# 基于贪心遗传算法求解校内快递配送问题

**摘要**

TSP问题是一个组合优化问题，具有十分重要的实际意义。本文在遗传算法的基础上加入贪心选择策略，通过构建“基因库”创建“基因片段”来初始化产生高性能的种群个体，并且进行贪心交叉、贪心变异等操作，对子代进行优化，若与父代相比，子代未得到优化，则不进行替换，否则，进行替换，产生新的个体。贪心遗传算法保证了个体前进的方向，降低了随机操作的错误率。

**关键词：**TSP问题；贪心策略；遗传算法；贪心遗传算法

# 一、引言

## 1.1 研究背景

随着互联网和智能手机的普及，快递订餐取得了迅速的发展。作为大学生群体聚集高消费水平强的大学校园，俨然成为第三方平台快递配送的主要区域。然而面对分布各异的大学校园，如何对校间各快递配送地点进行配送路径规划是快递员进行配送时需要常常考虑的问题。寻找最优配送路径的过程实际上就是TSP问题。目前，求解该问题的方法有穷举法，回溯法，分支限界法，贪心法，遗传算法，蚁群算法等，其中，穷举法、回溯法和分支限界法均 是基于搜索的确定性算法 ，复杂度随着问题规模的增加指数级增加 ；贪心法和智能计算方法是近似算法，不一定能够得到TSP的最短路径 。这些方法中，遗传算法得到的结果是 比较令人满意的，只是收敛速度较慢，当问题的规模比较大时，会出现局部最优解，即近似最优解。

## 1.2 TSP问题

TSP问题，又称为旅行商问题，是著名的数学领域问题之一，由威廉哈密顿爵士和英国数学家克克曼T.P.Kirkman于19世纪初提出。其内容是：假设有一个旅行商人，他要拜访n个城市，需要选择出将要走过的路径di(i=1,2,…,n)，限制条件为每个城市只能选择依次，且最终回到出发的城市。路径选择的目标为所走路径的最小值，即。

## 1.3 TSP的贪心策略

贪心算法，又称贪婪算法，是指在对问题进行求解时，总是作出在在当前看来是最好的选择，否则将来可能会后悔。对于贪心算法而言，每次选择的所得到的都是局部最优解。而对于选择贪心算法的问题来说，是具有无后性的，即在求解过程中，某个状态只与当前状态有关，不会影响到以后的状态。

针对TSP问题，使用贪心策略求解时，具体算法为从某一个城市出发，每次选择一个城市，直到所有的城市被走完，其核心为每次选择下一个城市时，只考虑当前状况，选择城市间距离最短的路径。

## 1.4 TSP的遗传算法

遗传算法（Genetic Algorithm，GA）最早是由美国的 John holland于20世纪70年代提出,该算法是根据大自然中生物体进化规律而设计提出的，是一种启发式搜索算法。是模拟[达尔文](https://baike.baidu.com/item/%E8%BE%BE%E5%B0%94%E6%96%87/23890?fromModule=lemma_inlink)[生物进化论](https://baike.baidu.com/item/%E7%94%9F%E7%89%A9%E8%BF%9B%E5%8C%96%E8%AE%BA/143686?fromModule=lemma_inlink)的[自然选择](https://baike.baidu.com/item/%E8%87%AA%E7%84%B6%E9%80%89%E6%8B%A9/1800930?fromModule=lemma_inlink)和[遗传学](https://baike.baidu.com/item/%E9%81%97%E4%BC%A0%E5%AD%A6/233918?fromModule=lemma_inlink)机理的[生物进化](https://baike.baidu.com/item/%E7%94%9F%E7%89%A9%E8%BF%9B%E5%8C%96/3311428?fromModule=lemma_inlink)过程的计算[模型](https://baike.baidu.com/item/%E6%A8%A1%E5%9E%8B/1741186?fromModule=lemma_inlink)，是一种通过模拟自然进化过程搜索[最优解](https://baike.baidu.com/item/%E6%9C%80%E4%BC%98%E8%A7%A3/5208902?fromModule=lemma_inlink)的方法。该算法通过数学的方式,利用[计算机仿真](https://baike.baidu.com/item/%E8%AE%A1%E7%AE%97%E6%9C%BA%E4%BB%BF%E7%9C%9F/2056561?fromModule=lemma_inlink)运算,将问题的求解过程转换成类似生物进化中的[染色体](https://baike.baidu.com/item/%E6%9F%93%E8%89%B2%E4%BD%93/195881?fromModule=lemma_inlink)基因的交叉、变异等过程。由于遗传算法的整体搜索策略和优化搜索方法在计算时不依赖于[梯度](https://baike.baidu.com/item/%E6%A2%AF%E5%BA%A6?fromModule=lemma_inlink)信息或其它辅助知识，而只需要影响搜索方向的目标函数和相应的适应度函数，所以遗传算法提供了一种求解复杂系统问题的通用框架，它不依赖于问题的具体领域，对问题的种类有很强的[鲁棒性](https://baike.baidu.com/item/%E9%B2%81%E6%A3%92%E6%80%A7?fromModule=lemma_inlink)。

TSP问题是经典的NP组合问题，采用遗传算法能够较方便的求出其近似解。首先对TSP问题进行编码和种群初始化；其次是确定适应度函数，在TSP问题中，适应度函数就是所选的路径的最小值；然后进行选择、交叉、变异等操作，产生新的个体，更新算法，推近算法向最优解方向前进。遗传算法是一种基于解空间求解的算法，每一次只能从大体上让求解域更接近最优解，相比于贪心策略，能够更好地提高算法的效率。

# 二、算法设计

## 2.1 问题描述

本文选取了西安部分高校32个快递配送地点，依次进行编号为 1 至 32 号。对快递配送地点坐标的选取，本文则通过高德地图生成的学校正门口的经纬度坐标，由于在地图上各点的经纬度相近，为了更加直观的显示仿真运算结果，对经纬度进行如下处理： （a）只保留经度小数点后的 1、2、3 位生成百位数字坐标；（b）只保留纬度小数点后的 1、2、3位生成百位数字坐标。如陕西师范大学长安校区的经纬度坐标为 （108.9515,34.20545），本文选取坐标为 （951,205）。设定完毕的快递配送地点坐标，如表 1 所示。

表1 各学校配送地点坐标

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 编号 | 学校 | X坐标 | Y坐标 |  | 编号 | 学校 | X坐标 | Y坐标 |
| 1 | 陕西师范大学雁塔校区 | 951 | 205 |  | 17 | 西安理工大学莲湖校区 | 912 | 272 |
| 2 | 陕西师范大学长安校区 | 891 | 153 |  | 18 | 西安石油大学雁塔校区 | 928 | 213 |
| 3 | 西北大学太白校区 | 875 | 220 |  | 19 | 西安石油大学明德校区 | 931 | 200 |
| 4 | 西北大学长安校区 | 874 | 145 |  | 20 | 西安石油大学鄂邑校区 | 649 | 104 |
| 5 | 西北工业大学友谊校区 | 915 | 243 |  | 21 | 西安邮电大学雁塔校区 | 948 | 214 |
| 6 | 西北工业大学长安校区 | 764 | 28 |  | 22 | 西安邮电大学长安校区 | 903 | 153 |
| 7 | 西安工业大学金花校区 | 998 | 265 |  | 23 | 西北政法大学雁塔校区 | 943 | 207 |
| 8 | 西安工业大学未央校区 | 987 | 379 |  | 24 | 西北政法大学长安校区 | 916 | 154 |
| 9 | 西安电子科技大学北校区 | 916 | 231 |  | 25 | 西安外国语大学雁塔校区 | 950 | 207 |
| 10 | 西安电子科技大学南校区 | 834 | 123 |  | 26 | 西安外国语大学长安校区 | 874 | 135 |
| 11 | 西安建筑科技大学雁塔校区 | 968 | 237 |  | 27 | 西安财经大学长安校区 | 934 | 115 |
| 12 | 西安建筑科技大学草堂校区 | 754 | 15 |  | 28 | 西安财经大学雁塔校区 | 960 | 221 |
| 13 | 西安交通大学兴庆校区 | 984 | 246 |  | 29 | 西安科技大学雁塔校区 | 962 | 233 |
| 14 | 西安交通大学雁塔校区 | 941 | 216 |  | 30 | 陕西科技大学西安校区 | 977 | 378 |
| 15 | 西安理工大学金花校区 | 993 | 254 |  | 31 | 长安大学雁塔校区 | 960 | 226 |
| 16 | 西安理工大学曲江校区 | 994 | 222 |  | 32 | 长安大学渭水校区 | 903 | 370 |

## 2.2 问题分析

经上述操作后，本文研究校间快递配送的路径规划问题可表示以西安部分高校的32个配送地点为研究对象，在高校之间寻找一个配送中心，以该配送中心为出发点，求历经32个配送地点后，最后回到配送中心的一条最短路径的TSP类规划问题。为简化问题，本文将配送中心建立在高校内。

求解本文配送中心的快递配送问题，可通过遗传算法求解得到近似最优解。为了提高遗传算法的收敛速度和求解质量，本文将贪心策略融入到遗传算法种群初始化和遗传算子中，着重考虑初始种群的质量和遗传操作的有效性，即一种基于贪心策略求解TSP的遗传算法。

## 2.3算法核心思想

### 2.3.1 贪心初始化个体

种群个体的初始化就是依据[编码规则](https://baike.baidu.com/item/%E7%BC%96%E7%A0%81%E8%A7%84%E5%88%99/12731449?fromModule=lemma_inlink)给出种群的初始解。算法在开始时都要进行种群的初始化，根据初始化方法的不同形式可以将其分成M类随机方法、定值设定法、两步式方法、混合方法和具体应用法，[随机数生成器](https://baike.baidu.com/item/%E9%9A%8F%E6%9C%BA%E6%95%B0%E7%94%9F%E6%88%90%E5%99%A8/760307?fromModule=lemma_inlink)是最常用的方法，然而，在面对大规模优化问题（[决策变量](https://baike.baidu.com/item/%E5%86%B3%E7%AD%96%E5%8F%98%E9%87%8F/774519?fromModule=lemma_inlink)超过100）时，这种[初始化](https://baike.baidu.com/item/%E5%88%9D%E5%A7%8B%E5%8C%96/100108?fromModule=lemma_inlink)方法效果不佳。在随机方法中，比较常用的是RNG。而定值设定法则比较偏向于在搜索空间中产生均匀分布的点。

针对TSP问题，个体质量的高低决定了算法的效率，如果群体中所有个体的适应值都较差，将会影响算法的全局性能。在TSP问题中，将n个城市点进行排序，根据“基因库”来构造“基因片段”，组成染色体串。对于32个高校快递配送地点问题，将距离配送中心最近的C个配送地点按距离由小到大排序进行编码，组成“基因库”An\*C，Aij元素为距离配送地点i第j近的配送地点的编码。构建“基因库”生成的初始种群是当前最好的解来构造种群个体，是一种贪心选择策略。按照“基因库”构造的“基因片段”，具有定义长度短、低阶、高适应值的特点。

### 2.3.2 贪心交叉算子

贪心选择策略下的遗传算法，经过交叉变异等基因操作，若子代个体优于父代个体，则进行替换，否则，不进行替换操作，降低了随机操作的错误率。交叉能够通过组合父代双亲的优良基因特性，生成新的更好的可行解。

贪心交叉算子选择父代第一个配送地点，然后再双方父代中对比剩下的配送地点，选择距离较近的配送地点，继续快递配送，如果该配送地点已经在配送过程中出现，则选择另一个父代的配送地点，如果两个配送地点都出现过，则随机生成未选择过的配送地点作为下一个配送地点。

### 2.3.3 贪心变异算子

由于遗传交叉的构造产生的个体不能保证该算法已经收敛于局部最优解，因此引入了贪心变异操作。采用相邻两位交换的局部寻优方法，减少了路径长度，从而找到局部最优点。若交换基因位改进了计算的结果，就对该基因进行交换操作，否则，不进行交换。由此可知，贪心变异算子是一个局部调整和改进的过程。

## 2.4 算法流程

（1）确定遗传算法的控制参数。即种群的规模大小，最大的遗传代数，贪心交叉的概率，贪心变异的概率等等。

（2）设置遗传代数计数器i，并设置其初始值为0。

（3）生成初始种群。通过构建“基因库”产生“基因片段”，按照贪心选择策略生产个体的基因。

（4）对种群进行遗传操作，产生新的种群。对种群采用贪心交叉算子、贪心变异算子，如若个体的性能得到改善，就替换父代个体，否则，不替换。

（5）将i赋值为i+1；并判断≥i G是否成立，若是，就每隔大约10代对种群进行移民操作。移民操作按贪心选择办法生成一些个体替换当前种群中较差的个体。G为一常数，一般

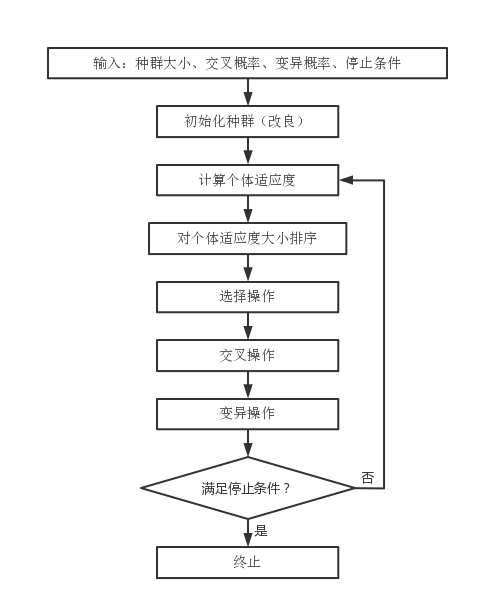
取为200左右，可以根据问题的复杂程度调整。

（6）判断i>=MG是否成立，若是，进化终止，输出结果；否则，转(4)。

# 三、实验仿真

## 3.1 遗传算法

遗传算法是一种随即检索方法，沿用了生物世界中“适者生存、优胜劣汰”的进化定律。遗传算法可以直接对种群空间中的多个解进行搜索，对目标函数的可导性和连续性没有约束，具有很好的全局搜索能力和鲁棒性。



## 3.1.1遗传算法实现

1. 染色体编码：整数编码法

在本问题的求解中，染色体编码采用整数编码法，即用每个城市的编号代表该城市，一条完整的路径就是染色体编码，如“[5, 17, 32, 30, 8, 7, 15, 13, 16, 11, 9, 3, 4, 26, 10, 20, 12, 6, 27, 24, 22, 2, 19, 23, 21, 29, 31, 28, 1, 25, 14, 18, 5]”作为一条染色体，多个个体相互之间组成一个种群。

1. 初始化种群

本文中采用改进的随机种群。首先随机生成初始种群，并计算初始种群中的个体适应度，为了获得初始更好的种群，如果随机交换染色体中两个城市的位置可以使总距离减小，则更新该染色体，将两个城市的编号调换。

1. 适应度函数

染色体上各点之间的欧氏距离之和即为适应度值。TSP问题为距离极小化问题，因此可以直接使用染色体代表路径的总距离作为适应度函数，距离越小适应度越高，存活率与总距离的倒数成正比。

1. 交叉操作

本文采用顺序交叉方法。在父代中选择个体a，选取个体a的交叉部分生成子代a’的部分基因判断，再将父代a中未被选中的基因复制到子代a’基因的空余部分，最终组成一条完整的基因。

1. 变异操作

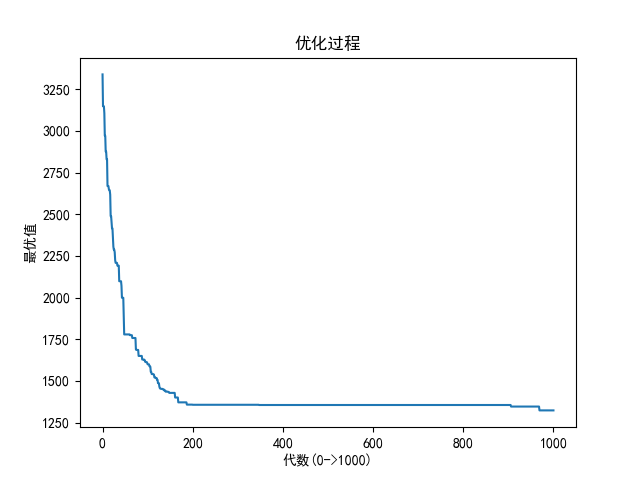
依据变异算子的概率，在染色体中随机选择两个不同位置的基因进行交换。

1. 终止条件

当迭代次数达到最大迭代次数时，则将进化过程中产生的适应度最大的个体作为遗传寻优的最优值输出，算法停止。

## 3.1.2遗传算法结果

本文中设置遗传算法的最大迭代次数为1000次，种群数量为200，迭代过程中总距离的变化如下图：

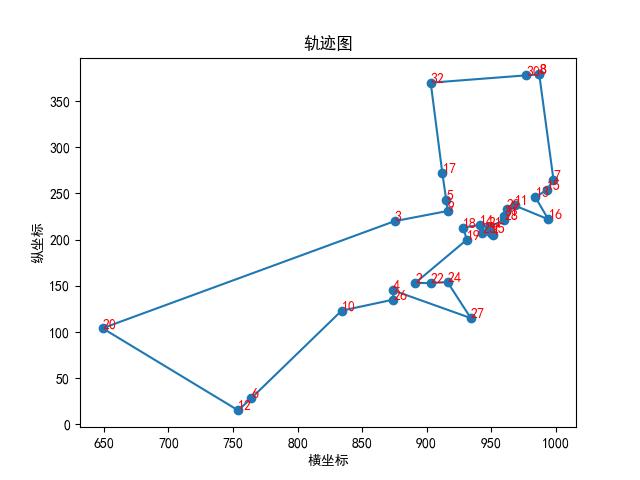


最终的最优路径为：

8->7->15->13->16->11->29->31->28->21->25->1->23->14->18->19->2->22->24->27->4->26->18->6->12->20->3->9->5->17->32->30->8

该路径对应的总距离值为1322.893

求解路径轨迹示意图如下：



## 3.2 贪心遗传算法

## 3.2.1 初始化个体设置

设种群的规模为M，基于始终选择距离商人所在城市最短且未走过的城市作为下一个要推销货物的城市，得到一条哈密尔顿回路的贪心策略；产生一个质量较高的正常染色体 ，然后进行 (M一1)次变异操作。得到的种群质量很高且其中的个体均为正常染色体。

## 3.2.2 种群规模设置

设种群的规模为M，在多次实验中以M+100的方式逐步增大种群规模，得到不同种群规模下贪心遗传算法的求解质量与速度。

## 3.2.3 交叉、变异概率设置

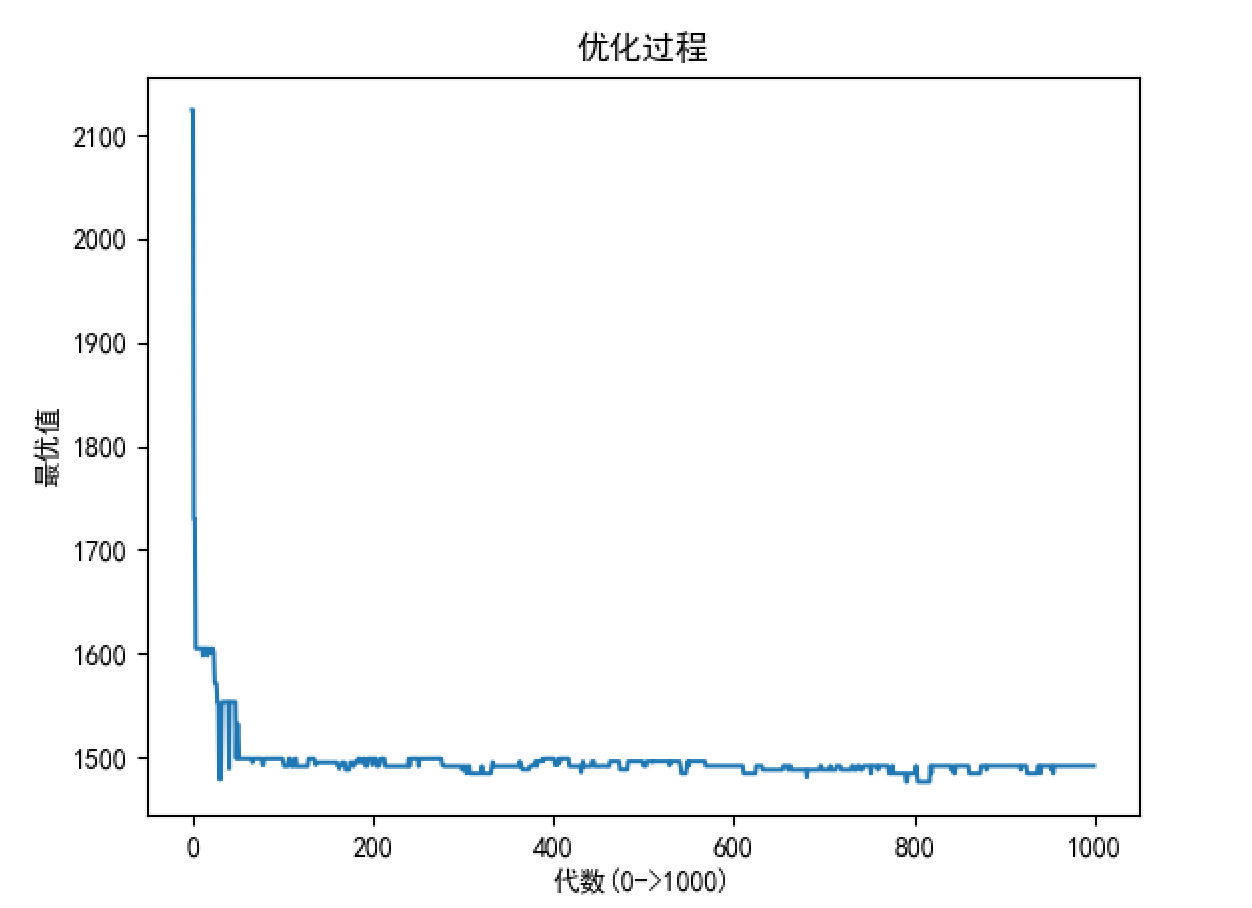
遗传算法中的交叉概率Pc和变异概率Pm的选择直接影响到算法的收敛性，贪心遗传算法中采用自适应交叉概率和变异概率，它们的计算公式如下：

其中fmax表示群体中的最大适应值；favg表示群体平均适应值；f表示要交叉的两个个体中较大的适应度值 ；f'表示要变异个体的适应值；k1，k2，k3，k4为常数，算法中取k1=0.5、k2=0.6、k3=0.01、k4=0.02。

当种群中各个体的适应度趋于一致或趋于局部最优时，使交叉概率和变异概率增加，相反，当群体适应度比较分散时，使交叉概率和变异概率减少。同时，对于适应度值高于群体平均适应值的个体，对应于较低的交叉概率和变异概率，使个体得以保护进入下一代;而低于平均适应值的个体,相对应于较高的交叉概率和变异概率。

## 3.2.4贪心遗传算法结果

本文中设置遗传算法的最大迭代次数为1000次，种群数量为200，迭代过程中总距离的变化如下图：



最终的最优路径为：0->24->22->20->13->17->18->8->4->16->28->10->30->27->15->12->14->6->7->29->31->2->1->21->23->3->25->9->19->11->5->30

该路径对应的总距离值为1433.617

求解路径轨迹示意图如下：

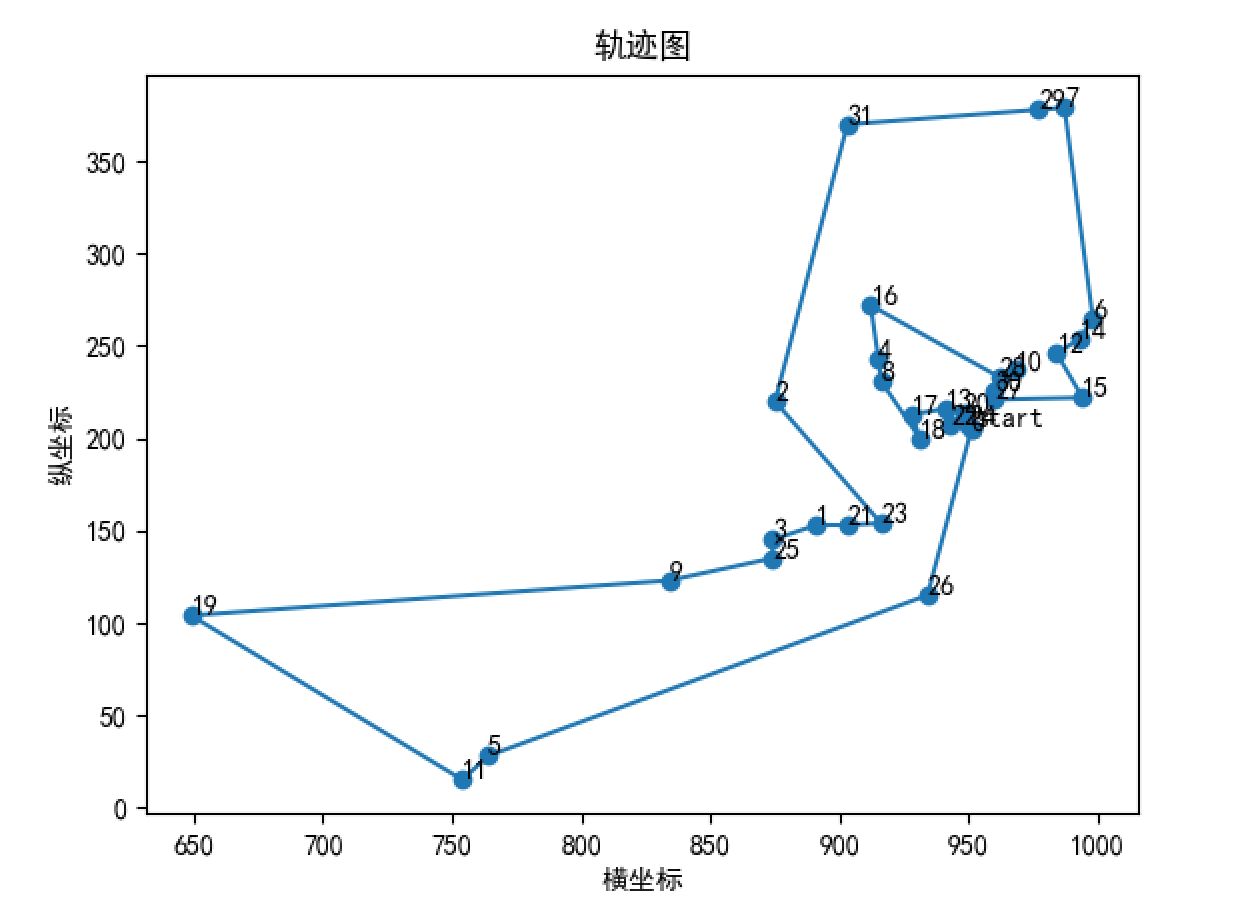


表2 种群数与最优值变化

|  |  |  |
| --- | --- | --- |
| 种群个数 | 最短距离 | 最优代数 |
| 100 | 1471.433 | 150 |
| 150 | 1452.436 | 200 |
| 200 | 1433.617 | 70 |

# 四、结果分析

## 4.1 算法对比

利用贪心遗传算法和基本遗传算法对上述实例分别进行了多次独立实验种群规模M=200，最大迭代次数1000代，实验结果统计如表3所示。

表3 实验结果比较

|  |  |  |
| --- | --- | --- |
| 比较指标 | 贪心遗传算法 | 基本遗传算法 |
| 平均迭代次数 | 100 | 180 |
| 最短距离 | 1433.617 | 1322.893 |

由表3可以看出：

1)贪心遗传算法的平均迭代次数优于基本遗传算法；

2)在求解结果方面，贪心遗传算法的求解具有不稳定性，其最好解与基本遗传算法的最好解基本一致。

由此，贪心遗传算法可以进一步提高基本遗传算法的求解速度，并且求解质量仍保持得较好。

## 4.2 优缺分析

优点：

1）贪心遗传算法利用贪心选择策略指导遗传操作，保证算法每步运行的有效性 。

2）贪心遗传算法因算法初始种群的生成、交叉、变异等环节都依据贪心选择原则指导操作，避免了传统遗传算法随机性操作的弊端 ，大大提高了算法的效率，加快了求解速度和提高了求解质量。

缺点：

贪心遗传算法的求解结果不完全稳定性，需要多次实验进行多次实验的修正结果。

# 五、总结与展望

基本遗传算法遗传进化机制是通过保存优良个体，使种群的整体性能得到提高，进而一步步地搜索寻优。本文在遗传进化过程中采用了贪心选择策略，即经过了交叉变异等基因操作之后，子代个体如果优于父代个体则替换，否则，不进行替换操作。这样做的目的是保证了个体进化方向， 降低随机操作的错误率。贪心遗传算法，该算法的运行机制是：利用贪心选择策略指导遗传操作，保证算法每步运行的有效性。移民向种群引入新的遗传物质，避免了种群封闭竞争的缺陷，使搜索不致陷入早熟收敛。由于算法初始种群的生成、交叉、变异等环节都依据贪心选择原则指导操作，避免了传统遗传算法随机性操作的弊端，算法相对较稳定。仿真试验结果表明，贪心遗传算法减少了GA计算工作量，大大提高了算法的效率。

贪心遗传算法是将基本遗传算法和贪心策略融合的产物。贪心策略的融入，改进了遗传算法的性能，提高了求解的质量和速度。

# 六、参考文献

[1] 程荣. 遗传算法求解旅行商问题[J]. 科技风，2017,16(37):40-51.

[2] 郭丰林. 基于遗传算法的旅游线路规划研究[J]. 现代营销 （经营版），2019(1):134.

[3]王秋芬,袁东锋,梁道雷.一种求解TSP的贪心遗传算法[J].制造业自动化,2013,35(02):71-74.

[4]范立南,吕鹏.基于改进遗传算法的校园快递配送路径规划[J].物流科技,2021,44(01):14-19.DOI:10.13714/j.cnki.1002-3100.2021.01.004.

[5]魏英姿,赵明扬,黄雪梅,胡玉兰.求解TSP问题的贪心遗传算法[J].计算机工程,2004(19):19-20+34.