

基于 Dijkstra 算法的一种最短路径优化算法

张福浩,刘纪平,李青元

(中国测绘科学研究院,北京 100039)

摘要:详细介绍了经典的 Dijkstra 算法,举例说明了该算法的实现方法以及该算法的缺点:即需要网络结点数平方级的内存;同时详细说明了一种基于 Dijkstra 算法的优化算法——邻接结点算法,该算法充分利用了网络拓扑信息中的弧段的连接关系,避免了使用含有大量无穷值的关联矩阵,使之更适合带有拐向限制设置的最短路径算法和大量结点的实际数据。实践证明,该算法可以节约大量的内存,对于结点数比较大的网络,或带有大量拐向限制设置的网络,具有较好的适用性。

关键词:网络分析;最短路径分析;Dijkstra

中图分类号:P208 文献标识码:A 文章编号:1000-3177(2004)74-0038-04

1 引言

网络分析是 GIS 中非常重要的空间分析功能之一,最短路径分析又是网络分析的核心算法,该算法的效率决定了网络分析的功能和效率。在大量的最短路算法中,Dijkstra 算法是一种最经典的方法,很多算法都是在该算法的基础上经过改进发展而来的。本文就是在详细研究该算法的基础上,结合矢量数据的拓扑关系,满足实际应用中的网络分析需求而优化的一种算法。本文首先详细介绍 Dijkstra 算法,并举例说明该算法的缺陷,并在此基础上提出一种新的优化方法。

2 Dijkstra 算法

2.1 Dijkstra 算法的原理

设 $D=(V,A,w)$ 是一个非负权网络, $V=(v_1,v_2,\dots,v_n)$ 。则 D 中最短 (v_i,v_j) 路的长满足方程:

u1=0
uj=min(uk+wkj) (j=2,3,...,n) (1)

如果 D 中从顶点 v_1 到其余各顶点最短路的长按照大小排序为:

u11 u12 ... u1n

这里 $i1=1,u_{i1}=0$,则由(1)式有

uij = min_k { uik + w_kij }
= min { min_{k < j} { uik + w_kij }, min_{k > j} { uik + w_kij } } (j = 2,3,...,n)

当 $k > j$ 时, $u_{ik} \geq u_{ij}$, 且 $w_{kij} \geq 0$, 从而

uij ≤ uik + w_kij

即

uij = min_{k > j} { uik + w_kij }

所以

uij = min_{k < j} { uik + w_kij }

容易证明:

u11 = 0

uij = min_{k < j} { uik + w_kij }

的解 $(u_{i1}, u_{i2}, \dots, u_{in})$ 中的 u_{ij} 是 D 中最短 (u_i, u_j) 路的长度, $j=1,2,\dots,n$ 。

这就是 Dijkstra 的理论证明,为了便于说明,采用下面的方法来举例说明。

在图 1 所示的一个简单网络中,根据其距离与结点之间的关系,形成邻接矩阵与距离矩阵,在此基础上,用着色的方法表示 Dijkstra 的计算方法,计算从 A 点到 D 点的最短距离如下:

第一步 A 着色, $d(A)=0, d(D)=\infty$ 。

第二步 $y=A$

d(B1)=min(d(B1), d(A)+a(A,B1))=min(∞,0+4)=4

d(B2)=min(d(B2), d(A)+a(A,B2))=min(∞,0+5)=5

d(B3)=min(d(B3), d(A)+a(A,B3))=min(∞,0+3)=3

$d(B3)=3$ 为最小,对 $B3$ 着色。

第三步 D 尚未着色,继续执行第二步。

收稿日期:2003-03-15
基金项目:本文得到国家“十五”重大科技专项课题“中国电子政务空间辅助决策示范工程”支持(编号:2002BA105A-01) 国家基础测绘项目“政府空间辅助决策系统建设”(1460070324211)
作者简介:张福浩(1973~),男,1996年毕业于同济大学测量专业,2003年获得北京大学硕士学位,现从事地理信息系统与电子政务的研究。

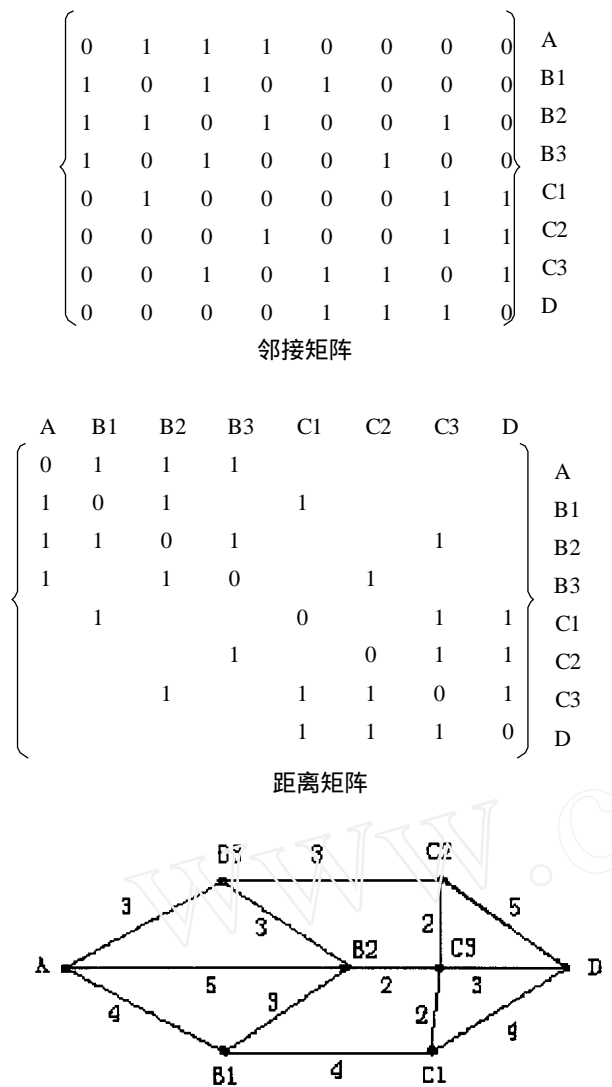


图 1 一个网络及其邻接矩阵与距离矩阵

第二步 令 $y=B3$

$d(B1) = \min(d(B1), d(B3) + a(B3, B1)) = \min(4, 3 + 3) = 5$

$d(B2) = \min(d(B2), d(B3) + a(B3, B2)) = \min(3, 3 + 3) = 6$

$d(C2) = \min(d(C2), d(B3) + a(B3, C2)) = \min(2, 3 + 3) = 6$

$d(B1) = 3$ 为最小,对 $B1$ 着色。

第三步 D 尚未着色,继续执行第二步。

第二步 令 $y=B1$

$d(B2) = \min(d(B2), d(B1) + a(B1, B2)) = \min(6, 4 + 3) = 5$

$d(C1) = \min(d(C1), d(B1) + a(B1, C1)) = \min(4, 4 + 4) = 8$

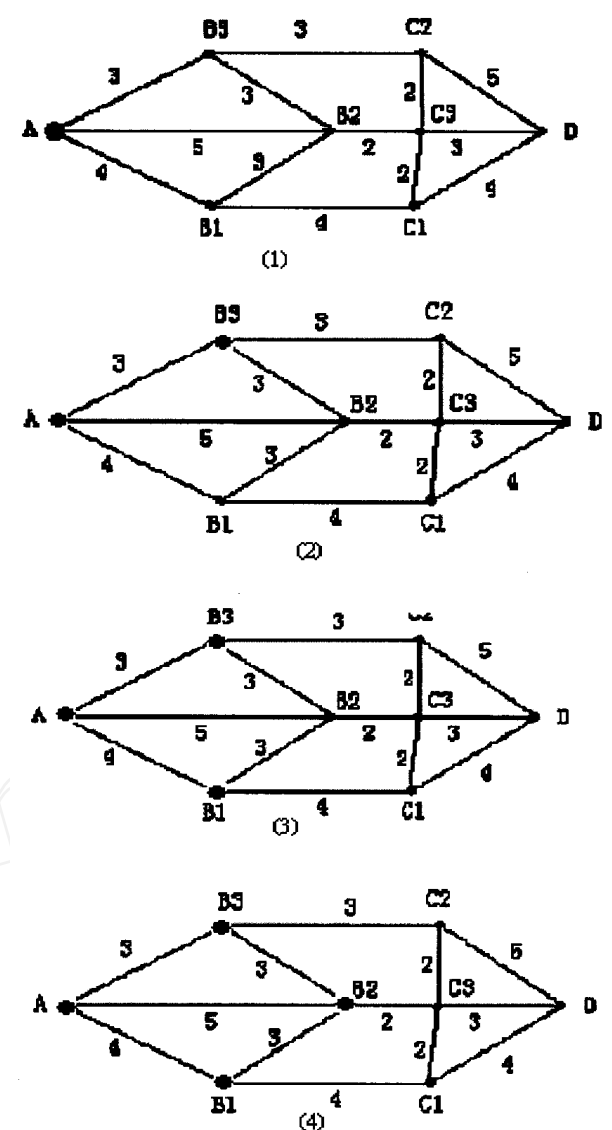


图 2 用着色的方法计算从 A 点到其他各点的最短距离

$d(B2) = 5$ 为最小,对 $B2$ 着色

第三步 D 尚未着色,继续执行第二步。

第二步 令 $y=B2$

$d(C3) = \min(d(C3), d(B2) + a(B2, C3)) = \min(2, 5 + 2) = 7$

$d(C2) = 6$

$d(C2) = 5$ 为最小,对 $C2$ 着色。

第三步 D 尚未着色,继续执行第二步。

第二步 令 $y=C2$

$d(D) = \min(d(D), d(C2) + a(C2, D)) = \min(4, 6 + 5) = 11$

$d(C3) = \min(d(C3), d(C2) + a(C2, C3)) = \min(7, 6 + 2) = 7$

$d(C3) = 7$ 为最小,对 $C3$ 着色。

第三步 D 尚未着色,继续执行第二步。

第二步 令 $y = C3$

$$d(D) = \min(d(D), d(C3) + a(C3, D)) = \min(10, 6 + 4) = 10$$
$$d(C1) = \min(d(C1), d(C3) + a(C1, C3)) = \min(8, 7 + 2) = 9$$

$d(C1) = 9$ 为最小,对 $C1$ 着色。

第三步 D 尚未着色,继续执行第二步。

$$d(D) = \min(d(D), d(C1) + a(C1, D)) = \min(10, 9 + 4) = 10$$

对 D 着色,停止计算。

得到从 A 到 D 的最短距离为 10。

2.2 Dijkstra 算法的缺点

使用 Dijkstra 算法基于网络的权矩阵求解最短路问题的计算机算法和程序中,运用了关联矩阵、邻接矩阵和距离矩阵的概念。在存储图形数据和运算时,需要定义 $N \times N$ 的数组,其中 N 为网络的结点数,当网络的结点数较大时,将占用大量的计算机内存。如果不对 Dijkstra 算法进行优化,该算法很难在实际中得到应用。

3 最短路的邻接结点算法

随着网络空间数据的增大,网络结点数的增加,直接使用 Dijkstra 算法是不现实的,要想处理大数据量的最短路径分析,必须进行优化。下面还以上面的例子说明基于 Dijkstra 算法的一种优化算法——邻接结点算法。

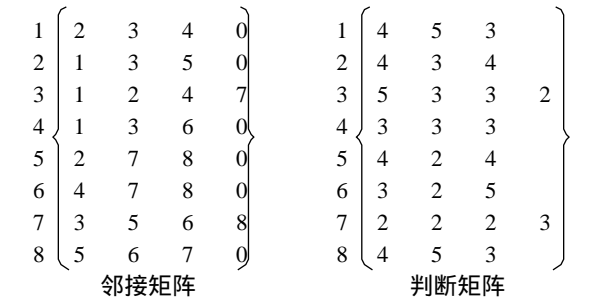


图3 网络的邻接矩阵与判断矩阵

- 首先做以下准备工作:
- (1) 依据最大相邻结点数的概念,求出网络的最大相邻结点数 $m = 4$ 。
 - (2) 构造邻接矩阵 Γ 。以结点为行,以相邻的结点为列,矩阵的行数为网络的实际结点数,列数为网络的最大相邻结点数 m 。与结点 i 相连的边号为矩阵的第 i 行。如果该结点的相邻结点数少于最大相邻结点数 m ,则用 0 代替。如图 1 网络的邻接矩阵和图 3 中的邻接矩阵。
 - (3) 构造判断矩阵 pJ 。对照邻接矩阵,用邻接矩阵里的

各个元素对应边号的边长代替同一位置的边号就构成了判断矩阵 pJ 。如图 1 网络的判断矩阵和图 3 中的判断矩阵。

(4) 在邻接矩阵 Γ 和判断矩阵 pJ 基础上求网络上任意两点间的最短路径。

然后,开始计算 A 到 D 的最短路径,步骤如下:

初始化临时标识向量 $T, Ti = 0, i = 1, 2, \dots, 8$ 。(着色)

查找判断矩阵的起始行中不为 0 的值,初始化距离向量 D 和标记向量 P :

$$D2 = 4,$$
$$P2 = 1,$$
$$D3 = 5,$$
$$P3 = 1,$$
$$D4 = 3,$$
$$P4 = 1,$$

计算距离向量中的最小值 $f_{\min} = 3$ 。

把对应的标识向量赋值 $T4 = -1$,并得到结点 $i = 4$ 。

查找邻接矩阵的 i 行中不为 0 的值,按照方程

$$uj = \min(uk + wkj) \quad (j = 2, 3, \dots, n)$$

循环计算与结点 i 相连的所有的其他结点与起点(结点 1)的距离,并判断该距离是否小于距离向量中对应的值,如果小,则:

赋值距离向量 D 和标记向量 P :

以第 4 行的结点 6 为例:

$$u6 = u4 + u46 = D4 + U46 = 3 + 3 = 6$$
$$6 < D6 =$$

则赋值距离向量 $D6 = 6$,标记向量 $P6 = 4$ 。

重复第 ~ 步:

计算距离向量中的最小值 f_{\min} 。标识向量值为 -1 的不计算在内, $f_{\min} = 4$;

把对应的标识向量赋值 $T2 = -1$,并得到结点 $i = 2$;

查找邻接矩阵的 i 行中不为 0 的值,按照方程

$$uj = \min(uk + wkj) \quad (j = 2, 3, \dots, n)$$

循环计算与结点 i 相连的所有的其他结点与起点(结点 1)的距离,并判断该距离是否小于距离向量中对应的值,如果小,则:

赋值距离向量 D 和标记向量 P :

以第 2 行的结点 5 为例:

$$u5 = u2 + u25 = D2 + U25 = 4 + 4 = 8$$
$$8 < D5 =$$

则赋值距离向量 $D5 = 8$,标记向量 $P5 = 2$ 。

重复第 ~ 步:

直到距离向量中的最小值为终止结点 8 时,该最小值就是要求的最短距离。

如果还没有计算到终止结点 8,最小值为 ,则不存在最短路径,即不连通。

通过标记向量 P ,可以很容易跟踪出经过的路径。

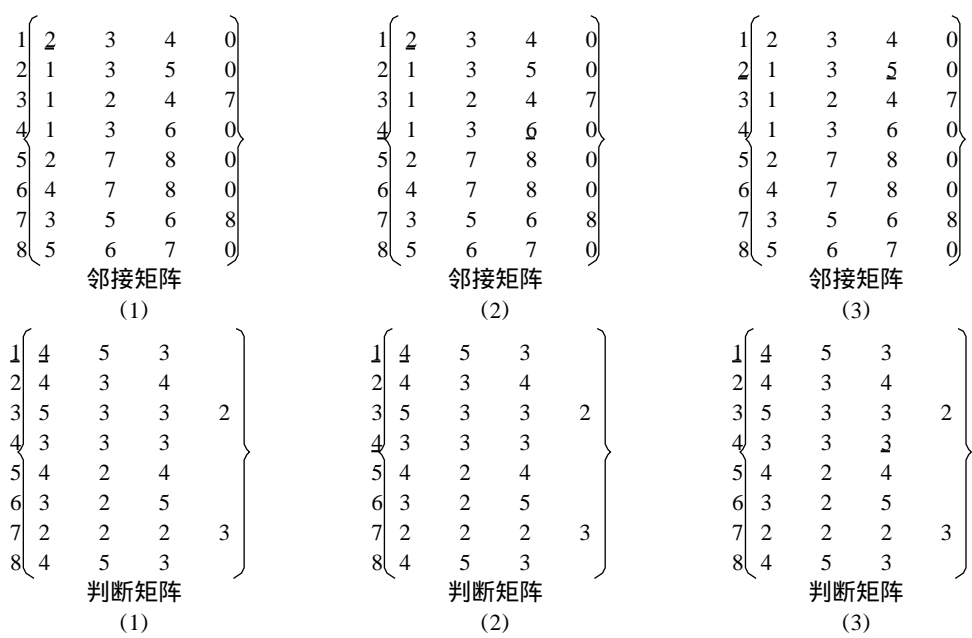


图 4 通过邻接矩阵与判断矩阵计算两点之间的最短距离

上述的邻接结点算法避免了使用含有大量无穷值的关联矩阵,使之更适合带有拐向限制设置的最短路径算法和大量结点(超过 8000 个结点)的实际数据。实践证明,该算法可以节约大量的内存,对于结点数比较大的网络,或带有大量拐向限制设置的网络,具有较好的适用性。经过笔者的应用验证,该优化算法不但节约内存而且效率很高。

改进的最短路径算法——邻接结点算法,该算法避免使用关联矩阵,利用了网络拓扑信息中的弧段的连接关系,并同时满足拐向限制设置和阻断的要求,节省内存,可以应用于结点数巨大的网络。但是,该优化算法目前只支持单幅地图的操作,多图幅的网络分析功能需要对数据作适当的预处理。

4 讨论

本文介绍了 Dijkstra 算法的原理,并在此基础上,提出了

参考文献

1 WU Qi - shi. The Application of Genetic Algorithm in GIS Network Analysis[J]. International Archives of Photogrammetry and Remo, 2000, 33: 1184 ~ 1191.

2 王杰臣. GIS 网络分析的图简化方法研究[J]. 测绘学报, 2001 (3) :263 ~ 268.

3 邵振峰. 配电网地理信息系统中的网络重构[J]. 测绘通报, 2001 (3) .

4 曾文. Design of MAPGIS Pipeline Management Development Platform[J]. 中国地质大学学报, 2002 ,27 (3) :250 ~ 254.

5 刘方爱. A Hierarchical Network HRN AND its Routing Algorithms[J]. 计算机学报, 2002 ,25 (12) :1397 ~ 1441.

6 王杰臣. 最短路径问题的一种改进算法[J]. 解放军测绘学院学报, 1999 ,16 (14) :282 ~ 285.

7 王杰臣. 图的节点 - 弧段联合结构表示法及其在 GIS 最优路径选取中的应用[J]. 测绘学报, 2000 ,29 (1) :47 ~ 51.

8 李霖. 变量查询代数及最短路径分析[J]. 测绘学报, 2000 ,29 (1) :59 ~ 63.

9 张锦明. 利用分区思路优化拓扑关系自动生成算法[J]. 测绘学院学报, 2000 ,17 (2) :119 ~ 122.

A New Way of Network Analysis Based on Dijkstra

ZHANG Furhao ,LIU Jirping ,LI Qingyuan

(Chinese Academy of Surveying and Mapping , Beijing 100039 , china)

Abstract : This paper introduces the classical arithmetic of Dijkstra , and it 's limitation , which needs geometrical progression memory with increase of network nodes. The paper emphasizes an optimization of shortest path - the algorithm of adjoining noddess , and its improvement , which takes full advantage of the linking relation of arc section in network topology. This avoids conjunction matrix which includes lots of infinitude and suits huge data which includes turning. It is proved that the method can save lots of memory and suit not only huge network but also network of tuning restriction.

Key words : GIS ;network ;network analysis ;shortest path analysis