

离散自由搜索算法

郭鑫¹, 孙丽杰¹, 李光明¹, 江开忠^{2*}

(1. 上海工程技术大学 化学化工学院, 上海 201620; 2. 上海工程技术大学 基础教学学院, 上海 201620)

(* 通信作者电子邮箱 kzjiang@sues.edu.cn)

摘要:针对离散组合优化问题, 给出一个自由搜索的算法。但是仅仅通过自由搜索算法求得的解, 往往存在交叉现象, 针对这个问题提出将离散自由搜索算法和交叉消除相结合的算法, 这样不仅大大地提高了自由搜索算法运算过程的收敛速度, 而且较大程度地提升了结果的质量。利用旅行商问题(TSP)标准库中的测试数据对所提算法进行了验证, 结果表明该算法比遗传算法性能提高了约1.6%。

关键词:旅行商问题; 智能算法; 自由搜索; 交叉消除

中图分类号:TP301 **文献标志码:**A

Discrete free search algorithm

GUO Xin¹, SUN Lijie¹, LI Guangming¹, JIANG Kaizhong^{2*}

(1. College of Chemistry and Chemical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China;

2. College of Fundamental Teaching, Shanghai University of Engineering Science, Shanghai 201620, China)

Abstract: A free search algorithm was proposed for the discrete optimization problem. However, solutions simply got from free search algorithm often have crossover phenomenon. Then, an algorithm free search algorithm combined with cross elimination was put forward, which not only greatly improved the convergence rate of the search process but also enhanced the quality of the results. The experimental results using Traveling Saleman Problem (TSP) standard data show that the performance of the proposed algorithm increases by about 1.6% than that of the genetic algorithm.

Key words: Traveling Saleman Problem (TSP); intelligent algorithm; free search; cross elimination

0 引言

旅行商问题(Traveling Saleman Problem, TSP)是一类复杂的组合优化问题,它是世界七大数学难题之一,是NP完全问题中的一种,具有很高的计算复杂性,广泛应用于交通、电路设计、机器人控制等领域。由于TSP的NP-Hard特性,不能找到有效的算法在多项式时间内求得问题的最优解,而只能采用近似算法以在有效时间内求得问题的满意解。一般而言,求解TSP的方法可分为精确算法和启发式算法。精确算法在理论上能保证求得最优解,但是计算时间呈指数复杂度,难以解决大规模问题。启发式方法可分为独立于问题的算法和问题相关的算法。独立于问题的算法,如遗传算法(Genetic Algorithm, GA)^[1]、模拟退火(Simulated Annealing, SA)^[2]、蚁群优化(Ant Colony Optimization, ACO)^[3]、禁忌搜索(Tabu Search, TS)^[4]等,虽然理论上能够求得最优解,但收敛缓慢,对于大规模问题计算时间比较长。而问题相关的启发式算法,如最近邻搜索(Nearest Neighbor, NN)^[5]、两、三段优化(2-Opt, 3-Opt)^[6-7]、LK优化(Lin-Kernighan, LK)^[8]等在求解TSP时效率非常高,不过由于此类算法过于依赖问题的特殊性,算法常常陷于局部最优解。

自由搜索(Free Search, FS)算法是由Penev和Littlefair在2005年基于“以不确定应对不确定,以无穷尽应对无穷尽”的思想^[9],提出一种新的高效群体智能优化算法,在国内周

晖等^[10-12]的研究非常深入,该算法借鉴动物个体存在的不同嗅觉、机动性等特征,提出了个体灵敏度和邻域搜索半径的概念,并利用动物释放信息素的机理,通过信息素和灵敏度的比较确定各自新的寻优位置,经过不断的搜索找到最优解。FS算法吸收了PSO、ACO、GA等算法的优点,在兼顾局部搜索和全局搜索、提高鲁棒性和自适应等方面有较大的创新。与粒子群算法和遗传算法相比不同的是,在FS算法中,动物群体的搜索行为没有被限定,而是通过概率描述,其原理简单,需要确定的参数不多,操作也很简便,并且具有较好的寻优能力。

由于FS算法刚刚兴起,有关FS算法以及其应用的报道还局限于连续解空间,鲜有关于FS算法应用于离散型解空间。本文将FS算法进行了改进,使其应用于一类离散型解空间,并且用于解决TSP,取得了很好的效果。

1 自由搜索算法

对优化数学模型:

$$\max f(\mathbf{x})$$

$$\text{s. t. } g_i(\mathbf{x}) \geq 0 (i = 1, 2, \dots, k); h_i(\mathbf{x}) = 0 (i = 1, 2, \dots, l)$$

其中 $\mathbf{x} = (x_1, x_2, \dots, x_n)$ 。

引入惩罚函数,可以转换为如下域约束的优化模型:

$$\max f(\mathbf{x})$$

$$\text{s. t. } x_{j_{\min}} \leq x_j \leq x_{j_{\max}} (j = 1, 2, \dots, n)$$

其中 $\mathbf{x} = (x_1, x_2, \dots, x_n)$ 。

收稿日期:2012-12-24;修回日期:2013-02-26。

基金项目:上海市教委学科建设项目基金资助项目(11XK11);上海工程技术大学内涵建设项目(nhky-2012-13)。

作者简介:郭鑫(1990-),男,山东平邑人,主要研究方向:生物数据挖掘;孙丽杰(1990-),女,山东诸城人,主要研究方向:医药数据挖掘;李光明(1990-),男,河南南阳人,主要研究方向:医药数据挖掘;江开忠(1965-),男,四川乐至人,副教授,博士,主要研究方向:知识发现、软硬件协同设计、搜索算法。

运用FS算法求解域约束优化模型步骤^[10]如下:

1) 初始化。

设初始种群 $S = \{x_1, x_2, \dots, x_N\}$, 第 i 个个体 $x_i = \{x_{i1}^{(0)}, x_{i2}^{(0)}, \dots, x_{in}^{(0)}\}$, 第 j 维分量 $x_{ij}^{(0)}$ 由下式之一产生:

$$x_{ij}^{(0)} = x_{jmin} + (x_{jmax} - x_{jmin}) \cdot \text{random}_{ij}(0,1)$$

$$x_{ij}^{(0)} = a_{ij}; a_{ij} \in [x_{jmin}, x_{jmax}]$$

$$x_{ij}^{(0)} = c_j; c_j \in [x_{jmin}, x_{jmax}]$$

其中: x_{jmin}, x_{jmax} 为变量 x_j 的搜索下界与上界; $\text{random}_{ij}(0,1)$ 为 $[0,1]$ 的随机数; a_{ij}, c_j 为常数。

2) 寻优搜索。

在算法搜索过程中, 每个个体在搜索半径 $R_{ij} \in [R_{min}, R_{max}]$ (第 i 个个体的第 j 维) 内随机行走 T 步, 更新个体位置。其进化过程中第 t 步位置如下:

$$x_{ij}^{(t)} = x_{ij}^{(0)} - \Delta x_{ij}^{(t)} + 2\Delta x_{ij}^{(t)} \cdot \text{random}_{ij}^{(t)}(0,1)$$

$$\Delta x_{ij}^{(t)} = R_{ij} \cdot (x_{jmax} - x_{jmin}) \cdot \text{random}_{ij}^{(t)}(0,1)$$

其中: $x_{ij}^{(t)}$ 为更新后的个体位置分量; t 为当前步, $t = 1, 2, \dots, T$, T 为每次行走的限制步数, $\Delta x_{ij}^{(t)}$ 为修改策略。

信息素定义为:

$$P_i = \max(f_i)$$

其中: $f_i = \max(f_i^{(t)})$ 是第 i 个个体直到第 t 步的最佳目标函数值, $f_i^{(t)} = f(x_i^{(t)})$ 是第 i 个个体第 t 步的目标函数值, $x_i^{(t)} = \{x_{i1}^{(t)}, x_{i2}^{(t)}, \dots, x_{in}^{(t)}\}$ 。信息素最大值和最小值为 P_{max} 和 P_{min} 。

灵敏度定义为:

$$l_i = l_{min} + \Delta l_i; \Delta l_i = (l_{max} - l_{min}) \cdot \text{random}_i(0,1)$$

其中: l_{min} 和 l_{max} 是灵敏度的最大值和最小值, 并设定 $l_{max} = P_{max}$, $l_{min} = P_{min}$ 。

在一轮搜索结束后, 确定下一轮搜索起点为:

$$x_{ij}^{(0)'} = \begin{cases} x_{ij}^{(0)}, & P_i < l_i \\ x_{ij}^{(0)}, & P_i \geq l_i \end{cases}$$

3) 终止搜索判断。

若满足下列条件之一, 迭代终止:

① $f_{max} \geq f_{opt}$, 目标函数达到目前函数全局最优解。 f_{max} 为寻优的结果; f_{opt} 为可接受的目标函数值。

② $g \geq G$, 搜索代数达到终止代数。 G 为迭代次数的限制数值, g 为当前迭代次数。

③ $(f_{max} \geq f_{opt}) \parallel g \geq G$ 。

未知的修正策略决定于勘查边界和个体在第 j 维解空间中的可变邻域值, 而与之之前的个体和群体经历无关, 并且允许以大于 0 的概率在解空间中进行搜索, 从而避免陷入局部最优。

灵敏度是FS算法的一个重要参数, 个体可以搜索任何区域, 可以在其自身的当前最佳邻域周围, 也可以在其他个体或群体发现的当前最佳邻域周围。增大灵敏度, 个体将趋近整个群体的当前最佳值, 局部搜索; 减小灵敏度, 个体可以在其他邻域进行搜索, 全局搜索。灵敏度是自由搜索算法中的一个创新, 在其他算法中没有类似的概念。

以上即为FS算法在连续型解空间的思想, 但目前还没有FS算法在离散型解空间问题的研究的报道; 连续FS算法的变量 x 是在实空间内连续取值, 它不能应用于在实空间的离散取值情况。本文将连续的FS算法进行了改进, 实现了FS算法在解决一类离散型解空间问题上的突破, 并将此方法应用于解决TSP这一NP-Hard问题上, 取得了很好的效果。

2 离散FS算法

对离散优化数学模型:

$$\max f(x)$$

$$\text{s. t. } x = (x_1, x_2, \dots, x_n), x_i \neq x_j,$$

$$x_i, x_j \in X = \{a_1, a_2, \dots, a_n\} (i, j = 1, 2, \dots, n)$$

约束条件说明, $x = (x_1, x_2, \dots, x_n)$ 中的各分量就是 $X = \{a_1, a_2, \dots, a_n\}$ 中所有元素的一个全排列。

利用FS算法思想并在FS算法的基础上, 求解上述离散优化数学模型如下:

1) 初始化。

对 $X = \{a_1, a_2, \dots, a_n\}$ 中的所有元素随机 N 次全排列, 得到初始种群 $S = \{x_1, x_2, \dots, x_N\}$ 。

2) 寻优搜索。

对于任意一个个体序列 $x_i = (x_{i1}, x_{i2}, \dots, x_{in})$, 随机产生一个长度 L_i ($1 < L_i \leq L_{max} < n$) 以及起点 x_{im} ($1 \leq m \leq n - L_i$) 的一段连续子列:

$$s_i = (x_{im}, x_{im+1}, \dots, x_{im+L_i-1})$$

对该连续子列顺序随机重新排序得到新的序列:

$$s'_i = (x'_{im}, x'_{im+1}, \dots, x'_{im+L_i-1})$$

从而个体 $x_i = (x_{i1}, x_{i2}, \dots, x_{in})$ 变为:

$$x'_i = (x'_{i1}, x'_{i2}, \dots, x'_{in})$$

信息素和灵敏度定义同上。

在一轮搜索结束后, 确定下一轮搜索个体为:

$$x'_i = \begin{cases} x'_i, & P_i < l_i \\ x_i, & P_i \geq l_i \end{cases}$$

3) 终止搜索判断。

终止判断条件: $g \geq g_{min}$, 即当进行 g_{min} 步的搜索时停止搜索。

得到此时的 $\max\{f(x)\}$ 即为用自由搜索算法求得的最大值, 其对应的个体序列即为最优序列。

3 路径交叉检测与消除

实验发现, 启发式算法在求解规模较大组合优化问题比如TSP时, 得到的结果的优劣具有一定的波动, 其在限定时间内得到的解所对应的组合排列常常存在在一定意义下的交叉情况^[13], 如图1所示。图1(b)路径的长度要比图(a)路径的总长度短, 因此存在交叉的解一定不是最优解, 消除交叉可明显提高解的质量, 但用上述已有的近似算法来消除路径交叉又需要以较长的时间消耗为代价。针对上述问题, 本文提出了针对TSP的路径交叉检测与消除算法, 可以有效检测并消除路径交叉从而进一步提高解的质量。

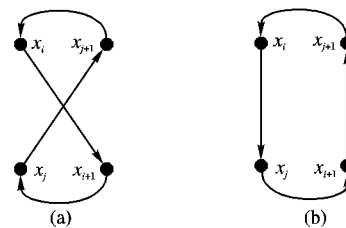


图1 消除交叉示意图

对于任意一个排列 $x = (x_1, \dots, x_i, x_{i+1}, \dots, x_j, x_{j+1}, \dots, x_n)$, 记 $C(x_i, x_{i+1})$ 表示子列 (x_i, x_{i+1}) 在某种意义上满足对称性质的费用函数 (比如, 节点之间的距离), 若

$$C(x_i, x_{i+1}) + C(x_j, x_{j+1}) > C(x_i, x_j) + C(x_{i+1}, x_{j+1})$$

则说明排列 x 中在该种意义下对应的路径上在节点 $x_i, x_{i+1}, x_j, x_{j+1}$ 之间存在交叉现象(如图1所示)。

若子列 (x_i, x_{i+1}) 与子列 (x_j, x_{j+1}) 之间存在交叉, 其中 $(i < j, i+1 < j)$, 只需把排列 x 中 (x_{i+1}, \dots, x_j) 这段子列在 x 中逆序排列就可消除 (x_i, x_{i+1}) 与 (x_j, x_{j+1}) 之间的交叉, 即把 $x = (x_1, \dots, x_i, x_{i+1}, \dots, x_{j-1}, x_j, x_{j+1}, \dots, x_n)$ 变为 $x' = (x_1, \dots, x_i, x_j, x_{j-1}, \dots, x_{i+1}, x_{j+1}, \dots, x_n)$, 并且在整个排列中不会产生新的交叉, 因为对 (x_{i+1}, \dots, x_j) 这段子列进行逆序操作, 不改变这段子列中原有的交叉, 也不改变 (x_1, \dots, x_i) 和 (x_{j+1}, \dots, x_n) 这两段子列各自内部的交叉。

在排列 $x = (x_1, \dots, x_i, \dots, x_n)$ 中, 对于任意一个节点 x_i ($i \in \{1, 2, \dots, n\}$), 可知子列 (x_i, x_{i+1}) 不需要与子列 (x_{i-1}, x_i) 和子列 (x_{i+1}, x_{i+2}) 判断是否交叉, 因为它们相邻, 本来就存在一个交点。所以只需与其余 $n-3$ 个子列判断是否存在交叉即可, 共需要判断 $(n-i-1)$ 次。因此, 对于 n 个节点的排列总的判断次数为:

$$(n-3) + (n-4) + \dots + 1 = \frac{(n-2)(n-3)}{2}$$

4 算法验证

为了验证本文提出的结合交叉消除的离散 FS 算法的性能, 选取 TSP 标准库 (<http://www.iwr.uni-heidelberg.de/groups/comopt/software/TSPLIB95/tsp/>) 中的 a280、d198、ts225、kroA200、Pr107、ch130、kroB150 作为测试数据, 与交叉消除遗传算法 (Genetic Algorithm, GA) 结果进行比较, 结果如表1。结果表明, 本文算法比交叉消除的遗传算法结果平均减小大约 1.6%。

表1 交叉消除离散 FS 算法与交叉消除 GA 结果比较

测试数据	交叉消除离散 FS 算法	交叉消除 GA	标准库提供的最优解
a280	2798.1317	2872.3599	2579
d198	16063.5183	16285.8039	15780
ts225	128252.7473	129487.2652	126643
kroA200	30045.7543	30964.4749	29368
Pr107	45088.0748	45591.2709	44303
ch130	6323.9210	6377.2781	6110
kroB150	27034.9448	27380.8561	26130

限于篇幅, 下面仅对 a280 数据的运行结果进行描述。a280 的 280 个点的坐标请参看图2。

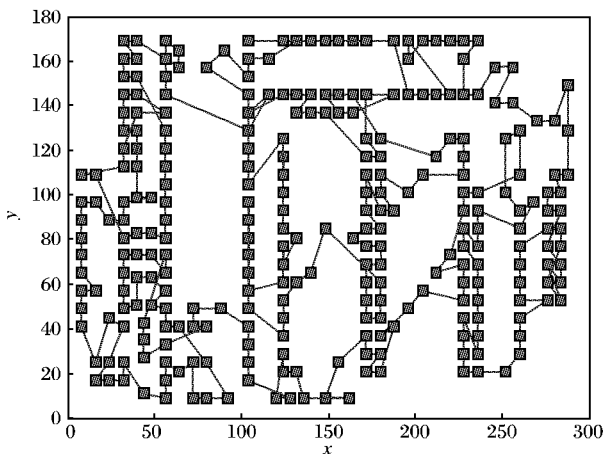


图2 离散 FS 算法的搜索结果

用离散 FS 算法进行最短路径的搜索结果是在第 500 代

时, 最短路径为 3536.4159, 路径中具有明显的交叉现象, 如图2所示。收敛速度较慢, 如图3所示。

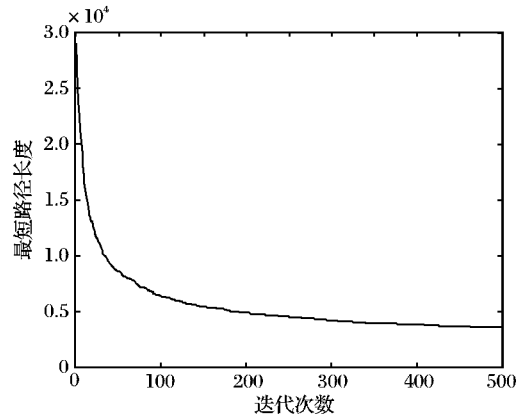


图3 离散 FS 算法收敛过程

结合交叉消除的离散 FS 算法在第 500 代时搜索到的最短线路长为 2798.1317, 并且从第 20 代开始, 搜索到的最短线路长度就稳定在 2800 附近, 收敛速度明显快于没有交叉消除的 FS 算法, 如图4和图5所示。

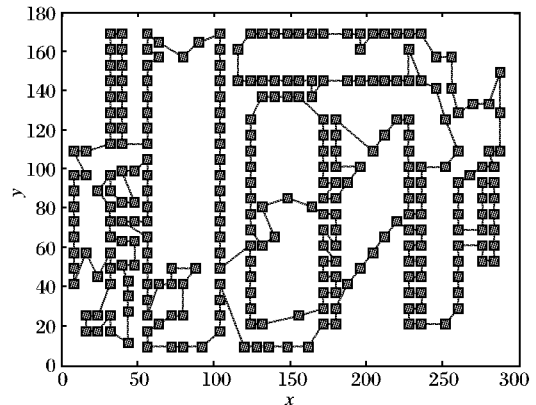


图4 结合交叉消除的离散 FS 算法的搜索结果

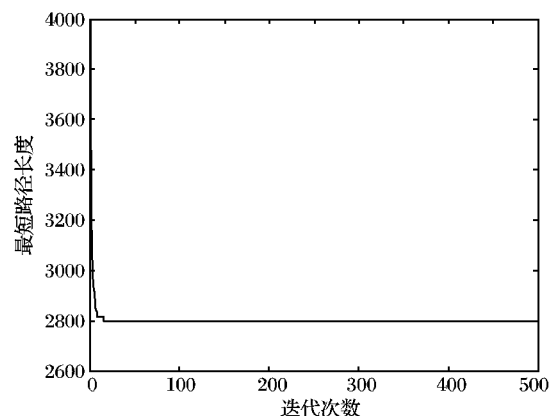


图5 结合交叉消除的离散 FS 算法的收敛过程

为了比较, 将交叉消除与遗传算法 (GA) 结合对标准数据库中的 a280 进行搜索, 在第 500 代时搜索到的最短线路长为 2872.3599, 并且从第 100 代开始, 搜索到的最短线路长度稳定在 2870 附近, 如图6所示。交叉消除的离散 FS 算法的搜索结果比交叉消除的遗传算法搜索结果性能略有提高。

5 结语

目前的 FS 算法主要应用在连续型解空间中, 本文给出一个 FS 算法在离散解空间的最优解搜索算法, 再把 FS 算法与交叉消除结合可使解的质量得到大幅度提高, 并且搜索速度 (下转第 1570 页)

待于今后继续深入研究。

参考文献:

- [1] 李昭智. NP-完全问题浅谈[J]. 天津理工大学学报, 1984(1): 27-33.
- [2] GAREY M R, JOHNSON D S. Computer and intractability: a guide to the theory of np completeness[M]. San Francisco: W. H. Freeman Company, 1979: 2-10.
- [3] 马炫, 刘庆. 求解多背包问题的人工鱼群算法[J]. 计算机应用, 2010, 33(2): 469-471.
- [4] 李娟, 方平, 周明. 一种求解背包问题的混合遗传算法[J]. 南昌航空工业学院学报, 1998(3): 31-35.
- [5] 张晓琴, 黄玉清. 基于禁忌搜索的启发式求解背包问题算法[J]. 电子科技大学学报, 2005, 34(3): 350-362.
- [6] 张玲, 张钺. 佳点集遗传算法[J]. 计算机学报, 2001, 24(9): 917-922.
- [7] 马良, 王龙德. 背包问题的蚂蚁优化算法[J]. 计算机应用, 2001, 21(8): 4-5.
- [8] 李将军, 叶仲泉, 宫子凤. 改进蚁群算法及其仿真研究[J]. 计算机应用, 2008, 28(12): 94-96.
- [9] 叶安新. 基于改进粒子群优化算法的 TSP 问题研究[J]. 计算机与现代化, 2011(4): 1-5.
- [10] YANG X S, DEB S. Cuckoo search via Lévy flights[C]// Proceedings of World Congress on Nature & Biologically Inspired Computing. Piscataway: IEEE, 2009: 210-214.
- [11] YANG X S, DEB S. Engineering optimization by cuckoo search[J]. International Journal of Mathematical Modeling and Numerical, 2010, 1(4): 330-343.
- [12] YANG X S, DEB S. Multiobjective cuckoo search for design optimization[J]. Computers & Operations Research, 2011, 10(9): 1-9.
- [13] YANG X S. Nature-Inspired Metaheuristic Algorithms [M]. 2nd ed. [S. l.]: Luniver Press, 2010: 4-9.
- [14] PAYNE R B, SORENSON M D, KLITZ K. The cuckoos[M]. Oxford: Oxford University Press, 2005.
- [15] BROWN C, LIEBOVITCH L S, GLENDON R. Lévy flights in Dobe Ju/'hoansi foraging patterns[J]. Human Ecology, 2007, 35(1): 129-138.
- [16] PAVLYUKOVICH I. Lévy flights, non-local search and simulated annealing[J]. Journal of Computational Physics, 2007, 226(2): 1830-1844.
- [17] REYNOLDS A M, FRYE M A. Free-flight odor tracking in Drosophila is consistent with an optimal intermittent scale-free search[J]. PLoS One, 2007, 2(4): e354.
- [18] MANTEGNA R N. Fast, accurate algorithm for numerical simulation of Lévy stable stochastic processes[J]. Physical Review E, 1992, 49(5): 451-458.
- [19] KENNEDY J, EBERHART R. Particle swarm optimization[C]// Proceedings of the IEEE International Conference on Neural Networks. Piscataway: IEEE, 1995: 1942-1948.
- [20] KENNEDY J, EBERHART R C. A discrete binary version of the particle swarm algorithm[C]// IEEE International Conference on Systems, Man, and Cybernetics. Piscataway: IEEE, 1997: 4104-4109.
- [21] 刘建华. 粒子群算法的基本理论及其改进研究[D]. 长沙: 中南大学, 2009: 77-98.
- [22] 蔡荣英, 李丽珊, 林晓宇, 等. 求解旅行商问题的自学习粒子群优化算法[J]. 计算机工程与设计, 2007, 28(2): 261-266.

(上接第 1565 页)

提高许多。FS 算法与遗传算法相比,性能略有提高。在进一步的研究中,我们将继续把 FS 算法用于更多组合优化问题中,并且与其他智能算法进行融合,弥补其对领域搜索半径敏感的缺陷,进一步提高 FS 算法的效率。

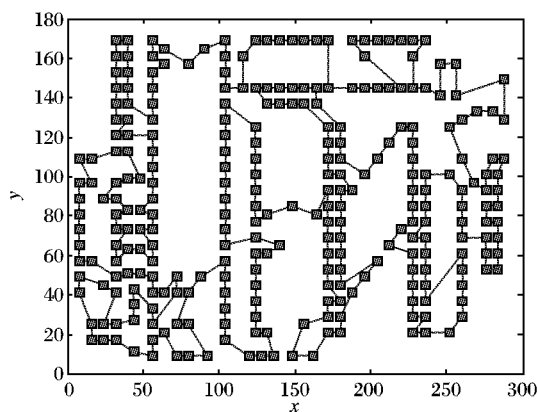


图6 结合交叉消除的遗传算法的搜索结果

参考文献:

- [1] TSAI C F, TSAI C W, YANG T. A modified multiple-searching method to genetic algorithms for solving traveling salesman problem [C]// IEEE International Conference on Systems, Man and Cybernetics. Piscataway: IEEE, 2002: 6-12.
- [2] KIRKPATRICK S, GELATF C D, VECCHI M P. Optimization by simulated annealing[J]. Science, 1983, 220(4598): 671-680.
- [3] TSAI C F, TSAI C W, TSENG C C. A new hybrid heuristic approach for solving large traveling salesman problem[J]. Information Sciences, 2004, 166(1): 67-81.
- [4] HE Y, QIU Y H, KIU G Y. A Parallel tabu search approach based on genetic crossover operation[C]// Proceedings of the 19th IEEE International Conference Cordon Advanced Information Networking and Applications. Washington, DC: IEEE Computer Society, 2005: 1550-1553.
- [5] ROSENKRANTZ D J, STEARNS R E, LEWIS P M. An analysis of several heuristics for the traveling salesman problem[J]. SIAM Journal of Computer, 1977, 6(1): 563-581.
- [6] CROES G A. A method for solving traveling salesman problems[J]. Operations Research, 1958, 6(6): 791-812.
- [7] BENTLEY J L. Fast algorithm for geometric traveling salesman problems[J]. ORSA Journal on Computing, 1992, 4(4): 387-411.
- [8] LIN S, KERNIGHAN B W. An effective heuristic algorithm for the traveling salesman problem[J]. Operations Research, 1973, 21(2): 498-516.
- [9] PENEV K, LITTLEFAIR G. Free search - a comparative analysis [J]. Information Sciences, 2005, 172(1/2): 173-193.
- [10] 周晖, 李丹美, 邵世煌, 等. 一种新的群集智能算法——自由搜索[J]. 东华大学学报: 自然科学版, 2007, 33(5): 579-583.
- [11] 周晖, 李丹美, 邵世煌, 等. 一种新的群集智能优化及其改进研究[J]. 系统工程与电子技术, 2008, 30(2): 337-340.
- [12] 周晖, 徐晨, 邵世煌, 等. 自适应搜索优化算法[J]. 计算机科学, 2008, 35(10): 188-191.
- [13] 伍国华, 马满好. 路径交叉检测与消除方法和邻节点置换方法改进 TSP 的解[J]. 计算机应用研究, 2011, 28(2): 485-487.