

# 一种求解 TSP 问题的单亲遗传算法<sup>\*</sup>

王 斌<sup>1</sup> 李元香<sup>1</sup> 王 治<sup>2</sup>

(武汉大学软件工程国家重点实验室 武汉 430072)<sup>1</sup> (上海贝尔阿尔卡特股份有限公司 上海200070)<sup>2</sup>

## A Partheno Genetic Algorithm for Traveling Salesman Problem

WANG Bin<sup>1</sup> LI Yuan-Xiang<sup>1</sup> WANG Zhi<sup>2</sup>

(State Key Laboratory of Software Engineering, Wuhan University, Wuhan 430072)<sup>1</sup> (Alcatel Shanghai Bell Inc.)<sup>2</sup>  
(wzfq@yeah.net)

**Abstract** In this paper, a kind of Partheno Genetic Algorithm(PGA)based on Path Representation scheme is proposed for solving Traveling Salesman Problem(TSP). This algorithm employs only mutation and selection operators to produce the offspring, instead of traditional crossover operator. A specific mutation operator is designed combining the insertion operator with inversion operator, which ensures its strong searching capability. This algorithm simulates the recurrence of nature evolution process, while providing fewer control parameters. Experiments based on Chinese 144 cities(CHN144)and 7 instances selected from TSPLIB are used to test the performance of this algorithm. They prove that it can reach the satisfying optimization at a faster speed. Especially, for the CHN144, the best path it finds is better than any other available one.

**Keywords** TSP, Combinatory operator, Evolution cycle, Partheno-genetic algorithm

## 1 前言

TSP 问题可描述为:给定一个城市的集合,寻找一条从集合中的某个城市出发,访问每个城市一次且仅一次,最后回到出发点的最短路径。这已被证明是一个 NP 难解问题<sup>[1]</sup>。求解 TSP 问题,遗传算法通常采用序号编码和非序号编码两种解表达方式。其中序号编码相对简单直接,其代表性的有“邻接表达”、“普通表达”和“路径表达”等几种编码方式,后者是最自然的表达方式。序号编码方式的杂交算子难于设计,杂交后解的合法性是需着重考虑的问题。虽然目前已提出了一些基于路径表达的杂交算子,如 PMX<sup>[2]</sup>、OX<sup>[3]</sup>和 CX<sup>[4]</sup>,但普遍计算额外开销很大,而且杂交算子的使用对群体的多样性存在很大影响<sup>[5]</sup>,容易使算法过早收敛。

本文提出了一种只使用变异算子和选择算子繁殖后代的单亲遗传算法(PGA),它使用“路径表达”作为染色体编码方式,在插入算子和倒位算子的基础上,设计了一种组合算子作为一个 PGA 算法的主搜索算子。相对传统遗传算法在进化的每一代都要进行个体的选择和淘汰操作;在算法的框架设计上,PGA 算法模拟了自然界存在的生物周期性进化、进化退化并存的现象,每进化若干代后才进行一次个体的选择淘汰并重组群体,以保证群体的稳定进化。计算实例证明,PGA 算法具有高的求解质量和求解效率,尤其是在求解中国144个城市(CHN144)TSP 问题时获得了优于目前最好解<sup>[6,7]</sup>的最短路径。

## 2 PGA 算法的相关概念

**定义1** 组合算子是为 PGA 算法新设计的一种算子,它有4步操作,其操作方式如下:假设父体是 S。

SETP1:从 S 中随机地取两个城市 c 和 d。

SETP2:如果将 c 和 d 间的城市(含 c 和 d)反序能使 S 的适应值增加,则将 c 和 d 间的城市反序,同时修改 S 的适应值。

SETP3:在 S 中随机地选择一个城市 c,再在 S 中随机地选取另一个城市 d。

SETP4:如果将 c 插入到 d 的前面能使 S 的适应值增加,则将 c 插入到 d 的前面,同时修改 S 的适应值。

从组合算子执行过程看,组合算子具有这样几个特点:(1)该算子作用后,所得后代的适应值大于等于父体的适应值。(2)当 SETP2和 STEP4中的条件都不满足时,组合算子等同于复制操作,反之则是变异操作。(3)该算子执行完毕,也计算出生成的子个体的适应值。

**定义2** 插入算子定义为如下操作步:假设父体是 S。

SETP1:在 S 中随机地选择一个城市 c,再在 S 中随机地选取另一个城市 d。

SETP2:将 c 插入到 d 的前面,同时修改 S 的适应值。

该算子不同于组合算子的是,它能使个体的适应值降低。

**定义3** 让群体进化指定的代数以后(只使用组合算子进行遗传操作产生下一代个体),然后通过竞赛选择对群体进行重组(保留历史最优个体),最后对重组后的群体中的每个个体使用插入算子进行遗传操作产生下一代个体,这整个过程称为一个进化周期。

## 3 PGA 算法的设计思想及描述

### 3.1 PGA 算法的设计思想

PGA 算法框架的设计是模拟自然界演化的周期性的特点。自然界的演化过程往往是进化和退化交替进行的,表现出

<sup>\*</sup>国家自然科学基金(69703011)资助项目,教育部骨干教师资助计划。王 斌 硕士研究生,研究方向为演化计算和并行计算。李元香 教授,博士生导师,研究方向为演化计算和并行计算。王 治 硕士,研究方向为演化计算,智能网管。

周期性的特点,它是一个循环往复的过程,但不是一种简单的回复,而是总的趋势上呈现的上升的特征,直线性的演化图景原则上是不会成功的,遗传算法是对自然过程的模拟,也应模拟这一现象。本文提出的 PGA 算法让群体的进化具有周期性,本文提出的组合算子有使群体不发生退化,保持进化的趋势的特点,插入算子有可能使群体发生退化的特点,所以 PGA 算法对一个进化周期的设计是:首先使用组合算子作为遗传操作算子对群体进化指定的代数(反映为群体的平均适应值升高或不降),再对群体进行重组(同时保留历史最优个体),最后使用插入算子对群体进行一代演化,使群体发生暂时的退化(反映为群体的平均适应值降低),再转入下一个进化周期。PGA 算法就是通过若干个这样的进化周期,最后找到最优解的,反映出群体进化的螺旋式上升的趋势。

遗传操作算子对于遗传算法能否成功地进行优化求解也是至关重要的。组合算子是 PGA 算法的主搜索算子,插入算子主要目的在于生成好的基因片段,倒位算子主要目的在于选择染色体中基因片段间的好的连接方式,但这两种算子对个体也具有破坏性,将这两种变异算子组合起来使用并加以控制,及通过这两种算子产生的变异能使个体的适应值提高时,则执行这两种变异操作否则不执行,这样就能快速地改进个体,提高算子的搜索能力。后面实例也验证了本文提出的组合算子具有优良的搜索性能。

### 3.2 PGA 算法描述

procedure PGA-TSP;

begin

    随机初始化群体 P; 计算 P 中个体的适应值;

    optimal\_indivi := P 中最优的个体; gen := 0;

    while gen < LS do //LS 是进化周期数

        begin

            // 一个进化周期开始

            k := 0;

            while k < EG do //EG 是一个进化周期所含的进化代数

                begin

                    对 P 中的每一个个体使用组合算子进行遗传操作,用生成的子个体替换父个体;

                    k := k + 1;

                end;

                // 保留最优个体

                if P 中最优个体好于 optimal\_indivi then optimal\_indivi := P 中最优的个体;

                // 在进入下一个进化周期前对群体进行重组

                S := {按竞赛度为 2 的竞赛选择策略从 P 中选择的 N-1 个个体}; //N 为群体规模

                P := S + {个体 optimal\_indivi};

                对 P 中的每一个个体使用插入算子进行遗传操作,用生成的子个体替换父个体;

                gen := gen + 1;

            end;

        end;

PGA 算法只使用了 N、LS、EG 共 3 个参数, N 为群体规模, LS 是进化周期数, 它用来控制算法的终止, EG 是一个进化周期所进化的代数。

### 4 实例验证

选用通用的 TSPLIB (ftp:ivr.uni-hei-delberg.de/pub/tsplib) 中的 7 个实例 和文 [6] 中提到的中国 144 个城市 (CHN144), 对 PGA 算法进行测试, 问题的规模从 52 个城市到 226 个城市。PGA 算法的参数选取为: N=16, EG=18000。

值得注意的是, 在算法执行前, 要计算每两个城市间的距离, 有两种计算方式, 第一种是由城市的坐标值计算出城市间的距离时 4 舍 5 入取整, 本文称之为基于整数运算, 第二种是保留后面的小数位数, 本文称之为基于浮点运算。应该指出的是, 一个算法按整数运算计算出来的城市间的距离值去求最优路径, 而另一个算法按浮点运算计算出来的城市间的距离

值去求最优路径, 这两种算法所得结果是不能相互比较的, 因为算法执行前对城市间的距离值的计算方法不一样, 计算出来的值也不一样。TSPLIB 给出的是基于整数运算的最优路径值, 文 [7] 求得的是基于浮点运算的最优路径值, 但文 [7] 把求得结果和 TSPLIB 给出的结果进行比较从而得出算法求得的结果好于已知最优解, 显然是不妥的。

本文对每个实例分别基于整数运算和基于浮点运算进行实验。基于整数运算, 对每个实例做连续 25 次实验, 每次实验让算法运行到找到 TSPLIB 给出的最优解 (CHN144 实例按文 [8] 给出的最优解 30347) 时中止, 记录算法运行的时间, 表 1 给出所用的最短时间, 最长时间和平均时间。对每个实例, 表 2 给出算法基于浮点运算所得结果, 并和文 [7] 给出的其中几个实例的浮点运算结果进行比较 (其他几个实例的浮点运算结果未见发表)。图 1 是文 [7] 给出的 chn144 最优路径图, 图 2 是 PGA 算法求得的最优路径图。浮点计算保留小数点后面 11 位, 时间单位为秒。(使用 PIII 500/256M 微机, 采用 PASCAL 语言编程)

表 1 PGA 算法求得已知最优解所用的时间(基于整数运算)

实例	已知最优解	最短时间	最长时间	平均时间
berlin52	7542	0.19	1.122	0.353
St70	675	0.381	21.601	7.797
pr76	108159	0.391	5.277	1.625
rd100	7910	0.801	20.360	5.707
pr144	58537	2.133	13.339	5.4098
Chn144	30347	2.835	137.110	92.615
Krob150	26130	13.350	223.540	136.154
pr226	80369	30.364	250.660	115.132

表 2 PGA 算法基于浮点运算求得的最优解

实例	PGA 算法求得的最优解	文 [7] 给出的最优解
berlin52	7544.3659019039	/
St70	677.1096092748	677.1096
pr76	108159.4382741325	/
rd100	7910.3962102149	/
pr144	58535.22176137	/
Chn144	30353.86099652386	30354.3062
Krob150	26127.3578887393	26127.3579
pr226	80370.2570528723	/

注: 文 [6] 采用并行演化算法给出的 chn144 浮点运算的解也为 30354.3062, “/” 表示文 [7] 没有使用这个实例, 也没有其他的文献给出该实例的基于浮点运算的最优解。

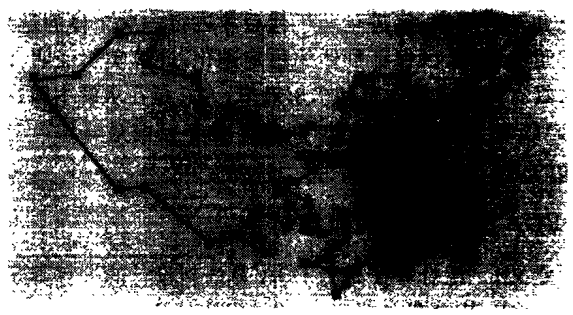


图 1 文 [7] 给出的 CHN144 最优路径图 (最短路径值 = 30354.3062)

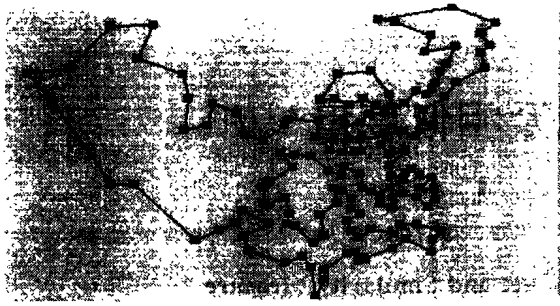


图2 PGA 算法求得的 CHN144 最优路径图(最短路径值=30353.86099652386)

**结论** 实验说明使用文中提出的组合算子作为主要遗传操作算子,和采用具有进化周期性特点的算法框架的 PGA 算法具有较高的求解质量,连续25次实验,每次算得的结果都能达到已知最优解,所用的时间也较短,100个以内的城市平均也就几秒钟,200个左右的城市平均2分钟左右,这个时间远少于基于杂交操作的遗传算法所用的时间。值得指出的是,对于中国144个城市的 TSP 问题,PGA 算法基于浮点运算求得的最优解明显好于已知基于浮点运算求得的最优解(文[6,7]给出的解)。

从实验看到,求解52个到226个城市的 TSP 问题所用的群体规模都是16个个体,而其他的遗传算法使用的群体规模都在50个个体以上,这说明 PGA 算法能有效地维持群体的多样性,从而不会发生算法的过早收敛,也说明文中提出的组合算子具有优良的搜索性能。PGA 算法使用了很少的参数,如果除开群体规模和算法的终止条件,则只使用了一个参数及

一个进化周期所进化的代数 EG。PGA 算法是所有的进化周期都固定为相同的进化代数 EG,下一步对算法的改进目标是使 EG 能自适应地调节,及一个进化周期的结束不由固定的参数来控制,一方面减少算法的参数,使算法更容易控制,一方面进一步加快算法的求解速度。

#### 参考文献

- 1 Garey M, Johnson D. Computers and Intractability. W. H. Freeman, San Francisco, 1979
- 2 Goldberg D E, Lingle R. Alleles, loci, and the Traveling Salesman Problem. In: Proc. of an Intl. Conf. on Genetic Algorithms and Their Applications, 1985. 154~159
- 3 Davis L. Job Shop Scheduling with Genetic Algorithms. In: Proc. of an Intl. Conf. on Genetic Algorithms and Their Applications, 1985. 136~140
- 4 Smith D. Bin Packing with Adaptive Search. In: Proc. of an Intl. Conf. on Genetic Algorithms and Their Applications, 1985. 202~206
- 5 徐宗本,高勇. 遗传算法过早收敛现象的特征分析及其预防. 中国科学(E辑), 1996, 26(4): 364~375
- 6 Jiang Rui, Szeto K Y, Luo Yu-pin, Hu Dong-Cheng. A path-splitting scheme based distributed parallel genetic algorithm for large traveling salesman problems. In: proc conf. on Intelligent Information processing (WCC2000-IIP2000), 2000. 478~485
- 7 吴斌,史忠植. 一种基于蚁群算法的 TSP 问题分簇求解算法. 计算机学报, 2001, 24(12): 1328~1333
- 8 郭涛. 演化计算与优化:[武汉大学博士学位论文]. 1999. 50~51

(上接第22页)

更多,尤其是当请求服务器对象的种类和数量增大时, LARD 算法通过动态划分 SAD 工作环境将同种对象请求分派到相同服务器上,其性能提高更加明显。由此我们可以得出结论,使用面向集群改进 LARD 算法的 SAD/SMD 进程管理服务具有负载均衡、可动态扩展以及提高系统处理能力等特点。

**结束语** 在部署了应用服务器集群的系统中,引入进程管理服务后使得服务器位置透明化:客户方通过统一的入口访问系统,由系统来定位具体的服务器;进程管理可以在服务器集群间迁移,管理 ORB 域中服务器的生命周期和状态。同时,基于 LARD 算法的定位策略在保证服务器运算稳定的情况下,较大地提高了集群系统的性能。通过进程管理服务,系统可以随时了解服务器的运行信息,并根据系统的运行情况动态增删服务器,动态调整系统负载均衡。集成了该 SAD/SMD 进程管理服务的 CORBA 平台,目前在省级电力系统中得到了广泛的应用。我们下一步的工作是在现有进程管理服务框架的基础之上,深入研究分布计算环境中 CORBA 集群冗余服务容错以及 SAD/SMD 和名字服务集成的问题。

#### 参考文献

- 1 Object Management Group. The Common Object Request Broker: Architecture and Specification 2.3, June 1999
- 2 Niderost B U, Gommans L, Kemmerling G, et al. Objectivity/Corba distributed database performance on a gigabit Sun-ultra-10 cluster. Real Time Conference, 1999. 442~445
- 3 Craske G, Tari Z. A property-based clustering approach for the CORBA Trading Service. Distributed Computing Systems, 1999. 517~525
- 4 Miller J A. CORBA-Based Run-Time Architectures for Workflow Management Systems. Journal of Database Management, Special Issue on Multidatabases, July 1996
- 5 Chien S-Y, Tsotras V J, et al. Efficient Management of Multiversion Documents by Object Referencing. The VLDB Journal, 2001. 291~300
- 6 Aron M, Sanders D. Scalable Content-Aware Request Distribution in Cluster-based Network Servers. USENIX 2000 Annual Technical Conference, San Diego, CA, June 2000
- 7 Vivek S, et al. Locality-Aware Request Distribution in Cluster-based Network Servers. Architectural Support for Programming Languages and Operating Systems, 1998