, Xi'an 710061, China)

Abstract: In this paper, an improved immune genetic algorithm for solving optimal path of travelling salesman problem is introduced. Stochastic method and greedy method are combined to produce the initial chromosomes populations. Based on affinity sorting, antibodies are selected to gain replicated subgroup. Roulette and clone selection are introduced in sequence to get high affinity antibodies. Next, vaccination and immune memory operations are implemented to revamp antibody. Using immune memory mechanism, city array generated by the new algorithm is proved to be globally convergent. Numerical experiments prove that the new algorithm is valid.

Keywords: travelling salesman problem; immune genetic algorithm; clone selection; vaccination

2010 MR Subject Classification: 90C39