

# 基于超地图原理的分布式空间数据模型

龚健雅 韩海洋

(武汉大学测绘遥感信息工程国家重点实验室, 武汉 430079)

**摘要** 分布式、部件化、交互性与互操作是因特网地理信息系统的4个主要特征。为了实现空间信息的分布式管理与处理,以便将大容量、多源地理信息纳入统一的框架管理,基于超地图与分布式数据库的原理,提出了满足上述特征因特网地理信息系统分布式地理信息空间数据模型,并以目录服务作为空间信息全球化的范式,同时定义了不同层次超链节点的目录服务数据结构,以便通过超链接实现分布式空间数据模型的关联,该模型已在GeoSurf中得以应用,并取得了良好效果。

**关键词** 分布式 地理信息系统 超地图 数据模型

**中图分类号**: P208 **文献标识码**: A **文章编号**: 1006-8961(2002)07-0688-05

## Hypermap-Based Distributed Geospatial Data Model

GONG Jian-ya, HAN Hai-yang

(National Key Lab for Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079)

**Abstract** With the development of internet technology, internet GIS have applied popularly. Distribution, interoperability, component are the main characters of internet GIS. In order to access, manage, process and distributed geodata, the paper proposes a distributed geospatial data model, which is based on hypermap concepts and distributed database principle. According to the distributed geospatial model, various geospatial data including electronic maps, images, pictures, videos, attributes, hyper documents can be constructed in the model, and they are linked by a hyper link. Every medium can have one or more hyper nodes, which link another medium. A hyper node recodes some information including the IP address and catalog of linked medium. After clicking a hyper node, we can move to another IP address and browse linked medium. Furthermore, the paper defines data structure of different type hyper nodes and discusses its implementation. An internet GIS software-GeoSurf based on the distributed hypermap geospatial model have developed by our university. Many application systems based on GeoSurf have been built. These applications has shown many advantages of the distributed hypermap geospatial data model.

**Keywords** Distribution GIS, Hypermap, Data model

## 0 引言

分布式地理信息(DGI)反映了地理信息具有空间分布性的天然特性。本文给出了它的描述性定义,即分布式地理信息系统是基于TCP/IP协议,以目录服务为中心的GIS。通过Internet人们可以获取包括图形、图象、数据集、分析操作、文本报告等各种

形式的地理信息,并可完成数据的浏览、查询、分析、制图等深加工过程,为公众或专业人士提供与地理信息相关的产品和服务,以实现数据与操作的最大限度共享。DGI主要包括客户机、服务器与网络通讯3部分,而互操作、分布式、部件化是其本质特征。

对于Internet GIS而言,分布式、部件化、交互性与互操作是4个重要特性。其中,分布式反映数据与空间操作的物理分布;部件化是系统构成方式;互

基金项目:国家杰出青年科学基金项目(49525101);国家自然科学基金项目(49871006);

国家重点基础研究发展规划项目(G20000077904)

收稿日期:2001-05-24;改回日期:2001-09-10

操作提供了不同格式空间数据集成与操作的共享或信息共享<sup>[1]</sup>。该 4 个特征中,分布式占主导地位,部件化为分布式的表现形式,而且基于分布式才为互操作提供了更为现实的实现途径。Internet GIS 通过部件化的构造与消息传输来表现分布式的结构,分布式包括数据的分布与对数据进行操作的分布。

就目前的 Internet GIS 而言,仅仅处于发展初期,距分布式 GIS 目标尚有相当大的距离,因为尚存在以下主要不足:

(1) 没有体现地理信息分布式的固有特性。尽管采集的地理数据在地域上是分布的,但是仍然采用集中、统一的空间数据库管理。这一方面是由于原始数据的数据格式的多样性,集中存储时,需要先做耗时耗力的格式转换工作;另一方面,不同地域数据的集中存储,不但消耗大量的人力、物力、财力,而且造成数据更新困难、现势性差。

(2) 由于上述不足,造成对数据的处理也必须集中进行。对 GIS 服务器而言,当接受大量请求时,可能不堪重负,并且无法在最佳的网络节点完成数据处理,以减少数据的网上传输和提高系统响应性能。

(3) 缺乏对分布式地理信息的有效组织。目前大部分的 Internet GIS 只能获取并处理单一图幅,而无法顾及图幅之间的联系,因而缺乏目录服务和工程的概念。这种形式无法胜任国家空间数据基础设施和国家省一级的基础地理信息系统网上信息发布的要求。

(4) 由于对超地图的研究与应用仍处于初级阶段,现有系统难以实现基于超链接的矢量图或栅格图自身与其他媒体深层次的交互<sup>[2]</sup>。

上述问题的出现,关键在于缺乏针对分布式地理信息特征的空间数据模型。对于传统的 GIS 而言,空间数据模型与数据结构的研究已经比较深入了,虽已有成熟的数据模型和数据结构可操作<sup>[3]</sup>。但是,由于 Internet GIS 发展时间短、没有形成完善的理论,迄今为止,还没有一个适合地理信息分布式存储与分布式处理的数据模型和数据结构,而现有的大部分 Internet GIS 还都是基于传统 GIS 的模型框架构造,故而不能适应进一步发展的需要。

本文将通过对分布式地理信息的特征、Internet GIS 的发展方向 and 目前这种方式局限性的分析讨论,结合超地图思想、分布式数据库理论与目录服务的概念,提出基于超地图理论的,以目录服务为核心的分布式地理信息空间数据模型,并讨论它的实现方式。

## 1 分布式地理信息空间数据模型

最近几年,基于超文本链接的浏览器逐渐成为 Internet 门户<sup>[4]</sup>。通过浏览器的超文本链接可以访问 Internet 上的分布式资源,但是由于单纯的文本协议不能满足 GIS 界的要求,因此 Laurini 和 Milleret-Raffort 借鉴超文本和超媒体的思想,于 1990 年提出超地图的概念<sup>[5,6]</sup>并应用到 GIS 中,从而推动了分布式 Internet GIS 的产生和发展,并成为未来 GIS 发展的重要趋势之一。

超地图理论可通过设置节点与超链,把不同的目标以超文本方式链接起来,使人们有可能获取分布式地理信息。但如同超文本一样,由于超链的最终结果将是庞杂无边的网络拓扑结构,不用说用户,就是组织者本身也很快迷失链接的结构关系,何况浏览器虽然支持文本或栅格地图的显示与链接,但并不支持矢量地图显示,因此如何定义矢量图节点的数据结构以及超链接方式,也是一个需要深入探讨的问题。

目录服务一般被视为分布式网络的中枢神经系统。所谓目录是指在公共和私有网络中用来存储特定对象的结构或者信息源,而目录服务即是包括这种信息源,并能向终端用户提供使用资源的服务。从根本上说,一种目录服务是一个机构中,包括人和机器在内所有信息资源的分布式数据库。它是网络环境下分布式 GIS 的一个重要组成部分,它可使得用户和管理员在不知道所想要的对象确切名字和位置的情况下,仅凭对象的一个属性或多个属性,便能够查找到它。

虽然超链的方式存在诸多不足,但是借鉴超地图的思想以及分布式数据库与目录服务的理论,仍可以设计一个分布式地理信息空间数据模型。该模型可针对分布式地理信息特性,通过对节点的定义,以超链来实现分布式地理信息的获取和以目录服务来实现分布式地理信息的组织管理,以及以分布式多服务器来实现对地理信息的处理。由于服务器端数据与处理都是分布的,而用户端操作则是集中的,因而保证了获取与处理地理信息的分布与集中的辩证统一。

### 1.1 分布式地理信息空间数据模型的一些概念与组织层次

分布式地理信息空间数据模型定义的概念有超地图节点、超地图工作区、超地图工程、超地图节点数据库、超地图主节点数据库、目录服务数据库、超地图图层、超地图超链、元数据库等。下面对这些概

念做进一步的定义。

(1) 超地图节点 通过超地图节点可以建立不同超地图工作区或不同媒体间的联系,超地图节点作为一种特殊的对象,它表现为对地图一定区域范围或某个空间对象的链接关系,它的属性包括链接对象的特征、与被链接对象的特征、超链关系(遵循TCP/IP协议)等。

(2) 超地图工作区 超地图工作区是一定区域内地物层的集合。通过将超地图工作区的一定区域或图形对象设置为超地图节点,可以建立该超地图节点到其他节点数据库超地图工作区的超链链接,也可以被其他节点链接。

(3) 超地图工程 对某一区域的超地图节点数据库,按照区域、专题属性可进一步细化为若干个超地图工程。超地图工程可是某个局部区域所有类别地理信息的组织者,也可反映某一专题,如水系、土壤等;超地图工程由多个超地图工作区或专题图层组成。

(4) 超地图节点数据库 与Internet相连的每一台主计算机称为一个网络节点,通过这样的节点可以访问Internet上的资源。网络节点如果作为数据库服务器,并获得了经Internet获取、访问数据库数据的权限,则这个网络节点对应的数据库服务器就称之为节点数据库。对于GIS而言,地理信息存储在空间数据库中,相应地,与网络节点对应的空间数据库服务器就称为超地图节点数据库。超地图节点数据库是上网发布某一区域范围或某一专题地理信息的集合体。它可以包括若干个超地图工程。

(5) 超地图主节点数据库 即所有超地图节点数据库包含的地理信息的空间区域并集。

(6) 目录服务数据库 目录服务是网络管理的基础,有了完整的目录服务,才可快速部署分布式计算环境。当网络安全服务也捆绑进目录服务中时,目录服务即能实现对策略的访问和控制。目录服务数据库目的就在于为分布式空间模型的各个节点提供超链及属性的目录服务,以便使整个模型成为有机的整体。超地图节点数据库的目录服务可提供模型的层状拓扑结构、超地图节点数据库的网络位置;超地图工作区或节点的目录服务因能对所有的链接和被链接对象进行有组织的管理,故便于不断更新和维护。

(7) 元数据库 提供超地图节点数据库、超地图工程或超地图工作区的基本信息。终端用户可以据此查询需要的地理信息,并确定其分布位置;服务器管理者可以据此合理组织超地图节点数据库和超

地图节点的链接。

分布式地理信息空间数据模型的网络拓扑结构有如下两个层次:其一是超地图主节点数据库和超地图节点数据库,可按照地理信息的层次关系来建立层次模型,并作为骨干网;其二是不同节点数据库包含的超地图工程、超地图工作区、超地图图层可通过超链与超地图节点来建立超地图联系,进而形成扩展网,而扩展网还可在原有的骨干网基础上,复合出更加复杂的网状模型。目录服务数据库与元数据库则提供骨干网和扩展网的层次关系、链接方式以及空间数据基本信息。

分布式的地理信息可以在不同的区域,通过建立多个超地图节点数据库来把全球的地理信息作为一个整体考虑,即按照区域范围、行政所属、专题类别来建立多个超地图节点数据库的层次关系,而不同层次的图幅则通过超链链接。与Internet的组织方式类似,任何一个组织或个人都可以将自己的空间数据库定义为一个超地图节点数据库,并按照一定的规则加入到这个层次的网络中,作为层状结构的一个节点。按照这样的框架组织的分布式地理信息,从超地图节点数据库的层次看,呈条理明晰的层状模型;而从具体某个图幅或图幅的节点看,则呈复杂的网状模型。同时对应于不同层次的目录服务数据库将完成相应层次的管理与更新。

## 1.2 分布式地理信息空间数据模型的实现

根据对分布式地理信息空间数据模型概念的介绍和模型层次的分析,可以构造出如图1所示的分布式地理信息空间数据模型。从图中可以看出,模型整体按照空间关系呈层次构造,便于对一定区域或专题的地理信息进行组织管理,管理人员也可以任选方式在根节点下创建或删除新的节点,并建立关联全局的目录服务;而模型局部由于引入超地图节点和超链的概念和错综复杂的网状构造,使超链突破了原有的层次关系,故可以充分表达分布式地理信息之间以及与其他媒体的关系,并能突破空间限制。基于分布式数据库与超地图理论的分布式数据模型本身也体现了地理信息空间分布与内在联系的辩证统一。

超地图节点数据库由超地图工程、目录服务数据库与元数据库组成,而每个超地图工程又由多个超地图工作区组成。这种超地图工程用来管理大型的空间数据,并建立超地图工作区之间的联系,而超地图工作区的数据则要求具有相同的坐标系和比例尺,并取决于用户的需要,一个超地图工作区可以是

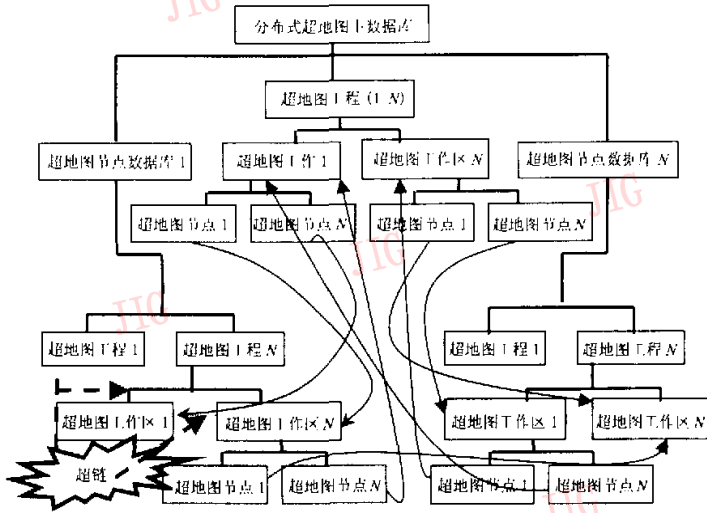


图 1 分布式地理信息空间数据模型

按行政区域或空间位置分割的一个多要素图幅,也可以是某个专题层。

超地图节点数据库的目录服务是整个模型的核心,因为目录服务的作用体现在:①确定超地图节点数据库在网络中的位置和在整个模型中的层次关系;②存储本超地图节点数据库的工程与工作区信息;③保留本超地图节点数据库每一个工作区图层的超地图节点的属性和超链结构以及建立到相应的元数据库的对应关系,以备检索。由此可见,超地图主节点数据库的目录服务包含了所有超地图节点数据库的网络位置以及体系的层次关系。

超地图节点数据库作为 Internet 上的节点,其构成如图 2 所示。

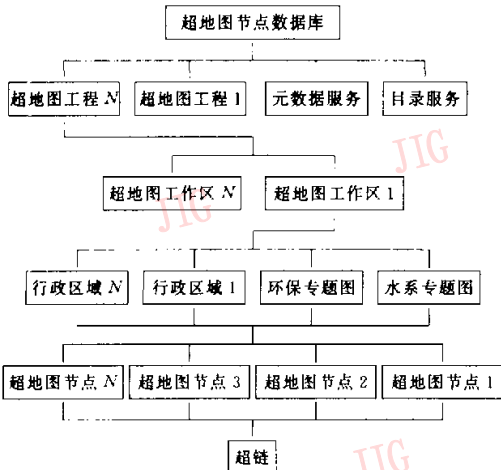


图 2 超地图节点数据库结构

## 2 分布式地理信息空间数据模型目录服务及其数据结构定义

由于分布式地理信息空间数据模型的显著特点是通过超地图节点与超链来建立不同图幅之间及与其他媒体的交互关系,并通过目录服务对超地图节点数据库的层次和所建立的超链加以组织与管理,因此,分布式地理信息的空间数据结构就集中在目录服务的数据结构的定义。

分布式地理信息空间数据模型目录服务主要包括工作区超地图节点与其他媒体的超链网状目录、节点数据库所有工程与工作区的层次目录、网络中所有节点数据库的层次目录等,下面逐一分析。

### (1) 建立超地图节点目录服务的数据结构。

在超地图的概念中,地图主要起图形参考和索引作用,即只要通过确定某一空间区域为超地图节点的热区,那么该超地图节点就可以链接到各种超媒体介质(包括文档、图象、矢量图形等)当中,但如何确定超地图节点的热区范围和链接到何处、何种媒体介质以及哪个对象,则取决于超地图节点的数据结构。

(2) 超地图节点类型包括文本、图象、音频、视频、矢量图形、虚拟现实等超媒体介质,这里主要定义与地理信息关系密切的矢量与栅格地图、虚拟现实超地图节点的数据结构,而对其他类型的结构不再赘述。

(3) 矢量图形超地图节点主要指在矢量地图中定义热区作为节点响应超链的空间范围. 对于热区的确定一般有设置矩形区域、设置多边形区域和设置点、线、面等空间对象 3 种方法. 对于最后一种方式, 还需要设置一定的区域容差范围.

在一个超地图工作区中, 当确定了超地图节点的热区范围后, 则需要给节点一个编码. 编码一般包含节点所在超地图节点数据库的 IP 数值, 以确保在整个模型中唯一. 与节点编码一致的是节点关键字, 即节点名称, 一般应该反映节点所链接对象的属性. 为了与所链接对象建立超链关系, 必须声明下级链接编码、下级链接的 IP 等字段. 由于下级链接的媒体类型可能是矢量图形、图象, 也可能是虚拟显示等媒体, 而不同的链接媒体对应着不同的获取和处理方式, 所以必须声明下级媒体类型以及该媒体的数据格式. 比如图象的数据格式可能是 BMP 或 GIF, 但是如果链接的是矢量图形就比较烦琐, 因为 GIS 矢量数据格式多种多样, 所以就必须先确定目标图形的数据格式, 然后针对其格式以及元数据信息, 采用相应的数据获取部件来获得存贮的数据, 经过处理后再显示给终端用户. 表 1 为矢量图形的超地图数据结构. 其他媒体与此相似.

表 1 中, 如果采用多边形区域来确定节点热区, 则使用 LinkRegion 字段, 否则令 LinkRegion 为空; 如果采用图形的地物标识码(OID 或 FratureID)来确定热区, 则用 ObjectOID 字段.

可以通过鼠标点击热区来调用矢量图形链接, 也可以将图形缩放到一定比例尺时, 自动与下级链接.

表 1 矢量图形超地图节点

字段名	字段类型	字段描述	字段说明
VnodeOID	长整数	节点编码	
VnodeName	字符	节点关键字	
VWSOID	长整数	所在工作区编码	
VobjectOID	长整数	对象标识	地物 OID
VlinkRegion	多边形对象	链接区域	Polygon(x[],y[])
VlinkScale	整数	调用链接的图形比例尺	
VlinkDB	字符	下级链接数据库编码	
VlinkWSOID	长整数	下级链接工作区编码	关联到目录服务查找

### 3 结 语

分布式地理信息的空间数据模型是在传统 GIS

空间数据模型基础上, 着重突出地理信息分布的特点来构建的. 其中, 基于超地图的节点数据库的目录服务是模型实现的核心, 因为它为分布式处理提供了强有力的索引和层次关系, 从而使整个模型更具条理性. 但同时必须认识到, 分布式地理信息的空间数据模型仅仅是分布式地理信息系统或 Internet GIS 设计、实现的基础. 为了实现对分布式异构空间数据的获取与处理, 还必须结合模型构造浏览器/分布式多服务器的数据获取、处理与表达框架(另文阐述)来将分布式数据处理服务器与超地图节点数据库对应, 以形成完善的分布式地理信息系统. 具体的实现已经在因特网地理信息系统软件 GeoSurfV3.0 中体现<sup>[7]</sup>.

### 参 考 文 献

- OGC. The OGC technical committee technology development process[M]. Open GIS Consortium, Way-land, Massachusetts, 1997.
- Limp WF. Weave maps across the Web[J]. GIS World, 1997, 10(9):46~55.
- 龚健雅, 李斌, 汤勤. 当代 GIS 的若干理论与技术[M]. 武汉: 武汉测绘科技大学出版社, 1999.
- Lennon J A. Hypermedia systems and applications: World Wide Web and Beyond Berlin[M]. New York: Springer, 1997.
- Hall W, Davis H, Hutchings G. Rethinking hypermedia: The microcosm approach[M]. Boston: Kuwer Academic Publishers, 1996.
- Korte G B. The GIS book [4th ed][M]. Santa Fe: On Word Press, 1997.
- 龚健雅. GeoSurfV3.0 (EB/OL). <http://www.rcgis.wtusm.edu.cn/wwwgis>.



龚健雅 1957 年生, 教授, 博士生导师, 长江学者奖励计划特聘教授. 现主要从事地理信息系统理论与方法、空间数据库管理系统、地理信息系统软件、地理信息系统应用等方面的研究.

韩海洋 1973 年生, 1996 年毕业于武汉测绘科学大学, 现为武汉测绘科学大学测绘遥感信息工程国家重点实验室硕博连读研究生. 主要从事 Internet GIS 和分布式计算研究. 现在加拿大学习.