

文章编号: 2095-2163(2019)-0025-05

中图分类号: TP301.6

文献标志码: A

一种求解 TSP 问题的协同进化算法

魏士伟

(桂林航天工业学院 计算机科学与工程学院, 广西 桂林 514000)

摘要: 传统的遗传算法 GA 在求解 TSP 问题时容易出现早熟和陷入局部最优等现象。为此本文提出了一种基于协同进化的遗传算法 (CEGA) 用于解决 GA 算法的缺陷。该算法通过定义个体的适应度值和个体间的差异度值, 将适应度值高和差异度大的个体分别放入 2 个不同的子群体。在进化过程中这 2 个子种群相互协同进化, 既保证了种群向最优解的方向移动, 又保持了种群的多样性。实验结果表明, 本文所提出的算法在解决 TSP 问题时, 具有收敛速度快、容易跳出局部最优等特点, 相较其他 GA 算法具有更好的性能。

关键词: 遗传算法; 进化算法; TSP; 协同进化; 资源调度

A new Co-Evolution based Genetic Algorithm designed for TSP problem

WEI Shiwei

(School of Computer Science and Engineering, Guilin University of Aerospace Technology, Guilin Guangxi 541004, China)

[Abstract] The traditional Genetic Algorithm (GA) is more prone to be of premature convergence and fall into local optimums when solving TSP problems. In order to overcome these defects, a new Co-Evolution based Genetic Algorithm (CEGA) is proposed in this paper. By defining the fitness value of individuals and the difference value between individuals, this new algorithm puts individuals with high fitness values into a sub-population and put individuals with large differences into another sub-population. These two sub-populations coevolves with each other during the evolution process, which not only ensures the movement of the whole population towards the optimal value, but also maintains the diversity of the population. The experimental results show that the algorithm proposed in this paper has characteristics of fast convergence and being easy to jump out of local optimums when solving TSP problems. The research demonstrates that the proposed algorithm has better performance than other GA algorithms.

[Key words] Genetic Algorithm; evolutionary algorithm; TSP; Co-Evolution; resource scheduling

1 概述

旅行商问题 (Traveling Salesman Problem, TSP) 是图论中一个著名的组合优化问题。现实世界很多行业中的优化问题, 如快递配送线路安排、飞机航线安排、应急灾害调度、车辆路径规划、印制电路的制作、卫星位置的调整、人类基因排序、晶体结构分析和数据串聚类等^[1-3], 都可以归纳为 TSP 问题。因此, 研究求解 TSP 问题的高效算法具有十分重要的意义。

求解 TSP 问题的传统算法通常有穷举法^[4]、分支限界法^[5]和动态规划法^[6]等。然而, TSP 已被证明是一个 NP 完全问题, 随着问题规模的扩大, 所需的计算时间呈指数级增加, 依靠这些传统算法求解 TSP 问题已经变得不可能。针对大规模 TSP 问题, 一些进化算法显示出其独特的优越性。例如, 蚁群算法^[7-8]、模拟退火算法^[9]、启发搜索式算法^[10]、遗

传算法 (GA) 以及一些混合智能算法^[11-14] 等。特别是遗传算法, 模拟了生物进化论的自然选择的机理和遗传学的生物进化过程, 能够使生物群体不断地向好的方向进化, 从而搜索到最优解。由于具有良好的全局最优解搜索能力、隐含的并行计算能力和灵活的应用方式, 在解决大规模复杂问题时遗传算法 (GA) 已经成为近期吸引学界高度关注的进化算法。例如, 陈林等人^[11] 利用贪婪思想产生初始种群, 通过改进的遗传算法来加快寻优速度, 从而优化遗传算法的质量和寻优效率。姜群等人^[12] 通过动态调整种群的交叉和变异概率控制了进化过程, 以便使遗传算法获得更好的性能。孙文彬等人^[13] 针对遗传算法求得的解质量不高等缺陷, 设计了一种基于遗传算法的多策略优化求解方法来处理 TSP 问题, 并取得了良好效果。张玉州等人^[14] 通过组合局部搜索算子集合并将其嵌入遗传算法, 从而形成混合遗传算法来求解 TSP 问题。

基金项目: 广西自然科学基金 (2016GXNSFAA380226); 广西高校中青年教师基础能力提升基金 (2017KY0866); 桂林航天工业学院物联网与大数据应用研究基金 (KJPT201809)。

作者简介: 魏士伟 (1979-), 男, 硕士, 讲师, 主要研究方向: 数据挖掘、智能计算、生物信息处理。

收稿日期: 2019-07-27

这些改进的遗传算法在研究 TSP 问题时,一定程度上都表现出良好的性能。然而,随着问题规模的进一步扩大,遗传算法依然会出现早熟收敛、容易陷入局部最优等问题。为此,本文提出一种基于协同进化的遗传算法,来改善算法易于早熟收敛的缺陷,强化搜索全局最优解的能力。实验结果表明新提出的算法在解决 TSP 问题时具备良好的性能。

2 TSP 问题描述及求解模型

TSP 问题要求商人把地图中的 n 个城市都拜访一遍,并且最后回到出发点。目标是选择一条最短的路径完成该项任务。在图论中,TSP 就是找一条哈曼顿回路,并且该哈曼顿回路的路径长度为最短。现假设有 n 个城市 $V = \{v_1, v_2, \dots, v_n\}$,第 i 个城市的坐标值为 $v_i = (x_i, y_i)$,城市 v_i 和城市 v_j 之间的距离为 $d(v_i, v_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$ 。假设一条哈曼顿路径为: $L = [l_1, l_2, \dots, l_n, l_{n+1}]$,其中 $l_i \in V, 1 \leq i \leq n, l_{n+1} = l_1 \neq l_2 \neq \dots \neq l_n$,则路径的总长度 D 可表示为:

$$D = \sum_{i=1}^n d(l_i, l_{i+1}), \quad (1)$$

例如,目前有 5 个城市为 $V = \{v_1, v_2, v_3, v_4, v_5\}$,假设一条遍历了这 5 个城市的路径为 $v_3 \rightarrow v_2 \rightarrow v_1 \rightarrow v_5 \rightarrow v_4 \rightarrow v_3$,则该哈曼顿路径 $L = [v_3, v_2, v_1, v_5, v_4, v_3]$,即 $l_1 = v_3, l_2 = v_2, l_3 = v_1, l_4 = v_5, l_5 = v_4$ 。在本文中规定,用 $\theta_{ij} (1 < i, j < n)$ 表示路径上的第 i 个节点途径的城市是 v_j ,即当 $\theta_{ij} = 1$ 时,哈曼顿路径上的第 i 个位置对应的城市是 v_j ;当 $\theta_{ij} = 0$ 时,表示哈曼顿路径上的第 i 个位置对应的城市不是 v_j 。对应给定的例子,可知 $\theta_{13} = 1, \theta_{22} = 1, \theta_{31} = 1, \theta_{45} = 1, \theta_{54} = 1$,除此之外其它的 θ_{ij} 值都为 0。哈曼顿路径上的节点 $l_i (1 \leq i \leq n)$ 可以由下式求出:

$$l_i = \sum_{j=1}^n (\theta_{ij} v_j), \quad (2)$$

当一条哈曼顿路径确定以后,其路径的总长度可以表示为:

$$D = \sum_{i=1}^n d(l_i, l_{i+1}) = \sum_{i=1}^n \sqrt{\left(\sum_{j=1}^n \theta_{ij} x_j - \sum_{j=1}^n \theta_{(i+1)j} x_j\right)^2 + \left(\sum_{j=1}^n \theta_{ij} y_j - \sum_{j=1}^n \theta_{(i+1)j} y_j\right)^2} = \sum_{i=1}^n \sqrt{\left(\sum_{j=1}^n x_j \cdot (\theta_{ij} - \theta_{(i+1)j})\right)^2 + \left(\sum_{j=1}^n y_j \cdot (\theta_{ij} - \theta_{(i+1)j})\right)^2}, \quad (3)$$

此外,由于每一个城市都必须访问一次,并且一个城市只能遍历一次(除了选为起点的城市),有如下的 2 个约束条件:

$$\sum_{j=1}^n \theta_{ij} = 1, \quad (4)$$

$$\sum_{i=1}^n \sum_{j=1}^n \theta_{ij} = n, \quad (5)$$

又因为选为起点的城市必是路径的终点,因此有:

$$\theta_{1j} = \theta_{(n+1)j}, \quad (6)$$

求解 TSP 问题的目标是使得哈曼顿路径的长度 D 为最短,即该问题的求解模型可描述为:

$$\begin{cases} \min D = \\ \min \left\{ \sum_{i=1}^n \sqrt{\left(\sum_{j=1}^n x_j \cdot (\theta_{ij} - \theta_{(i+1)j})\right)^2 + \left(\sum_{j=1}^n y_j \cdot (\theta_{ij} - \theta_{(i+1)j})\right)^2} \right\}, \\ \text{s. t.} \\ \sum_{j=1}^n \theta_{ij} = 1, \\ \sum_{i=1}^n \sum_{j=1}^n \theta_{ij} = n, \\ \theta_{1j} = \theta_{(n+1)j}, \\ 1 \leq i, j \leq n. \end{cases} \quad (7)$$

3 一种求解 TSP 问题的协同进化算法

对于复杂的大规模的 NP 问题,经典的遗传算法在求解过程中容易出现早熟收敛和陷入局部最优的问题。为此,本文提出了一种基于协同进化的遗传算法并将其用于求解 TSP 问题。对此展开研究论述如下。

3.1 编码与解码

为了简化遗传算法的运算过程,本文采用整数编码方案进行求解。具体来说,将 TSP 中要访问的城市给定一个唯一的整数编号 $1, 2, 3, \dots, n$ (假设共有 n 个城市),即 $V = \{v_1, v_2, \dots, v_n\}$ 。一条染色体就是一个访问城市序列的方案 $L = [l_1, l_2, \dots, l_n, l_{n+1}]$,其中 $l_i \in V, 1 \leq i \leq n, l_{n+1} = l_1 \neq l_2 \neq \dots \neq l_n$ 。由于城市的坐标位置是已知的,因此,当染色体确定后,可以根据式(3)求出哈曼顿路径的长度。染色体不同所求得的哈曼顿路径的长度就不同。研究的最终目的就寻找一条长度最短的哈曼顿路径。对于 TSP 问题,采用这种整数的编码方案可以使本文的编码和解码工作变得非常简单。

3.2 个体选择策略

评价个体(染色体)好坏的标准是适应度函数,适应度函数的值越大,表明该个体越好,越接近最优解。对于 TSP 问题,研究需要的是一条最短的哈曼顿路径,因此适应度函数可以定义为:

$$F(L) = \frac{1}{D} = \frac{1}{\sum_{i=1}^n d(l_i, l_{i+1})}, \quad (8)$$

遗传算法通过轮盘赌的方式总会以大概率的方式把适应度高的个体选择到下一代群体中。随着进化代数的不断推进,群体中个体的适应度值会越来越高,但其多样性就会降低,群体的同质化现象严重,很容易出现早熟现象,从而导致算法陷入局部最优的现象发生。为了解决这一问题,本文设计了一种个体选择策略,群体中一部分个体是适应度值高的个体,另一部分是差异度高的个体。适应度值高的个体可以保证种群向着全局最优解的方向前进,差异度值高的个体可以保证群体的多样性在进化过程中维持不变,从而避免早熟现象的发生。任意2个个体之间的差异度可以定义如下:

$$\begin{cases} \delta(p_i, p_j) = \frac{1}{n} \sum_{i=1}^n \|l_{ik} - l_{jk}\|, \\ \|l_{ik} - l_{jk}\| = \begin{cases} 1, l_{ik} = l_{jk}, \\ 0, l_{ik} \neq l_{jk}. \end{cases} \end{cases} \quad (9)$$

其中, $p_i = l_{i1}l_{i2}l_{i3} \cdots l_{in}$ ($1 \leq i \leq \text{popsize}$) 是群体中的一个个体(染色体),群体的大小为 popsize 。

具体的个体选择策略为,将群体分为2个子群体 subpopA 和 subpopB ,各子群体中的个体数量各占群体总数的 $1/2$ 。接着即需确定 subpopA 中的个体,可使用轮盘赌的方式将适应度值高的个体选择进来,待 subpopA 一旦确定后,就可以根据 subpopA 中的每个个体从种群中查找出和其差异度最高的个体而组成 subpopB ,从而完成个体的选择。

3.3 协同进化技术

选择出来的个体分别组成了2个子群体 subpopA 和 subpopB ,通过各个子群体的内部的交叉变异操作可以生成下一代新的种群个体。同时,子种群 subpopA 和 subpopB 之间也可以通过相互的交叉和变异操作,即通过协同进化的方式产生下一代种群新个体,对此进化方法可详述如下。

(1) 子种群 subpopA 和子种群 subpopB 各自的独立进化。对于每个子群体,分别以随机的方式从子种群中选择 $1/3\text{popsize}$ 个个体,其后以概率 P_c 进行两两交叉操作,接着又以概率 P_m 进行变异操作从而生成新的个体,再将其放入下一代群体中。

(2) 子种群 subpopA 和子种群 subpopB 的协同进化。从 subpopA 子种群中随机选择 $1/6\text{popsize}$ 个个体,从 subpopB 中选择同样数目的个体,并以概率 1 进行交叉操作,后续则以概率 P_m 进行变异操作从而生成新的个体,再将其放入到下一代群体中。

采用这种协同进化技术可以让个体在向最优解靠近的同时,还能保持种群的多样性,防止种群随着

进化代数的增加过早出现早熟现象,并且各个子种群之间通过相互交叉变异更有可能产生出更优的解,从而找到最优解。下一节将给出协同进化算法的详细描述。

3.4 协同进化算法的详细描述

综合前述研究后可得,多精英协同进化遗传算法 MECGA 的基本思想如算法 1 所示。

算法 1 协同进化算法

Step 1 令进化代数 $t = 0$, 设置的种群大小 popsize 、交叉概率 P_c 、变异概率 P_m 的初始值,并进行种群初始化操作生成初始化种群 $\text{POP}(t)$;

Step 2 对 $\text{POP}(t)$ 中的个体根据式(8)计算适应度值,以轮盘赌的方式从中选择 $1/2\text{popsize}$ 个个体组成子群体 subpopA ;

Step 3 对 subpopA 中的每个个体,根据公式(9)找出 $\text{POP}(t)$ 中与之差异度最大的个体并将其放置于 subpopB 中。

Step 4 从 subpopA 中随机选择 $1/3\text{popsize}$ 个体,并以概率 P_c 进行交叉操作,生成的新个体再以 P_m 概率进行变异操作,然后将其插入到下一代种群 $\text{POP}(t+1)$ 中;

Step 5 从 subpopB 中随机选择 $1/3\text{popsize}$ 个体,并以概率 P_c 进行交叉操作,生成的新个体再以 P_m 概率进行变异操作,然后将其放入种群 $\text{POP}(t+1)$ 中;

Step 6 从 subpopA 中随机选择 $1/6\text{popsize}$ 个体,并让其与从 subpopB 中选择出的相同数目的个体进行交叉操作,新生成的个体将以概率 P_m 进行变异操作,再将其放置于下一代群体 $\text{POP}(t+1)$ 中。

Step 7 判断群体 $\text{POP}(t)$ 中最优个体的适应度值是否大于下一代群体 $\text{POP}(t+1)$ 中最优个体的适应度值(精英保留策略),若不大于,则将其放置于 $\text{POP}(t+1)$ 。

Step 8 令 $t = t + 1$ 。

Step 9 判断是否满足终止条件,若是,则结束;否则,转到 Step 2。

4 实验与结果分析

为了验证协同进化算法在解决 TSP 问题时的良好性能,研究将其与传统的 GA 算法、带有精英保留策略的遗传算法 GAE 做了实验对比和分析。本研究中,进行实验验证和分析时的部分参数取值详见表 1。

表1 实验中参数的取值

Tab.1 The value of parameters in the experiment

参数	GA	GAE	CEGA
种群规模	200	200	200
进化代数	100	100	100
选择算子	RR	RR	RR
交叉概率 P_c	0.65	0.65	-
变异概率 P_m	0.05	0.05	0.05
精英个数 M	0	1	25
运行次数	20	20	20

实验在标准的测试集 TSPLIB 数据库 (<http://elib.zib.de/pub/mp-testdata/tsp/tsplib/tsp/index.html>) 中选择了 6 个数据集进行测试。针对文中选定的 3 种对比算法运算求得的最短哈曼顿路径的实验结果见表 2。

表2 对比算法的实验结果

Tab.2 Experimental results of the compared algorithm

序号	数据集名称	哈曼顿路径长度/ 10^5		
		GA	GAE	CEGA
1	att48.tsp	1.139 4	1.137 7	0.612 5
2	berlin52.tsp	0.230 6	0.231 5	0.134 7
3	kroa100.tsp	1.388 3	1.388 4	0.965 3
4	eil101.tsp	0.028 6	0.029 0	0.022 1
5	ch130.tsp	0.396 5	0.396 4	0.306 4
6	ch150.tsp	0.465 8	0.468 3	0.374 1

从表 2 中可以看出,GA 和 GAE 算法求得的最短路径的长度值相差不多,而且受到所处理问题规模的限制,这 2 种算法都容易陷入局部最优解而缺乏跳出局部机制能力,由其求得的结果并不理想。而 CEGA 算法由于设计了精英种群和普通协同进化的思想,既可以保持向最优解不断靠近的压力,又保持了种群的多样性,具备跳出局部最优的能力,所求解的质量要好很多。在此基础上研究又发现,对于所有的测试案例,CEGA 算法的结果总是好于 GA 算法和 GAE 算法的结果。特别是 att48.tsp 测试数据集,CEGA 算法求得的最短路径长度只有 GA 算法和 GAE 算法的 50% 左右。

为了进一步验证本文所提出的算法的寻找最优解的性能,分别在 att48.tsp 和 berlin52.tsp 数据集上对比分析 3 种对比算法的进化过程。如图 1 和图 2 所示。

从图 1、图 2 中可以看出,随着进化代数的推进,GA 和 CGA 算法在进化的前期都能引导种群不断向最优解的方向移动,所求的最小值也在不断下降,但是在进化的后期(种群大概进化到 60 代以

后),这种下降的趋势越来越不明显,几乎丧失了进一步寻优的能力。而 CEGA 算法却表现出较好的进化态势,群体寻找到的最优值随着进化代数的逐渐演进而不断向最优解靠近,即使到进化的后期,这种趋势也十分明显。分析可知,这一结果即得益于 CEGA 算法的良好设计,由于将群体分为 2 个子种群,一个子种群中的个体具有较高的适应度值,而另一个子种群具有较高的差异度值,这些子种群之间又相互协同进化,使得种群既有向最优解前进的驱动力,又能够保持种群的多样性,从而使群体具备了跳出局部最优解的能力。至此,图 1、图 2 中的结果可以看出,无论是最终求得的解的质量,还是求解过程中的进化趋势,CEGA 算法都优于 GA 和 GAE 算法。

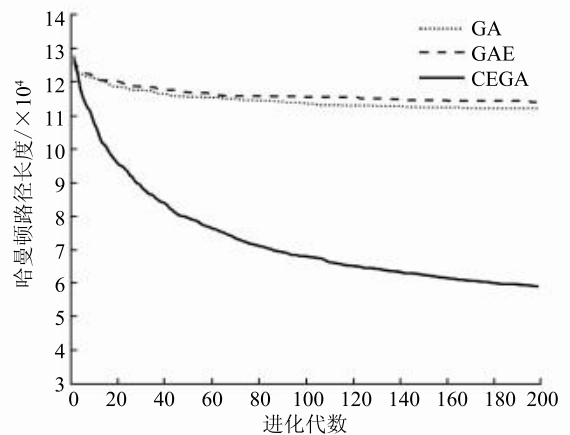


图1 算法在 att48.tsp 上的进化过程图

Fig. 1 Evolutionary process diagram of the algorithm on att48.tsp

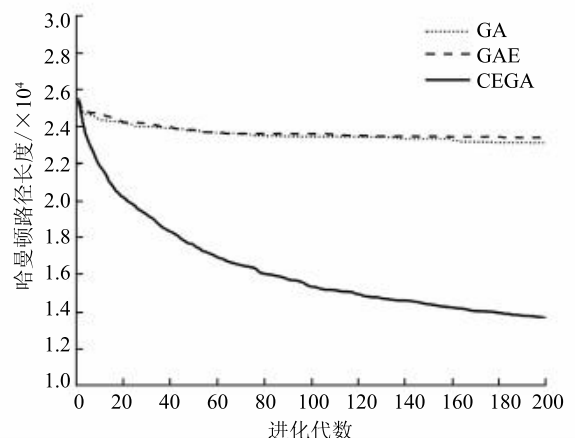


图2 算法在 berlin52.tsp 上的进化过程图

Fig. 2 Evolutionary process diagram of the algorithm on berlin52.tsp

此外,关于本文提出算法 CEGA 在 att48.tsp 数据集和 berlin52.tsp 数据集上求得的旅行线路图,即

哈曼顿路径,见图3和图4。

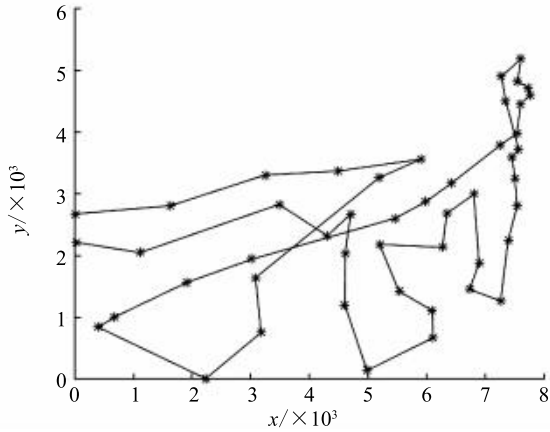


图3 CEGA求得att48.tsp上的旅行线路

Fig. 3 Travel routes on att48.tsp by using CEGA

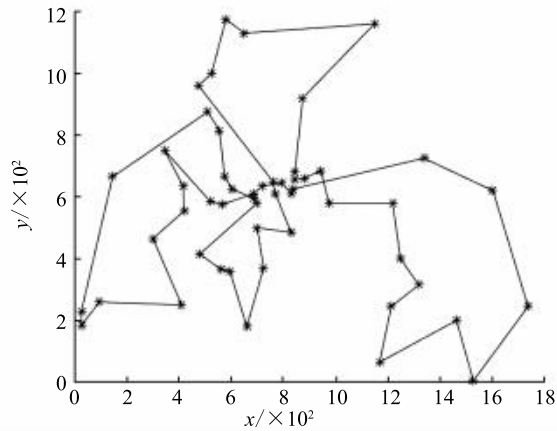


图4 CEGA求得berlin52.tsp上的旅行线路

Fig. 4 Travel routes on berlin52.tsp by using CEGA

5 结束语

传统遗传算法在解决 TSP 这一 NPC 问题时容易出现早熟和陷入局部最优等问题。为此,本文提出了一种基于种群协同进化的遗传算法 CEGA。该算法将种群中的个体依据适应度值和定义的差异度值分到了2个子种群中,一个子种群中的个体具有较高的适应度值,而另一个子种群中的个体具有较

大的差异度值。在进化过程中,这2个子种群相互协作,共同进化,使种群既能够向靠近最优解的方向移动,又能保持种群的多样性。算法具有了易于跳出局部最优解的能力。为了验证所提出的算法的性能,研究则在 TSPLIB 数据库中的测试集上进行了多组实验。实验结果表明,CEGA 算法在解决 TSP 问题时,其所求得的解的质量总是好于 GA 算法和 GAE 算法,而且在跳出局部最优解的能力方面有着良好表现,算法的稳定性也更高。

参考文献

- [1] 申艳光,张玲玉,刘永红.基于混合遗传算法的物流路径优化方法研究[J].计算机技术与发展,2018,28(3):192-196.
- [2] 吴新胜,姜婷,赵梦超,等.基于群智能混合算法的应急物流路径优化研究[J].四川理工学院学报(自然科学版),2018,31(4):68-73.
- [3] 唐健,史文中,孟令奎.基于遗传算法的时相关动态车辆路径规划模型[J].武汉大学学报(信息科学版),2008,33(8):875-879.
- [4] 许玲.优化穷举法求旅行商问题的近似最优解[J].微型机与应用,1998(10):20,33.
- [5] 陈涛,张思发.分支限界法求解实际 TSP 问题[J].计算机工程与设计,2009,30(10):2431-2434.
- [6] 来学伟.动态规划法在 TSP 问题中的应用[J].吉林化工学院学报,2017,34(3):65-67.
- [7] 王宝生,屈宝存.蚁群算法在求解 TSP 问题中的改进研究[J].电子设计工程,2014,22(22):14-18,21.
- [8] 康岚兰,李康顺.蚁群算法在求解 TSP 问题上与遗传算法的对比研究[J].计算机系统应用,2008,17(10):60-63.
- [9] 周君,贾昆霖.求解旅行商路径规划问题的改进模拟退火算法[J].电子科技,2017,30(7):62-64,68.
- [10] 邓志刚,何琦,肖媛媛,等.一种基于 TSP 问题的启发式搜索算法研究[J].科技广场,2010(9):29-31.
- [11] 陈林,潘大志.改进遗传算法解决 TSP 问题[J].智能计算机与应用,2016,6(5):17-19,23.
- [12] 姜群,晏雨.改进的遗传算法在 TSP 问题中的应用[J].重庆理工大学学报(自然科学),2012,26(9):96-99.
- [13] 孙文彬,王江.一种基于遗传算法的 TSP 问题多策略优化求解方法[J].地理与地理信息科学,2016,32(4):1-4.
- [14] 张玉州,梅俊,徐廷政.一种求解 TSP 问题的混合遗传算法[J].安庆师范大学学报(自然科学版),2018,24(3):77-81.