



全国警用空间信息服务互联互通的关键技术

伍胜¹,杨崇俊²

(1.西南大学 计算机与信息科学学院,重庆 400716;2.中国科学院 遥感应用研究所遥感科学国家重点实验室,北京 100101)



伍胜(1976-),男,重庆人,讲师,博士,主要从事网络地理信息系统方面的研究工作。

基金项目:中央高校基本科研业务费专项资金项目(XDJK2120131037)资助
E-mail:wushengcq@126.com

收稿日期:2011-08-20

摘要:针对大型空间信息服务平台构建过程中的数据互联互通和信息集成问题,从集成框架体系结构、技术路线和关键技术3个方面进行研究,提出了分布式地理空间信息集成框架的5层体系结构模型和原子空间信息服务概念;通过研究分布式空间查询路由算法和虚拟四叉树模型等关键技术,探索实现分布式环境下的空间信息的互联互通。

关键词:地理空间信息集成;警用地理信息平台;地理空间信息服务网络;分布式资源查询路由;虚拟四叉树模型

Key Technologies of Service Integration in Police Geographic Information System

WU Sheng¹, YANG Chong-jun²

(1. Institute of Computer & Information Science, Southwest University, Chongqing 400716, China; 2. The State Key Laboratory of Remote Sensing Information Sciences, IRSA, CAS, Beijing 100101, China)

Abstract: Focus on integration framework architecture design, technology roadmap and key techniques aspects, this paper present the researches conduct on how to realize data interoperability and information integration in process of building large scale GIS service platform. In this paper, it proposed a five-layer-architecture for the integration framework of distributed geospatial information, and declared the conceptual of fundamental geospatial information web services. By raising the distributed geospatial query routing algorithm and virtual quadtree model, it promoted the step to achieve interoperability of spatial information.

Key words: distributed geospatial information integration; Police GIS; geospatial information services network; distributed resource query routing; virtual quadtree

0 引言

地理信息系统经过近半个世纪的研究和发展,逐渐从专业研究领域走向社会生产领域,从桌面系统走向网络环境,呈现出网络化、分布式和面向服务的发展趋势,由原来的封闭、紧耦合的单机系统演变成成为开放、松耦合的地理信息服务。在大型的省部级信息化平台建设过程中,各地各部门自主建设的地理信息系统往往呈现出一种在

广域网范围内、横向跨区、纵向跨级的松散分布状态,难以实现GIS数据、功能和服务的互联互通。

为了有效地组合这些资源,实现一体化的管理和使用,避免重复建设带来的资源浪费,迫切需要对这些网络上松散分布的地理空间信息资源进行组织、管理和集成,对上层应用开发提供统一、透明的分布式地理空间信息服务访问接口。

在计算机科学领域,针对分布式计算已进行了多年研究,涉及并行计算、高性能计算、网格计算、云

计算和分布式文件系统等多个领域,但针对地理空间信息这一特定对象的分布式信息集成还有待进一步研究。日益完善的网络基础设施、网络通信技术和被广泛应用的HTTP服务,为基于WebGIS的分布式空间信息系统集成提供了良好的支撑环境;Web Services技术的发展和成熟,为分布式的空间信息集成提供了新技术手段;Web Services与GIS的结合成为分布式GIS服务的主要研究方向和实现途径;OGC,ISO/TC 211等国际标准



化组织对 GIS Web 服务进行了基础研究,提出了一整套相关的技术规范,并得到了业界广泛的认可和采纳,为异构空间信息集成提供了有力支持。

国家科技部和公安部共同组织研发的警用地理信息平台(简称 PGIS 平台)正是这样一个横向覆盖全国各地行政区域,纵向贯穿部、省、市多级管理机构的大型警用地理信息平台,其部署节点分布于全国各地公安机关。为满足公安部门的基本业务需要、支撑跨区域的协同打击犯罪工作、方便系统内部的垂直领导与监管体制以及机构内部的空间信息资源共享(实现“全国一张图”),迫切需要实现全国各个 PGIS 结点之间的信息互联互通。同时,伴随着 PGIS 平台的不断推广,也为分布式地理空间信息集成研究提供了一个良好的实验环境。

本文将主要从集成框架体系结构、技术路线和关键技术 3 个方面对分布式地理空间信息集成问题进行研究。提出地理空间信息集成框架的 5 层体系结构模型;在此基础上,探讨如何利用 Web 服务技术实现遵循 OGC 规范的原子地理空间信息服务;研究如何依据服务管辖的地理空间范围关系实现跨区域的联网服务注册,构建地理空间信息服务网络的基本框架;分析并设计空间查询路由算法,实现在分布式网络环境中的空间资源查找与定位。通过以上内容研究与探讨,建立地理空间信息集成框架原型。

1 集成框架设计

1.1 集成框架概念设计

主要目的:在现有计算机网络提供的物理连接基础之上,基于空间位置关系建立一个虚拟的分布式地理空间信息集成框架,在此基

础上汇集松散分布的地理空间信息服务资源,屏蔽异构空间数据格式和异构 GIS 系统的差异,隐藏服务结点的位置分布关系,使其在集成框架内部能互联互通、协同服务,为上层应用提供可透明访问的一体化地理空间信息服务网络(Geographic Information Service Network, GISN)(如图 1 所示)。以此实现横向跨区、纵向跨级的地理空间信息的集成共享和地理空间信息的规模化、社会化服务。

是用户完全感觉不到后台分布式系统的存在,网络中所有的可用数据资源和服务资源都无缝地集成在一起。解决地理空间信息服务无缝集成问题,本质上是要解决地理空间服务的目录组织问题;⑤解决分布式空间分析和计算结果的缝合问题。GIS 系统中需要处理大量的矢量和栅格地图数据,针对这些具有空间属性的分布式计算结果,如何通过信息融合技术实现重新组合,也是分布式集成框架需要解

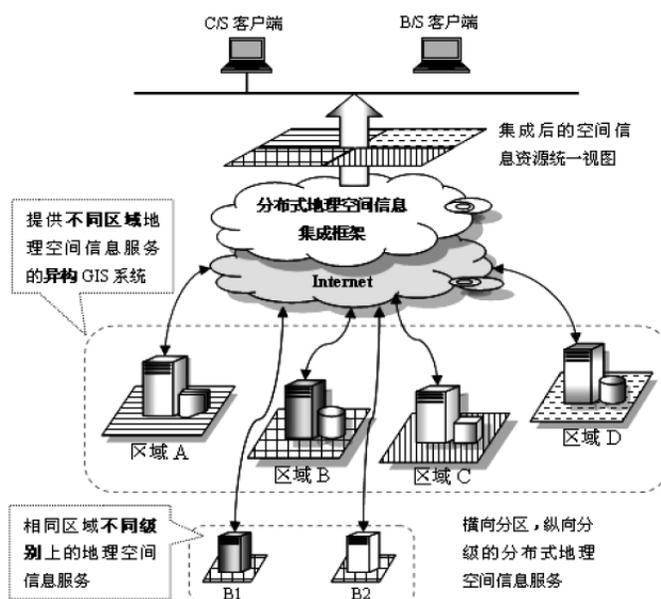


图 1 分布式地理空间信息集成框架概念设计图

Fig.1 Concept of distributed geospatial information integration framework

1) 实现目标

实现空间信息集成需要解决如下基本问题;①解决异构系统集成问题。只有消除异构才能实现系统间的互操作,才能实现服务功能的叠加;②解决海量空间数据的存储与组织问题。通常的分布式存储解决了数据的物理存储问题,但同时割离了数据之间的地理空间逻辑关系。分布式空间数据的组织问题就是要在集成框架中,把物理分布的空间数据组织成逻辑统一的全域空间数据视图;③解决地理空间信息系统分布式部署的问题;④解决地理空间信息服务的无缝集成问题。集成框架的最理想的效果

决的问题;⑥解决集中访问带来的系统压力问题。空间资源逻辑上集成可能会带来用户访问的集中,用户访问的集中最终会带来系统的访问压力问题。

2) 设计原则

设计原则是对整个框架设计的限制和指导性意见:①兼容性原则。集成框架作为整个结构中的上层建筑,不能对已建或待建的 GIS 系统具有很强的侵入性,只能采用兼容或对现有系统做扩展的策略来实现信息的集成。②面向服务的原则。采用 SOA 架构实现系统良好的松耦合、平台无关性和协议无关性等。③分层设计原则。通过分



层简化系统设计的复杂度,同时降低编程实现时的难度。④可伸缩性原则。整个集成框架结构应该是开放的,局部系统可随意地进入或离开集成框架,并且同时保持局部的可用性,局部可用性不依赖于集成框架的整体可用性。而整个集成框架系统是由局部系统松散耦合在一起的,单个局部系统的加入或离开对整个集成框架结构没有任何影响。其中每个子系统的结构是完备的,子系统与整个集成框架是同构的,子系统与集成框架呈现出一种分形的关系。⑤透明性原则。一是访问透明性,对用户屏蔽底层的分布式环境。二是框架结构透明性,对于每个要加入到集成框架内的子系统来说,它不必了解集成框架系统的整体结构,只需感知相邻结点的存在,集成框架对其内部所有资源的定位过程都是一个空间路由查找过程。

1.2 集成框架结构设计

概念设计阶段,对集成框架的结构设计拟定了分层结构的设计原则及各层完成相对独立的功能,下层为上层服务,上层通过接口调用下层的功能,共同完成对分布式环境下的地理空间信息集成。从信息的联通性看,分布式地理空间信息的集成过程可以看作是分布式地理空间信息的联网服务过程。其体系结构可以和经典的计算机网络体系结构作类比,集成框架从总体结构上一共分为5层,其结构如图2所示。

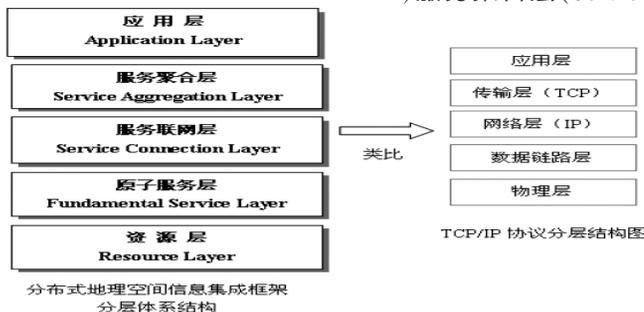


图2 分布式地理空间信息集成框架分层体系结构图

Fig.2 Five-layer architecture of distributed geospatial information integration framework

1)资源层(Resource Layer)

主要功能是执行完成具体的空间数据访问和空间分析与计算工作。该层由各类异构的原始地理空间信息资源组成,主要形式是各种空间数据文件、各类空间数据库、多种空间数据访问引擎以及各类GIS平台提供的空间信息服务功能。

2)原子服务层(Fundamental Service Layer)

原子服务层是分布式系统中功能单元的访问接口,所有上层的GIS功能都是通过对原子服务层的访问来实现的;原子服务层是底层异构地理空间信息的屏蔽层。通过开放的Web服务模式,可以完全屏蔽下层的操作系统,开发环境,编程语言,通信协议等差别,实现多系统间的信息交互;Web服务与OGC规范的结合,使得Web服务层能够屏蔽下层的异构GIS环境,实现系统间的空间互操作。

3)服务联网层(Service Connection Layer)

通过分布式地理空间信息服务资源目录的联网注册机制,把所有分散的空间信息原子服务连接成一个逻辑上的地理空间信息服务网络,实现GIS原子服务的联网功能。与TCP/IP协议体系结构中的IP层一样,服务联网层是集成框架的核心层,需要通过联网服务注册机制完成GISN的组建,同时需要实现和资源的路由查找算法。

4)服务聚合层(Service Aggre-

gation Layer)

在服务聚合层内部实现基于GISN的分布式空间资源查询路由机制,针对查询结果设计并实现分布式数据的并行访问程序框架,对于返回的多个计算结果,研究实现空间数据缝合技术,最终实现对上层屏蔽底层的分布式网络结构,把服务联网层的逻辑整体结构变成实际意义上的整体结构,把服务网络变成一个可以直接通过程序调用,并获得最终返回结果的服务实体。

5)应用层(Application Layer)

应用层是集成框架分层体系结构中的最上层,包含了在GISN上运行的空间信息应用系统,这些应用系统通过调用下层服务来完成对分布式空间信息资源的访问,并提供相应的人机交互接口,空间信息可视化组件或是计算结果的信息展示平台等,为最终的业务用户提供空间信息服务。

集成框架的5层体系结构分别从不同层次上解决地理空间信息集成所面临的问题,每层处理相应的问题,下层为上层服务,下层是保证上层功能实现的基础,共同实现对分布式异构地理空间信息的集成。

2 集成框架关键技术

集成框架的5层结构,从逻辑上划分了分布式环境中地理空间信息互联互通的基本问题域。在具体实现过程中,还涉及到很多具体关键技术需要研究解决。

2.1 分布式查询路由算法

在地理空间服务网络中“服务域”可分为两类:域和子域。域和子域是相对的逻辑关系,同一空间信息服务即是域也可能是其他域的子域。与地理空间服务域划分概念及其相似的是运行于Internet上的



域名空间服务 (Domain Name Service, DNS), DNS 服务几乎是伴随着网络的诞生就开始不断发展和完善, 已经成为网络最重要的基础设施之一。与 DNS 采用名称空间作为域划分的理论有所不同, 在地理空间服务域的划分理论中, 所依据的不是名称空间 (Name Space) 而是与 GIS 关系密切的地理空间 (Geographic Space) 作为域的划分标准。但从结构上看, 两者是一致的, 都是倒立的树状结构。除了逻辑结构的相似性以外, 更为重要的是, DNS 的系统结构和空间信息服务集成框架的系统结构一样, 都是运行在网络环境下的分布式系统。所以, DNS 成熟的系统结构和技术实现方案, 对地理空间信息服务的联网集成框架具有很强的指导性和很高的参考价值, 在一定程度上减小了集成框架理论探讨和具体系统实现过程中的技术风险, 提高了对系统可用性原则的保证。

所有服务结点上的服务注册中心联网形成了一个分形结构的树状服务网, 每个结点只关心其上级父亲结点和所管辖的下级结点 (从实现角度考虑, 参照了 DNS 域名系统的实现方法^[1], 在每个节点内部设置了默认的根节点访问路径, 作为缺省配置参数。每个结点除了知道其上级父亲结点和所管辖的下级结点外, 还知道系统的根结点位置), 在整个网络内没有一个结点了解整个网络的结构, 所有的资源定位过程都是一个需要反复迭代的查询路由 (Query Routing) 的过程^[2], 如图 3 所示。

2.2 基于虚拟四叉树模型的构建分布式栅格地图服务

当前的互联网环境中, 绝大多数的空间信息服务都是通过切块栅格地图的方式来实现的。所以, 栅格地图服务也成为空间信息集成

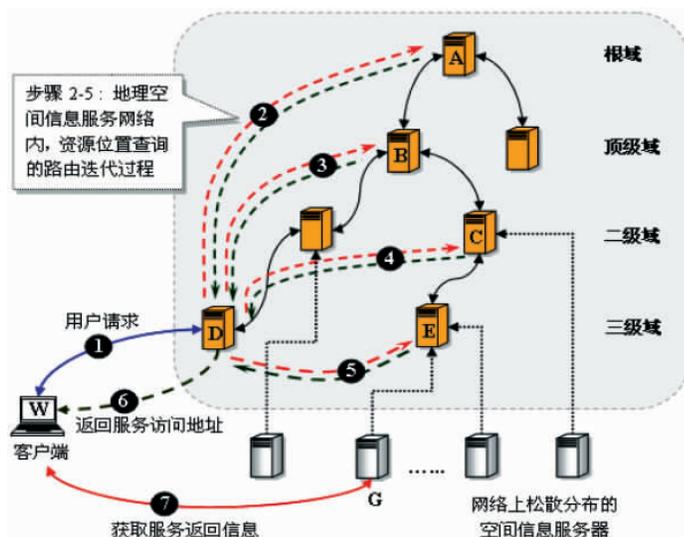


图 3 分布式空间信息服务查询路由步骤图

Fig.3 Distributed geospatial service query routing steps

研究工作中, 需要实现互联互通的首要对象。结合 PGIS 项目的需要, 对分布式环境下的栅格地图服务做了深入研究, 在空间信息服务集成框架的体系结构基础之上, 提出了基于虚拟四叉树模型的空间数据组织理论, 解决多服务器间、跨区域、跨多比例尺的栅格地图数据集成, 实现高速无缝的地图漫游。

四叉树模型通常是作为一种地理空间域的区划分模型^[3]。其主要思想是对给定的地理空间按四等分方式——直递归划分, 直到规定的切分次数后终止。基于文件系统的集中式四叉树模型, 它将空间数据的组织和存储紧密耦合在一起, 通过数据的目录层次结构来表达数据的层次关系。这种四叉树模型结构简单, 易于实现。但这也导致了这种四叉树结构本身缺乏灵活性和扩展性等固有缺陷。

1) 数据的导入过程几乎是单向的。数据一旦融入到大树中, 就和大树连为一体, 很难再把数据导出。特别是当该树下又挂载了多个子树结点时, 会出现牵一发动全身的问题。

2) 数据扩展效率低。当有更详细的数据来扩展原四叉树时, 必须将新数据全部复制到挂载点所在

的目录下, 涉及大量数据迁移。

3) 侵入式的融合方式。当新的子树与原树存在空间位置重叠时, 挂接操作将导致新旧数据间的相互覆盖, 发生数据丢失现象。

虚拟四叉树模型 (Virtual QuadTree, VQT) 将数据的组织方式和数据的存储方式分离。它仅建立空间数据的四叉树组织框架, 并不关心数据的存储方式和存储位置。空间数据通过挂载的方式附着到虚拟四叉树上, 由此建立起多级空间数据的金字塔模型。

VQT 具有两种类型的结点: 信息结点 IQNode 构成 VQT 的主干框架 (如图 4 所示), 但不包含具体的空间数据; 数据结点 DQNode 具体存储了空间数据。它通过注册的方式挂载到 IQNode 结点下, 从而融入到 VQT 结构之中, 形成多级海量空间数据的金字塔模型。虚拟四叉树的注册机制使四叉树的子

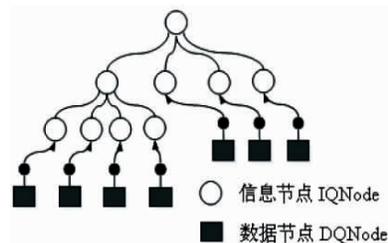


图 4 虚拟四叉树模型

Fig4 Virtual QuadTree Model

树挂载、多树融合算法变得简单,无需数据的复制,也不会引起数据的覆盖而破坏原始四叉树结构(如图5,图6所示)。VQT具有以下优点:

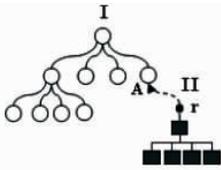


图5 VQT子树挂载
Fig.5 Mount sub tree

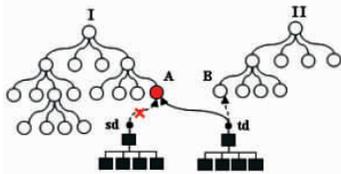


图6 VQT多树融合
Fig.6 VQT multi tree merge

1) 分离了数据组织和数据存储,增强了四叉树的灵活性和扩展性。子树挂载、多树融合无须复制数据,使树结构更易于扩展、更新和融合,适合大规模数据的集成。

2) 通过数据的注册机制,屏蔽了数据的存储位置,屏蔽了本地数据的访问和远程数据访问的差异,支持分布式网络环境下的多级海量空间数据的组织。

3) 通过数据的注册机制,实现了分布式环境下多级空间数据的松散耦合。

3 集成框架的实现

在PGIS项目中主要实现了对分布式WMS服务和栅格切片地图服务的集成,通过集成框架从多个服务中心获取其管辖范围内的空间数据,提供给嵌入浏览器中的GIS客户端使用,对用户提完全透明的分布式空间数据访问。建立

了跨省范围内的地理信息数据访问基础平台。这个地理信息基础平台设计支持全国各级机构不同比例尺的地图不同分辨率的影像和矢量栅格地图。使用数据包括:全球范围500m分辨率影像数据和1:1 000 000地理数据,全中国范围15m分辨率影像数据和1:250 000地理数据,部分城市1m分辨率影像数据和1:2 000地理数据。在实际生产工作环境中,实现部局与宁夏、内蒙、天津等多区域的地理空间信息共享和互联互通,为上层应用提供了跨省的多区域栅格地图无缝快速地图漫游。

4 结论

本文主要从系统体系结构、技术实现方案和关键技术研究3个方面对分布式地理空间信息集成问题进行了比较深入的研究。提出了地理空间信息集成框架的5层体系结构模型的基本思想;在此基础上,结合Web服务技术和OGC规范提出原子地理空间信息服务的概念;并依据服务管辖的地理空间范围关系实现跨区域的服务注册联网,探讨了如何建立地理空间信息服务网络;通过分布式空间查询路由算法实现在分布式网络环境中的资源查找与定位;最后利用空间信息结果集合进行缝合,为用户提供透明、无缝的地理空间信息服务访问,实现分布式环境下的地理空间信息集成。

分布式环境下的空间信息集成是一个涉及内容较广的研究课题,本文的设计思想主要是从具体应用需求的角度出发对这个问题做了研究和探讨,并没有包含分布式环境

下空间信息集成各方面的所有问题。另外,就本文提出的集成框架5层体系结构,在实现过程中只是完成了其主体的框架原型,还需要在后续具体的实践过程中结合问题对设计做相应修改和调整。

参考文献

[1] YANG ChongJun, WU Sheng, REN YingChao. Loose Architecture of Multi-level Massive Geospatial Data Based on Virtual Quadtree [J]. Science in China Series E -Technological Sciences 2008, 51 (zk1): 114-123, DOI: 10.3724/SP.J.1005.2008.00225.
ISSN:1006-9321CN:11-3754/NL

[2] Samet, Hanan. The quadtree and related hierarchical data structures [J]. ACM Computing Surveys, 1984, 16 (2): 187-260.

[3] Jibor Dostálek, Alena Kabelová. DNS in Action A detailed and practical guide to DNS implementation, configuration, and administration [M]. 2006: Packt Publishing Ltd.

[4] 陈军. 试论中国NSDI建设的若干问题 [J]. 遥感学报, 1999, 3(2): 94-97.

[5] 李德仁, 龚健雅, 李京伟, 等. 中国空间数据基础设施建设 [J]. 测绘通报, 2002 (5): 70-74.

[6] 伍胜. 分布式地理空间信息集成框架 [D]. 北京: 中国科学院遥感应用研究所, 2006.

[7] ISO/TC211. ISO/Geographic information/ Geomatics [EB/OL]. [2009]. <http://www.isotc211.org/>.

[8] Consortium, Open GIS. Open GIS Consortium Website [EB/OL]. [2009]. <http://www.opengeospatial.org/>.

[9] 王方雄. 基于原子服务的网格空间信息服务互操作研究 [D]. 武汉: 武汉大学, 2005.