

数字城市框架下地理信息服务语义分析

谭永滨¹, 朱海红¹, 李霖^{1*}, 毛凯¹, 李成名², 印洁²

(1. 武汉大学资源与环境科学学院, 湖北 武汉 430079; 2. 中国测绘科学研究院, 北京 100830)

摘要:构建数字城市服务公共平台,能够有效地解决异源异构数据共享的问题,实现地理信息互操作。该文以数字城市服务公共平台下的地理信息服务为例,基于本体论,利用属性枚举的语义表达方法构建地理信息服务本体,以明确且形式化的方法描述服务语义,力图解决地理信息服务语义共享的问题。最后给出数字城市公共平台部分地理信息服务的形式化语义描述,在 Protégé 软件中完成地理信息服务本体的构建,同时结合 HermiT1.3.4 实现推理过程,并以地理编码服务为例构建了一组本体结构。

关键词:数字城市;地理信息服务;语义;服务本体

中图分类号:P208 文献标识码:A 文章编号:1672-0504(2012)02-0005-04

0 引言

随着地理空间数据获取手段的不断进步,地理数据资源不断完善,为了避免数据管理混乱以及数据重复采集造成的人力、物力大量浪费现象,人们对如何整合海量地理数据、最大效益地发挥地理信息资源的作用,增强我国地理信息公共服务能力提出了进一步的需求。因此,建设一个可自由交换异源异构数据的、开放的、具有功能互操作性的地理信息公共平台已成为我国数字城市建设的重要基础。

目前,众多学者在利用 Web 服务技术解决异构数据交换与功能互操作方面进行了大量的研究和探索,已有许多成功的案例。但是随着 Web 服务的不断增多,服务提供者也越来越复杂,如何让服务使用者快速、准确地在海量的服务群中发现满足实际需求的 Web 服务已成为亟待解决的问题。国内外已有学者致力于该方面的研究^[1-4],他们通过引入本体论,为地理信息服务增加语义信息,并利用语义信息提高服务发现的准确性和完整性。郑亮等^[5]提出了一种基于空间服务语义模式的服务发现方法,明确表达空间数据中隐含的知识,以此克服数据源之间语义的异构问题,并通过实验证明了利用语义模式能够显著提高地理信息服务发现的查全率和查准率。李霖等^[3]以国家基础地理信息中陆地水系要素为例,通过定义其本体属性来明确及规范化地表达概念语义,以实现陆地水系要素的形式化表达。此方法值得借鉴用于表达服务概念语义。

本文以数字城市公共平台构建过程中所涉及的地理信息服务域为例,从构建地理信息服务本体入手,试图通过具体分析得出服务明确的形式化语义,进而实现语义共享。

1 地理信息服务语义分析

1.1 形式化本体语义

为了解决信息领域内知识共享和重用的问题,研究人员在信息领域中引入本体论,形成了信息本体。Studer^[6]归纳了信息本体包含的四方面:概念化(Conceptualization)、明确(Explicit)、形式化(Formal)和共享(Share)。如果仅使用自然语言或文字表达地理信息服务,易产生模糊性,因此需采用形式化本体明确地理信息服务的语义信息。同时,服务提供者与服务使用者不同的行业背景也会导致二者对具有相同含义的服务产生不同的理解,从而产生服务异质的现象。Buccella 等^[7]将本体的异质性分成 4 类:概念模型异质性、空间模型异质性、结构异质性和语义异质性。由于 Web 服务技术自身特点已解决概念模型、空间模型和结构的异质性,因此本文只讨论服务的语义异质性问题。

为了实现分类化管理海量地理信息服务,本文建立了服务概念间的语义关系。常用的语义关系包括上下义关系和总分关系,本文采用 kind-of 和 part-of 分别表示概念间的上下义关系和总分关系。关于本体的结构定义尚缺乏统一的标准,一般根据具体的上下文确定。目前比较广泛使用的方法有传统的四元素表示方法和较新的六元组表示方法^[8]。

收稿日期:2011-08-18; 修订日期:2011-10-25

基金项目:国家公益性项目(201010246-03,201110237-02);中央高校基本科研业务费专项资金资助项目

作者简介:谭永滨(1985-),男,博士研究生,主要研究方向为地理语义分析与地理本体研究。*通讯作者 E-mail:lilin@whu.edu.cn

本文针对地理信息服务的特点,参考其他学者的定义,将地理信息服务本体结构定义如下:

定义 1 形式化服务本体结构 $O = (C, S_C, R, H_C)$ 。其中 C 表示服务概念集合; S_C 表示每个概念 c 的语义信息集合; R 表示概念间关系集合; H_C 表示分类关系,是父概念和子概念之间的上下位及总分的层次关系。

1.2 语义表达方式

本体的语义表达实质上是从一个概念空间到另一个概念空间的映射^[3]。目前用于表达概念语义空间映射的方法众多,常见的语义表达方式包括比喻法(Metaphor)^[9-11]、代数法(Algebraic)^[12-14]、概念集成/绑定(Conceptual Integration/Conceptual Blending)^[15]、属性枚举^[3]、一阶谓词逻辑(First-Order Predicate Logic)和描述逻辑(Description Logic)方法^[16]。

比喻法是利用概念之间相似性的特征将陌生概念映射到人们熟悉的概念中,其难点在于如何找到符合要求的熟悉的概念;代数法是通过定义类型、操作和规则来表达概念的语义,但是缺少优秀的建模工具且易学性较差;概念集成/绑定法是通过概念组合方式和结构来表达语义;属性枚举法是通过逐一列举概念本体属性的方式表达概念语义,此方法的关键在于如何列举出代表本体概念本质的属性项;一阶谓词逻辑与描述逻辑表达方法可明确地、形式化地表达语义且改善推理问题的复杂性,但二者均不够直观、简明地表达服务语义的信息。

基于公共平台的多用户性以及用户知识背景的复杂性,需要用简明扼要、易于理解且行之有效的方法表达服务概念的语义信息,本文采用属性枚举方法表达服务概念语义。

1.3 OWL-S 服务描述语言

OWL-S 是在 OWL 的基础上专门为描述 Web 服务而产生的描述语言^[17],它允许在用户查找服务或服务自动发现的过程中结合服务本体,根据服务内容描述而非特定关键字来明确满足条件的服务集,以提高查全率和查准率。根据 OWL-S 服务描述语言的定义,Web 服务的描述信息主要包含 3 部分: ServiceProfile 描述服务所能提供的一般性、功能性及辅助性描述信息; ServiceModel 描述服务的操作细节信息,包括原子过程、简单过程和组合过程; ServiceGrounding 描述服务的具体实现规范信息,包括访问服务的传输协议、传送信息的格式等内容。

由于本文针对地理信息服务的语义进行分析,因此只考虑 ServiceProfile 的内容。ServiceProfile 关联类与属性的相关信息如图 1 所示。



图 1 ServiceProfile 的关联类和属性^[17]
Fig. 1 Selected classes and properties of the ServiceProfile

构建形式化本体需要采用一组结构化的属性项表达本体概念的语义。服务概念包含众多属性,既有一般描述属性,如服务提供者、服务名、服务文本描述等与具体功能无关的属性;也有功能性属性,如服务类型、服务功能性描述、输入参数、输出参数等属性。根据形式化本体的特点,要求在保证语义完整性的前提下剔除多余的属性,保留概念的本体属性^[3]。确定概念语义本体属性的方法是提取该概念所有实例对象都应具有的属性项,且其属性值不因对象状态的变化而改变,始终保持为常量。本文采用文献^[3]的定义形式描述概念的语义信息。

定义 2 $S_c = (c, A)$,其中 c 为本体概念, A 为概念的本体属性集合。

1.4 服务概念语义描述

根据地理信息服务的特点,本文将数字城市公共平台的基本服务分为六大类,包括目录服务、数据发布服务、数据操作服务、专题制图服务、数据分析服务和地理编码服务。由于地理编码是空间数据和非空间数据共享整合的重要手段,地理编码系统在数字城市网格化建设中也越来越重要^[18],本文提取出地理编码的相关内容作为独立的基本服务项。

通过分析公共平台地理信息服务的特点,本文将服务概念的本体属性归纳为 4 个层面:服务类型、用户群、功能性、操作对象。必须指出的是,并不是任意一个服务本体概念都必须包含以上所有的本体属性,而应视特定服务的具体情况而定。

(1)“服务类型”属性关注服务是否产生功效,根据这一特性,将其值域细分为数据服务和功能服务。

(2)“用户群”属性关注公共平台的用户群体,结合徐开明等^[19]的研究,根据用户的性质和对地理信息的主要用途将数字城市用户划分为三大用户群

体:公众用户包括利用地理信息从事经营活动的企业与使用地理信息的城市民众;行业用户包括需要解决与地理信息相关业务的行业;政务用户包括管理、规划城市、发布公开信息以及公共平台管理的政府职能部门。

(3)“功能性”属性主要强调公共平台提供的各类功能。根据地理信息系统的基本功能^[20],结合公共平台的特点,从数据产生到成果发布的过程将功能性属性值域划分为数据供应、数据管理、数据查询、数据处理、数据统计、数据分析和地图制作。

(4)“操作对象”属性关注调用地理服务所需使用的数据。其中,地理信息系统的所有功能都是围绕着数据进行的,而服务的操作对象为服务信息。因此该属性值域可划分为地理数据、服务信息和地理对象描述信息,其中地理数据又进一步划分出地理编码数据。

2 实例分析

本文以公共平台中的地理编码服务为例,根据归纳的本体属性将服务本体形式化描述如下:

$O(\text{地理编码服务}) = (\{\text{地理编码服务, 地理编码发布, 地理编码管理, 地理编码查询, 地理编码逆查询}\}, \{\text{(地理编码服务, \{数据服务 \cup 功能服务/服务$

类型, 所有用户/用户性, 数据供应 \cup 数据管理 \cup 数据查询/功能性}\}, (\text{地理编码发布, \{数据服务/服务类型, 政务用户/用户性, 数据供应/功能性, 地理编码数据/操作对象}\}, (\text{地理编码管理, \{数据服务/服务类型, 政务用户/用户性, 数据管理/功能性, 地理编码数据/操作对象}\}, (\text{地理编码查询, \{功能服务/服务类型, 所有用户/用户性, 数据查询/功能性, 地理编码数据/操作对象}\}, (\text{地理编码逆查询, \{功能服务/服务类型, 所有用户/用户性, 数据查询/功能性, 地理对象描述信息/操作对象}\}), \{\text{“part-of”, “kind-of”}\}, \{\text{(“part-of”, (地理编码服务, 地理编码发布), (地理编码服务, 地理编码管理), (地理编码服务, 地理编码查询), (地理编码服务, 地理编码逆查询))}, (\text{“kind-of”, (O(数据发布服务), 政务数据服务, 地理编码发布), O(数据发布服务), 地理数据管理, 地理编码管理), (O(数据操作服务), 查询检索服务, 地理编码查询), (O(数据操作服务), 查询检索服务, 地理编码逆查询))}\})).

表 1 总结了公共平台七大基本服务域内服务概念的语义描述。此外,本文还使用 Protégé4. 1 软件完成数字城市公共平台地理信息服务本体的构建,并使用 HermiT 1. 3. 4 推理机完成各个服务类别的分类,图 2 描述了公共平台服务关系。

表 1 公共平台服务概念语义
Table 1 Service concepts semantics of sharing platform

基本服务	基本用户群	服务类型	子服务	子用户群	功能性	操作对象	与基本服务关系	与其它服务关系	其它服务
目录服务	行业用户 \cup 政务用户	数据服务	服务管理	政务用户	数据管理	服务信息	kind-of	无	无
		数据服务	服务元数据管理	政务用户	数据管理	服务元数据	kind-of	无	无
		功能服务	服务检索	行业用户 \cup 政务用户	数据查询	服务信息	part-of	kind-of	查询检索服务
		数据服务	地理元数据管理	政务用户	数据管理	地理元数据	kind-of	无	无
		数据服务	地理数据管理	政务用户	数据管理	地理数据	kind-of	无	无
数据发布服务	所有用户	功能服务	地理数据查询	行业用户 \cup 政务用户	数据查询	地理数据	part-of	kind-of	查询检索服务
		数据服务	公众地图服务	公众用户	数据供应	地理数据	kind-of	无	无
数据操作服务	所有用户	数据服务	专题数据服务	行业用户	数据供应	地理数据	kind-of	无	无
		数据服务	政务数据服务	政务用户	数据供应	地理数据	kind-of	无	无
		功能服务	查询检索服务	所有用户	数据查询	数据信息	kind-of	无	无
		功能服务	地图基本功能服务	公众用户	数据处理	地图数据	kind-of	无	无
专题制图服务	所有用户	功能服务	高级地图功能服务	行业用户 \cup 政务用户	数据处理	地理数据	kind-of	无	无
		功能服务	个性化地图服务	公众用户	地图制作	地理数据	kind-of	无	无
		功能服务	行业专题图制作	行业用户	地图制作	地理数据	kind-of	无	无
数据分析服务	所有用户	功能服务	城市专题图制作	政务用户	地图制作	地理数据	kind-of	无	无
		功能服务	公共空间分析	公众用户	数据分析	地理数据	kind-of	无	无
		功能服务	高级空间分析	行业用户 \cup 政务用户	数据分析	地理数据	kind-of	无	无
地理编码服务	所有用户	功能服务	专题统计服务	行业用户 \cup 政务用户	数据统计	地理数据	kind-of	无	无
		数据服务	地理编码发布	政务用户	数据供应	地理编码数据	part-of	kind-of	政务数据服务
地理编码服务	所有用户	数据服务	地理编码管理	政务用户	数据管理	地理编码数据	part-of	kind-of	地理数据管理
		功能服务	地理编码查询	所有用户	数据查询	地理编码数据	part-of	kind-of	查询检索服务
		功能服务	地理编码逆查询	所有用户	数据查询	地理对象描述信息	part-of	kind-of	查询检索服务



图 2 地理信息服务本体概念关系

Fig. 2 Concepts relationship of geographic information service ontology

3 结论

随着地理信息系统技术、网络技术以及计算机技术的不断发展,地理信息在城市建设与大众生活中的作用显得越来越重要。由于缺乏统一标准规范,造成地理数据格式各异,成为当前制约社会化的数据共享及交换的瓶颈。通过本文的分析可以看出,明确且形式化地描述概念语义是实现数据共享及功能互操作的关键。本文引入本体论,利用属性枚举的方法提取地理信息服务的本体属性来描述概念语义,进而构建地理信息服务本体,以解决服务集成与语义异质的问题。进一步的工作是利用地理信息服务本体进行本体匹配,解决基于形式化本体的服务检索以及服务自动组合的问题。

参考文献:

[1] HAKIMPOUR F, TIMPF S. Using ontologies for resolution of semantic heterogeneity in GIS[C]. Citeseer, 2001.

[2] BIANCHINI D, DE ANTONELLIS V. Ontology-based semantic interoperability tools for service dynamic discovery[J]. Interoperability of Enterprise Software and Applications, 2006, 323-333.

[3] 李霖, 朱海红, 王红, 等. 基于形式本体的基础地理信息语义分析——以陆地水系要素类为例[J]. 测绘学报, 2008(2): 230-235.

[4] 安杨, 边馥苓, 关佳红. 基于 Ontology 的网络地理服务描述与发现[J]. 武汉大学学报(信息科学版), 2004(12): 1063-1066.

[5] 郑亮, 李德仁. 基于空间服务语义模式的地理信息服务发现[J]. 测绘科学, 2011, 36(2): 127-129.

[6] STUDER R. Knowledge engineering: Principles and methods[J]. Data & Knowledge Engineering, 1998, 25(1-2): 161-197.

[7] BUCCELLA A, CECHICH A, FILLOTTRANI P. Ontology-driven geographic information integration: A survey of current approaches[J]. Computers & Geosciences, 2009, 35(4Sp. Iss. SI): 710-723.

[8] 黄茂军. 地理本体的形式化表达机制及其在地图服务中的应用研究[D]. 武汉大学, 2005.

[9] WAY E C. Metaphor as a mechanism for reorganizing the type hierarchy[J]. Knowledge-Based Systems, 1992, 5(3): 223-232.

[10] GOLFARELLI M, PROLI A, RIZZI S. M-FIRE: A metaphor-based framework for information representation and exploration[A]. WEBIST 2006: Proceedings of the Second International Conference on Web Information Systems and Technologies-Internet Technology/Web Interface and Applications[C]. 2006. 332-340.

[11] NEUMAN Y, NAVE O. Metaphor-based meaning excavation[J]. Information Sciences, 2009, 179(16): 2719-2728.

[12] CENGARLE M V. Semantic Typing for Parametric Algebraic Specifications[M]. Berlin 33: Springer-Verlag Berlin, 1995. 261-276.

[13] KUHN W. An Algebraic Interpretation of Semantic Networks[M]. Berlin: Springer-Verlag Berlin, 1999. 331-347.

[14] HOSAIN S, JAMIL H. An algebraic language for semantic data integration on the hidden Web[A]. 2009 IEEE Third International Conference on Semantic Computing (ICSC 2009)[C]. 2009. 237-244.

[15] VEALE T, O'DONOGHUE D. Computation and blending (computational feasibility of the conceptual integration theory of Fauconnier and Turner)[J]. Cognitive Linguistics, 2000, 11(3-4): 253-281.

[16] 陆建江, 张亚非, 苗壮, 等. 语义网原理与技术[M]. 北京: 科学出版社, 2007.

[17] MARTIN D, BURSTEIN M, HOBBS J, et al. OWL-S: Semantic Markup for Web Services. <http://www.w3.org/Submission/2004/SUBM-OWL-S-20041122/>. 2004-11-22.

[18] 江洲, 李琦. 地理编码(Geocoding)的应用研究[J]. 地理与地理信息科学, 2003, 19(3): 4-7.

[19] 徐开明, 吴华意. 地理信息公共服务平台的用户分类及服务分类[J]. 地理信息世界, 2009(3): 12-16.

[20] 胡鹏, 黄杏元, 华一新. 地理信息系统教程[M]. 武汉: 武汉大学出版社, 2005.

(下转第 19 页)

参考文献:

- [1] 陆家驹,李士鸿. TM 资料水体识别技术的改进[J]. 环境遥感, 1992,7(1):17-23.
- [2] 黄海波,赵萍,陈志英,等. ASTER 遥感影像水体信息提取方法研究[J]. 遥感技术与应用,2008,23(5):525-528.
- [3] 李小曼,王刚,田杰. TM 影像中水体提取方法研究[J]. 西南农业大学学报,2006,28(4):580-582.
- [4] 丁莉东,吴昊,王长健,等. 基于谱间关系的 MODIS 遥感影像水体提取研究[J]. 测绘与空间地理信息,2006,29(6):25-27.
- [5] 韩栋,杨晓梅,纪凯. 小卫星遥感影像自动提取水体方法研究[J]. 测绘科学,2008,33(1):51-54.
- [6] 徐涵秋. 利用改进的归一化差异水体指数(MNDWI)提取水体信息的研究[J]. 遥感学报,2005,9(5):589-595.
- [7] 王净,李亚春,景元书. 基于 MODIS 数据的水体识别指数方法的比较研究[J]. 气象科学,2009,29(3):342-347.
- [8] 权维俊,郭文利,叶彩华. 基于 TM 卫星影像获取北京市水体密度指数与植被覆盖指数的方法[J]. 南京气象学院学报,2007,30(5):610-616.
- [9] MCFEETERS S K. The use of Normalized Difference Water Index(NDWI) in the delineation of open water features[J]. International Journal of Remote Sensing, 1996, 17(7):1425-1432.
- [10] 莫伟华,孙涵. 基于 EOS/MODIS 卫星数据的洪涝灾害遥感监测应用技术研究[D]. 南京信息工程大学,2006.
- [11] 梅安新,彭望球,秦其明,等. 遥感导论[M]. 北京:高等教育出版社,2001. 98-100.
- [12] 孙家柄. 遥感原理与应用[M]. 武汉:武汉大学出版社,2003. 142-145.
- [13] 王桥,张兵,韦玉春,等. 太湖水体环境遥感监测实验及其软件实现[M]. 北京:科学出版社,2008. 66-67.

Comparison and Research on Different Indices for Water Extraction Based on CCD Images from HJ Satellite

FAN Deng-ke¹, LI Ming¹, HE Shao-shuai²

(1. School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079;

2. Beijing Landview Mapping Information Technology Co. Ltd, Beijing 100096, China)

Abstract: Based on CCD images captured from HJ satellite, this paper presents a comparison and research on capabilities of different indices in extracting inland lakes' water, relating to the requirement of water environmental monitoring by remote sensing technique for East Lakes in Hubei Province. Referring to optimal bands used for water extraction from typical sensors, such as TM and MODIS, images from HJ satellite were utilized as data source to analyze and assess its availability on water extraction. Radiation in reflectance from all four bands was composed to evaluate the notable variation trend. Simultaneously, three water identification indices with their corresponding threshold were investigated to retrieve feature scenes and find the best solution. Experiments show that CIWI(Combined Index of Normalized Difference Water Index) can get the best image inspection, the highest discrimination and the lowest misidentification ratio, which could cope with fast water identification issues.

Key words: HJ satellite; inland lakes; water extraction index

(上接第 8 页)

Semantic Analysis of Geographic Information Service under the Digital City

TAN Yong-bin¹, ZHU Hai-hong¹, LI Lin¹, MAO Kai¹, LI Cheng-ming², YIN Jie²

(1. School of Resource and Environment Science, Wuhan University, Wuhan 430079;

2. Chinese Academy of Surveying & Mapping, Beijing 100830, China)

Abstract: Construction of Digital City service sharing platform can solve the problem of heterogeneous and isomerism geographic data effectively in order to achieve the interoperability of geographic information system. By exempling of geographic information services within the sharing platform, service semantic descriptions were presented explicitly and formally by employing the properties enumeration description method based on ontology theory, and then try to overcome geographic information services semantic sharing barrier. Four ontology properties of geographic information services were extracted, including service-type, service-user-group, functionality and operated-object. Based on these properties, service taxonomy was achieved, which included six categories: Catalogue Service, Data Publishing Service, Data Manipulation Service, Thematic Mapping Service, Data Analysis Service and Geocoding Service. Finally, the geographic service ontology of sharing platform was constructed in Protégé and reasoned by using Hermit 1.3.4, and an example of the formal description of Geocode Service was presented based on ontology properties.

Key words: digital city; geographic information service; semantic; service ontology