

城市 GIS 地名定位工具的设计及应用

曾文¹, 鄢军霞²

1. 中国地质大学信息工程学院, 湖北武汉 430074

2. 中国地质大学研究生院, 湖北武汉 430074

摘要: 为了满足城市 GIS 中高效查询、规划设计、应急处理及社会化服务的应用需求, 采用地名数据库、管理及搜索引擎、维护及搜索工具三层架构, 设计和实现了一个通用的城市 GIS 地名定位工具。该定位工具兼顾点状要素和区域范围, 提供建库、维护、搜索、移图等功能, 支持快速模糊匹配。结果表明: 城市地名数据库便于建立和更新, 不仅使 UGIS 系统能快速将关注区域移动到指定地点, 还是一种要素类的表达方式, 可以为高级的查询和分析提供数据基础。

关键词: 城市地理信息系统; 地名定位; 地名数据库; 模糊查询。

中图分类号: P208; TP311

文章编号: 1000-2383(2006)05-0725-04

收稿日期: 2006-05-10

Design and Application of an Urban GIS Placename Location Tool

ZENG Wen¹, YAN Jun-xia²

1. Faculty of Information Engineering, China University of Geosciences, Wuhan 430074, China

2. Graduate School, China University of Geosciences, Wuhan 430074, China

Abstract: In order to meet the requirements of urban GIS applications for information query, planning, design, emergency response and society-oriented service, the authors designed and implemented a general-purpose urban GIS placename location tool. Adopting three-layer architecture, this tool consists of a placename database, a management & search engine and an interface toolbox for maintenance and query. With placenames standing for point features or rectangular regions, this placename location tool includes functions such as database setup & maintenance, placename search, quick fuzzy matching and map redirection. A placename database, which is convenient to establish and update, can not only quickly lead a UGIS user to places of interest, but also serves as a presentation of feature classes, and hence provides a data foundation for advanced query or analysis.

Key words: urban geographic information system (UGIS); placename location; placename database; fuzzy query.

0 引言

各类城市 GIS 应用经常需要根据地名快速定位到特定地点, 显示该地点的地形和专题数据。一些重要的 GIS 基础软件提供了地址匹配 (address matching; Dueker, 1974; Drummond, 1995) 工具来支持这种需求, 但是这种通用工具在中国的使用受到以下因素的限制: (1) 地址数据库的建立和维护不便, 工作量巨大; (2) 当前中文地址表达缺乏统一标

准, 使得地址数据库的内容难以规范; (3) 中文词法和语法分析还存在一些尚未解决的问题, 中文地址的识别和匹配的效率和准确性都有待提高。另一种可以用于定位的工具是城市基础地形图。城市基础地形图一般对重要地点有文字标注, 但是这种标注常常止于图形显示, 而不包含实质性的地名属性信息; 另外不同应用系统对于涉及本专题的重要地点有不同的理解, 有些与专题密切相关的重要地名, 基础地形图还无法包容 (张军和周玉红, 2004)。本文设计和实现了一种易于使用和维护的城市地名定位机

制. 该机制基于独立的地名数据库, 通过模糊匹配技术, 建立高效便捷的定位工具, 满足城市 GIS 应急处理、辅助设计、分析规划, 以及社会化服务的需要.

1 总体设计

城市 GIS 地名定位工具采用数据层、应用层和表示层 3 层架构(毕硕本等, 2003)实现(图 1).

(1)数据层: 实现地名数据库的创建和维护. 地名数据库核心表结构如表 1.

表 1 中地名编号是地名内容字段存放一个地名描述字符串, 它就是定位时需要进行匹配查询的字段; 地名类别字段用于对地名进行分类, 如将地名划分为金融机构、学校、政府部门、居民区、重要设施等, 这样的分类使用户可以局限在某个类别中查询地名, 也可以不区分类别进行全局查询; X 坐标和 Y 坐标字段确定地名中心点的位置; 范围高度和范围宽度 2 个字段确定地名影响范围, 如果它们的内容

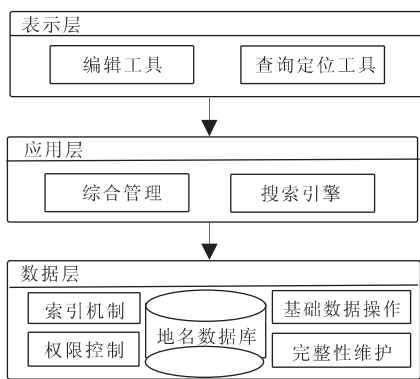


图 1 地名定位工具的三层架构

Fig. 1 Three-layer architecture of the placename location tool

表 1 地名数据库核心表结构

Table 1 Structure of the core table in placename database

字段名称	类型	长度	小数位数
地名编号	长整型	10	
地名内容	字符串型	512	
地名类别	短整型	5	
X 坐标	双精度型	15	6
Y 坐标	双精度型	15	6
范围宽度	双精度型	15	6
范围高度	双精度型	15	6
图幅索引号	长整型	10	
要素类号	长整型	10	
实体号	长整型	10	

都为 0, 则表示是点特征地名, 否则是矩形范围地名(包括影响范围是水平线段或垂直线段的情况); 图幅索引号、要素类号和实体号 3 个字段建立地名与基础地形图之间的对应关系, 为基础地形和地名数据之间的双向动态更新提供基础. 数据层同时支持文件型数据库、SQL Server 和 Oracle 数据库. 针对不同的数据库, 在索引建立、查询优化及权限控制等方面采取了不同的策略.

(2)应用层: 是一系列组件构成的管理及搜索引擎, 提供地名数据库初始化、批量地名引入、地名数据库维护、地名搜索等功能. 组件技术的应用使得 C/S(Client/Server)和 B/S(Browse/Server)架构都可以进行调用. 系统的核心组件主要有以下 4 个: (a)地名库表建立组件: 建立地名库表, 完成初始化; (b)批量地名添加引入组件: 根据实际业务的需要, 提供 3 种方式批量添加地名, 基础地形图中的地名注记加入地名库表; 从点要素数据层提取空间位置和地名描述到地名库表; 从外部数据库表成批读取地名; (c)地名搜索组件: 主要实现地名的精确查询以及模糊查询功能; (d)地名库维护组件: 响应表示层的调用, 对地名数据库条目进行添加、删除或修改.

(3)表示层: 以 C/S 或 B/S 模式, 提供编辑录入工具, 以及地名界面表现(文字或旗标、边界线)、多地名选择、根据地名移图等功能. 系统使用者可以通过鼠标逐个手工录入. 这时地名又被称为“定位标签”. 用户可以随时记录重要地点, 以便于以后的查询.

2 关键技术及其实现

2.1 兼顾点状要素和区域范围的地名数据库

一个地名根据它所对应的空间位置, 可分为由单独的 (x, y) 坐标定位的点特征地名, 以及由一个中心点坐标和高度、宽度所界定的矩形范围地名. 这种设计使得地名既可以对应于空间中的一个点(如地名点), 也可以对应于一个矩形范围(如企业、小区等)(图 2).

2.2 地名数据库批量引入建库技术

城市地名数据库往往比较庞大, 建库效率是必须考虑的重要问题. 本文设计和实现的系统支持基础地形图信息提取、点要素文件转换以及外部数据表引入 3 种批量建库手段, 使得地名建库的工作量大大减少.

依照国家规范制作的基础地形图, 一般包含以

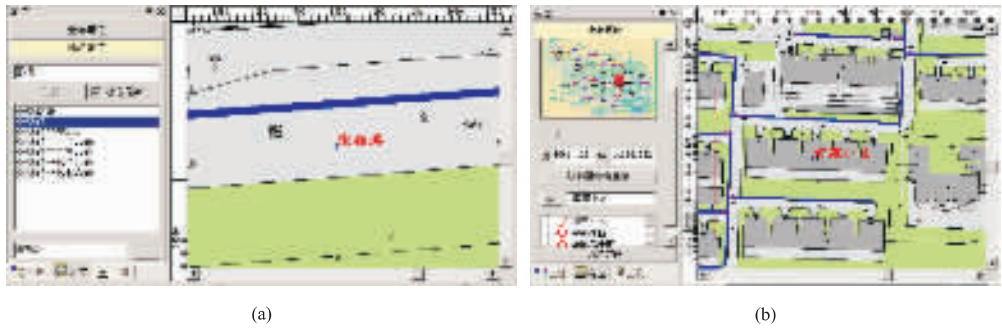


图 2 点特征地名和矩形范围地名

Fig. 2 Point featured and rectangle bounded placenames

文字方式注记地名的数据层,可以利用这类信息建立地名数据库.但是通常地形图数据侧重于地名的图面表现,不完全符合地名数据库的需要,常见的情况有 2 种:一种情况是一些文字注记并非地名;另一种情况是一些地名被人为分割成多个离散的文字注记.因此,系统在引入地形图数据作为地名时,要对数据进行必要的检查、筛选及合并.系统还提供交互式工具便于用户进行事后调整.

点要素文件批量引入功能使得道路交叉口、大口径阀门、事故多发地等信息都可以便利地转化为地名;而用户日常管理的设备卡片、大用户记录等数据,则可以作为外部数据表转入地名定位系统.

2.3 模糊搜索技术

用户在实际使用城市地理信息系统时,对地名的描述往往是不准确、不完备的,所以需要提供宽容性来处理不准确的地名描述(张森和韦明,2000).系统运用模糊搜索技术来解决这一问题,通过模糊搜索技术,操作者不必输入精确的地名,而是依据印象输入其中的关键字词,系统列出所有与这些字词匹配的地名供选择.系统采用的模糊搜索技术有 2 种模式:一种是计算字符串匹配程度;另一种是利用数据库 SQL 查询语言中的模糊匹配功能.前者允许用户设定某一匹配度,系统调用独立的字符串匹配度计算函数来筛选地名;后者将用户的输入转化为带通配符的 like 子句,用 SQL 语句直接查询地名数据库表,得到地名列表(沈兆阳和李劲,2001).用户在键入字词的同时,系统实时搜索地名数据库,随时列出匹配的地名.这样,用户在很多情况下只需要键入 2 至 4 个字就能方便地指定想查找的地名(图 3).

字符串匹配度函数判断 2 个字符串的匹配程度,返回一个 $[0, 0, 1, 0]$ 区间内的数值.设 s 是用户提交的模糊地名字符串,字符串 d 是数据库中某一

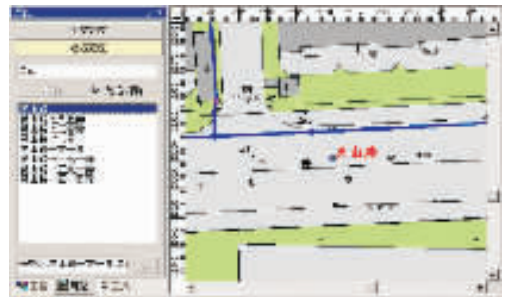


图 3 地名库的模糊查询

Fig. 3 Fuzzy query to the placename database

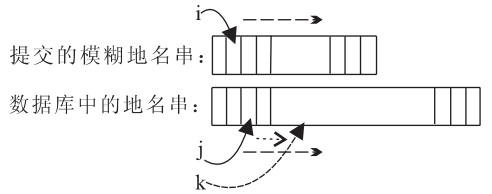


图 4 字符串匹配过程

Fig. 4 Process of string matching

个地名条目.设定 2 个指针 i, j , 分别在 s 和 d 中从头至尾移动,不断比较,寻求匹配,使用变量 $count$ 记录当前获得的匹配字符个数.同时设定一个辅助指针 k ,对于每一个 $s[i]$, k 从 j 出发,在 d 中往后移动并比较.如果发现了与 $s[i]$ 相同的字符,则 $count$ 累加,并把 j 移到 k 后面; i 也后移一个位置,继续比较;如果 k 移到 d 尾部仍不能找到与 $s[i]$ 相同的匹配,则放弃匹配 $s[i]$,将 i 后移, k 回到 j 处,再移动继续查找.这一过程反复执行直至 i 移到 s 的尾部.函数返回的字符串匹配度是最终的 $count$ 值与 s 的长度之比.在匹配过程中针对中文等双字节字符需要做特殊处理.遇到这类字符,要将两个字节一起比较,指针每次的移动也是以双字节为单位(图 4).

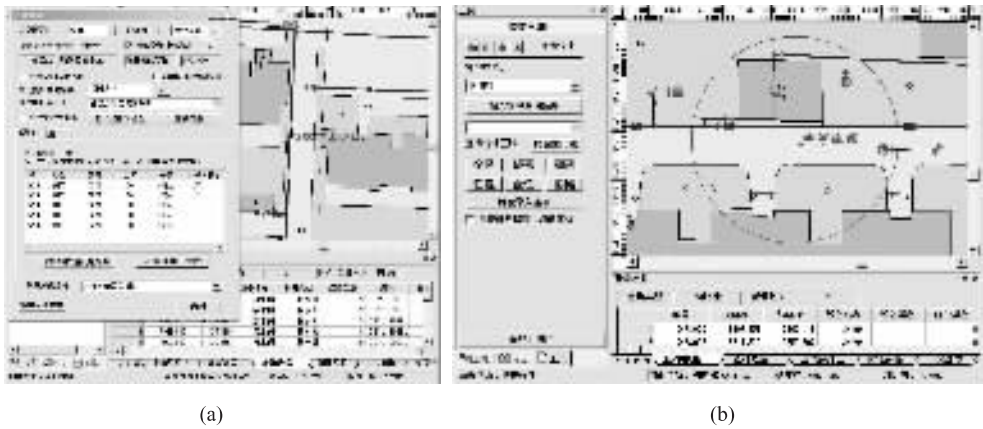


图 5 地名定位在管网事故处理(a)和空间分析(b)中的应用

Fig. 5 Application of placename location in handling pipeline accidents (a) and spatial analysis (b)

3 地名定位工具的应用

城市 GIS 地名定位工具被广泛应用于各类城市 GIS 应用系统,在快速查询、应急处理、设计分析中都有很好的应用。在市政设施信息系统中,当发生爆管等事故时,市民在事故报警电话中往往只能描述大致地点,事故处理中心需要利用地名数据模糊查询和定位工具,迅速准确地将 GIS 图形切换到事故现场,以便及时生成事故处理方案(图 5a)。

在 GIS 空间分析中,地名定位工具也得到了很好的应用。如果把地名看作一种空间要素或实体,对它实施各类空间分析运算,则地名数据库将发挥很大作用。例如,要查找某一地点周边一定区域的消防栓,就可以将特定地名作为点要素,并基于它产生缓冲区,接着利用此缓冲区对消防栓空间数据进行叠置分析(图 5b)。

4 结论

城市 GIS 地名定位工具作为 UGIS 系统的一个部件,使系统能快速将关注区域移动到指定地点,数据建库及维护的便利性以及快速模糊查询功能是它得到成功应用的关键。同时,城市地名数据库还可以作为一种要素类的表达方式。这样的理解使得地名定位的应用领域更为广阔。下一步的研究和开发工作重点在于:(1)扩展地名定位工具,以表达线状地物和不规则区域;(2)建立地名与基础地形、专题数据之间的双向动态联系机制,在基础地形和专题数据发生变化时动态变更地名数据库,将用户对地名

库的修改动态反馈到基础地形和专题数据之中。

References

- Bi, S. B., Wang, Q., Xu, X. H., 2003. The principle and method of GIS software engineer. Science Press, Beijing, 116—181(in Chinese).
- Drummond, W. J., 1995. Address matching: GIS technology for mapping human activity patterns. *Journal of the American Planning Association*, 6(3): 240—251.
- Dueker, K. J., 1974. Urban geocoding. *Annals of the Association of American Geographers*, 64(2): 318—325.
- Shen, Z. Y., Li, J., 2001. Conformity and application of SQL Server 2000 and XML. Tsinghua University Press, Beijing, 33—42(in Chinese).
- Zhang, J., Zhou, Y. H., 2004. Database technique on urban planning. Wuhan University Press, Wuhan, 104—105(in Chinese).
- Zhang, S., Wei, M., 2000. Fuzzy search technique on databases. *Electronics and Automation*, (5): 23—24(in Chinese with English Abstract).

附中文参考文献

- 毕硕本,王桥,徐秀华,2003. 地理信息系统软件工程的原理与方法. 北京:科学出版社, 116—181.
- 沈兆阳,李劲,2001. SQL Server 2000 与 XML 整合应用. 北京:清华大学出版社, 33—42.
- 张军,周玉红,2004. 城市规划数据库技术. 武汉:武汉大学出版社, 104—105.
- 张森,韦明,2000. 数据库的模糊查询技术. 电子与自动化, (5): 23—24.