

技术方法

面向地理信息服务链的工作流技术应用研究

杨仕勇

(山东省国土资源信息中心,山东 济南 250014)

摘要:通过分析借鉴当今服务链技术与工作流技术研究成果,充分考虑地理信息数据量大、数据源异构性的特点,对适于地理信息服务链的过程建模方法、过程定义语言、建模工具以及工作流引擎的体系结构、调度算法、应用框架进行重新设计,实现了地理信息服务的可视化组合建模和服务链的自动执行。

关键词:地理信息服务链;Web服务;工作流;工作流建模;工作流引擎

中图分类号:P208

文献标识码:B

0 引言

随着Web服务在地学领域的成功运用,地理信息系统应用进入了地理信息服务发展阶段。为组合松散耦合的地理信息服务,实现服务过程增值并推动地理信息服务的产业化和社会化进程,服务链技术应运而生。

目前服务链的研究,主要集中于服务的描述与发现层面,对适于地理信息服务链的工作流技术研究较少;在将工作流技术引入地理信息领域的实践中,多数系统缺乏对地理信息服务的支持;同时,其过程描述语言,多为采用面向服务的商业或国际标准,引擎运用商业流程的表单文档处理,对多源和海量的地理空间数据不能很好的支持。

Web服务技术满足互联网的智能化服务对分布式开放性和平台无关性的要求^[1]。工作流技术实现了应用逻辑和过程逻辑的分离,在过程逻辑建立中可不考虑应用和资源的异构性,但没有解决分布式异构环境中资源的互操作问题^[2],而Web服务恰好解决了对分布式异构资源的互操作能力,两者的结合提供了逻辑和应用两个层面的互操作能力。

该文将借鉴现有研究成果,重新设计地理信息服务链的描述语言与工作流引擎,实现服务的可视化组合和服务链的自动执行。

1 地理信息服务链过程建模研究

工作流过程模型定义工作流的过程逻辑,包括过程建模方法和过程定义语言。

1.1 过程建模方法

过程建模方法是对业务逻辑的直观描述,当前还没有统一的平台。地理信息服务链的原材料为分布在网上的地理信息服务,面向的用户群一般为普通用户,需求也较为简单。该文以简便、图形化为原则,采用基于活动网络的建模方法探讨地理信息服务链的过程建模。

活动网络模型的基本结构是由节点和边构成一个有向图。为了丰富模型的语义表达,将节点扩充为开始节点、结束节点和普通节点;把边扩展为控制边、数据边和循环边。

开始节点是工作流过程模型的逻辑起点。一个模型只允许一个开始节点,不参与活动的执行,只表达流程的开始状态,并作为流程的入口激活后继节点;结束节点与开始节点相呼应,是模型的逻辑结束点。当工作流运行至结束节点时,流程结束;普通节点代表实际的地理信息服务,一个模型中至少有一个普通服务。

控制边体现过程的控制逻辑、节点间的时序关系。它的转移代表节点状态的转移和整个过程的演

* 收稿日期:2010-04-29;修订日期:2010-09-26;编辑:程光锁

作者简介:杨仕勇(1973—),男,山东临朐人,工程师,主要从事土地管理及测绘工作;E-mail:yyy_ysy@126.com。

进;数据边的引入是为解决控制流与数据流不一致时,辅助控制边建模元素准确提供节点数据;有循环和条件分支时,使用循环边建模。

1.2 过程定义语言

过程定义语言是过程可被计算机理解的形式化描述。现比较有影响的由微软和IBM共同提出的面向Web服务的工作流过程定义语言WSBPEL,定义了用来描述基于流程及相关方之间互操作的业务流程行为的模型和方法,但其过于复杂,给开发者和用户都带来不便^[3]。在借鉴现有过程定义语言的前提下,该文提出一个描述地理信息服务链的过程定义语言:GI-Service Chaining Process Modeling Language,简称GCPML。

GCPML采用基于XML的方式描述过程,根元素为<processes>,包括一个或者多个<process>元素。一个<process>元素意味着一个服务链过程定义;<process>由<activities>组成,<activities>是活动的容器,包括一个开始活动、一个结束活动和一个或者多个普通活动。每个活动又有一个<transitions>和<parameters>元素组成(图1)。

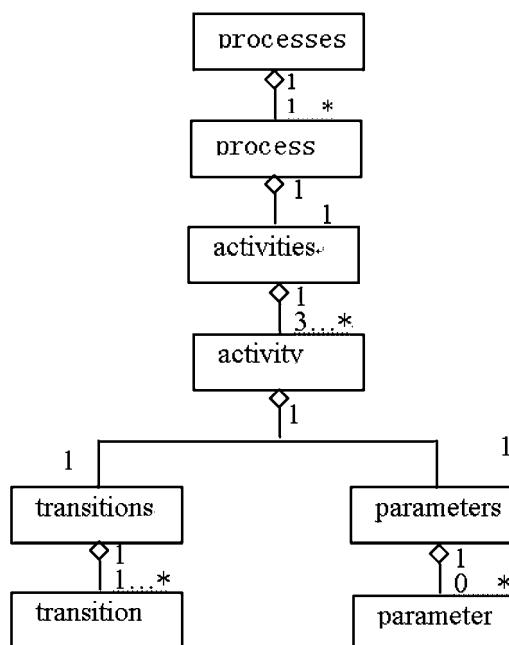


图1 GCPML主体结构图

(1)<process>代表一个服务链流程, name 是它的唯一属性。

(2)<activity>是过程的主要元素,包括:开始

活动、结束活动和普通活动。普通活动即具体的地理信息服务,有name, type, X, Y, wsdl, interface, method, priority等8种常见属性,其中type分别用值0,1,2;分别表示开始活动、结束活动和普通活动。除属性外,activity元素还包括两个子元素:parameters和transitions。

(3)<parameters>只包含一个ownedMethod属性,代表该参数集所属的方法名称,此属性须与activity中的method一致。parameters包含零个或多个parameter子元素。

(4)<parameter>用于存储方法输入参数的值。有两种情况:一是直接输入值或者是参数资源地址;二是前一个活动方法的输出作为参数值,需要绑定相应的活动和方法的名称,用parameter元素的activityBind(活动名称)和path(方法名称)属性来设置。

(5)<transitions>包括一个或者多个transition子元素。

(6)<transition>代表边的转移,实现对控制流转移的形式化建模。包括forward(转移到下一个活动的名称)、type,X,Y等常见属性,其中type分别用0,1,2表示控制边、循环边和数据边。在链接边transition元素中还包括条件子元素<transconditions>,判断边转移的条件。

元素parameter和transition共同决定了数据流的转移。当parameter为空时,意味着transition所连结的两个活动之间没有数据流。

1.3 建模工具设计

建模工具是联系建模方法与过程定义语言的纽带。用户在建模工具上拖拽建模元素构建服务链流程,同时通过相关属性的设置来保存流程描述文件(GCPML文档)。也可以导入GCPML文档,在图形化界面下对文档进行编辑。

1.3.1 主要功能模块

用例图是一个较好的解决手段,包括两个主要功能模块:一是模型绘制。它提供给用户一个可视化过程建模的界面,用户在拖拽操作中完成基于活动网络的地理信息服务链流程建模,产生图形方式表达的服务链模型。基本图元包括:开始节点、结束节点、普通节点以及控制边、数据边和循环边。对各图元元素,通过菜单或双击鼠标来设置属性。对于控制边还可以设置相应转移条件属性,便于动态决

定实际路由的选择。模块支持拷贝、粘贴、删除等编辑功能。模型绘制工具产生服务链流程图的同时，在数据层产生模型相关的数据，既包含节点、链接边的位置信息，又包括相关的属性信息，实现模型的重现和便于工作流引擎提取相关信息。二是 GCPML 语言转换器。它负责将数据层的工作流模型转化成 GCPML 格式的数据流，并将此数据流以 XML 文件形式导出保存。也可以将符合 GCPML 格式的 XML 文件导入建模工具，通过数据和模型图元的映射，实现图形化编辑。

1.3.2 类图设计

MVC (Model – View – Controller) 设计模式适合既有数据层又有强调图形表现层的建模工具的设计开发^[4]。该文采用 MVC 模式探讨相关类模块的设计。将模型分为节点(普通节点、结束节点和开始节点)和链接边两大类；视图类包括菜单、工具栏、绘图画布控件等；控制器类设计为每个模型类设置了一个对应的类型：辅助控制器类、链接边、普通节点、开始节点和结束节点。

2 面向地理信息服务链的工作流引擎设计

本着够用、灵活和低成本的原则^[5]，该文选择轻量级工作流引擎作为研究内容。

2.1 工作流引擎体系结构

地理信息服务链工作流引擎体系结构遵从“SOA 三角形”模式(图 2)。

工作流引擎接收 GCPML 格式的 XML 文件，通过内部的流程解析器将过程的活动、控制信息等存入模型库。流程执行管理器根据模型库中提供的流程数据控制流程的执行状态和服务活动代理的流转秩序。主要有由 4 部分组成：

(1) 过程解析模块。它将 GCPML 文件中的元素映射为内存中的实体类，交由工作流引擎处理，包括过程、活动和转换等。实体类的相关属性通过解析 GCPML 文件元素的属性值得到。

(2) 模型管理模块。模型管理对数据库进行操作，包括将解析后的实体映射为数据库中的模型记录，对数据库表进行维护等。

(3) 活动执行模块。该文采用代理的模型来实现活动与具体地理信息服务的交互。服务链中的活动是地理信息服务在本地的代理。在工作流体系结

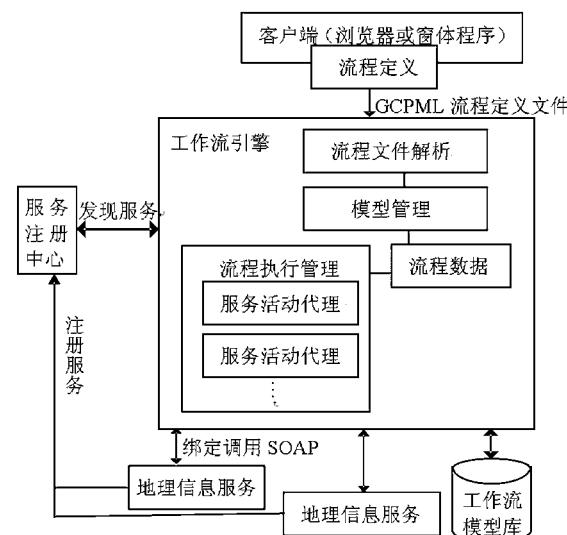


图 2 面向地理信息服务链的工作流引擎体系结构

构中，活动执行模块是其最底层实现。引擎通过服务的 wsdl 地址生成服务代理，并通过反射技术和相关流程数据来激活代理。

(4) 运行控制模块。负责过程的执行、挂起和活动的激活、执行等控制状态的转换。工作流的相关数据和控制数据是其运行控制的依据。

2.2 工作流引擎模型设计

传统工作流引擎模型包括数据模型和控制模型^[6]，数据模型包括组织模型和信息模型。组织模型是对工作流管理系统中涉及的人员、机构等信息的建模，该文不涉及此类信息。

2.2.1 数据模型设计

数据模型主要是对工作流引擎中用到的过程实例、活动实例、方法、应用数据和转换边的数据结构的定义(表 1)。实体的数据结构用关系数据库的二维表来建模。

由于 Web 服务返回的是对象，对地理信息服务的异构性，该文采用“黑箱”方式处理，即不管数据服务的内部结构如何，都作为一个对象序列化到数据库中。当把此数据服务对象作为输入参数时，根据反射技术，转化为相应输入参数的数据类型。

2.2.2 控制模型设计

(1) 过程实例状态控制

过程实例的状态包括：就绪、执行、挂起、完成、异常和结束。当过程解析器对 GCPML 文件完成解

析后,过程就处于就绪状态,得到执行消息,过程实例迁移到执行状态。在执行状态下,用户可以挂起过程,一段时间后,重新回到执行,或者用户在挂起状态下取消过程执行,结束过程;用户也可以在执行状态下直接取消执行,过程结束。在执行中,可能发生异常,过程自动结束。过程结束或完成后释放资源。

表 1 数据模型结构参数

表名	字段名	字段类型	说明
过程实例表	processName	varchar	过程实例的名称
	activityName	varchar	活动名称
	transitionID	int	转换边 ID
	activityPriority	int	活动优先级
	processState	varchar	过程状态
活动实例表	activityName	varchar	活动名称
	wsdlURL	varchar	服务的 wsdl 地址
	interface	varchar	服务接口
	methodName	varchar	接口提高的方法
	activityState	varchar	活动状态
方法表	methodName	varchar	方法名称
	activityBind	varchar	绑定活动的名称
	path	varchar	绑定活动的方法名称
	parameterValue	varchar	参数值
	activityName	varchar	方法所属活动名称
应用数据表	ID	int	应用数据标示
	activityName	varchar	活动名称
	methodName	varchar	方法名称
	returnValue	BLOB	方法返回值
置换边表	ID	int	转换边标示
	startActivity	varchar	源活动
	endActivity	varchar	目标活动
	condition	varchar	转换条件

(2) 活动实例状态控制

活动实例状态包括：就绪、激活、执行、挂起和完成。过程实例状态进入执行时，依据流转规则，相应的活动实例先后被激活和执行。当过程被挂起后，状态为执行的活动实例停止执行，并将状态设为挂起；当过程由挂起恢复执行后，状态为挂起的活动重新设置为激活状态，由引擎调度规则调度执行。活动完成后释放相应的资源。当用户取消过程后，活动实例被销毁。

2.3 工作流引擎调度算法

该文设计的引擎调度算法步骤如下(图3)：

(1) 遍历过程实例表 process, 找到活动优先级 activityPriority 为 0 的开始活动节点及其对应的所有转换边 ID。

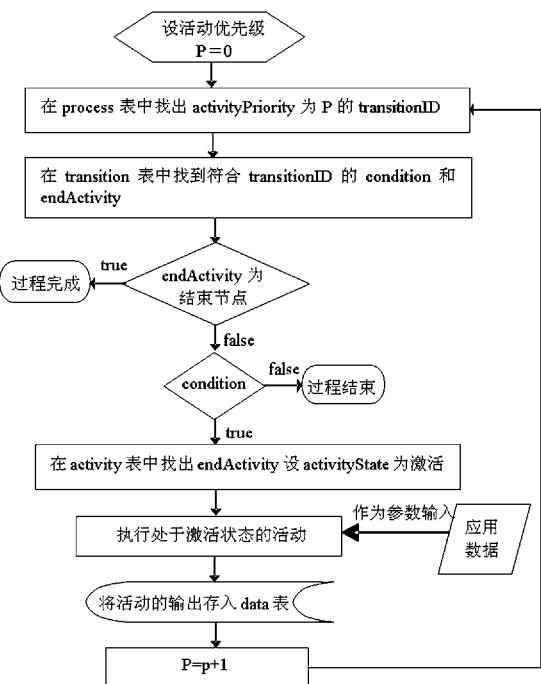


图3 引擎调度算法步骤图

(2) 在转换边表 transition 中根据转换边 ID 找到转换条件 condition 和目标活动名 endActivity。若 endActivity 为结束节点，则过程完成，释放相关资源；若转换条件符合，则于活动实例表 activity 中找到相应目标活动节点，设活动状态 activityState 为激活，直至遍历所有目标活动。若目标活动不是结束节点且没有目标活动被激活，则过程发生异常，过程结束。

(3) 执行处于激活状态的活动(若活动为结束节点,则结束流程),此时设活动状态为执行,在方法表 method 中找到该活动方法输入参数,调用函数并将函数返回的数据存入应用数据表 data 中,设活动状态为完成。当活动的输入参数为绑定前一活动的输出时,需根据参数的绑定活动和方法在应用数据表中定位输入参数值 returnValue。

(4) 在过程实例表中找出活动优先级 activityPriority + 1 的转换边 ID, 重复步骤 2,3 和 4。

2.4 工作流引擎应用框架设计

为使工作流的业务组件和蕴含业务规则的协调逻辑相分离,该文采用访问者设计模式设计工作流引擎框架(图4)。主要有4个模块组成:

(1)数据库操作组件(DatabaseManager)。主要

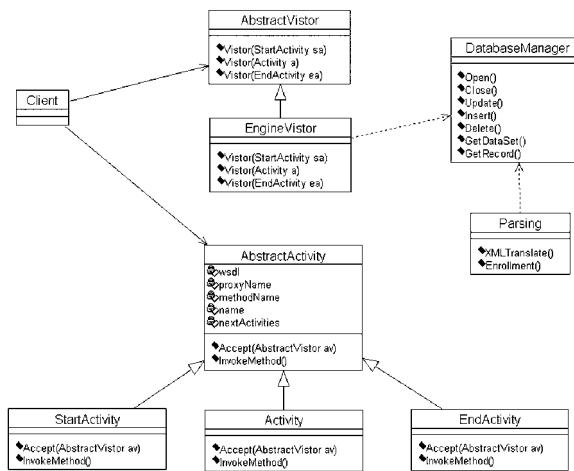


图4 工作流引擎框架类结构图

维护数据模型中定义的各类关系表,配合业务协调逻辑组件对业务活动的调度。

(2) 解析组件(Parsing)。主要包括两个操作:XMLTranslate 和 Enrollment。前者将 GCPML 文件解析成内存中的实例对象,后者将相关的模型信息注册到数据库中。

(3) 业务组件。通过对业务活动模型的抽象,提取出一个公共接口:AbstractActivity,包括服务的 wsdl 地址、服务接口、接口方法和下一个活动的集合等属性,方法主要包括 Accept 和 InvokeMethod,前者是入口方法,接收一个 AbstractVistor 类型的参数;后者代表具体完成的任务。StartActivity、Activity 和 EndActivity 分别代表开始活动、普通活动和结束活动,是抽象类 AbstractActivity 的继承实现。

(4) 业务协调组件。它封装了工作流引擎调度算法,代表访问者类。AbstractVistor 是抽象类,定义了一组公共接口来实现对业务逻辑的封装,可通过不同的继承实现来定制不同的业务逻辑。EngineVistor 是抽象类 AbstractVistor 的实现,通过实现多态函数 Vistor 来协调开始活动、普通活动和结束活动的相关流转逻辑。

3 应用示例

根据以上工作流建模方法和工作流引擎技术,设计一个地理信息服务链工作流应用软件(图 5),主要有两大功能:服务链流程建模和工作流引擎执行输出。应用场景:利用基础地理信息数据服务和警力分布专题信息数据服务底图,进行叠加分

析服务,了解警力在地图上的位置信息,实现对警员灵活调动的辅助决策。

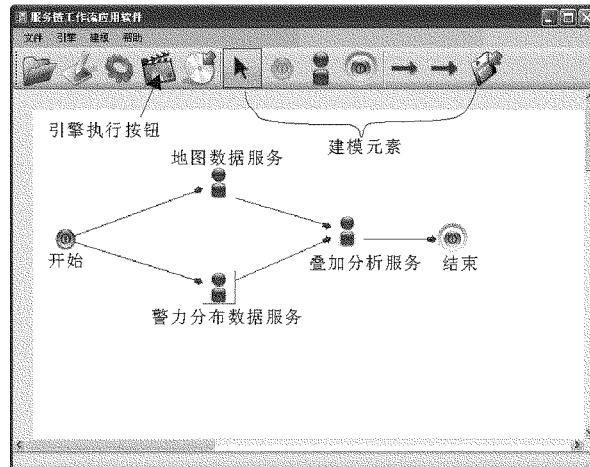


图5 服务链流程建模

3.1 建立服务

服务原材料采用微软新推出的基于 SOA 的框架产品 WCF 来构建,包括:底图数据服务、警力分布数据服务(托管在 IIS 上)和叠加分析服务。

3.2 服务链流程建模

打开建模界面,设计服务链的过程逻辑(图 5),并设置服务相应属性。警力分布数据服务;Name 为“警力分布数据服务”;wsdl 地址为 http://mb/PoliceDistributeData/Service.svc?wsdl, Interface 为“DataClient”;Method 为“GetData”;Priority 为“2”。对于方法“GetData”的输入参数在 ListViewGrid 中设置,包括选择数据的范围(经纬度坐标)和图层等信息。双击人形图标可设置相应服务的属性信息。

3.3 输出结果

在软件中打开建模工具生成的 GCPML 文件(略),点击引擎执行工具栏按钮,如果返回一个栅格地图图像,单击浏览图片按钮可在图片框中呈现;如果返回的是矢量底图数据,单击工具栏的存储按钮,可将活动的数据存到硬盘上。本例返回的是一张叠加图,图中五星代表警力的空间分布情况(图 6)。如果要对整个区域内的警力分布情况进行了解,可在服务属性设置对话框中,对流程中服务的输入参数进行重新设置即可。



图6 示例输出

4 结语

地理信息服务是地理信息系统发展的一个重要

方向。网络的广泛应用使地理信息系统可以服务的方式进行开发和包装，并以标准的接口对外提供数据和功能服务。服务链满足这一需求，但仍处于理论研究阶段。该文采用工作流技术对相应的理论和实践进行了初步探索，实现了地理信息服务的可视化组合建模和服务链的自动执行。下一步需在服务发现与工作流引擎完备性方面继续探索研究。

参考文献：

- [1] 余敏. 基于 Web Services 平台下的协同工作流系统研究与实现 [D]. 武汉理工大学硕士论文, 2006.
- [2] 范玉顺. 工作流管理技术基础 [M]. 北京: 清华大学出版社, 2002.
- [3] 萧峻妹, 张钢, 孙达. 一种基于 XML 的工作流过程定义语言 [J]. 微处理机, 2007, 28(3): 47-48, 52.
- [4] 韩怿冰. MVC 模式浅析 [J]. 科技信息, 2007, (25): 57-57.
- [5] 蔡明彦, 邱桂明. 基于 JAVA 的集成型工作流引擎的设计 [J]. 汕头大学学报(自然科学版), 2004, 19(3): 40-43.
- [6] 路春光, 孟丽丽, 等. 基于 WEB 的柔性工作流引擎的设计 [J]. 微计算机信息(管控一体化), 2006, 22(5-3): 21-23.

Research on Workflow Application for Geographic Information Service Chain

YANG Shiyong

(Shandong Information Center of Land and Resources, Shandong Jinan 250014, China)

Abstract: Through analysis and references of advanced achievements of the service chain and workflow technology, thinking of the characteristics of large amount of geographic information datas and heterogeneity of data sources, workflow management system which is suitable for the geographic information service chain, including the process modeling method, process define language, modeling tool, model structure, adjusting measuring method and application framework of workflow has been redesigned. It has realized visual combination of geographic information services and the chain's automatic performance.

Key words: Geographic information service; web services; workflow; workflow modeling; workflow engine