

TSP 问题的一种快速求解算法

王建忠¹, 唐 红²

(1 中国石油大学, 山东 青岛 266555; 2 重庆邮电大学, 重庆 400065)

摘 要: 针对现有遗传算法收敛速度慢和容易“早熟”两大缺陷, 提出一种新的遗传算法求解 TSP 问题. 该算法融入不同的启发式算法分别作为交叉和变异算子, 并引入外部最优个体集, 以改善局部搜索能力弱的情况和增加群体多样性. 通过对该算法的测试, 论证了其有效性. 以此为基础, 利用递归分治策略, 将上述算法并行化. 实验结果表明, 该算法可以得到理想的解和收敛速度, 具有一定的普适性.

关键词: TSP; 遗传算法; 启发式算法; 并行

中图分类号: TP301.6

文献标识码: A

文章编号: 1000-7180(2011)01-0007-04

A Fast Algorithm for TSP

WANG Jian-zhong¹, TANG Hong²

(1 China University of Petroleum, Qingdao 266555, China;

2 Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: A novel genetic algorithm for TSP was proposed for the defections of traditional GA which are converging very slowly and easy to premature. It merged two heuristics as crossover and mutation respectively which can improve the local search ability of GA, and then imported the exterior best-individual set to increase the diversity of the population. And further tests for the proposed Algorithm are applied. Based on these, the algorithm is parallelized by the recursively divide-and-conquer method. The simulation results indicated that it can get high quality solution and cost litter running time.

Key words: TSP; Genetic Algorithm; heuristic; parallel

1 引言

遗传算法(Genetic Algorithm)是模拟生物在自然环境中的遗传和进化过程而形成的一种自适应全局优化概率搜索算法^[1], 但传统的遗传算法在解决大规模优化问题时显得效率并不高, 所以大多数遗传算法在应用时都对传统遗传算法作了修正和完善.

TSP(旅行商)问题已经被证明是一个 NP-hard 问题, 由于 TSP 问题代表一类组合优化问题, 因此对其近似解的研究一直是算法设计的一个重要课题. 文献[2]也给出了较为满意的结果, 但其算法设计完全是基于并行模式下的, 存在很大的局限性; 因

此, 一个更为重要和普遍的研究分支是在传统的遗传算法基础之上加入某种局部搜索机制和多样性保持策略^[1,3-6]. 文献[5]对遗传算法中的各算子进行改进, 提出了离散轮赌选择、扩展的边匹配交叉以及 Dmutation 变异算子, 改进后的算法在小规模 TSP 测试相对传统遗传算法收敛要快, 但仍然达不到理想的收敛速度. 文献[5]探讨了在传统遗传算法中利用弹性策略改善群体多样性, 从测试结果看, 算法能有效跨过局部最优障碍, 但其操作过程复杂, 牺牲了算法的求解时间. 本文针对现有遗传算法存在的缺陷, 提出一种新的遗传算法, 并针对 TSP 的具体特征, 融入贪心算法和 Lin-Kernighan 启发式算法分别作为交叉和变异算子改善其局部搜索能力, 引入

外部最优个体集保持了群体多样性. 以此为基础, 利用递归分治策略, 将上述算法并行化. 实验结果表明, 该算法可以得到理想的解和收敛速度, 具有一定的普适性.

2 TSP 问题定义

TSP 是典型的组合优化问题, 给定 n 个城市以及各城市之间的距离, 要求找到一条遍历所有城市且每个城市只被访问一次的路线, 并使得总路线距离最短, 或表述为在有 n 个结点的完全图中找出最短的 Hamilton 回路. 其数学描述为

$$\min \sum_{i \neq j} d_{ij} x_{ij} \quad (1)$$

$$\text{s. t. } \sum_{j=1}^n x_{ij} = 1 \quad i = 1, 2, \dots, n \quad (2)$$

$$\sum_{i=1}^n x_{ij} = 1 \quad j = 1, 2, \dots, n \quad (3)$$

$$\sum_{i,j \in S} x_{ij} \leq |S| - 1, 2 \leq |S| \leq n - 2,$$

$$S \subset \{1, 2, \dots, n\}, x_{ij} \in \{0, 1\}, i, j = 1, \dots, n, i \neq j \quad (4)$$

式中, d_{ij} 表示两个城市 i 和 j 之间的距离, $|S|$ 表示集合 S 中元素个数. 式(2)、(3)表示了旅行商恰好经过每个城市一次且仅一次, 式(4)则约束了旅行商在任何一个城市的真子集不形成回路.

针对上述 TSP 问题的具体特征, 提出一种新型混合遗传算法 NHGA(a Novel Hybrid Genetic Algorithm)求解 TSP.

3 NHGA 算法设计

NHGA 算法设计主要包括了四个遗传操作算子设计:

(1) 选择算子. 采用简单易行的比例选择策略, 个体按照式(5)所示的概率规则来进行选择:

$$P\{T_i(\mathbf{X}) = X_i\} = \frac{f(X_i)}{\sum_{k=1}^n f(X_k)} \quad (5)$$

(2) 交叉算子. 结合到具体问题的特征, 为加快遗传算法的收敛速度, 融入贪心算法, 从而形成贪心交叉算子. 主要计算步骤描述如下所示.

Step1: 从 $\mathbf{X}(n)$ 中独立选择 N 个母体, 得到种群 $\mathbf{X}(n) = (X_1, X_2, \dots, X_N)$;

Step2: 随机选取两个个体:

$$X_i(n) = (v_1, v_2, \dots, v_n) \quad (6)$$

$$X_j(n) = (v'_1, v'_2, \dots, v'_n) \quad (7)$$

Step3: 随机生成一个 $1 \sim n$ 中的一个值 v_p , 查找在式(6)、(7)中的位置, 分别记为 p_i 和 p_j ;

Step4: 将 v_p 作为 $X_i(n+1)$ 的第一个基因位, 比较 v_p 到下一位置 $((P_i + 1) \bmod n)$ 和 $((P_j + 1) \bmod n)$ 基因值的距离, 若 $d_i \geq d_j$ 则取式(7) $((P_i + 1) \bmod n)$ 位置的基因值作为 $X_i(n+1)$ 的下一基因, 并在式(6)中删除 $((P_i + 1) \bmod n)$ 位置的基因值. 否则, 取式(6), 并执行类似操作;

Step5: 直至 $X_i(n+1)$ 取满 n 个基因为止.

(3) 变异算子. 在 NHGA 算法设计中将 Lin-Kernighan 启发式搜索算法^[7] 作为变异策略, 以增加算法的局部搜索能力. 变异算子的主要计算步骤如下所示.

Step1: 从 $\mathbf{X}(n)$ 中独立的选择 N 个母体, 得到种群 $\mathbf{X}(n) = (X_1, X_2, \dots, X_N)$;

Step2: 随机选取 1 个个体: $X_i(n) = (v_1, v_2, \dots, v_n)$, 执行 LK 操作, 得到 $X_{i+1}(n) = (v_1, v_2, \dots, v_n)$.

(4) 最优个体集操作算子. 其基本思想: 设置一个最优外部个体集, 存储当前个体中一定数量的相对最优个体, 并对原种群中的相似个体进行隔离. 文献[2]对两个典型的 TSP 测试实例 att532 和 pr2392 的测试中发现: 局部最优解与全局最优解的距离在问题规模的 $1/5$ 左右, 以此作为设计最优个体集操作算子的依据. 具体步骤如下所示.

Step1: 在随机生成的初始群体 $\mathbf{X}(0)$, 依据各个个体的适应度对其进行降序排序, 记忆前 M 个个体 ($M \approx \frac{1}{5}N$);

Step2: 在经历完变异运算之后得到的 N 个个体与 Step1 中保留下来的 M 个个体合并, 得到一个含有 $(M+N)$ 个个体的新群体, 对这 $(M+N)$ 个个体按照式(8)求出两个个体 X_i 和 X_j 之间的海明距离:

$$\|X_i - X_j\| = \sqrt{\sum_{k=1}^N \delta_{ij}^2} \quad (8)$$

式中, $i = 1, 2, \dots, M+N-1, j = i+1, \dots, M+N$; δ_{ij} 是 x_{ik} 和 x_{jk} 的示性函数, 当 $x_{ik} = x_{jk}$ 时, $\delta_{ij} = 1$, 否则 $\delta_{ij} = 0$;

Step3: $\|X_i - X_j\| > L$ 时, 比较个体 X_i 和 X_j 的适应度大小, 并对其中的适应度较低的个体施以罚函数:

$$F_{\min(x_i, x_j)} = \text{Penalty} \quad (9)$$

Step4: 依据 $(M+N)$ 个个体的新适应度对各个个体进行降序排序, 记忆前 N 个个体;

Step5: 转到 Step2, 直至满足停机条件.

相比其他算法,NHGA能得到较为理想的解和收敛速度,其设计思想也具有一定的普适性.

4 NHGA 并行设计

4.1 递归分治策略

分治策略是一种解决大规模复杂问题行之有效的方法,其思想是将较大的问题划分成一个个相互独立的子问题来求解,最后将子问题的解结合或归纳起来作为整个问题的解.

4.2 划分原则

问题的划分可以采取不同的划分原则.本文从简单易操作的角度出发,将“最近邻居”作为划分TSP问题的依据.

在随机产生的初始环游中,取定一个参数值 L ,对任意两点 p_i 和 p_j ,若使得 $d(p_i, p_j) < L$ 时, p_i 和 p_j 连接成一个“超点”,不断往复,直至条件不满足,此时一个“超点簇”便形成了,类似地可以形成若干“超点簇”.即问题被划分为若干个子问题.如图1所示.

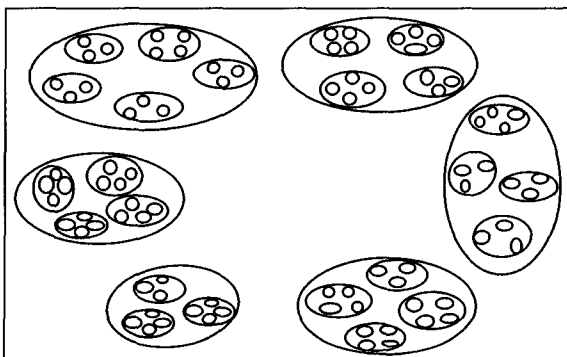


图1 城市之间根据距离形成“超点簇”

4.3 递归分治

如图1所示,在每一个簇内部可以运用混合遗传算法找到Hamilton路,而簇与簇之间可以并发执行.假设问题规模很大,可以利用递归原理,在每一个簇内进一步运用“最近邻居”来划分子问题,而形成一系列子簇.同理,子簇之间可以应用混合遗传算法并发的执行.这在某种程度上解决了可扩展性的问题.如图2所示.

4.4 NHGA 并行化

依据上述问题划分原则,从问题可扩展性的角度出发,将NHGA算法并行化得到PNHGA(Parallel NHGA).该算法的主进程伪代码如下:

```
int main()
{
    divide();
```

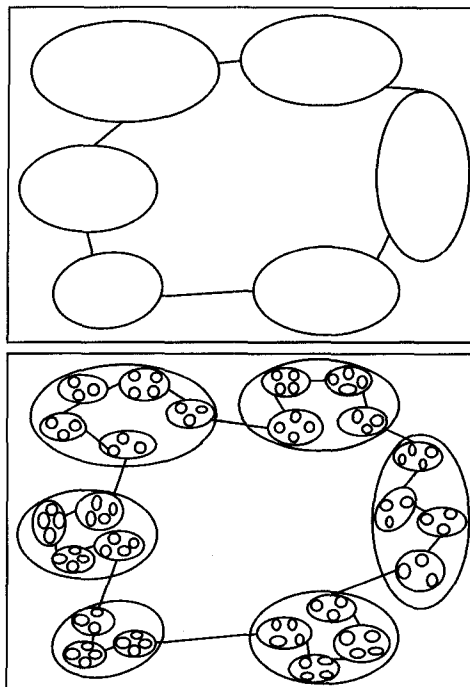


图2 递归分治过程

```
send (partition1);
generation=0;
InitialPopulation();
CountFitness();
FindBestWorst();
Evolve();
while(generation<GENERATION)
{
    generation++;
    SelectOperator();
    CrossoverOperator();
    MutationOperator();
    CountFitness();
    FindBestWorst();
    Evolve();
}
receive (collect result which transfer from the other
processors);
Output ();
return 0;
}
```

从进程代码与之类似,只需要将接受函数(receive)和发送函数(send)交换位置即可.

在对各个簇中所得的环游进行合并时,先找到簇内最长的边,并将其去掉便得到簇内的最短路径,其中路径的两个端点构成了簇内的两个“边界”城市.通过在簇与簇之间运行NHGA,可得到簇之间

的最短路径,从而完成整个 TSP 问题的结果归并。

5 实验测试结果

实验在两台计算机(硬件环境:Pentium D 2.8 GHz,512 MB DDR,软件环境:Red Hat Linux 9.0,MPICH2)下执行,定义并行加速比(Speed Up)如下。其中, T_p 是 NHGA 在并行模式下的运行时间,而 T_s 是 NHGA 在串行模式下的运行时间。

$$\text{Speed Up} = \frac{T_s}{T_p} \quad (10)$$

测试结果如表 1 所示,从表 1 中可以看到,本文提出的算法(PNHGA)具有求解质量性能高、运行时间短的特点。

表 1 并行实验测试数据

TSP	Opt	Best Result	Average Result	Running Time / s	Speed Up
Burma14	33.23	33.23	33.23	1.13	1.83
Bayg29	9 074	9 074	9 074	2.07	1.78
Att48	33 524	33 524	33 524	3.43	1.81
Eli51	426	426	426	3.67	1.79
Eli101	629	629	634	8.77	1.80
Ch130	6 110	6 110	6 225	11.72	1.82
Tsp225	3 919	3 919	4 376	24.26	1.81
Pcb442	50 778	50 778	53 481	31.05	1.74
Rat575	6 773	6 775	6 894	40.34	1.73
Pr1002	259 045	259 048	268 470	92.55	1.69
Fl1440	20 127	20 127	24 158	62.65	1.68
Fl1557	22 249	22 249	22 249	138.12	1.66

在实际的测试中我们发现,对于测试实例 Ch130,如果在串行模式下运行,极易陷入局部最优。但如果运行在并行模式下,由于使用了分治的策略,就不会存在陷入局部最优的缺点。所以,递归分治的策略能在某种程度上解决 TSP 的可扩展性问题。

这里从算法解的质量来对比算法性能。PNHGA 与其他算法解的质量的比较如表 2 所示,其中 NG 表示未给出。

表 2 求解质量对比

TSP	Burma14	Bayg29	Att48	Eli51	Ch130
Opt	33.23	9 074	33 524	426	6 110
PNHGA	33.23	9 074	33 524	426	6 110
SI ^[8]	31.88	9 336	34 707	451	NG
GEA ^[4]	NG	NG	33 604	427	6 227

从表 2 看出,PNHGA 相比其他遗传算法在时间和求解质量上均占有显著的优势,并且相比其他算法更具扩展性。

6 结束语

本文通过对传统遗传算法的改进,设计了一种新的遗传算法 NHGA。对 TSP 典型实例的测试,得到如下结论:融入了启发式算法能使得算法局部搜索能力显著增强,能快速收敛;外部最优个体集的引入又很好地保持群体多样性,逃离局部最优。最后利用递归分治技术将 NHGA 并行化,有效降低了算法计算时间,这为今后遗传算法在具体应用中的设计提供了一条新思路。

参考文献:

- [1] Hung Dinh Nguyen, Kunihiro Yamamori, Moritoshi Yasunaga, et al. Implementation of an effective hybrid GA for large-scale traveling salesman problems[J]. IEEE Transactions on Systems, Man, And Cybernetics, 2007,37(1): 92-99.
- [2] 邹鹏,周智,陈国良,等. 求解 TSP 问题的多级规约算法[J]. 软件学报,2003,14(1): 35-42.
- [3] Tsai C F, Tsai C W, Tseng C C. A new approach for solving large traveling salesman problem[C]// Proceedings of the 2002 Congress on Evolutionary Computation. Honolulu, HI: IEEE, 2002(8):1636-1641.
- [4] Firbes J. Burkowski, proximity and priority: applying a gene expression algorithm to the traveling salesperson problem[C]// International Parallel and Distributed Processing Symposium. Nice: IEEE, 2003,30(5): 803-816.
- [5] 周涛. 基于改进遗传算法的 TSP 问题研究[J]. 微电子学与计算机, 2006, 23(10): 104-110.
- [6] 王雷. 基于并行遗传算法的弹性 TSP 问题研究[J]. 微电子学与计算机, 2005, 22(8): 130-133, 137.
- [7] Lin S, Kernighan B W. An effective heuristic algorithm for the traveling-salesman problem[J]. Operations Research, 1973,21(2):498-516.
- [8] He Xiaoxian, Zhu Yunlong, Hu Hechun. A new algorithm for TSP based on swarm intelligence[C]// Proceeding of the 6th World Congress on Intelligent Control and Automation. Dalian: IEEE, 2006:3241-3244.

作者简介:

王建忠 男,(1973-),博士生,讲师,中国计算机学会会员。研究方向为并行计算、仿真技术。

唐红 女,(1957-),博士,教授。研究方向为网络管理、并行计算。