

Web 环境下实现空间数据表达的框架研究

刘荣高, 庄大方, 刘纪远

(中国科学院 地理科学与资源研究所, 北京 100101)

Framework for Representation of Spatial Data in Web

LIU Rong-gao, ZHUANG Da-fang, LIU Ji-yuan

(Institute of Geographical Sciences and Natural Resources, Chinese Academy of Sciences, Beijing 100101, China)

Abstract The development of Internet promotes the requirement for acquisition of spatial information from Internet. Because of some characters of spatial data such as object relationship, data integration and very large volume, it is a challenge to represent of the spatial data in Web environment. In this paper, we discuss issues about the architecture of implementing WebGIS. Through analysis of the processes from spatial data to display map, the architecture of WebGIS can be concluded as four types. We proposed an open architecture integrating all four types. The test results in Geolibview prototype showed that using vector format is better than image format in high interactive environment, but if only need to display spatial information, image type is better.

Key words geospatial visualization; system architecture; WebGIS

摘要: Internet 的迅速发展, 促进了通过网络获得空间信息的要求。由于空间矢量数据具有数据量大、对象关系复杂等不同于文本、图像的特殊性, 在数据传输较慢且要求高交互操作的 Web 环境下, 空间矢量数据的可视化是一个难题。本文讨论了 Web 中实现空间矢量数据可视化的体系结构。分析由数据库中的空间数据到客户端地图显示的整个过程, 认为 WebGIS 的实现可以分为服务器端方式、客户端方式、中间数据转换方式和客户服务器分布计算方式等 4 种方式, 每种方式各有优点, 应用于不同的环境。我们提出一种能够同时实现这 4 种方式的开放框架结构, 方式的选择由客户端传送的参数决定。通过对比, 认为在交互操作较高的环境中, 以矢量数据方式传输较好, 而对于交互操作较少的空间信息表示, 以图像方式传输较好。

关键词: 空间数据可视化; 体系结构; WebGIS

1 前言

地理信息系统(GIS)的定义较多, 从应用角度, 表现在能够收集、管理、分析与表达关于人类

生活世界空间的信息系统, 分析功能是其重要特征^[1]; 从数据管理角度, GIS 是一种能够处理和表达空间数据类型的信息系统^[2], 核心是空间数据的处理与表达。在 Web 环境下实现 GIS 功能的

收稿日期: 2000-04-10; 修回日期: 2001-02-13

基金项目: 中国科学院知识创新工程重大项目(KZCX1-Y-02)

作者简介: 刘荣高(1970-), 男, 贵州锦屏人, 中国科学院地理科学与资源研究所博士后, 博士, 主要从事面向空间数据应用的核心技术研究。

技术称为 WebGIS, 这既包括了 Web 对 GIS 功能的限制也包括了其提供的扩展功能。

由于 Web 环境的限制, 要求在 WebGIS 中实现单机 GIS 软件的所有功能是不可能的, 但应具备这些最基本的功能: 能够对空间、属性数据进行查询检索与显示; 用户能够在客户端对图形进行平移、放大、缩小等操作。在这些功能的基础上, 可提供一些较常用的功能: 多层数据的叠加, 用户能控制显示的层面及显示顺序; 客户端能够改变信息表达的风格, 如可改变图例颜色、图符等。在一些特殊的应用中有更高的要求, 如野外军队的应用, 要求在网络中断的情况下, 客户端仍可获得有限的信息服务, 不因网络中断而完全停止。实现这些功能的关键是空间数据的表示, 因此 WebGIS 也称为基于 Web 的制图 (Web-based map)^[3]。

尽管在 WebGIS 方面做了许多研究工作^[4,5], 然而, 实现的系统通常只针对某些特定的应用, 单纯采用矢量数据或栅格数据的传输方式, 系统不具有良好的扩充性及适应性。一个理想的 WebGIS 体系结构应考虑空间数据与 Web 平台的特点, 在现有 Web 平台上进行扩展, 以实现空间数据的有效可视化。

2 WebGIS 的特殊性

WebGIS 的特殊性表现在 Web 环境的特殊性和 GIS 空间数据的特殊性, 前者与单机的计算环境的不同, 后者区别于数字文本、图像信息。

2.1 Web 环境的特殊性

Web 环境的最大优点在于以一致的用户界面进行信息共享, 用户可获得一些特殊功能而不必付出额外费用、平台独立性等, 这正是 Internet 得到迅速发展的原因。

Web 环境由浏览器(客户端)、网络、服务器 3 部分组成。浏览器与 Web 服务器通过 HTTP 协议进行信息传输, 与单机应用有很大的差别: 网络数据传输速率大大低于本地数据的传输, 要求尽可能减少冗余数据传输; 客户端机器千差万别且需求复杂, 要求应用程序具有平台独立性与可扩展性; 多用户同时访问, 要求实现多连接管理; 高度交互性, 要求服务器端尽快响应, 在服务器端应尽可能减少运算量, 将较多的数据运算分散到客户端, 减轻服务器端的负担。

2.2 空间数据的特殊性

空间数据与文本、图像不同的一些基本特点如下:

1. 空间数据有复杂的结构, 一个空间对象可以是在空间中任意分布的, 由单个点或多个点构成的数据集。与文本数据不同, 强调对象在空间的位置。也与图像不同, 空间对象可以是 0 维、1 维、2 维甚至多维数据, 对象间存在空间关系, 图像信息是一个平面整体。

2. 空间对象不只一种表示, 如一个居民区可以看着一个点或一个区, 取决于应用目的与比例尺大小。由于应用目的不同, 对信息的需求也不同, 如林业部门关心林地的分布, 农业部门关心耕地使用状况, 而交通部门则强调公路网络设施。不同的应用强调不同的空间对象实体, 不可能通过同一视图满足所有需求, 对同一空间数据集需实现信息的多重表达。

3. 由于数据的积累越来越多, 提供服务的部门越来越广, 用户对于数据信息的层次要求也不同。如对于土地资源的利用情况, 存在国家部门、省级部门与地方县级部门的不同层次要求。上级部门要求宏观的、空间分辨率较低的信息, 而下级部门则希望了解空间分辨率较高、范围较小的信息。因此, 同是一个数据集, 作为不同的用户、不同的应用目的, 要求得到的空间信息分辨率也不同。用户对空间信息要求的层次性需要对不同层次的信息细节进行表达, 形成从整个空间信息的概览到特定区域的详细视图。

4. 空间数据常常数据量巨大, 信息的多维性使得空间数据的检索计算复杂, 运算量大, 数据处理是一种耗时的操作。

5. 空间数据格式与表示复杂, 差别较大, 导致空间数据共享困难。

2.3 客户端空间数据的操作

Web 环境中, 用户通过浏览器获得信息。计算机显示空间有限, 对于空间信息, 不能将所有需要表达的信息一次全部表示。同时, 由于空间数据的尺度效应, 不同尺度的信息通过缩放操作来实现, 同一层次的不同位置的空间信息由移动操作获得。而文本、数字信息的详细程度由参数的选择来确定或由滚动条移动来实现。

放大操作使空间对象变大, 但视觉范围变小, 在视觉范围内的空间对象数目减少。纯几何图形的放大只简单改变空间对象的放大倍数, 所有几何对象使用相同的放大因子, 对象间的相对位置、

相对大小、空间几何关系、细节程度保持不变,纯几何的缩小操作与此类似。在这种情况下,如果比例尺变化较大,获得的信息不能满足该比例尺下人对客观世界认知的规律。如果需要得到基本符合制图规律的空间数据表示,需要实现智能放大技术^[6]。

3 实现WebGIS的框架结构

3.1 空间数据可视化过程

矢量空间数据经过制图过程形成图像显示给用户,通过下列步骤实现:

- 1 矢量数据库:管理空间矢量数据,如对数据的增加、删除等。
- 2 数据检索:对用户发出的数据请求命令进行解释,根据提供的数据检索参数,提取需要数据。
- 3 图素生成:将检索到的数据集转变中立的几何图形元素序列,是一个将空间对象转为几何对象的过程。
- 4 渲染过程:渲染过程从几何图形元素中生成显示命令,常转换为栅格图或系统图形接口的API函数。
- 5 图像显示:显示过程是通过合适的设备使得经过渲染的图展现给用户。

3.2 WebGIS实现的体系结构

上述制图5个过程,除图像显示必须在客户端,数据库必须在服务器端外,其他3个过程可在服务器端与客户端完成。这样,WebGIS的实现体系结构可分为4种类型:服务器端方式、客户端方式、中间数据转换方式、客户服务器分布计算方式。

3.2.1 服务器端方式

服务器端方式^[7]是最早且仍在使用的一种。在这种方式中,所有数据的处理在服务器端完成,Web服务器通过ASP、CGI、ISA PI或NSAPI等接口与GIS应用程序通信,实现空间数据处理,客户端只需提交请求与显示响应结果。其优点在于:在网络数据流量不大的情况下,用户能够访问大的和复杂的数据集;服务器端可充分利用现有软件资源;客户端不需附加软件。但是,客户端的每一个请求,不管该请求如何小,必须返回服务器处理,处理结果再通过Internet返回,客户端数据不能重用。

3.2.2 客户端的实现方式

通过HTTP(或基于TCP/IP)的协议,服务器只提供文件服务功能,向客户端提供请求数据,不做任何处理,所有处理过程在客户端实现。最常用的方法是利用JAVA或控件(如OCX)直接访问服务器数据。优点包括:将所有复杂的处理过程分散到客户端,减轻服务器负担,实现请求的快速响应;用户可更多控制数据的使用;Web服务器不需要额外软件支持。缺点是:Web端没有附加软件,不能实现数据的有效索引,大数据量时很难检索;对于多种数据源的访问不能提供一致的接口。

3.2.3 中间数据转换方式

服务器检索出空间数据后,生成一种中间矢量格式数据,客户端通过HTTP下载后通过插件显示,典型的如VRML格式。优点是:用户下载数据后可以任意放大、缩小、移动等操作;服务器端可以数据库中的数据格式进行有效的屏蔽,实现访问多种数据源的一致性访问及数据的保密性;下载的数据能在其他软件中使用。缺点是:所有显示数据进行一次传输,不能根据用户操作的实际情况对数据进行选择,数据冗余量较大;放大、缩小等操作只是简单的几何放大,不能实现智能放大。

3.2.4 客户服务器分布计算方式

在服务器端实现空间数据的管理与检索,客户端对返回数据实现图素生成、渲染与显示,将制图过程步骤在客户与服务器端分配。特点是:通过适当的索引和检索参数,可实现不同层次数据的传送;客户端对得到的数据,可根据应用的要求进行处理、显示。如果用户的操作不需要新的数据,如层间顺序、显示风格等的改变,则不必向服务器端发出请求;通过适当的缓冲机制,可实现客户端的数据重用,不必对所有数据进行重新传输,减少冗余数据的传输。这些特点可满足大量空间数据可视化对多层信息显示、有效数据传送、智能缩放等技术的要求。

4 实现WebGIS体系结构的实例

上面4种方式都具有一定的优点,其中服务器端生成图像的方式对只提供空间信息显示而交互操作较少的情况下较好,特别对大数据量非常实用(如表1所示)。中间数据转换如VRML方式是当前发展的一种趋势,但在数据量大且客户与服务器端数据交换较多的情况下,数据的传输量

非常大,而且不能解决数据的重用问题,而客户端方式与分布式计算方式本质上是相同的,前者是后者的一种特殊情况。

表 1 不同操作的数据传输量

Tab 1 Data transmission size for difference operation

| 范 围 | 多边形数目 | 数据量/kB | 传输数据量/kB |
|--------------|-------|--------|----------|
| 1 10 万(市中心) | 233 | 2 232 | 2 232 |
| 1 20 万(放大一倍) | 664 | 477 | 2 515 |
| 1 10 万整屏右移 | 72 | 1 026 | 615 |
| 1 10 万整屏上移 | 258 | 1 332 | 888 |
| 1 10 万整屏下移 | 198 | 2 227 | 742 |

当前,尚没有发现一种体系结构能够同时包括上述 4 种方式。我们在空间数据浏览器(GeoSpatial Data Library Viewer,简称 GeoLibViewer)原型中使用的体系结构(如图 1 所示)成功解决了这一问题。客户端程序由界面管理、数据下载管理 2 模块组成,服务器端由开放网络服务器及数据格式驱动模块组成。客户端由移动代码 JAVA 对服务器端的数据进行访问及显示,服务器端由开放网络服务器负责响应客户端数据下载模块的数据请求,通信协议为无状态通信协议 HTTP。

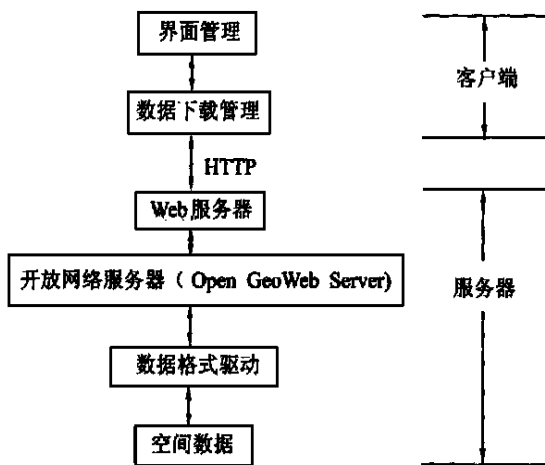


图 1 GeoLibViewer 的体系框架结构

Fig 1 Architecture of GeoLibViewer prototype

界面管理: 界面管理负责数据的显示及与用户的交互。提供给用户最基本的供用户操作空间信息的最简单的操作,包括放大、缩小、移动及复原等几何类型操作,同时包括空间对象与属性相互查找的操作,专题图图层的增加、删除与顺序变化。操作类型的选择由工具条、菜单和对话框方式提供。

数据下载管理: 负责管理及缓冲从服务器端下载的数据,管理专题图制图需要的符号,向界面管理器提供数据,当缓冲器中没有满足请求条件的数据时,向服务器请求该数据。

开放网络服务器: 以 Windows NT 的服务方式运行,当 Web 服务器接到 HTTP 的请求后,调用 DLL 处理 HTTP 的参数,并以进程通信的方式传递至 GeoWeb 服务器,服务器进行处理后将数据返回。HTTP 是非状态协议,在每次请求时不保存上一次的操作状态,而为响应每次请求而访问数据需要打开与关闭文件,这是一种费时的操作,从效率方面考虑,有些系统(如 OGD)采用保持状态的协议。在多用户同时操作的情况下,保存所有用户的状态对服务器是一种巨大的负担,并不适合高交互的 Web 环境。在此方案中,提高数据访问效率的方法是在完成数据访问后,并不马上关闭文件,而是保存文件打开状态一定的时间,时间的长短由服务器访问的繁忙程度决定。

数据格式驱动模块: 将不同的数据格式通过相对应的驱动程序转换为符合开放 GIS 简单特征(OpenGIS Simple Feature)规范的格式,对象由对应的点、线、面及对应的 ID 组成,使用统一的坐标及投影方式。

客户端与服务器端的通信由 HTTP 完成,其格式如下:

http://ServerName//GetServerData DLL? RequestType= " Image ";

LayerName= " FileName or ProjectName ", Extent = ...

请求类型(RequestType)决定系统服务时采取何种方式,并返回相应类型的数据。这种符合 HTTP 方式的通信协议实际使得服务器与客户端的耦合作用非常松散,客户端并不需要知道服务器如何运作,而服务器端也无须了解客户端如何处理数据,这使得服务器与客户端均具有很好的扩充性与移植性。LayerName 表示需要访问的数据名称,可能是目录服务中的名字,Arc/Info 格式的 Coverage 名、Map 名或 Arcview 文件的工程文件名等。名称及对象的对应关系由服务器的元数据管理器决定。后面的参数根据前面的类型不同而不同,如需要以图像方式返回,则规定需要图像的范围,而对于矢量对象的访问,可根据范围、ID 值等访问。

数据下载管理屏蔽界面程序从服务获得数据

的细节。界面模块通过函数调用获得需要的数据。这些函数包括6个核心函数,可分别访问图像、2D数据(点、线、多边形)、栅格数据(GRID与DEM)及中间数据格式如VRML数据,在这6个核心函数的基础上,提供更高级的20多个函数以更方便使用。属性数据的访问则通过对象的ID,由JDBC直接访问服务器端的数据库。

在这种体系结构的原型上,对贵州省贵阳市1998年1:10万土地利用现状数据进行测试,全部数据多边形1750个,数据量6863kB,比较浏览操作过程中,以矢量与图像传输所需的数据总量,结果如表1。

对于数据量较大的空间数据服务,假设服务器的处理速度足够快且需要的处理操作不复杂,客户端的反应时间仅仅由数据的传输速度决定,很明显,传输的数据量越大,反映时间就越长。采用JPEG图像传输,每屏数据量约140kB。矢量数据通过压缩后,数据量能减少约40%~50%,在操作50次后,图像传输的数据量将超过矢量数据。分布计算方式的缺点是第一次传输时数据量较大,而随着操作的不断增多,数据逐渐积累,可减少数据传输冗余与分散数据传输时间。这是采用一个大比例尺多边形数据直接生成多个小比例尺专题图的情况,如果采用数据优化措施,如多边形的公共边不重复传输、服务器保存多比例数据等,矢量传输的数据会更小,但同时服务器与客户端程序的复杂程度会增加。目前,我们已经研究出一种方法,能够将矢量数据压缩至1/10以下,这种情况下,矢量数据的传输会明显优于图像方式。

5 结 论

根据空间数据到图像由矢量数据库、数据检索、图素生成、渲染过程、图像显示5个步骤组成;根据在客户端或服务器的不同实现,WebGIS的体系结构可分服务器端方式、客户端方式、中间数

据转换方式、客户服务器分布计算方式4种,我们提出了一种能够实现上述所有方式的体系结构,选择何种方式由客户端的参数进行选择,在服务器端通过格式驱动的方式,实现对不同格式数据的透明访问。通过对原型的实验表明,对于用户交互操作较频繁的应用,分布式计算方式是比较理想的。而对于只是向用户提供空间信息显示的应用,采用图像传输的方式更好。但对于某些必须需要矢量数据在客户端完成分析操作的应用,如交通导航,以图像方式传输则无能为力。

参考文献:

- [1] MARTIN D. Geographic Information Systems[M]. London: Routledge Ltd, 1996.
- [2] GÜTINGER H. An Introduction to Spatial Database Systems[J]. VLDB Journal, 1994, 3(4): 357-399.
- [3] Open GIS Consortium. User Interaction with Geospatial Data[EB/OL]. <http://www.opengis.org/wwwmap/98-060.pdf>, 1999-06-02/2000-03-10.
- [4] KÄHKÖNEN J, LEHTO L, et al. Interactive Visualization of Geographical Objects on the Internet[J]. International Journal of Geographic Information Science, 1999, 13(4): 385-396.
- [5] GOUND, MORIN P. Solving the Geospatial Data Barrier: Open Geospatial Datastore Interface (OGDD)[J]. CIG Geomatica Journal, 1997, 51(3): 278-287.
- [6] FRANKAU, TOMPFS. Multiple Representations for Cartographic Objects in a Multi-scale Tree—An Intelligent Graphical Zoom[J]. Computers and Graphics: Special Issue on Modelling and Visualization of Spatial Data in GIS, 1994, 18(6): 823-829.
- [7] FOOTEKE, ANTHONYPK. WebGIS, NCGIA Core Curriculum in GIScience[EB/OL]. <http://www.ncgia.ucsb.edu/giscc/units/u133.html>, 1998-12-12/2000-03-10.