

DOI :10.3876/j.issn.1000-1980.2009.05.015

数字流域环境下信息共享服务研究

许国艳,李晓芳,王志坚

(河海大学计算机及信息工程学院,江苏 南京 210098)

摘要:为了提高数字流域环境下信息共享的能力,基于 SOA 体系结构,引入发布/订阅技术的思想,提出了一种数字流域环境下信息共享服务框架.该框架在面向服务的平台中提供了:事件驱动支持,满足了平台的松散耦合通信与协同需求;一种松散耦合的信息共享服务方式,达到了时间解耦、空间解耦和同步解耦,增强了平台的可复用性和可扩展性;一种主动的协同信息共享服务方式,更好地满足了信息共享的灵活性、主动性.具体探讨了服务生成和服务订阅模型和流程,通过领域本体的引进增强了服务描述和订阅描述的能力,解决了语义的问题.

关键词:数字流域;信息共享服务;发布/订阅;面向服务的体系结构;领域本体

中图分类号:TP391 **文献标识码:**A **文章编号:**1000-1980(2009)05-0573-05

数字流域就是从整个流域的全局出发,用数字化的手段刻画整个流域,综合处理流域的空间、地理、气象、水文和历史信息,为我国解决江河水利与水害问题提供高科技手段^[1-3].

在数字流域环境下,水资源信息作为国家重要的基础性公共信息资源,应该为流域内用户所共享^[4].由于传统的软件系统大多在封闭的网络环境下运行,相互之间缺乏资源共享,形成信息“孤岛”,从而造成系统重复开发,导致信息和资源的浪费.如黄河水利委员会,防洪减灾系统、水资源保护系统、水量调度系统、办公自动化系统等都是独立开发,自成体系,能很好地解决单一的专业问题,但较少考虑其他系统的需求和整个流域应用的集成.所以,在数字流域建设中,如何将分布在流域各个网络结点上的独立信息资源集成为一个整体进行资源共享,已成为数字流域建设的关键.

面向服务的体系结构(service-oriented architecture,SOA)、领域本体技术和发布/订阅技术的发展,为信息共享提供了共享策略和方式上的新思路.本文将发布/订阅技术应用于 SOA 中,结合领域本体的语义能力,解决数字流域信息共享问题.

1 相关技术

1.1 SOA

SOA^[5]来源于早期的基于构件的分布式计算方式.服务是将构件描述成提供相关服务的物理黑盒封装的可执行代码单元,通过一致的已发布的接口进行访问.接口采用中立的方式进行定义,独立于实现服务的硬件平台、操作系统和编程语言,这使得构建在各种系统中的服务可以以一种统一和通用的方式进行交互.在 OMG 和 IONA 的推动下,目前 SOA 已成为一个广泛认可的规范.SOA 体系结构如图 1 所示.

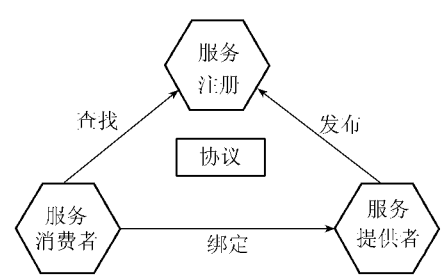


图 1 SOA 体系结构

Fig. 1 Service-oriented architecture

1.2 领域本体

领域本体(domain ontology)是专业性的本体,提供了某个专业学科领域的概念以及概念间的关系.目前,本体建模的研究已经进入了实际应用阶段,许多研究领域都建立了自己标准的本体.领域本体构建的基本方式有 3 种:手工构建领域本体,复用已有的领域本体和半自动、自动构

收稿日期:2009-04-08

基金项目:国家自然科学基金(60573098) 河海大学自然科学基金(2008430011)

作者简介:许国艳(1971—),女,内蒙古赤峰人,副教授,硕士,主要从事软件构件技术及数据集成研究.

建领域本体. 领域本体的构建方法都是在具体领域本体建立过程中总结出来的, 具有特定的领域特色, 其成熟度由高到底依次是: 七步法、MATHONTOLOGY 法、IDEF5 法、TOVE 法、骨架法、SENSUS 法和 KACTUS 法^[6-7].

1.3 发布/订阅系统

发布/订阅系统^[8-11]是一种使分布式系统中的各参与者, 能以发布/订阅的方式进行交互的中间件系统, 如图 2 所示. 在发布/订阅系统中, 信息的生产者和消费者之间所交互的信息被称为事件. 生产者将事件发送给发布/订阅中间件. 消费者则向发布/订阅中间件发出一个订阅条件, 表示对系统中的哪些事件感兴趣, 如果不再感兴趣, 也可以取消订阅. 发布/订阅中间件保证将生产者发布的事件及时、可靠地传送给所有对之感兴趣的消费者. 信息的生产者称为发布者 (publisher), 信息的消费者称为订阅者 (subscriber).

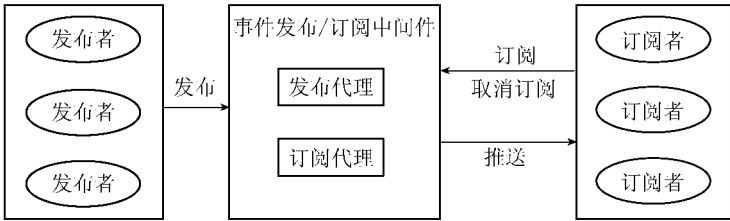


图 2 发布/订阅系统
Fig. 2 Publishing/subscription system

2 数字流域环境下信息共享服务机制

数字流域环境下信息资源共享^[12]的关键是机制. 为了能方便地共享分布、异构的数字流域环境下的信息资源, 又能保证数据信息的安全, 采用以服务的形式共享信息资源, 以订阅的方式主动推送信息资源的方式. 在这种信息共享服务机制下, 在具体的数字流域建设中, 信息共享服务的方式主要有:

- a. 用户访问. 用户通过调用服务来实现相应的信息访问, 信息的提供是被动的.
- b. 用户下载. 用户通过调用服务来实现相应的数据信息下载, 信息的下载是被动的.
- c. 用户订阅. 用户根据需求和信息共享的权限协定, 提前订阅服务, 并提交服务执行的事件触发条件. 一旦触发条件满足, 服务自动执行, 信息主动提供给用户. 触发条件可以是用户登录, 可以是时间的约定, 可以是某个参数达到阈值, 可以是数据信息有更新等.

信息资源共享采用一站式登录认证, 由资源共享服务门户提供入口, 统一建立信息共享服务认证体系和机制. 用户必须申请注册, 注册过的用户享受调用、下载、订阅服务. 访问权分系统和用户, 系统必须是数字流域有关部门或单位建设的系统, 不同的系统赋予不同的访问权或标识. 用户必须是使用相关系统的有权用户.

信息资源按照权威部门权威信息的原则, 由相应的信息中心和分中心实施管理, 及时更新信息资源. 服务的管理主要包括服务提供者角度的管理、服务使用者角度的管理和服务注册机构角度的管理.

3 数字流域环境下信息共享服务框架

随着 SOA 及相关技术的发展, 基于 SOA 的信息共享平台日益成熟, 能很好地为分布、异构的系统提供统一的信息访问接口. 发布/订阅技术是一种松散耦合的异步消息交换方式, 将发布/订阅技术应用于信息提供者和信息使用者之间共享信息的方式, 能更好地满足信息共享的需求. 为了提高信息服务的表达能力, 为用户提供更准确的信息, 可采用领域本体作为信息的描述. 基于 SOA、采用领域本体描述信息、增加发布/订阅信息通信方式, 能为数字流域环境下信息资源共享提供更松散、灵活、主动、准确的服务.

基于此, 本文提出基于 SOA、采用领域本体描述信息、增加发布/订阅信息通信方式的数字流域环境下信息共享服务框架. 框架主要包括数据源层、领域本体、Web 服务层、服务订阅层和信息共享服务门户, 具体如图 3 所示.

数据源是数字流域的存储数据信息, 数据源的类型包括业务数据库和空间数据库. 数据源的存储可以采用数据中心和数据分中心两级的数据信息存储体系. 数据中心存储基础数据库, 数据分中心存储水资源信

息、水文信息、基础地理数据等专业数据信息。

领域本体作为数据源的元数据,包括领域概念、概念关系和规则,也是完成信息共享的统一视图。通过研究领域本体的建立方法,建立流域本体库,作为流域信息资源的元数据,并通过语义标注解决信息资源共享的语义冲突问题。流域本体是流域基础数据库和防汛数据库的元数据,主要包括:防洪工程基础数据本体、基础地理数据本体、遥感影像数据本体、实时工情险情数据本体、气象信息数据本体、实时水雨情数据本体、流域水文数据本体、流域河道断面数据本体、流域社会经济数据本体和流域水资源保护数据本体等。

Web 服务层包括模型服务、空间信息服务、业务信息服务和 UDDI 注册中心。模型服务、空间信息服务和业务信息服务封装成 Web 服务规范,发布在 UDDI 注册中心。用户可以通过信息共享服务门户调用和下载 Web 服务,也可以通过服务订阅层订阅 Web 服务,满足条件的 Web 服务将主动推送给用户。

服务订阅层包括订阅管理、订阅匹配和订阅推送,主要完成用户对 Web 服务的订阅和向用户推送 Web 服务 2 个功能。

信息共享服务门户为用户提供统一的用户接口,完成服务用户的管理、服务目录的管理、服务的调用、服务的下载、服务的发布和服务的订阅。

4 具体实现

在数字流域信息共享中,存在的问题包括数据结构异构、语义异构以及不同场景下用户需求信息方式的不同。本文提出的数字流域环境下信息共享服务框架,利用 SOA、领域本体和发布/订阅技术 3 方面的优势,来解决数字流域环境下信息共享存在的问题。信息共享时的数据结构异构问题通过 Web 服务封装处理解决,语义异构问题通过流域本体建立以及标注解决,共享方式问题通过服务订阅技术解决。共享的信息可以按照用户的需求主动提供给用户。

4.1 流域本体构建

结合流域特点,由领域工程方建立适用于流域信息资源共享、可操作的本体库。将流域本体用于 Web 服务的描述,在现有的 Web 服务标准(WSDL 和 UDDI)上增加语义信息,采用流域本体直接对 WSDL 文件进行标注的方法。流域本体构建分为流域概念的获取、流域概念的组织、流域类模型的构建、流域本体的评价和流域本体的修剪 5 个主要过程。

流域本体编制主要根据国家、水利部颁布的有关技术标准进行编制。流域概念的获取主要包括术语选择、抽取本体概念。流域概念主要从流域词典、数据库表结构及数据、现有规划及系统获取。流域概念的组织是本体构建中的主要环节,主要是构建本体类层次以及为每一个类别填充实例。在完成流域概念的获取、流域概念组织的基础上,主要完成流域本体构建。最后,进行流域本体的评价和流域本体的修剪。

4.2 服务生成与发布

服务生成根据服务工厂模式设计思想,生成用户实际需要的 Web 服务。服务工厂通过抽取实际开发 Web 服务的模式,形成一般意义上的开发模式。服务工厂的设计基于 Web 服务的开发流程,服务生成主要包括 3 个子功能,即服务构件接口描述、服务构件行为描述和服务封装。服务构件接口描述用于描述服务构件接口的语法信息,服务构件的拥有者在特定的领域本体库的范畴内描述服务构件的接口信息。服务构件行为描述用于描述服务构件的内部行为,同样,服务构件拥有者在特定的领域本体库的范畴内描述服务构件的内部行为信息。最后,把服务构件封装成符合 Web Service 规范的服务构件。具体服务生成流程描述如下。

输入 服务名称,查询条件,对应本体信息集;

输出 组装完整的查询语句,以 Web 服务的形式发布在 Web 服务器上。

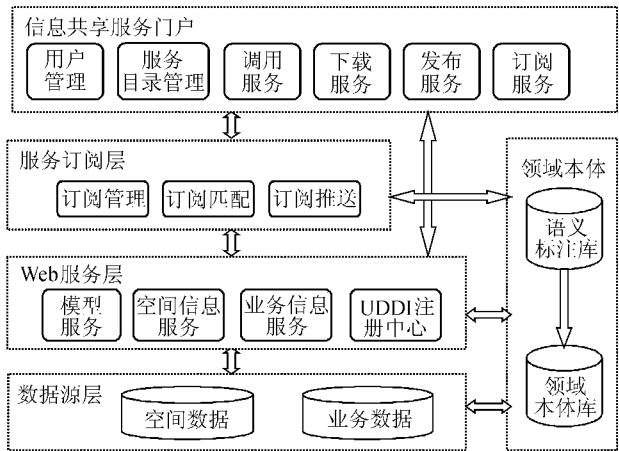


图 3 数字流域信息共享服务框架

Fig. 3 Framework of information sharing services under digital basin environment

算法过程 获取类服务模板 获取方法模板 利用方法模板将待替换变量循环替换为用户提供的服务定制条件 生成临时类便于循环放置服务中的方法 ;Web 服务类生成 ;Web 服务发布在 UDDI 注册中心.

4.3 服务订阅

流域信息共享采用 Web 服务方式实现.在此基础上,为了实现信息服务的定制和推送,满足信息的主动提供,在 Web 服务中结合“发布/订阅”技术,更好地支持了信息的主动共享,可以实现时间、空间和控制流 3 个方面的解耦.

在数字流域环境信息共享服务框架下:(a)发布/订阅中间件由发布代理和订阅代理 2 部分组成.在 Web 服务中实现发布/订阅的通信机制,发布代理由注册中心完成,事件的发布即是 Web 服务的发布.订阅代理主要完成订阅匹配、订阅的推送和订阅管理.(b)采用基于主题和基于内容相结合的订阅模型.具体服务订阅模型如图 4 所示.基于主题的订阅可以完成对某一主题的 Web 服务的订阅,并将主题组织成一种层次结构.订阅者在订阅上层主题时,即涵盖了下层的各主题.基于内容的订阅是由订阅者声明的查询和谓词决定的,事件的内容为多个“属性=值”的集合.为了增强主题和事件内容的表达能力,增强语义信息和一定的语义推理能力,采用领域本体描述 Web 服务的接口信息和订阅信息,具体采用 OWL 语言进行描述,使用 Jena 推理机进行推理.

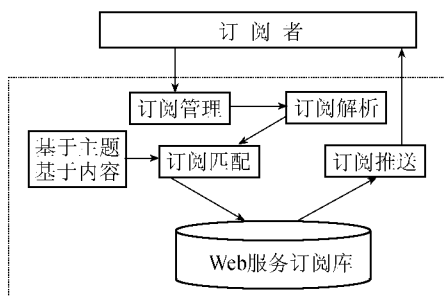


图 4 服务订阅模型

Fig. 4 Service subscription model

订阅管理负责对订阅者需求的订阅进行处理,包括对订阅信息注册、注销、查询等管理;订阅解析在描述模型的支持下,对订阅进行分解,形成平台统一描述;订阅匹配基于主题,或基于内容,或基于主题和内容的结合,完成 Web 服务的过滤和 Web 服务组合检测功能,形成 Web 服务订阅库;订阅推送完成事件的通知,将订阅的 Web 服务信息推送给订阅者.

5 小 结

本文针对数字流域环境下信息共享的需求,基于 SOA、引入发布/订阅技术的思想,提出了一种数字流域环境下信息共享服务框架,分析探讨了服务生成和服务订阅的模型和流程,通过领域本体的引进,增强了服务描述和订阅描述的能力,解决了语义的问题.本文提供的一种松散耦合的信息共享服务方式,可达到时间非耦合、空间非耦合和同步非耦合,增强平台的可复用性和可扩展性.提供的一种主动的协同信息共享服务方式,可更好地满足信息共享的灵活性、主动性.

参考文献:

- [1] 张勇传,王乘.数字流域 数字地球的一个重要区域层次[J].水电能源科学,2001,19(3):1-3. (ZHANG Yong-chuan, WANG cheng. Digital valley-an important regional level of digital earth[J]. Hydroelectric Energy, 2001,19(3):1-3. (in Chinese))
- [2] 刘家宏,王光谦,王开.数字流域研究综述[J].水利学报,2006,37(2):240-246. (LIU Jia-hong, WANG Guang-qian, WANG Kai. Review on advancement of study on digital riverbasin in China [J]. Journal of Hydraulic Engineering, 2006,37(2):240-246. (in Chinese))
- [3] 张秋文,张勇传,王乘,等.数字流域整体构架及实现策略[J].水电能源科学,2001,19(3):4-7. (ZHANG Qiu-wen, ZHANG Yong-chuan, WANG Cheng, et al. Framework and technology of digital valley [J]. Hydroelectric Energy, 2001,19(3):4-7. (in Chinese))
- [4] 丁悦,毛文跃,刘枫,等.“数字流域”环境下信息资源的共享途径[J].广东水利水电,2008(3):88-90. (DING Yue, MAO Wen-yue, LIU Feng, et al. Information resources sharing ways in digital basin [J]. Guangdong Water Resources and Hydropower, 2008(3):88-90. (in Chinese))
- [5] CAREY M J, BLEVINS M, TAKACSI-NAGY P. Integration, web services style [J]. IEEE Data Engineering Bulletin, 2002,25(4):17-21.
- [6] 韩婕,向阳.本体构建研究综述[J].计算机应用与软件,2007,24(9):21-23. (HAN Jie, XIANG Yang. A survey on ontology building [J]. Computer Applications and Software, 2007,24(9):21-23. (in Chinese))
- [7] 李恒杰,李军权,李明.领域本体建模方法研究[J].计算机工程与设计,2008,29(2):381-384. (LI Heng-jie, LI Jun-quan, LI Ming. Research on modeling method of domain ontology [J]. Computer Engineering and Design, 2008,29(2):381-384. (in Chinese))

- [8] 马建刚,黄涛,汪锦岭,等.面向大规模分布式计算发布订阅系统核心技术[J].软件学报,2006,17(1):134-147. (MA Jian-gang,HUANG Tao,WANG Jin-ling,et al.Underlying techniques for large-scale distributed computing oriented publish/subscribe system [J]. Journal of Software,2006,17(1):134-147. (in Chinese))
- [9] 苑洪亮,尹刚,王怀民.支持发布订阅的 Web 服务应用集成框架[J].计算机与数字工程,2008,36(3):81-83. (YUAN Hong-liang,YIN Gang,WANG Huai-min. A web service application integration framework supporting publish/subscribe [J]. Computer & Digital Engineering,2008,36(3):81-83. (in Chinese))
- [10] 刘家红,吴泉源.一个基于事件驱动的面向服务计算平台[J].计算机学报,2008,31(4):588-599. (LIU Jia-hong,WU Quan-yuan. An event-driven service-oriented computing platform[J]. Chinese Journal of Computers, 2008,31(4):588-599. (in Chinese))
- [11] 董飏,陈金辉,孙亚民.大规模发布/订阅系统中的服务质量[J].计算机科学,2008,35(8):284-286. (DONG Biao,CHEN Jin-hui,SUN Ya-min. Quality of service in large scale publish/subscribe systems [J]. Computer Science, 2008,35(8):284-286. (in Chinese))
- [12] 周晓峰,王志坚.“数字流域”中数据共享机制研究与实现[J].计算机工程与设计,2003,24(9):8-10. (ZHOU Xiao-feng,WANG Zhi-jian. Research & realization on data shared of “digital basin” [J]. Computer Engineering and Design,2003,24(9):8-10. (in Chinese))

Information sharing services under digital basin environment

XU Guo-yan, LI Xiao-fang, WANG Zhi-jian

(College of Computer & Information Engineering, Hohai University, Nanjing 210098, China)

Abstract : In order to improve the information sharing under the digital basin environment, a framework of information sharing services under the digital basin environment was put forward based on the SOA architecture and the publishing/subscription technique. It provided an event-driven communication pattern in the service-oriented platform, and the loose coupling demand of communication and coordination was satisfied. A loose coupling of the sharing information service mode was provided. As a result, the time decoupling, space decoupling and synchronous decoupling were reached. The reusability and expansibility of the platform were enhanced. A kind of active coordination information sharing service mode was provided so that the flexibility and initiative of the information sharing were satisfactorily achieved. The service production and service subscription were discussed. By introducing the domain ontology, the description ability of services and subscription was enhanced. Accordingly, the semantic problems were solved.

Key words : digital basin ;information sharing service ;publishing/subscription ;service-oriented architecture ;domain ontology