第19卷第3期　合肥工业大学学报(自然科学版)　Vol.19№3
1996.9　JOURNAL OF HEFEI UNIVERSITY OF TECHNOLOGY　Sept.1996

用遗传算法求解最短路径问题

曹鲁寅　罗　斌　　钦明浩
(安徽大学)　　(合肥工业大学)

摘要　文章应用遗传算法求解图论中的最短路径问题,并提出了该算法在解决这一问题中的一些处理方法.使用该算法可以很快地求出一批最短路径集。文中最后给出了算法运行结果及总结。

关键词　最短路径;遗传算法;邻接矩阵

中图分类号　O157.5①

1　最短路径问题和遗传算法

所谓最短路径问题就是在给定的起始点 S 到终止点 t 的通路集合中,寻求长度最小的通路,这样的通路称为 S 点到 t 点的最短路径。

在寻找最短路径问题上,有时人们不仅要知道两个指定的顶点间的最短路径,还需要知道某个顶点到其它任意顶点间的最短路径。用遗传算法解这类问题,没有太多的约束条件和有关解的限制,因而可以很快地求出任意两点间的最短路径以及一批次短路径。

遗传算法(Genetic Algorithm,简写 GA),是新近发展起来的一种模拟生命进化机制的搜索和优化方法,是把自然遗传学和计算机科学结合起来的优化方程。1975年,Holland 在其专著中指出了 GA 的概念和方法,因其有很强的解决问题的能力和广泛的适应性,因而近年来渗透到研究与工程的各个领域,取得了良好的效果。

下面介绍遗传算法的几个基本概念:

(1)染色体(Chromosome):在使用遗传算法时,需要把问题的解编成一个适合的码子。这种具有固定结构的符号串即是染色体,符号串的每一位代表一个基因。符号串的总位数称为染色体的长度。一个染色体就代表问题的一个解,每个染色体也被称为一个个体。

(2)群体(Population):每代所产生的染色体总数称为群体,一个群体包含了该问题在这一代的一些解的集合。

(3)适应度(Fitness):对群体中每个染色体进行编码后,每个个体对应一个具体问题的解,而每个解对应于一个函数值。该函数值即适应函数,就是衡量染色体对环境适应度的指标,也是反映实际问题的目标函数。

① 收稿日期:1996-05-05

在前代群体的基础上产生新一代群体的工作称为遗传操作，基本的遗传操作有：

(1)选择(Select)：按一定的概率从上代群体中选择 M 对个体作为双亲，直接拷贝到下一代，染色体不发生变化。

(2)交叉(Crossover)：对于选中进行繁殖的两个染色体X，Y，以X、Y为双亲作交叉操作，从而产生两个后代X′、Y′。

(3)变异(Mutation)：对于选中的群体中的个体(染色体)，随机选取某一位进行取反运算，即将该染色体码反转。

用遗传算法求解的过程是根据待解问题的参数集进行编码，随机产生一个种群，计算适应函数和选择率，进行选择、交叉、变异操作，如果满足迭代收敛条件，此种群为最好个体，否则，对产生的新一代群体重新进行选择、交叉、变异、操作，循环往复直到满足条件。

2 求最短路径问题的遗传算法的表示与实现

用遗传算法求解一个优化问题，就是对该优化问题存在许多解 x_i，计算每个 x_i 对应的适应函数 f_i，优化的过程就是要寻找这样的 x_m，使得与之对应的 $f(x_m)$ 最大或最小。

2.1 最短路径问题的图论描述

求最短路径问题，用图论术语描述如下：在图 $G(V,A)$ 中，V 表示顶点集合，$V=(v_1,v_2,\cdots,v_n)$ 对 G 中的某一条边 (v_i,v_j)，相应地有一个数 $d(v_i,v_j)$，如果 G 中不存在边 (v_i,v_j)，则令 $d(v_i,v_j)=\infty$，如把 $d(v_i,v_j)$ 认为是边 (v_i,v_j) 的长度(也可认为是边 (v_i,v_j) 的费用或权)，则路的长度定义为组成路的各条边的长度的总和。

顶点 v_i，v_j 之间是否有边相连，由邻接矩阵来决定。

邻接矩阵 A：对一个具有 V 个顶点，e 条边的图 G 的邻接矩阵 $A=[a_{ij}]$ 是一个 $v\times v$ 阶方阵，其中 $a_{ij}=1$，表示 v_i 和 v_j 邻接，$a_{ij}=0$，表示 v_i 和 v_j 不相邻接(或 $i=j$)。

2.2 染色体编码

对于一个给定的图模型，将图中各顶点按顶点号自然排序，然后按此顺序将每个待选顶点作为染色体的一个基因，当基因值为1时，表示相应的顶点被选入该条路径中，否则反之。此染色体中的基因排列顺序即为各顶点在此条通路中出现的先后顺序，染色体的长度应等于该图中的顶点个数。

2.3 适应函数 $f(i)$

对具有 n 个顶点的图，已知各顶点 (v_i,v_j) 的边长度 $d(v_i,v_j)$，把 v_{i1} 到 v_{jn} 间的一条通路 $v_{i1},v_{i2},\cdots,v_{in}$ 的路径长度定义为适应函数：

$$f(i)=\sum_{r=1}^{n-1}d(v_{ir},v_{ir+1})$$

对该优化问题，就是要寻找解 x_m，使 $f(x_m)$ 值最小

2.4 选择操作

选择作为交叉的双亲，是根据前代染色体的适应函数值所确定的，质量好的个体，即从起点到终点路径长度短的个体被选中的概率较大。

2.5 交叉与变异操作

将被选中的两个染色体进行交叉操作的过程是先产生一个随机数，确定交叉点，位于染色体的第几位基因上，然后在此位置进行部分基因交换。变异操作是将染色体中某位基因逆变，即由1变为0，或反之。变异的意义为在某条路径上去掉或增加某顶点，但这样做的结果并不一定能使路径的长度减少。也就是说有可能使各代中产生的比较好的方案在遗传过程中丢失，迟缓了获得最优解的速度。

为了使算法尽可能快地获得更好的解，改善遗传算法的收敛性，在变异操作时，增加了个体求优的自学习过程。即在某位基因变异后，计算新产生的染色体的适应函数值，若适应函数值更小，即获得的路径更短，则保留；否则，保持原来的解不变。如果有连续 $N/3$ 次没有得到更好的解，则该过程结束，其中 N 表示从起点到终点的顶点数。

解最短路径问题的遗传算法如下：

generate $[p(c)]$;
evaluate $[p(h)]$;
repeat noof Generations times:
Select $p(h)$ form $p(n-1)$;
Crossover and mutation $p(n)$;
learning $[p(n)]$;
evaluate $[p(n)]$;
end;

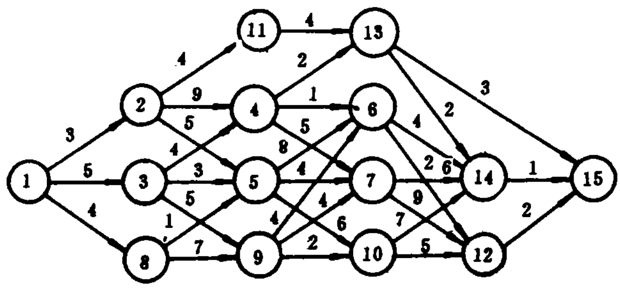

图1 15个顶点加权有向图

其中：select $p(n)$ form $p(n-1)$ 表示从前一代群体中选择一对双亲，用于交叉、变异操作。$p(n)$ 代表第 n 代群体。

generate $[p(n)]$ 表示在程序开始时要首先产生一个群体。

evaluate $[p(n)]$ 表示计算每个个体适应度。

3 算法实现与结果

3.1 原始数据对解的影响

交叉率 p_c 不可选择过小，否则，延缓获得最优解的过程，本题选择 $p_c=0.9$。

变异率(p_m) 的选择对规模大的优化问题影响很大，本题选 $p_m=0.009$。

群体中的个体数的选取是算法中一个很重要的参数，群体中的个体数目越大，算法就越能找到更好的解。个体数目过小，有可能找不到最优解。

3.2 算例

(1)求图1中从1点到15点间的最短路径。图中 v_i 到 v_j 的数值代表两点间的路径长度。对该例用本算法求得的最短路径集为：

$$1\to3\to4\to6\to12\to15;$$
$$1\to2\to11\to13\to15;$$
$$1\to2\to11\to13\to14\to15;$$
$$1\to3\to4\to13\to15;$$
$$1\to3\to4\to13\to14\to15;$$

最短路径的长度为14。

(2)图2给出了60个顶点的例子

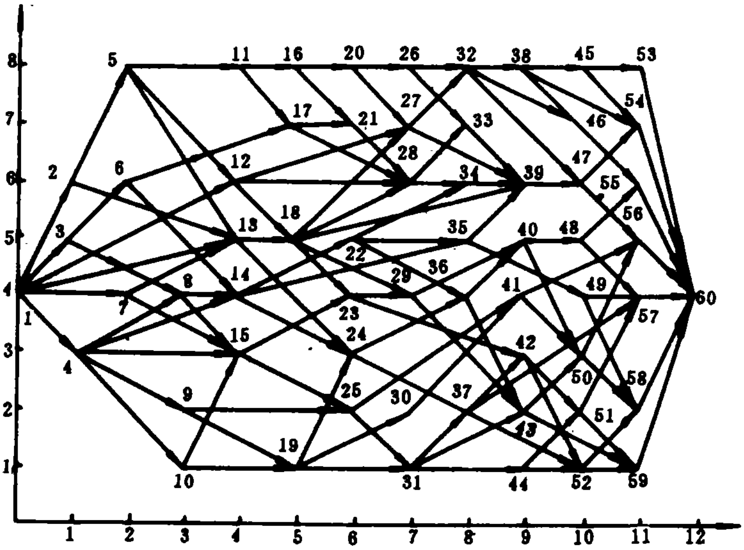

图2　60个顶点加权有向图

上图用坐标给出了60个顶点的位置，若两点 v_i,v_j 之间有连线，则权值 $d(v_i,v_j)$ 取两点间的距离；否则 $d(v_i,v_j)$ 取一个大数。求解过程中，迭代次数与计算结果关系见表1。

对图2，求得的最短路径为：

$$1\rightarrow4\rightarrow14\rightarrow35\rightarrow49\rightarrow57\rightarrow60$$

最短路径长度为12.936。

遗传算法在最初几次迭代中，个体的出现是良莠并存的，个体的适应度也不高，随着迭代次数的增加，适应度高的个体将被遗传出来。表1中的结果也说明了这点。

表1　遗传算法优化过程

迭代次数	路　径	适应函数
3	1→2→3→4→14→17→19→22→24→26→27→36→40→48→49→5157→60	1100.048
8	1→4→10→16→41→56→60	205.890
12	1→13→18→29→40→48→57→60	13.009
17	1→4→14→35→49→57→60	12.936

4　总　　结

通过用遗传算法解题，可知

遗传算法明显的优点：(1)算法是使用参数的编码集参数的选择十分方便；(2)遗传算法是在点群中寻优；(3)它仅使用问题本身所具有的目标数据进行工作，而不需其它任何先决条件或辅助信息；(4)它使用的是随机规则。

本文是遗传算法一个应用,在用该算法解题时,增加了个体的自学习过程,因而克服了简单的遗传算法存在的收敛速度慢的缺点,改善了算法的收敛性。目前,对GA的理论研究还在深 入,应用领域也在不断地开拓,相信用遗传算法解决的问题将越来越多。

参 考 文 献

1 [美]E. 米涅卡·网络和图的优化算法. 中国铁道出版社,1984

2 David B Fogel. An Introduction to Simulated Evlutionary Optimization. IEEE Trans. Neural Networks. 1994.5(1)

3 Qi Xiaofeng. Palmieri Francesco. Theoretical Analysis of Evolutionary Algorithms with an Infinite Population Size in Continuous Space. Part I: basis properties IEEE Trans. Neural Networks. 1994.5(1)

4 陈根社,陈新海. 遗传算法的研究与进展. 信息与控制,1994,23(4)

A GENETIC ALGORITHM FOR FINDING SHORTEST PATHS

Cao Luyin *Luo Bin* *Qing Minghao*

(Anhui University) (Hefei University of Technology)

Abstract This paper presents the application of genetic algorithm to finding the shortest paths in a graph. A number of issues of genetic algorithm for solving this problem are presented. A series of shortest paths can be obtained quickly by using this algorithm. At the end of this paper, the calculation resultes of the algorithm and the conclusion are shown.

Key Words genetic algorithm, shortest paths, adjacency matrix

姓名 曹鲁寅 出生于 1963 年 4 月 1987 年毕业于 安徽大学电子 系 学位 硕士 职称 讲师

主要研究方向 VLSI 计算机辅助设计

联系地址 安徽大学电子系 邮编 230039

(本文责任编辑 涂捷)