

网络地理信息系统

与空间元数据

李宏伟 李中原
黄建都 陈书涛 编著



黄河水利出版社

网络地理信息系统与空间元数据

李宏伟 李中原 编著
黄建都 陈书涛

黄河水利出版社

内 容 提 要

本书介绍了网络地理信息系统和空间元数据的基本概念、基本原理、基本方法, 共分为 11 章, 主要内容包括地理信息系统和互联网、服务器端 GIS 操作、客户端 GIS 操作、可扩展标识语言、信息网络、分布式对象和 OpenGIS、Web 上的元数据和元数据标准、数据仓库和对地理信息系统新技术的展望等。本书以基本概念、基本理论为核心, 以网络 GIS 建设必须解决的问题为主线, 力图给读者一个清晰的网络地理信息系统的总体构架。另外, 书中给出了大量的实例和底层开发应用程序, 可作为程序员开发在线地理信息系统时的参考素材。

本书面向地理信息系统开发人员和广大用户, 也可作为大专院校地理信息系统专业及相关专业高年级学生、研究生参考用书。

图书在版编目(CIP)数据

网络地理信息系统与空间元数据/李宏伟等编著.
郑州: 黄河水利出版社, 2004.4
ISBN 7 - 80621 - 765 - 7

I.网… II.李… III.计算机网络-应用-地
理信息系统 IV.P208-39

中国版本图书馆 CIP 数据核字(2004)第 016891 号

出 版 社: 黄河水利出版社

地址: 河南省郑州市金水路 11 号 邮政编码: 450003

发行单位: 黄河水利出版社

发行部电话及传真: 0371 - 6022620

E-mail: yrcp@public.zz.ha.cn

承印单位: 黄委会设计院印刷厂印刷

开本: 787 mm × 1 092 mm 1/16

印张: 10.5

字数: 243 千字

印数: 1—2 000

版次: 2004 年 4 月第 1 版

印次: 2004 年 4 月第 1 次印刷

书号: ISBN 7 - 80621 - 765 - 7 / P · 31

定价: 25.00 元

前 言

自 20 世纪 60 年代加拿大着手开始建立加拿大地理信息系统(CGIS)至今, GIS 已经走过了 40 年的发展历程。历经 40 年的发展, GIS 取得了十分辉煌的成就, 已经广泛应用于政治、经济、军事、社会、环境、人类生活等各个方面。可以说 GIS 无处不在。

GIS 的发展与计算机技术的进步密切相关。今天, 随着计算机技术的发展, 人类已进入以 Internet 和 WWW(World Wide Web)为代表的网络时代, GIS 通过 WWW 的扩展, 已经真正成为一种大众使用的工具并进入了千家万户。从 WWW 的任何一个节点, Internet 用户都可以浏览 WWW 站点中的空间数据, 进行各种空间查询和分析, 制作专题地图。因此, 在网络环境下, 地理信息资源的分布和共享问题成为 GIS 关注的焦点。

本书围绕网络地理信息系统和空间元数据问题展开论述, 主要内容包括地理信息系统和互联网、服务器端 GIS 操作、客户端 GIS 操作、可扩展标识语言(XML)、信息网络、分布式对象和 OpenGIS、Web 上的元数据和元数据标准、数据仓库和地理信息系统新技术展望等。该书以基本概念、基本理论为核心, 以网络 GIS 建设必须解决的问题为主线, 力图为读者构建一个比较清晰的网络地理信息系统的总体架构。书中给出了大量的实例和底层开发应用程序, 可作为程序员开发在线地理信息系统时的参考素材; 本书注重基本概念的分析, 能够帮助初涉 GIS 领域的读者奠定良好的学习基础。“纸上得来终觉浅, 绝知此事要躬行”, 希望本书的出版能在 GIS 领域激起点点浪花, 吸引更多的 GIS 开发者和用户来关注网络 GIS 的应用和发展。

GIS 研究和开发领域还面临着许多亟待解决的问题。从技术层面上分析, GIS 要解决“4W”问题, 即 Where、When、What Object、What Time; 从应用层面上看, GIS 要做到“4A”服务, 即 Anybody、Anything、Anywhere、Anytime。由此可见, 地理信息系统建设任重而道远, 构建新一代 GIS 是摆在广大地理信息系统工作者面前的十分艰巨的任务。

黄河水利出版社马广州编辑以及硕士研究生赵姗、常小慧同学, 为本书的出版提供了许多宝贵的资料, 做了大量的工作。在此, 作者对他们表示衷心的感谢。

本书不足之处, 欢迎批评指正。

(E-mail: laob-811@sina.com)

作 者

2003 年 8 月

目 录

前 言

第 1 章 概 述	(1)
1.1 GIS 的特征	(1)
1.2 Internet 的兴起	(6)
1.3 分布式信息系统的优点	(6)
1.4 分布式信息实例	(7)
第 2 章 地理信息系统和互联网	(9)
2.1 在线 GIS 的优势	(9)
2.2 新媒体引发的争论	(9)
2.3 在线 GIS 举例	(10)
2.4 空间数据采集工程	(15)
2.5 单机 GIS 和在线 GIS 之间的差别	(16)
2.6 在线 GIS 实施	(17)
第 3 章 服务器端 GIS 操作	(19)
3.1 Web 服务器	(19)
3.2 服务器处理	(21)
3.3 在线地图制图	(29)
3.4 高级脚本语言的使用	(30)
3.5 实现地理查询	(33)
第 4 章 客户端 GIS 操作	(36)
4.1 影像地图	(36)
4.2 客户端操作中 JavaScript 的应用	(38)
4.3 Java Applets 的应用	(46)
4.4 客户端 GIS 操作实例	(52)
第 5 章 标识语言	(54)
5.1 XML: 结构思想	(56)
5.2 文档类型定义	(58)
5.3 XML 命名空间	(59)
5.4 XML 模式	(61)
5.5 XQL: XML 查询语句	(65)
5.6 获得 DTDs 和其他规范	(69)
5.7 未来发展	(69)
5.8 进一步阅读的资料	(69)

第 6 章 信息网络	(71)
6.1 信息网络的概念	(71)
6.2 信息网络的功能	(72)
6.3 信息网络的组织	(72)
6.4 与信息网络相关的问题	(73)
6.5 实用信息网络	(80)
第 7 章 分布式对象和 OpenGIS	(84)
7.1 标准化组织	(84)
7.2 在线对象及其元数据: CORBA	(85)
7.3 地理标识语言	(96)
7.4 OpenGIS	(104)
第 8 章 Web 上的元数据	(107)
8.1 Dublin core	(108)
8.2 PICS: 因特网内容选择平台	(111)
8.3 资源描述框架(RDF)	(114)
第 9 章 元数据标准	(120)
9.1 美国的元数据研究	(120)
9.2 澳大利亚标准	(124)
9.3 欧洲的情况	(134)
9.4 元数据的作用	(135)
9.5 今后的主要研究工作	(136)
第 10 章 数据仓库	(137)
10.1 什么是数据仓库	(137)
10.2 地理数据仓库	(139)
10.3 数据仓库的结构	(139)
10.4 建数据仓库的问题	(142)
10.5 组织和实施	(143)
10.6 数据挖掘	(145)
10.7 数据仓库实例	(146)
10.8 地理数据仓库的前景	(151)
第 11 章 空间信息新技术	(152)
11.1 全球 GIS 综览	(152)
11.2 智能系统	(155)
11.3 移动计算	(158)
11.4 从网络 GIS 到虚拟世界	(159)
参考文献	(161)

第1章 概述

GIS 是获取、存储、理解、显示有组织的空间信息的计算机系统。GIS 的产生和发展源于许多不同应用目的的学科,包括地图学、地理学、地质测量、环境管理、城市规划等。目前, GIS 已成为所有这些专业领域的必不可少的工具。

地理信息系统建立后,其功能和应用范围正在发生重大的变化。地理信息已不再孤立于单机上,一个新的在互联网上存储与访问地理信息的环境正在形成。我们将讨论这种变化,它所带来的冲击是非常巨大的,如因特网提供了一种强大的信息访问能力。不过,更重要的影响是因特网为不同来源的数据综合集成提供了可能性,而这在以前是难以想像的。为了能在这种新的环境中工作,管理人员、开发商与用户都必须学习一些新的基本技术。

许多在线制图的商业开发正在积极推进。我们的目的不是描述或罗列其中的细节,而是考察其发展特点,在考虑把地理信息发布到网上之前,对信息的类型及其具有的功能有一个清晰的轮廓。首先,简要回顾一下 GIS 中的一些概念。

1.1 GIS 的特征

1.1.1 GIS 数据

在 GIS 中,地理数据由图层组成。一个图层是具有共同主题或类型的地理索引数据集,例如海岸线、道路、地形、城镇、公共用地。为了制作一张地图, GIS 用户需要选择基本地图,通常是关键图层集合(如海岸线、道路),再叠加选择的图层。通常,图层分为以下三种类型:

(1)矢量数据。矢量图层由点、线、多边形对象组成。这类数据通常存储在数据库的数据表中,表中的每个记录都包含着对象在空间的属性,包括位置。

(2)栅格数据。由规则的像元阵列来表示空间对象分布,阵列中的每个数据表示对象的属性特征。栅格数据如卫星影像、数字航空摄影相片等。影像记录的光谱数据是每个像元所对应的地表区域内所有地物光谱辐射的总和。栅格数据记录的是数据本身,而位置数据可以由属性数据对应的行列号转换为相应的坐标。

(3)数字模型。计算任何地方地表属性的值,例如根据数字高程模型(DEM)可进行插值运算。实际应用中,数字模型通常被转换成矢量数据或栅格数据进行显示和使用。

正如上面提到的那样,矢量数据通常存储在数据库表中,这就可以根据属性进行大范围的搜索或建立索引。这些数据也可以通过建立空间索引来搜索和查找。例如:四叉树就是这样一种结构——它按照任意的尺度把一个区域分成四个象限,任何数据项可根据它所在的象限进行索引,这个过程适应于本地快速检索。

叠置分析是 GIS 中经典的分析工具。某些在线服务也允许数据层在不同的背景上进行叠置(见图 1-1)。

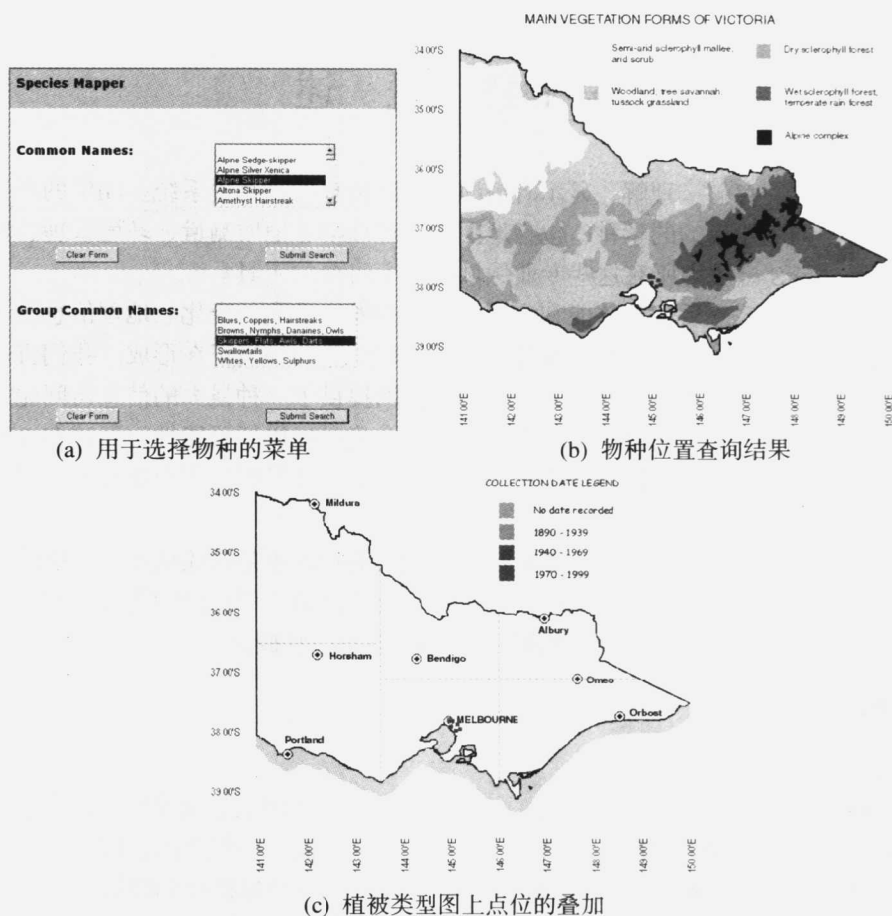


图 1-1 在线空间查询和地图系统示例

1.1.2 GIS 功能

尽管通常多数情况下 GIS 是与地图生产密切相联系的,但是 GIS 包含着更丰富的形式和功能。最普通的(但绝不是惟一的)是 GIS 输出地图的功能。这些地图既可以是给定地区的单图层图,也可以是查询结果的显示。GIS 中的查询分为两个类别。

第一类问题包括某单一图层的特征。例如:

- (1)对象 X 在哪里? 如:北京在哪里?
- (2)对象 X 多大或多长? 如:长江有多长? 中国的面积多大?
- (3)具有某一特定属性的所有对象在哪里? 如:大庆油田在哪里?

第二类问题包括了两个或多个图层的叠加。例如:

- (1)在指定类型的区域内找到什么对象? 如:在某个省或地区里有什么矿藏?
- (2)两种对象之间重叠的区域是什么? 如:国家公园里某种林型的分布面积? 一条规划的高速公路能穿过哪些农场?
- (3)什么环境因素会影响稀有物种的分布?

叠置过程通常包含新数据层的建立。设想由多边形组成的两个图层进行叠置(如选举区和灌溉区),那么叠置结果应是两个图层中多边形交叉构成的新多边形。

GIS 的两种重要功能是数据分析和模型分析。数据分析工具包括一组比较、分析数据层的技术,如变量的空间相关性(如:医疗条件与地域污染水平的关联)、最邻近分析(如:鸟巢紧靠在一起的海鸟物种的数量)、形状的分形维(如:森林地类界)。

GIS 中的数据模型通常就是图层,通过其他图层的操作运算得到。例如:在地形图上,将水流和降雨以及土壤类型等输入到一个模型中,可以输出一幅土壤湿度图。在传统方法中,这些图层是通过综合软件包处理的。同时,很显然,这些数据集可能来自不同的地区,在线系统允许每一个组件从不同的结点进行动态访问。

1.1.3 GIS 对象模型

GIS 的影响之一是引入了一种全新的探究空间数据的方法。尽管地图学领域在 GIS 出现之前就一直致力于数字化多年,但重点仍只是在地图生产上。例如,一张道路图,是由绘制到纸张上的线划构成的,道路和城镇的名称注记则构成一个文本数据集。在地图上表达一个城镇的点位和该城镇的名称之间没有直接的关系,只有地图被印刷出版时这些关系才变得十分明显。

当科学家将空间数据用于研究目的时,面向地图生产方法的缺点就突现出来了,用以下几个例子来说明:

(1)某地区的河流数据存储为若干分离的线段,每条河流数据由许多独立的线段构成,没有迹象表明这些线段有着惟一的特征;河流被分成若干段是因为地图表达的需要,而不管其他的特征要素被绘制在地图上的什么地方。

(2)在一个城镇的地名词典里,每一个地名注记位置都会有一些小小的误差。这是因为每个地名注记只是放在地图上相对于城镇比较美观、方便的位置,而不一定是精确的实际位置,所以出现位置偏移是难免的。

GIS 带来的最大变化是把地理特征作为明确的对象存储,所以一条道路就被存储成有特定属性的对象。这些属性包括它的名称、类型、质量特征和一系列表达道路所穿过路径的坐标。例如,一张地图包含下列对象(见图 1-2、表 1-1):一条道路连接两个城市,城市有两个属性(名称和规模),道路也有两个属性(名称和等级)。

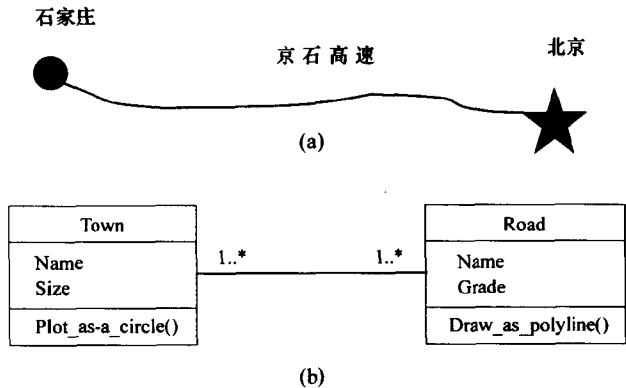


图 1-2 一个简单的地理对象

注: (1)(a)所表示的对象绘制在地图上; (b)表示应用 UML 绘制的两个对象类的简单模型, 图框表示对象类, 与每个对象相关, 模型定义了两个特征属性和绘制方法。 (2)连接两个类的线表示两者之间的关系。 (3)1..*表示每条道路是与一条或更多个城镇相连。

表 1-1 图 1-2 所表达的对象数据

对 象	名 字	类 型	属 性
1	北 京	城 市	首 都
2	石 家 庄	城 市	省会城市
3	京石线	道 路	高速公路

在这种面向对象方法中，每一个地理对象属于某个对象类。在图 1-2 中，对象北京和石家庄都属于城市类，每个都是城市类的一个实例(见图 1-3)。类的概念在 GIS 中是十分重要的，这暗示着一种处理地理数据的方法。

每个对象都有与之相关的属性值。对城市类，至少包括名称和位置(经纬度)。当然，还可以包含许多其他描述城市的属性值，如人口、所在省份名、地址以及电话区号等。定义一个类的重要意义在于对添加到 GIS 中的其他任何城市对象类，我们拥有所需提供的精确数据列表。

GIS 中的面向对象方法有很多优点。首先它对每个地理实体进行了完全封装，这不仅包括数据，而且扩展来处理数据的方法。例如，城市对象可能包括将城市绘制在地图上的方法，在这种情况下，可以将一个城市画成一个圆点，圆点的大小取决于城市人口规模。面向对象方法的另一个优点是我们可以定义不同对象类之间的关系。比如我们前面描述的城市类其实属于更一般意义上的对象类，我们称其为居民区。这个具有普遍意义的对象类不仅包括城市，也有城镇、军事基地、研究机构、厂矿、农田和其他人类建筑。

城市类继承了居住地的一些属性，最基本的特性包括位置以及当前人口等。城市类的其他特性，如它与城市政府机关联系的紧密程度，却是独有的，不能为全部居住地分享。居住区实际上构成了一个层级系统(见图 1-3)，中心城市继承了来自城市的特性，城市继承了来自城镇的属性，城镇继承了来自居民区的属性(见表 1-2)。因而，一个中心城市至少有六个属性。

对象模型通常用对象类术语和对象之间的关系来描述(见图 1-3)。对特定的数据对象模型和处理，UML(一致性模型语言)提供了一种标准方法(Larman, 1998)。

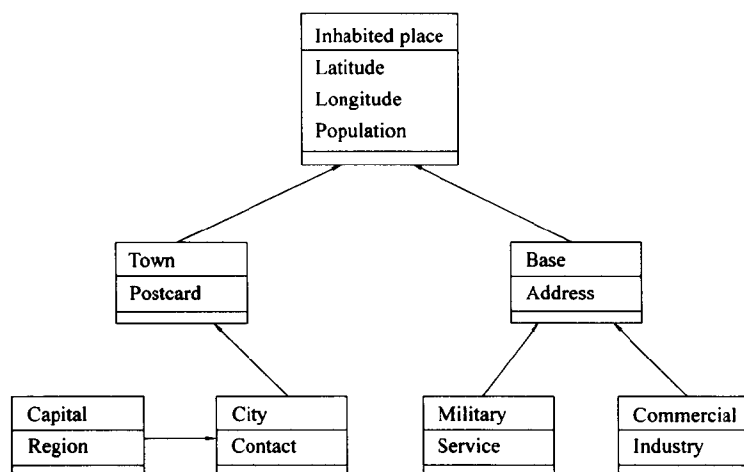


图 1-3 用 UML 勾画的一个简单数据模型

图 1-3 用 UML 勾画的一个简单数据模型，表明了几个地理对象类之间的关系。每个对象的属性得以表达，但方法却被忽略。箭头表明了对象类之间的总体对个别的关系。例如一个城市是一种特定的城镇类，并继承了所有城镇的属性。同样地，一个中心城市也继承了这些属性，如邮政编码、相互联系、区域等。

表 1-2 居住地类体系

类	属 性	方 法
居住地	位置、人口	画一个小圆
城镇	邮政编码	画一个圆
城市	与政府机关的联系	画一个阴影圆
中心城市	区域服务	画一个大阴影圆

对象类关系之间联系的另一种重要类型是整体和部分的关系。例如，一幅地图由几个要素构成，有边界、比例尺和可能的几个层，每个图层包含许多要素。图 1-4 是一个简单的地图类实例，表明了整幅地图作为部分的组合。因此，当绘制一幅地图时，我们需要依次绘出每个要素。

制图活动的一个引人入胜的领域是面向对象数据库。传统的关系数据库以表和表之间关系存储数据。随着多媒体技术和万维网技术在过去十几年的发展，不仅限于数据存储，而且还要存储诸如视频、影像、文本等多媒体信息，反映动态变化，改进查询机制。

与此同时，一种将数据作为具有相关对象查询语言的完整对象的思想出现了。在关系模型中，一个复杂对象可能被分解为许多表；而在对象模型中，对象(数据和方法)作为一个实体被存储。一个典型例子是提取某地与许多其他事物的距离地图和相关方法(见图 1-4)。

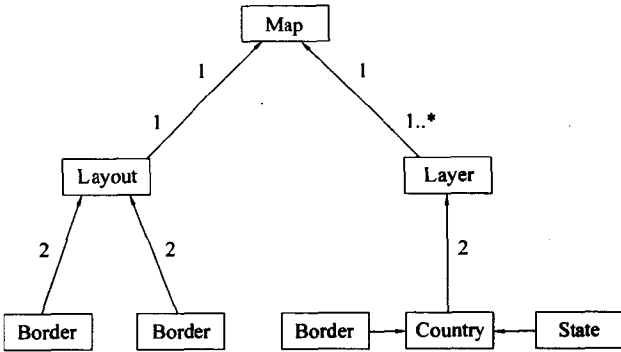


图 1-4 一个地图类是一个由几部分构成的组合

面向对象数据库模型对 GIS 数据极具吸引力。可是，商业化的现实却是不同的。企业数据库存储大量的信息，需要异常的稳健性，在转换和优化方面需要嵌入非常多的经验。面向对象数据模型拓展了一条很漫长而前景广阔的路，尽管目前它所占的市场份额还相对较小。

1.2 Internet 的兴起

Internet 是一个巨大的通讯网络,它把全世界大约两亿台计算机连接在一起。Internet 在过去的几年里增长迅猛,引起了所谓的“信息爆炸”。从用户的视角看,该过程强调几个关键的需求:

(1)自组织。为确保用户能轻松快捷地获得信息,Internet 搜索引擎技术得以快速发展,很好地满足了网络信息服务。可是由于垃圾信息的膨胀,如何获得有价值信息的问题就十分突出,基于用户兴趣和优先权的自组织方法是惟一可行的解决方案。

(2)稳健性。确保资源持久实用,确保联系不会过时。不是在一个中心采集信息,一个很重要的原则是保持信息站点应该是基本的信息源。数据集的复制很快就会过时,所以确保与保持数据集的站点的持续联系是更为有效的,而不只是简单地拷贝它。

(3)质量。确保信息是有效的、正确的,数据是可更新的和精确的,软件能正确地运行。

(4)标准化。确保信息的形式和内容规范,便于实际应用。

Internet 的兴起,特别是 WWW 在 20 世纪 90 年代的发展,引起了信息发布的革命。一旦某个组织或部门在网上发布了信息,任何人在任何地方都可以获取它。这意味着人们获取相关信息比以前更简单、更快捷、更有效率。如在线填写一张表格,远程用户即可查询检索数据库。GIS 以前要求专业化,通常花费很贵,需要特定的设备,而现在只需要一个标准的 Web 浏览器就可以进行远程访问了。

自 1999 年美国国家超级计算机应用中心(NCSA)发布第一个多媒体浏览器(Mosaic)以来,WWW 上的信息量呈指数增长。信息的爆炸首先是由数据提供者推动的,他们认识到发布信息能吸引网民的注意。随着信息量的增长,用户开始按需推动这一进程。

在线信息的爆炸是个问题。在成千上万条信息中找到一条自己需要的信息简直就像大海捞针一般。有发展潜力的搜索信息的方法包括智能代理技术的应用——筛选和记录相关的数据项。推动元数据标准形成自索引文档,信息网络提供了一种组织信息源的方法(第 6 章将有详细讨论)。

Internet 由 Internet 协会(ISOC)负责管理,包括检查、监督新标准和协议的发展和执行,WWW 则由 WWW 协会(W3C)负责管理,并已经有效地发布了许多标准。

1.3 分布式信息系统的优点

或许 Internet 的最大影响是其无缝地融合不同来源信息的能力(Green, 1994)。这种能力拓展了数据共享和操作的广阔前景,也引起了对协作的更大需求。

作为一个地理信息系统,WWW 有许多重要的优势。在线地理信息系统发展最大的实际问题之一是需要采集数据的容量陡增。来自供应商采集的数据集需要较长的周期,大多数系统需要专业的、必须是开发者自己采集的数据。开发者之间不可避免地缺少交流,导致了許多重复性劳动。Internet 有潜力减少这些问题的困扰。

从原理上讲,单个数据集或数据层的提供能够将他们的产品在线发布。这些方法不仅加快了 GIS 的发展,而且也有助于减少重复劳动。事实上,数据的在线分布对在线 GIS 开发者来说提供了更丰富的信息。在线信息共享把各种数据采集工作分配到许多部

门来共同完成,简化了信息的更新。

在线 GIS 的另外一个优点是,它扩展了 GIS 开发商和用户可用的具有发展潜力的工具。正如我们在第 2 章中将要看到的,对 GIS 在线发布可以有很多选择,最终是减少了开发者的经费投入。设计特定应用目的的小型 GIS 不仅是可能的,而且是有效的,如为在线文档或数据库提供地理接口。

鉴于上述优势,对执行 GIS 在线出版商来说,通常意义上的编排工作是没有必要的。GIS 服务可以通过租用来自专业 GIS 站点的信息来实施,如某旅游部门提供有关旅行、食宿等,要提供一幅街道图来表明数据库中每个站点的位置。这对旅行者来说是非常有用的服务,借此他们可以在陌生的城市,很方便地找到他们想居住的旅馆。无须发展一系列的地图本身,旅游机构就可以安排连接到提供所需的街道图的位置。这样的安排使得 GIS 获得了许多新的商业机会,主要是通过网络提供地理信息服务。在第 11 章我们将看到一些新的商业模型的应用。

对 GIS 用户来说,GIS 的发展前景是十分令人鼓舞的。Internet 给 GIS 带来了成千上万的不懂得使用必要设备 and 专业软件的用户。除了提供免费的有价值的服务外,还有访问完全 GIS 解决方案的需求。例如,用户不需购买完整的 GIS 软件,却可以从供应商那里买到他们需要的 GIS 服务,对于长期固定用户,可以在网上付款购买。另一方面,非固定短时间用户,可以用同样的方式购买一些特定的服务,就像以前买纸质地图一样,按需购买和接受供应商的服务。

为了将上述可能性变为现实,在线 GIS 的管理经营者需要构建一种合作和共享的氛围。Internet 是一个理想的合作媒体和桥梁,它使得在多种尺度上的信息交流和共享达到了一种前所未有的程度。然而,起初 Internet 的迅速增长却导致了許多混乱。由于商业利益的驱动,人们看到了 Internet 作为传统市场竞争手段的延展,可以由此获得更大的利润空间。这在一定意义上阻碍了信息资源的共享。为了解决这个问题,有关组织和国家要求就数据记录达成协议,制定相应的标准,提供质量保证、版权所有、承担法律责任等。在后面的章节(特别是第 6 章)我们将详细地讨论这些问题。

1.4 分布式信息实例

万维网(WWW)具有连接诸多不同来源信息的能力,并形成一种协作效应,使整体信息资源大于部分之和。最早证明 Internet 这种能力的是一家名为虚拟旅行者的服务机构,由 Brandon Plewe 于 1997 年所创。通过构建一个中心地理索引访问不同国家的旅游信息,重要的是,这些信息是跨越成千上万个站点分布的,而且地理索引并不指向某个具体国家,只是指向有关特定主题的详细信息源。这样系统允许用户点击进入任何一个国家,通过简单的鼠标操作即可获得有关旅游、天气以及其他有用信息的详细报告。

很多例子表明,众多领域都强调分布式在线信息系统的优势,很多领域一开始就强烈依赖于 Internet,如生物工程技术。在该领域,大量在线资源是进行基因图编纂、蛋白质排序、酶特性分析、分类学记录。同时,也提供许多大型应用软件库和对数据进行处理及说明的在线服务。多数著名站点也提供了领域广泛的其他各类信息,包括文献目录、电子新闻组和教育素材,所有这些资源都可以在线访问。有些大的机构,如欧洲分子实

验室(EML)实际上就是一个跨网络的国际合作组织。研究人员必须将他们的数据提供在线仓库,作为参入科学研究和合作的条件,这是一种典型的网络环境下科技资源共享的实例。

许多环境信息也已经在线共享。环境科研人员已经做了很多努力,借助在线信息关注环境资源领域的进展,如林业和生物多样性问题。无论何种情况,都会出现新的可能性引导人们和有关机构重新思考工作的方式,从更广阔的意义上观察问题,筹划加强国际合作的计划。

我们认为,在线 GIS 代表着一种全新的工作环境。本书力图达到两个目的:第一,向读者介绍在线 GIS 的基本技术;第二,展示在线地理信息如何协同的模式。

为此,需要考虑几个不同的方面。首先,必须理解 GIS 在 Web 上的处理机制。其次,转向考虑数据如何组织、访问、搜索查找、维护、获取和处理;包括研究相当复杂的标准和规范,并对它们进行重新定义,以适应未来在线 GIS 的应用。XML 构成了目前网络标准的基础,如关键标准 RDF 就是一种适于 Web 的元数据标准。还有一个十分重要的概念是分布式对象以及它怎样适合于 OpenGIS 框架体系。考虑上述所有细节,我们现在来研究 GIS 的元数据标准的内容,探讨以怎样的方式在线开展工作。

到目前,有一个问题我们一直避而未谈,即关于已有 WebGIS 软件的详细讨论。主要原因是随着软件供应商对相关标准的发展,软件工具适应的范围和特性变化太快。

本书第 2 章~第 5 章介绍发展和实现在线 GIS 的技术方法。这些技术细节和范例提供给开发者,并帮助他们理解关键技术点,也提供给读者必要技术背景资料。第 2 章围绕在线 GIS 阐述主要技术点,给出一些方案概述。第 3 章描述在 Web 服务器上进行的各种开发和操作。第 4 章介绍在标准的 Web 浏览器上执行操作运算的 GIS 工具。第 5 章涉及关键的开发语言——XML。很多年来,SGML 国际标准变得越来越没有生命力了,除在像美国军队这种少数几个庞大的组织外,几乎不再使用。某种程度上是因为它的高度复杂性和处理工具的高昂开销。但是随着 Web 的日益增长,对一个有较好组织性的研究的需求变得迫切起来。HTML,这种 Web 语言实际上成了一种 SGML DTD,显然,一种完全的 SGML 标准简化版是迫切需要而且有优势的。因此,XML 出现了,并且改变了绝大多数其他网络标准。SGML 也成为编写空间元数据标准的一种选择,但是,现在它正在被 XML 所替代。

第 6 章~第 9 章介绍探讨在因特网分布式环境中协同地理信息发展的一些关键技术。第 6 章开始讨论信息网络的本质特性。信息网络是一个信息可以在许多不同站点协同分配的系统。第 7 章介绍了一些在分布式信息系统中可以共同互操作的标准,最著名的是 OpenGIS。同时,特别关注与分布对象相关的标准,在计算机网络中,对象的位置实质上是透明的。第 8 章介绍元数据的概念体系框架,研究 RDF 和适于 Web 的网络标准。第 9 章介绍几种全世界应用的空间元数据标准,该章建立在第 5 章的 XML、第 7 章的分布式对象模型和第 8 章的元数据原则基础之上。第 10 章着重介绍分布式信息装入数据仓库的方法,介绍在这种系统中数据挖掘的基本思想。

本书的最后部分,第 10 章和第 11 章,展望在线 GIS 的发展前景。第 10 章讨论了一些与地理信息相关的最新技术。第 11 章介绍了一个全球地理信息系统的内在可能性,以及由在线 GIS 引发的可能性。

第2章 地理信息系统和互联网

2.1 在线 GIS 的优势

万维网正以极快的速度发展成为 GIS 的标准平台。大量的地理信息服务已经出现在万维网上(Green, 1998), 在万维网上应用的大量环境信息给我们留下了深刻的印象。开发商已经研制出各自的 GIS 软件在线版本, 提供各种在线服务, 包括空间查询检索和地图制图等。

在线 GIS 系统较之单机 GIS 系统有几个潜在的优势, 包括:

- (1)全球范围的访问。一个基于 Web 的信息系统可以访问世界上任何地方。
- (2)标准化的界面。每个 Web 用户都有一个浏览器, 所以基于 Web 的任何系统可以被任何人访问, 而无须昂贵的、专门的设备。
- (3)更便捷、更低廉的维护。在万维网上, 可以直接访问信息源, 没有必要集中在一个中心进行数据分类。

在这一章中, 我们讨论如下问题:

- (1)讨论发展在线 GIS 的一些技术问题。
- (2)简要分析目前在线 GIS 的一些实例。
- (3)说明一些实现系统基本类型的简单方法。
- (4)讨论在线 GIS 的未来发展。

2.2 新媒体引发的争论

Internet, 特别是万维网, 对所有各种计算都是一种新的环境。相对于单机环境, 它提出了许多在单机环境下根本不存在的新问题和新挑战。

在下面的在线 GIS 讨论中, 我们主要解释实现标准 GIS 操作的方法, 关注 Web 环境的特性、用户界面和地理数据及其处理的分离。

2.2.1 Web 环境

Web 应用超文本转换协议(HTTP)实现了互联网的通讯。HTTP 是一种客户/服务器协议。客户端为用户浏览器程序, 它传送一个查询请求到一个 Web 服务器, 然后等待得到回应。在 HTTP1.0 版本中(目前仍为大多数服务器所应用), 这些事务处理建立在无联系的、单一的查询基础上。一个用户即使对相同的服务器进行了一系列查询, 该服务器通常并不保留过去的查询历史记录或建立客户“话路”。这跟几个其他的 Internet 协议如 FTP 和 Telnet 相反, FTP 和 Telnet 服务器建立了与客户的“login”对话。

运行在开放计算机上的交互式 GIS 软件隐含了一种假设, 即程序的当前状态是用户跟 GIS 软件交互的直接结果。

HTTP 和相关的客户/服务器软件具有相同的技术优势。由 HTTP 提供的任何服务都可以立即应用于任何运行合适客户程序的人和计算机。

WWW 浏览器和客户的重要特征包括:

- (1)允许浏览所有主要网络协议(FTP, Telnet 等)。
- (2)允许使用文本和影像、图片,是真正意义上的电子图书。
- (3)综合运用各种显示工具,能阅读影像、声音、脚本、动画等信息。
- (4)允许用户自己的本地数据和来自 Web 上的任何地点的信息的无缝集成。
- (5)表单式界面支持用户与以表的形式出现的文档交互,包括按钮、菜单、对话框等,并通过复杂的查询返回到服务器。
- (6)影像图接口允许用户交互地查询一幅地图,允许用户以与 GIS 类似风格的形式,通过点击一张世界地图而获得不同国家的信息。

(7)授权许可特性提供了各种安全保证,如限制对特定信息的访问,需要提供访问密码等。

(8)SQL 网关允许用户通过对服务器数据库的查询,实现资源共享的目的。这种网关已在许多数据库中实施(如美国和欧洲的植物志和动物志、DNA 序列等)。

(9)具有在服务器上运行脚本和程序的能力,具有传递查询结果到 WWW 的能力。

(10)动态计入文件的能力,以极快的速度建立和传递文档的能力。

这些技术提供了足以进行简单空间查询的功能。最近的两项 Web 技术创新允许浏览器和客户做更多的事:

(1)来自 SUN 公司的 Java 语言是一种功能强大的程序语言(详见第 4 章),具有高级图形特征,不久将用于完全三维空间,这种语言是独立于计算机的和安全的。

(2)SVG 和 X3D 丰富了 XML 顶层的图形语言(详见第 5 章),这将使 XML 用于矢量图形和 3D 构建,对 GIS 发展极富潜力(见第 4 章关于 SVG 的进一步讨论)。

2.2.2 用户界面的分离

Web 环境将用户界面同数据处理站点分离开来。考虑到系统的反应速度,这种分离对任意操作可能会产生一些问题,我们来看下面的例子:

(1)最普通的一个 GIS 操作是用计算机鼠标点击来定义一个多边形,标准 Web 浏览器处理单一鼠标点击事件作为一种激发传递查询到服务器,来自服务器的响应极其缓慢而不能维持相关的程序。

(2)地理查询经常是上下文关联敏感的。例如,一个用户可以希望有一个流行的地图菜单,很快地显示关于空间对象特征的数据。

2.3 在线 GIS 举例

目前大多数在线地理信息服务可以分为两类,即空间查询系统和地图建立程序。本书中的大多数例子都是限于这两种功能,只是在第 11 章有些例外。只有少数服务要求提供更高级的 GIS 功能,如数据分析和模拟。当然这些功能在在线 GIS 环境下也是可行的。正如我们在后面章节中将要看到的,Internet 环境下实现 GIS 高级应用的标准和工具正在发展之中。

2.3.1 空间查询

最常用到的在线 GIS 功能是基于空间位置的查询,包括选取空间对象,如国家或省,

或一些随机点。对象最常见的是按文本界面选择, 如一个列表或一个数据库查询表单。随机点通常是通过影像地图、表单影像或 Java Applet 选取(见第 4 章)。Web 浏览器是基于超文本概念的, 允许更复杂查找的系统还为数不多。下面是一些实例。

2.3.1.1 虚拟旅行家(VT)

虚拟旅行家(Plewe, 1997)是一个为旅行者和旅游业提供世界各地详细旅游信息的在线 GIS 服务系统。从一张世界地图开始, 用户可以点击地图进入, 直到选择一个他们喜欢的国家或地区为止。

VT 是最早的在线超媒体范例。它通过一张索引地图进行导航, 连接世界上各个提供旅游资源信息的国家, 而且由专门的机构负责对信息进行更新(见第 6 章)。

2.3.1.2 Pierce 县数据库

美国华盛顿州 Pierce 县创建了一种公共在线 GIS 服务, 称之为 MAP—Your—Way!™。该系统提供了一种对众多县级单元公共数据库的灵活的、易于访问的界面(见图 2-1)。任何用户都可以创建一幅自己习惯的、任何县的综合特征地图。该系统是由 ESRI 制图软件工具实现的。

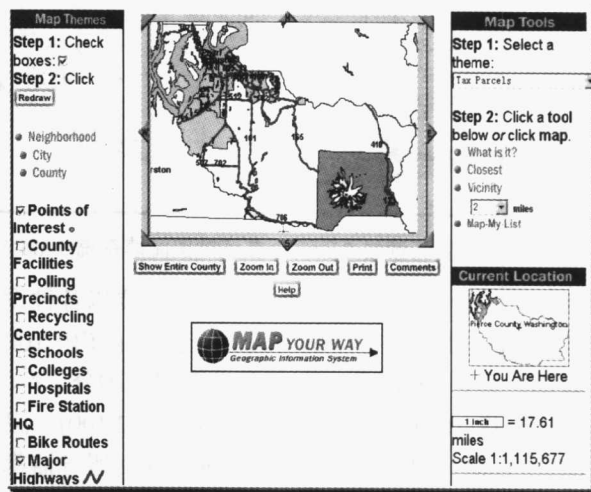


图 2-1 Pierce 县 MAP—Your—Way!™ 在线信息系统界面

这些站点包括对如下关键问题的授权许可: 数据容限、解释、空间精度、可靠性以及对商业应用的警示。

2.3.2 地图建立和传输

对于构建用户定义的地图在线工具, 已经很普及。下面是一些例子。

2.3.2.1 澳大利亚的交互式制图环境

纸质地图的麻烦之一, 是用户需要的信息可能跨两张或更多张地图。这可能是由于用户研究的地域覆盖两张或多张地图所表达的范围, 也可能是用户想达到综合显示那些非正常分幅的数据所需。GIS 克服了这个问题, 允许用户按需交互地组合研究区域的任何数据层来建立自己的地图。然而, 多数潜在的地图用户不会直接访问单机版 GIS。在线地图创建服务器通过在线界面, 给用户提地图分层数据和绘图运算法则来解决这个