

一种基于海量电子地图的空间拓扑分块检索算法^①

沈永增, 沈旭东, 吕天剑

(浙江工业大学 信息工程学院, 杭州 310023)

摘要: 路网拓扑信息是进行 GIS 空间分析的数据基础。目前由于实际的道路数据量越来越庞大, 而且嵌入式系统又受到自身条件的限制, 因此不适宜把全部的路网拓扑数据读入到内存中, 针对内存大小受到限制的问题, 通过在嵌入式电子硬盘上对路网拓扑数据进行“分块”存储, 并给出了在进行最优路径查找时, 在不同的拓扑分块之间进行数据块检索的方法。

关键词: 海量电子地图; 路网组织; 分块数据检索; 最优路径

Topological Retrieval Algorithm for Blocks Base on Mass E-Map

SHEN Yong-Zeng, SHEN Xu-Dong, LV Tian-Jian

(College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China)

Abstract: Road network topology information is the basis for GIS spatial analysis. Currently as the actual amount of road is increasingly growing, and embedded systems has been its own constraints, it is not suitable for reading all of the road network topology data into memory, against the problem of limited memory size. In this paper, by means of storing road network topology as “blocks” in the embedded electronic drives, and give the way to retrieve data block between different topological blocks when finding the optimal path.

Key words: mass e-map; road network organization; block data retrieval; optimal path

1 引言

由于城市路网数据的海量特性, 对于嵌入式系统, 在进行最优路径查找时, 不可能把全部的数据一次性的读入内存中, 因此有必要对海量电子地图拓扑数据进行“分块”处理, 使得每次读入内存的数据量为有限的分块数据大小。所以在进行最优路径查找时, 如何组织这些数据块, 以便在算法的运行过程中能够高效的检索合适的分块数据, 进行快速的调入、调出, 是值得研究的一个问题。

本文通过对 MapInfo 电子地图文件格式进行深度解析, 给出由矢量路网数据转换为拓扑数据的一般方法^[5], 并对路网拓扑数据进行“分块”存储, 最后给出对分块拓扑数据的快速检索算法, 实现最优路径的快速查找。

2 路网组织

2.1 路网模型

图 1 显示了道路路网的基本的集合元素和拓扑关系。图中节点是指道路段与道路段的交点或者道路的终起点, 形状点是指同一非直线道路上的转弯点, 它和节点一起构成一条道路。

假设把形状点包含在道路的形状属性数据里面, 则可定义如下路网模型:

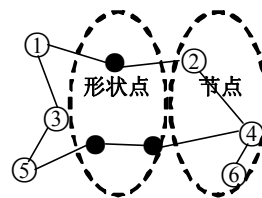


图 1 路网中的几何元素和拓扑关系

^① 基金项目:浙江省科技厅项目(2007C30008)

收稿时间:2010-10-16;收到修改稿时间:2010-11-13

$$\begin{cases} R_w=(P, R) \\ R=\{<x, y>|x, y \in P, \text{And } C(x, y)\} \end{cases}$$

其中, 式中 R_w 表示整个道路路网;

P 表示整个路网节点集;

R 表示整个路网道路集合, 包含有序对 $<x, y>$ 元素, 其中谓词 $C(x, y)$ 表示从节点 x 到节点 y 存在一条有向通路, 如图 1 中所示, 如果分别由 x 和 y 代表节点 1 和节点 2, 则 $<x, y>$ 就表示从节点 1 到节点 2 有一条有向道路。

2.2 路网拓扑生成算法

本文采用 MapInfo 格式电子地图作为开发, 本身不具备构建对象拓扑结构的能力, 为了便于路径导航, 必须由 MapInfo 矢量电子地图构建出道路网的网络拓扑结构, MapInfo 原始道路数据格式如图 2 所示 (其中某条道路):

```
Pline 10
100.91637 22.34897401
100.91387 22.33000401
100.91499 22.317368
100.91988 22.30617299
100.92056 22.299736
100.91832 22.29456499
100.90803 22.28630401
100.8997 22.27795799
100.89252 22.26750399
100.88938 22.260246
Pen (2, 2, 15761473)
```

图 2 Mapinfo 格式道路数据

路网拓扑结构利用节点表达路段与其它路段之间的连通性, 根据上面定义的路网模型结构, 构建路网拓扑结构主要有以下几个部分: 提取和处理节点、路段信息、从而建立拓扑结构。具体方法如下:

(1) 构建节点和路段数据结构

```
Class Node //节点数据结构
{public:
int NodeID; //节点编号
int NodeLongitude; //节点经度
int NodeLatitude; //节点纬度
int NodeArcsNum; //相连路段个数
int NodeArcsID; //相连路段编号
}
Class Road //路段数据结构
{public:
```

```
int RoadID; //路段编号
int StartID; //首节点编号
int EndID; //尾节点编号
int ArcLength //路段长度
String RoadName; //道路名度
}
```

其中节点和路段 ID 是索引, 这样即使路段名重名也能区分它们。节点经纬度坐标原本是 double 型, 考虑到嵌入式系统存储数据能力有限, 而路网数据又非常的庞大, 所以在保证数据精度的前提下将 double 类型的经纬度值乘以 1000000 转变为占 4 个字节的 int 类型数据, 这样大大节省了存储空间。

(2) 算法设计

拓扑构建算法分为三步: 相交路段交点求取, 设置“捕捉距离”合并相近节点, 拓扑结构重建。

① 遍历原始路网数据的所有节点和形状点, 进行数据压缩 (去除冗余数据, 提高计算速度), 求出路网中所有的交叉节点, 对路网道路弧段进行重新规划, 原始路网的不规则路段拓扑关系主要分为以下几种情形如图 3 所示, 在求交点时必须考虑到实交和虚交两种情形, 具体算法如图 4 所示。

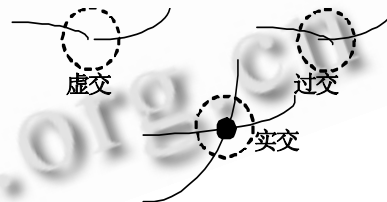


图 3 电子地图原始道路拓扑关系典型情形

② 在过交的拓扑关系情况下, 同一个交叉路口通常在地图上会显示多个点的描述, 对于图 3 中的过交情况, 在计算时会出现 3 个节点同时描述一个交叉点, 且 3 个节点的经纬度坐标是不一样的, 如图 5 所示, 这样将会导致求取的拓扑结构发生错误。

设置“捕捉距离”合并相近节点具体算法如图 6 所示, 捕捉距离的设置比较关键, 设置的过大, 就把两个不同的节点合并成了一个节点, 设置的过小, 节点冗余依然会大量存在, 因此需要设置合适的捕捉距离来合并相近节点, 本文设置的距离为 30M, 经过实验对比效果比较好。

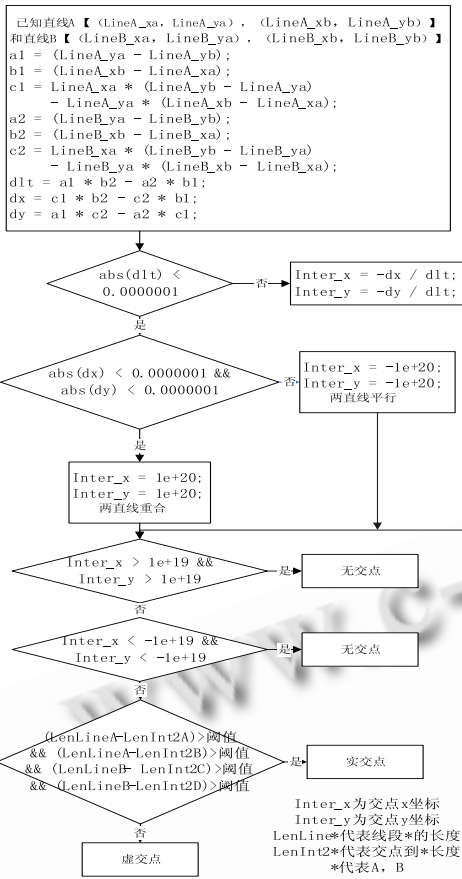


图4 路网求交点算法

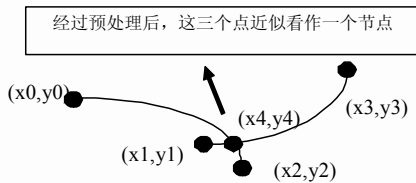


图5 节点冗余情形

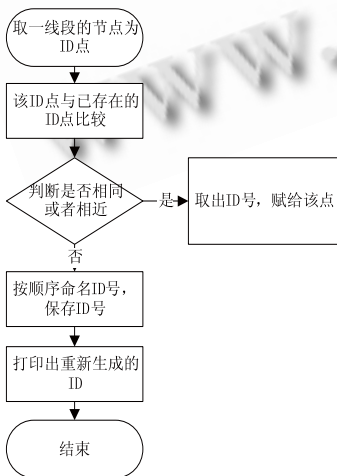


图6 合并冗余节点算法

③ 拓扑数据重建, 经过上面两步的预处理, 全部路网的节点以及交叉点全部计算出来了, 交叉点就是新产生的节点, 拓扑构建算法如下:

(a) 根据新加入的节点, 重新计算出所有路段数据结构中的首尾节点编号以及路段长度, 组成新的路网节点集和路网道路集合。

(b) 从第一个节点开始, 遍历所有路段寻找出所有与第一个节点相连接的路段, 并记下路段长度权值, 这样就组成了以第一个节点为中心的星型拓扑结构, 同理, 依次找出第二个节点, 第三个节点……直至遍历完所有的节点, 组成新的网路拓扑结构, 拓扑结构文本表示方式如图7所示, 图中表示的是节点4和节点5的拓扑结构, 其中节点4可以到达4个其它的节点, 分别为节点3, 5, 37, 21, 距离分别为189M, 589M, 554M, 254M。

```

4 x=120727062;y=30797431
target point counts 4
(3)=189.226700
(5)=589.625793
(37)=554.055664
(21)=254.305222
5 x=120725202;y=30802488
target point counts 3
(4)=589.625793
(20)=238.061737
(6)=1626.042969

```

图7 构建的路网拓扑结构

3 拓扑分块及检索算法

由于电子地图路网数据海量的特性, 本文针对嵌入式系统内存相对较小的特点, 对最优路径的查找提出了一种新的方法, 对构建的路网拓扑数据进行分块处理, 分块限制大小读入路网拓扑数据, 使得每次读入内存的数据只是整个路网拓扑数据的一小部分。拓扑数据分块如图8所示, 图示拓扑分为48块, 例如第19块包含有以A为节点的拓扑关系。

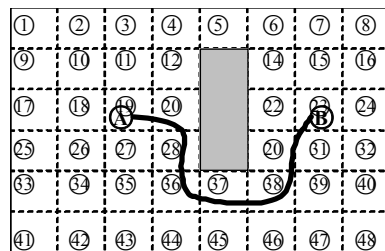


图8 拓扑数据分块图

分块拓扑数据检索算法如下:

假设: 分块 13, 分块 21, 分块 29 为障碍物分块, 即与其他分块拓扑的不连通区域, 现要求检索出从“分块 19”中的 A 点出发到“分块 22”的 B 点适合最优路径查找的分块拓扑网格。利用 A*启发式搜索的思想^[6], 对选择所有网格中经过哪个网格进行评分, 评分等式记为 $F=G+H$ 。

G 表示为把 A 作为起点, 沿着产生的路径, 移动到路网上指定网格的所产生的移动耗费。H 表示为从指定的网格移动到终点 B 的预估产生的移动耗费, 这是启发式搜索的思想。

令水平或者垂直移动网格的耗费为 10, 对角线方向耗费为 14, 因为对角线方向的距离约为水平方向或者垂直方向的 1.4 倍。H 值和 G 值的大小计算有很多种方法, 这里采用曼哈顿距离的方法。例如从“分块 1”到“分块 2”消耗记为 10。

检索最优分块拓扑数据流程如下:

- ① 把起始网格 (分块 19) 添加进开启 (OPEN) 列表。
- ② 计算开启 (OPEN) 列表中所有网格的 F 值, 寻找其中 F 值最小的网格作为当前网格, 把当前网格存放入关闭 (CLOSE) 列表中。
- ③ 对与当前网格相连的 8 个网格拓扑中的任意一个网格:
 - 如果为障碍物网格 (不可连通) 或已经在关闭 (CLOSE) 列表中, 略过。
 - 如果该网格不在开启 (OPEN) 列表中, 将其添加进开启 (OPEN) 列表中, 当前网格是这一网络的父节点。
 - 如果已经在开启 (OPEN) 列表中, 根据 G 值判断是否有更好的网格, 更低的 G 值代表着更好的网格。
- ④ 当发生如下条件时, 停止。
 - 把目标点所在的网格添加进了开启 (OPEN) 列表, 此时, 所有最优分块拓扑被找到。
 - 没有找到目标点所在的网格, 但是开启 (OPEN) 列表已经空了, 此时没有网格可以找到最优路径。
- ⑤ 从目标网格开始, 沿着每一个网络的父节点网格寻找, 直到找到起始网格, 这些所有的网格就是我们需要找的最优分块拓扑。

流程示意图如图 9 所示, 其中每个网格的左下角为 G 值, 右下角为 H 值, 左上角为 F 值。

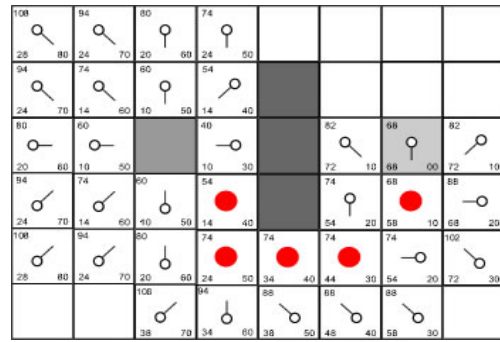


图 9 分块拓扑检索图示

4 改进的最优路径查找算法

文中采用前面所述的拓扑分块方法对最优路径的查找算法进行改进。方法为:

- (1) 载入用户设置的起点、终点、障碍点, 确定节点所在的拓扑分块网格, 使用 2.2 节所述的方法检索出最合适的网格拓扑数据块。
- (2) 在 1 中确定的所有网格拓扑块中查找最优路径, 如果没有找到最优路径, 再适当的加载进与最优网格相连的网格拓扑分块。在没有找到最优路径的情况下, 加载进来多少相连的网格拓扑块需要根据用户实际应用情况定义。

本文采用嵌入式系统测试环境对算法程序进行测试, 环境如下: 操作系统为 Linux 2.6.30, CPU 为 S3C2440 400Mhz, 内存为 64Mhz, 以杭州市地图为例节点数为 7043。

表 1 算法性能分析对比

起始节点	目标节点	占用内存(K)		花费时(ms)	
		改进算法	传统算法	改进算法	传统算法
1	100	24	810	2340	4458
1000	3000	62	810	2564	4860
46	5500	110	810	2780	5572
4200	7000	83	810	2604	4943

图 10 是该算法执行结果的一个图示。(杭州市地图, 节点数为 7043, 查询时间 2.4s)。



图 10 基于拓扑分块的路径寻优算法结果示意图

(下转第 46 页)

先部署稀疏网络结构,能够以最少的节点、最佳的部署方案监测信息以及传递信息,WSN网络的成本最低;非均匀部署的簇头结构,保证簇头节点间能量消耗的均衡性,降低人工替换簇头结点的次数,解决了网络中能量均衡问题,有效地延长了网络的生存时间;簇头结点未以最大通信半径进行数据传递,存在一定的能量冗余,有利于网络容错性机制的实施,保证网络数据传输的可靠性。仿真实验结果表明,该机制与均匀分布的部署方案相比,在降低网络节点数的同时,降低了网络的能耗,提高了网络的生存时间。

参考文献

- 1 康绍忠,蔡焕杰,冯绍元.现代农业与生态节水的技术创新与未来研究重点.农业工程学报,2004,20(1):1-6.
- 2 汪耀富,蔡寒玉,李进平,等.不同供水条件下土壤水分与烤烟蒸腾耗水的关系.农业工程学报,2007,23(1):19-23.
- 3 张利琼等.无线传感器网络在农业生产中的应用.现代农业科学,2008,15(3):49-53.
- 4 冯友兵,张荣标,沈敏,等.面向精确灌溉的无线传感器网络构建.农业机械学报,2009,40(1):56-59.
- 5 Wang XQ, Yang YT, Zhang ZL. A virtual rhomb grid-based movement-assisted sensor deployment algorithm in wireless sensor networks. Proc. of the First International Multi-Symposiums on Computer and Computational Sciences (IMSCCS2006). Hangzhou, China: IEEE Computer Society, 2006.491-495.
- 6 Wang XQ, Yang YT, Song YB. ϵ -redundant movement-assisted sensor deployment based on virtual rhomb grid in wireless sensor networks. Proc. of the 2006 IEEE International Conference on Mechatronics and Automation (ICMA 2006). Luoyang, China: IEEE 2006.775-779.
- 7 Zhang H, Hou JC. Maintaining sensing coverage and connectivity in large sensor networks. Ad Hoc & Sensor Wireless Networks, 2005,1(1-2):89-124.
- 8 Heinzelman W, Chandrakasan A, Balakrishnan H. Energy-efficient communication protocol for wireless Microsensor networks. Proc. of the 33rd Hawaii International Conference on Systems Science. Hawaii, USA, 2000.3005-3014.
- 9 Manish B, Timothy G, Anamha PC. Upper bounds on the lifetime of sensor networks. Proc. of IEEE International Conference on Communication. Hangzhou, China: IEEE Computer Society, vol.3, Helsinki, Finland, June 2001. 785-790.

(上接第89页)

5 结语

本文针对 MapInfo 电子地图的特点,研究了一种通用的电子地图分块拓扑结构求取方法,文中详细阐述了使用拓扑分块方法求取最优路径算法,实际测试表明该算法在路径寻优中可以只读取一部分拓扑数据,大大的节省内存占用开销,对于研究嵌入式系统海量电子地图的路径寻优具有一定的参考价值。

参考文献

- 1 王小国,王庆,万德钧.基于物理统一存储大规模数字导航地图道路网拓扑自动生成算法.中国惯性技术,2009,17.
- 2 王宏勇,卢战伟.嵌入式 GIS 最短路径分析中 Dijkstra 法改进.测绘学院学报,2005,22.
- 3 付梦印,李杰,邓志红.限制搜索区域的距离最短路径规划算法.北京理工大学学报,2004,24.
- 4 陆锋,卢冬梅,崔伟宏.道路网络限制搜索区域时间最短路径算法.中国图象图形学报,1999,4.
- 5 陈和平,黎小琴,顾进广,陈彬,沈磊.基于矢量地图数据的路径规划算法研究.计算机工程与应用,2010,46.
- 6 刘浩,鲍远律.A*算法在矢量地图最优路径搜索中的应用.计算机仿真,2008,4.