

Analysis of Research on Spatial Information Service Retrieval

Lijun Duan¹, Ying Song²

¹School of Computer, Hubei University of Education, Wuhan Hubei

²School of Information Engineering, Hubei University of Economics, Wuhan Hubei

Email: duanlijun@hue.edu.cn

Received: Sep. 29th, 2017; accepted: Oct. 13th, 2017; published: Oct. 20th, 2017

Abstract

Spatial information service retrieval is one of the main problems that need to be solved in the application of space information. In this paper, we systematically analyzed the research status of spatial information service retrieval. We introduced the concept scope of spatial information service retrieval and object form in retrieval, and made the classification for the existing methods. Based on the analysis of the principle and application of typical methods, we summarized the characteristics of each category and pointed out the problems in future research.

Keywords

Spatial Information, Service, Retrieval, Analysis

空间信息服务检索研究分析

段丽君¹, 宋 莺²

¹湖北第二师范学院计算机学院, 湖北 武汉

²湖北经济学院信息工程学院, 湖北 武汉

Email: duanlijun@hue.edu.cn

收稿日期: 2017年9月29日; 录用日期: 2017年10月13日; 发布日期: 2017年10月20日

摘 要

空间信息服务检索是空间信息应用中需要解决的主要问题之一。本文对空间信息服务检索的研究现状进行了系统的分析。介绍了空间信息服务检索的概念范围和对象形式, 将现有方法进行了分类, 通过分析

其中典型方法的原理和应用情况,总结了每类方法的特点,指出了未来研究中面临的问题。

关键词

空间信息, 服务, 检索, 分析

Copyright © 2017 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

近年来,伴随着智慧地球与智慧城市建设的落实与深入,现代地理信息技术的处理手段也日益丰富。以新一代实时地理信息系统为代表的技术,能够高效地接入天空地环境中各种传感网传送实时地理空间信息,并能有效利用各类空间信息数据和传感器,面向政府、行业和公众提供不同层次的服务,为地理国情监测和智慧城市提供基础平台。目前我国已成功发射了近百颗卫星,形成了较为完善的海洋、气象和环境等系列的遥感卫星体系。其中的高分辨率立体测图卫星资源三号每年生产将近 200 TB 空间地理信息数据,为国家地理测绘、国情监测提供了可靠的数据保障。同时,网络基础设施的普及、Web 技术的发展和面向服务架构(Service-Oriented Architecture, SOA)的出现,也促进了地理信息以网络服务形式进行访问、处理和共享[1]。开放地理空间信息联盟(Open Geospatial Consortium, OGC)作为行业国际化组织引领着空间地理信息标准及定位基本服务的发展,主要指导地理信息行业软件和数据及服务的标准化工作,它制订了一系列重要的有关空间数据和服务的标准,如 Web 地图服务(Web Map Service, WMS)、Web 要素服务(Web Feature Service, WFS)、Web 覆盖服务(Web Coverage Service, WCS)等,这些标准促进了各种开源地理信息服务框架的出现,也大大降低了空间信息服务发布的门槛。目前,地理位置图片、GPS 定位数据、地名、GIS 矢量格式数据、遥感卫星影像等各种类型的地理信息都以空间信息服务的形式被发布到空间数据基础框架(Spatial Data Infrastructure, SDI)的分布式环境中,它们的格式不尽相同,甚至内容迥异,因此,如何在海量的空间信息服务中进行高效、准确的检索就成为了空间信息应用中需要解决的主要问题之一。

本文阐述了空间信息服务检索的概念范围和对象形式,结合近年来空间信息服务检索的研究与应用情况,对典型的空间信息服务检索方法进行分析,旨在帮助学习研究人员进一步了解空间信息服务检索的研究进展,为后续研究工作的开展做好铺垫。

2. 空间信息服务检索概述

2.1. 概念范围

如前文所述,当今社会对地理数据和地理信息服务的需求快速增长,与此相对的是,地理信息的获取效率和效果并没有得到显著提高,应用最广的信息检索方式依旧是搜索引擎这种相对简单的方法,面对着日益增长的空间数据和逐渐开放的地理信息源,这种矛盾日益突出。地理信息检索是关于访问地理信息资源的一门技术,可以被看作是传统信息检索的一个分支,它包含了传统信息检索所有核心研究领域,唯一不同的是更强调对地理信息的检索。地理信息检索技术研究主要包含两方面,即检索工具与检索机制[2]。空间信息服务检索属于地理信息检索的内容之一,它主要检索反映地理实体空间分布特征的信息,如定位数据、多光谱卫星影像数据等描述实体的位置、形状及实体间的空间关系、区域空间结构

的信息。为得到更好的检索结果, 目前空间信息服务检索研究工作的重点内容主要集中在检索工具的设计与改良, 检索机制与方法的创新上。

2.2. 检索的对象形式

网络服务作为一种最基本的组件来支持分布式异构环境下的应用组合, 能够推动资源的共享、复用和集成。面向服务架构 SOA 的出现促进了大量地理信息开始以网络服务形式进行封装。Web 服务作为 SOA 架构重要的实现方式之一, 它只需要定义接口就能实现服务应用间的互操作, 而与具体实现语言和平台无关。由于 Web 服务的优点明显, 所以越来越多的空间信息服务以 Web 服务的形式被发布, 目前空间信息检索的对象绝大部分也是基于 Web 技术的服务。因此, 本文也主要分析和比较以 Web 服务为主要的研究对象的空间信息服务检索方法。

3. 典型空间信息服务检索方法

3.1. 基于通用注册中心的检索

此类方法属于早期的空间信息检索方法, 该类方法主要借助 UDDI (Universal Description Discovery and Integration) 等通用的注册中心来实现服务的检索。该类方法基于集中式拓扑网络体系, 所有服务必须登记并发布在 UDDI 注册中心中才能够被服务请求者发现。而且由于 UDDI 只支持基于关键字的 Web 服务查找, 随着 UDDI 内容的激增, 从中检索出理想的服务往往困难且低效。更重要的是, UDDI 是与领域无关的, 不能很好地体现空间信息检索的领域特点。因此, 此方面的研究往往将重点集中到对 UDDI 的功能扩展上。文献[3]就对 UDDI 注册中心进行扩展以实现其对 GIS 领域的支持, 并在此基础上提出了一个空间信息网格体系结构, 如图 1 所示。

UDDI 只能基于关键字查询, 其检索效率低, 检索结果精度差。通过 UDDI 等通用注册中心来检索空间信息服务往往不能取得满意的结果, 因此目前在空间信息检索中很少有研究将这些通用注册中心作为研究对象, 最新的研究成果也较少。

3.2. 基于专用服务目录的检索

为适应地理技术的应用需求, OGC 提出了网络目录服务(Catalog Service for the Web, CSW)。CSW 是专门管理对地观测数据的标准协议框架, 也是最常用的一种空间信息目录服务, 主要针对 Web 网络, 它支持 ebRIM、ISO 等应用纲要(Application Profile), 提供了数据、服务和其他资源的元数据的发现、浏览和查询的接口。随着地球空间信息产业的不断发展, 越来越多的地理信息资源被发布到 CSW 中, 针对网络目录服务的检索需求也日益增多[4]。目前在 SDI 内, 地理数据和服务的检索都主要依赖于 CSW。但由于分布式网络环境的复杂性和异构性, 以及跨学科领域的语义差别, CSW 的应用受到了较大限制。解决此问题, 许多研究都将工作重点集中在 CSW 的接口功能扩展上。Yue P.等扩展了 CSW 的 ebRIM 信息模型, 以增强目录服务的地理信息处理和知识发现能力, 更好地支持服务发现和编排[5]。为了实现对公开服务接口的空间资源的统一化集成和发现, 文献[6]提出了一种基于 CSW 的空间资源统一化发现服务实现框架。文献[7]通过分析地表覆盖信息的特征扩展了 CSW 目录的地理信息元数据标准, 提出了全球地表覆盖信息元数据模型, 在此基础上设计了一个全球地表覆盖信息元数据目录系统软件架构, 以实现此应用方面空间服务的检索, 如图 2 所示。

由于 CSW 的查询接口是基于 XML (Extensible Markup Language) 实现的, 而且查询条件一般也是基于文本相似或数值条件的约束。因为服务的语义不能得到显式表达, 所有此类方法在文本相互匹配时一般会产生语义缺失, 容易导致匹配失败, 也无法有效解决异构问题。

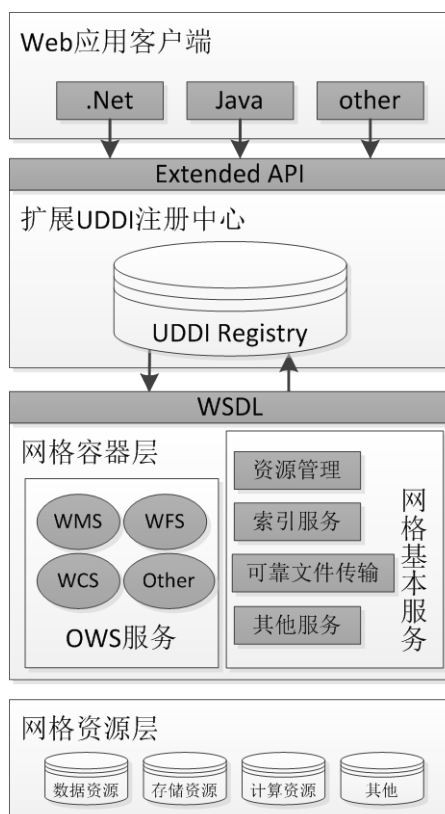


Figure 1. Spatial information grid structure based on extended UDDI
图 1. 基于扩展 UDDI 的空间信息网格结构

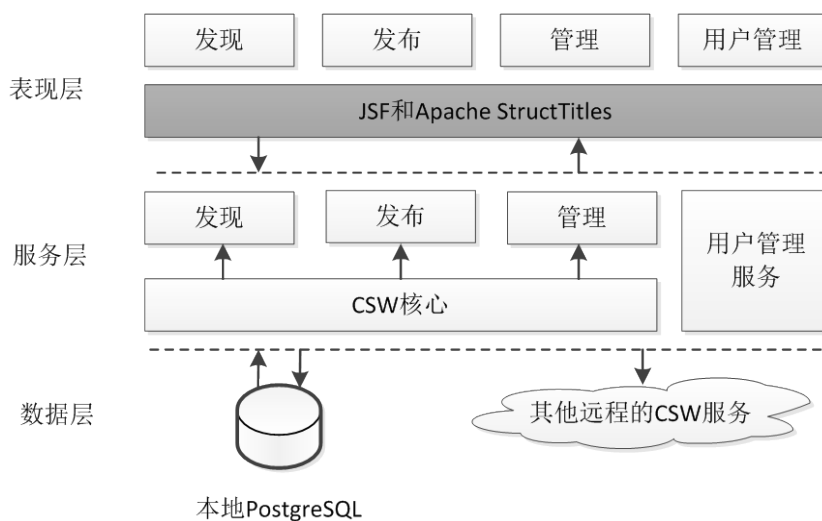


Figure 2. Software architecture of global surface coverage information metadata directory
图 2. 全球地表覆盖信息元数据目录系统软件架构

3.3. 基于本体的检索

本体(Ontology)是语义技术的重要组成部分和技术基石,是实现逻辑推理的基础和保障。其本质是针对那些被广泛认可的概念而提出的形式化描述和规范。目前本体技术在信息技术研究领域内被广泛地应

用于电子商务、信息检索和 Web 服务等方面。由于有本体的存在, 可以使来自多个数据源的异构信息在语义层面上得到统一, 从而使得 Web 服务之间的共享和互通更加容易。因此, 构建一个地理领域相关本体是这类空间信息服务检索方法的必然选择。文献[8]构建了一个空间信息本体模型, 通过它来完成对网络地理信息的推理以及对相关系统的描述和定位。文献[9]使用一系列不同类型的本体来表示领域概念, 并基于这些本体设计了一个地理空间知识基础设施模型, 以此来进行地理空间资源的智能搜索。文献[10]将地理数据空间本体表示为一个五元组, 通过使用该本体来实现地理数据的智能关联, 并按照图 3 所示的检索应用流程在地球系统科学数据共享平台中进行对其进行了性能验证。

目前围绕地理本体的空间信息检索研究已经进行得比较深入, 但制约该类方法的最大因素还是地理本体的构建问题。因为虽然本体是对客观世界的抽象, 但其构建过程却受到太多主观因素的影响。就本体学习而言, 对于同一个目标领域, 不同的人因为理解的不同, 提炼出基本术语和概念可能不尽相同, 由此提取的概念间关系也会不同, 这样就会造成构建的同一领域本体普遍存在着明显差异。因此, 这类方法之间的本体模型往往无法通用, 方法性能差异也比较大。

3.4. 基于索引的检索

虽然地理空间信息中通常包含着矢量数据、栅格数据和文本描述等不同类型的信息, 但这些信息之间一般存在着信息耦合和内在关联, 通过发掘和利用这些关联信息来实现对空间服务的有效检索也成为了一种主要的研究手段。为实现基于 SPARQL 的地理语义空间查询, 文献[11]利用已有的数据组织方法和空间索引设计了一个地理空间四元组模型, 并基于该模型构建了地理语义空间索引, 其组织的实验表明该方法能快速定位空间节点, 实现有效的空间查询。为提高空间查询的效率, 文献[12]设计由几何对象集中空间对象的最小界定矩形上的分层索引组成 R 树索引, 通过最小边界矩形来匹配每个空间几何体, 并设计如图 4 所示的空间信息服务平台, 以实现空间对象的高效检索。为获得更为高效和全面的检索结果, 文献[13]分析了多源地理空间数据的多类型特征和多维度关系, 设计了一种多源地理空间数据关联模型, 该模型的构建流程框图如图 5 所示。为提高此模型构建效率, 文中提出了一种基于特征索引的分块构建方法来支持支持地理空间信息查询、分析及综合展现等多种地理空间应用。

一般而言, 此类方法综合运用了空间数据库的索引技术、Web 技术、语义技术和地理标记语言(GML)等数据交换技术, 其方法成本和计算复杂度一般较高, 但其往往能取得较为满意的结果, 尤其是能提高查询的响应速度。

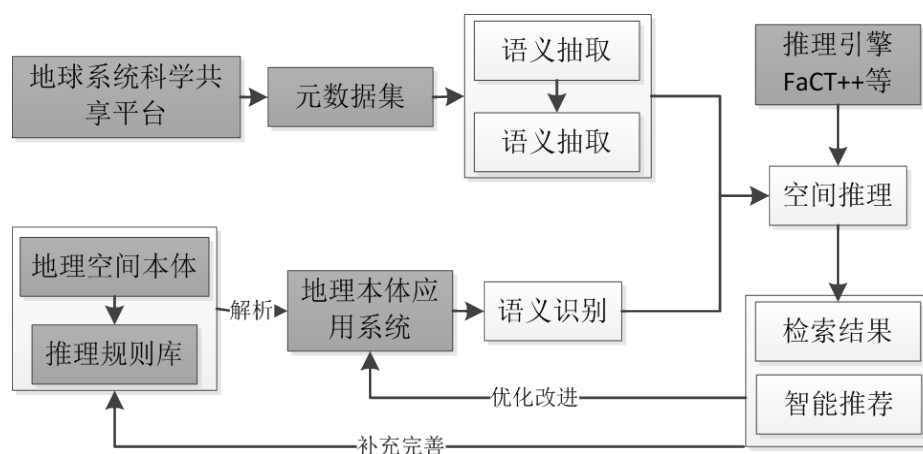


Figure 3. Application of Geodata spatial ontology in semantic retrieval

图 3. 地理数据空间本体在语义检索中的应用流程

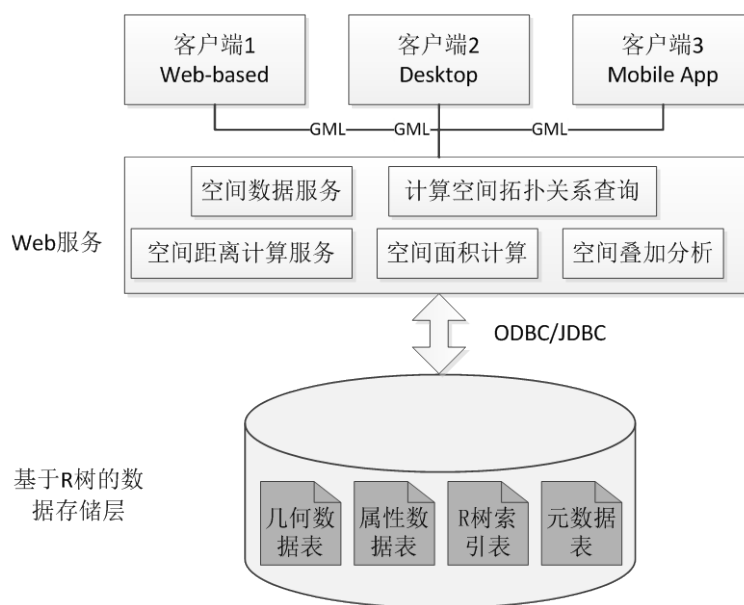


Figure 4. The three-tier architecture of the space information service platform
图 4. 空间信息服务平台的三层架构

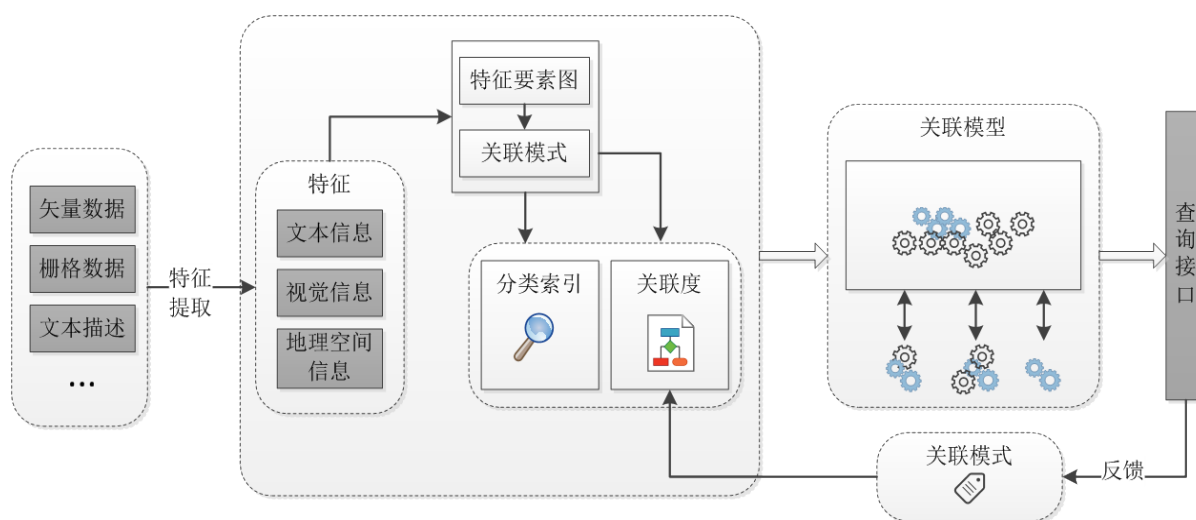


Figure 5. The flow diagram of building relational models
图 5. 关联模型构建流程框图

4. 总结

从空间信息服务检索研究的趋势来看, 研究方法的技术综合性越来越明显, 特别是随着推荐算法和人工智能技术的发展, 空间信息服务检索方法在检索精度、检索效率和个性化需求方面的表现取得了一定的进展。但在大数据技术的驱动和影响下, 检索方法面临的对象规模和算法要求都有层次性的提升, 这也是空间信息服务检索研究中必须要考虑和解决的问题。

致 谢

本文受到湖北省自然科学基金项目(2016CKB714)资助。

参考文献 (References)

- [1] 杜武. 基于排序学习的空间信息服务检索关键技术研究[D]: [博士学位论文]. 武汉: 武汉大学, 2017.
- [2] 王振峰. 基于本体的地理事件信息检索 [D]: [博士学位论文]. 武汉: 武汉大学, 2009.
- [3] 赵姗, 郭建忠, 成毅, 等. 基于 UDDI 扩展的地理空间信息网格注册中心设计[J]. 测绘信息与工程, 2008, 33(2): 36-38.
- [4] 邵远征. 网络环境下对地观测数据的发现和标准化处理[D]: [博士学位论文]. 武汉: 武汉大学, 2011.
- [5] Yue, P., Di, L., Wei, Y. and Han, W. (2013) Intelligent Services for Discovery of Complex Geospatial Features from Remote Sensing Imagery. *Journal of Photogrammetry and Remote Sensing*, **83**, 151-164. <https://doi.org/10.1016/j.isprsjprs.2013.02.015>
- [6] 杨中, 李国庆, 于文洋, 等. 基于 CSW 的空间资源统一化发现服务设计实现[J]. 计算机工程与设计, 2014, 35(5): 1816-1822.
- [7] 徐生磊, 李松年, 孙亚琴. 基于 CSW 的全球地表覆盖信息共享[J]. 测绘通报, 2017(1): 74-78.
- [8] 许泉立, 易俊华, 杨昆. 基于地理本体的空间信息检索机制初探[J]. 测绘地理信息, 2015, 40(1): 65-68.
- [9] Stock, K., Stojanovic, T., Reitsma, F., Ou, Y., Bishr, M., Ortmann, J. and Robertson, A. (2012) To Ontologies or Not to Ontologies: An Information Model for a Geospatial Knowledge Infrastructure. *Computers & Geosciences*, **45**, 98-108. <https://doi.org/10.1016/j.cageo.2011.10.021>
- [10] 王东旭, 诸云强, 潘鹏, 等. 地理数据空间本体构建及其在数据检索中的应用[J]. 地球信息科学学报, 2016, 18(4): 443-452.
- [11] 段红伟, 孟令奎, 黄长青, 等. 面向 SPARQL 查询的地理语义空间索引构建方法[J]. 测绘学报, 2014, 43(2): 193-199.
- [12] 张宁, 刘渊, 沙宗尧. R 树索引环境下空间信息服务平台的开发及应用[J]. 计算机应用, 2014, 34(S2): 135-139.
- [13] 吴焯, 陈萃, 熊伟, 等. 面向高效检索的多源地理空间数据关联模型[J]. 计算机学报, 2014, 37(9): 1999-2010.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2325-2286, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: sea@hanspub.org