

QoS支持下的GIS服务链最优化问题研究

宋现锋, 刘军志

(中国科学院研究生院资源与环境学院 北京 石景山区 100049)

【摘要】针对当前OGC互操作标准未提供QoS支持的现状,对OGC服务标准进行了服务质量扩展。在此基础上实现了一种基于遗传算法、具有全局QoS约束的最优化服务选择方法。该方法采用了树型染色体编码方式及基于罚函数法的适应度函数。通过与整数规划得到的最优解进行对比,验证了该方法解质量的有效性,并通过改变候选服务数和任务数等试验条件,证明了该方法的时间复杂度近似线性。

关键词 遗传算法; GIS服务; 开放地理信息联盟; 最优化; 服务质量

中图分类号 TP75

文献标识码 A

doi:10.3969/j.issn.1001-0548.2010.02.031

Optimization of GIS Web Service Chaining Based on QoS

SONG Xian-feng and LIU Jun-zhi

(College of Resource and Environment, Graduate University of Chinese Academy of Sciences Shijingshan Beijing 100049)

Abstract In the process of GIS service chaining, the quality of service (QoS), as well as functionality of service, is very essential for Web GIS users. Nevertheless, OGC standards for GIS services have not yet supported QoS currently. This paper proposes a QoS extension to OGC interoperation specifications related to GIS services as well as an optimal approach of GIS web service chaining with global QoS constraints. This constrained optimization question is solved using a genetic algorithm with a penalty-based fitness function and a tree coding structure modeling the flowchart of GIS service chaining. With an experimental example, the effectiveness of this approach is verified in comparison with Integer programming method. By changing examination conditions such as the number of candidate services or task nodes, it is found that this approach has near linear time complexity.

Key words genetic algorithm; GIS Web service; open geospatial consortium; optimization; quality of service

随着网络技术和空间信息技术的发展,越来越多的GIS服务可以通过网络获得。单个GIS服务一般提供比较单一的功能,而用户的需求往往是复杂和多样化的,就需要把多个GIS服务组合成GIS服务链以实现更高级的功能。OWS-4项目的earth observation(EO) workflows将分布于不同地点的WMS、WFS、WCS、WPS等服务,通过BPEL语言串起来构成服务链,协同完成空间数据处理任务^[1-3]。文献[4]进行了基于工作流技术的动态GIS服务链研究,提出抽象GIS服务、具体GIS服务以及二者间的动态映射等。这些工作大都围绕可行性和功能要求进行,服务质量(quality of service, QoS)作为GIS服务的一个重要属性,没有得到足够的重视,在OGC的GIS服务标准中也没有QoS的定义。随着GIS服务数目的增多,对应某一功能,通过目录服务(catalog service,

CS)得到的服务可能有多个,需要通过QoS对候选服务进行选择,使服务链满足用户指定的QoS约束,并使用户目标最优化。本文对OGC服务标准提出QoS扩展,并在此基础上对基于遗传算法的、具有全局QoS约束的服务选择方法进行了研究。

1 问题描述

ISO19119服务体系结构规范中提出了服务链的基本概念,并根据用户控制程度的不同,将服务链分为用户自定义(透明)链、流程管理(半透明)链和集成服务(不透明)链^[5]3种类型。其中,流程管理链利用工作流技术实现了逻辑层和执行层的分离,用户在逻辑层上构建工作流,然后通过服务选择把逻辑工作流转换成可执行工作流,并送入工作流引擎,由工作流引擎管理和监控工作流的运行。这种类型

收稿日期: 2008-09-23; 修回日期: 2009-05-17

基金项目: 国家自然科学基金(40771167); 中国科学院知识创新工程重要方向项目(KGCX3-SYW-408)

作者简介: 宋现锋(1969-),男,博士,副教授,主要从事Web GIS方面的研究。

既可以让用户指定服务链的业务流程, 又隐藏了底层复杂的实现细节, 是3种类型中最灵活的一种。

本文以流程管理链作为研究对象。设一个 workflow 在逻辑层有 m 个任务, $T = \{t_1, t_2, \dots, t_m\}$ 。经过目录服务查询后, 一个任务对应多个可选的服务, 其中 $S_i = (s_{i1}, s_{i2}, \dots, s_{in_i})$ 为任务 $t_i (1 \leq i \leq m)$ 的候选服务集, n_i 为 t_i 的候选服务数, 候选服务集中的服务功能相同, 服务质量不同。服务选择为每个 $t_i (1 \leq i \leq m)$ 从 S_i 中选择一个服务 $s_{ik} (1 \leq k \leq n_i)$, 构成可执行 workflow $p = \{\langle t_1, s_{ij_1} \rangle, \langle t_2, s_{ij_2} \rangle, \dots, \langle t_m, s_{ij_m} \rangle\}$ 。选择需要满足两个条件: (1) 满足所有的QoS约束, 符合SLA(服务等级协议)。局部约束(如某个特定任务的费用)和全局约束(如流程总的响应时间)都需要被满足。(2) 对于指定的某个或某几个QoS参数达到最优化。例如用户可能需要在保证费用低于一定限制的条件下使响应时间最小, 或者使响应时间和成功率的加权和最小。

2 OGC服务标准的QoS扩展

服务质量(QoS)是进行服务选择的基础, 而目前的OGC服务标准没有提供对QoS的支持, 因此需要对现有OGC服务标准进行扩展。QoS描述了一个服务满足用户需求的能力, 现有的QoS研究大部分是针对标准Web服务进行的^[6], 如考虑服务执行时间、费用、可用性等因素。GIS服务除了标准服务的QoS外, 还有特殊的需求, 如对于空间数据服务来说, 数据准确度和数据完整性是用户进行GIS服务选择的重要指标^[7]。考虑到GIS服务的特殊性, 对OGC服务标准QoS扩展的数据结构定义如下:

QoS是每种OGC服务都具有的一个属性, 所以可以把QoS的定义放到ows命名空间的owsCommon.xsd文件中。QoS信息在描述服务详细元数据的函数操作中返回, 如WFS的QoS信息应该由describe feature函数返回, 而WPS的QoS信息应该由describe process函数返回。

表1 QoS数据结构

名称	定义	数据类型	取值范围	出现次数
响应时间	传输时间和处理时间之和	ows:UoM	[0, ∞]	1
费用	调用一次服务所需要支付的费用	ows:UoM	[0, ∞]	1
成功率	网络计算服务被调用执行的成功率	xsd:double	[0,1]	1
信誉等级	用户对网络计算服务评价等级的平均值	xsd:integer	[1,5]	0或1
完整性	空间数据服务提供的数据对真实地理现象描述的完整程度	xsd:double	[0,1]	0或1
准确度	空间数据服务所提供的数据与真实地理现象的接近程度	xsd:double	[0,1]	0或1

3 工作流的QoS计算

为了根据QoS完成服务的选择, 需要从单个服务的QoS属性值计算整个工作流的QoS, 即服务组合的QoS。服务组合的QoS依赖于参与组合的各个服务的QoS以及组合的流程结构。组合的基本模型分为

顺序、并行、选择及循环4类, 表2列出了这几类基本模型的组合QoS计算函数^[8]。由于只有空间数据服务才具有数据准确度和完整性属性, 这两个属性上的约束通常是局部的, 所以未给出它们的计算公式。

选择结构表达式中 P_i 表示第 i 个服务组件的执行概率, 并且 $\sum_{i=1}^m P_i = 1$ 。

表2 组合的基本模型及QoS属性计算函数

QoS属性	顺序	选择	并行	循环
响应时间	$\sum_{i=1}^n T(t_i)$	$\sum_{i=1}^m P_i T(t_i)$	$\max\{T(t_i)\}$	$kT(t)$
费用	$\sum_{i=1}^n C(t_i)$	$\sum_{i=1}^m P_i C(t_i)$	$\sum_{i=1}^n C(t_i)$	$kC(t)$
成功率	$\prod_{i=1}^n S(t_i)$	$\sum_{i=1}^m P_i S(t_i)$	$\prod_{i=1}^n S(t_i)$	$S(t)^k$
信誉	$\frac{1}{n} \sum_{i=1}^n R(t_i)$	$\sum_{i=1}^m P_i R(t_i)$	$\frac{1}{n} \sum_{i=1}^n R(t_i)$	$R(t)$

4 基于遗传算法的服务选择方法

对QoS进行全局优化的服务选择是一个NP-Hard问题, 现有的解决方案主要有整数规划和遗传算法两种。整数规划在线性约束、线性目标函数条件下可以精确求解^[9], 但现实中很多约束和目标函数是非线性的, 而且整数规划的计算量随问题规模增长急剧增加。遗传算法作为一种随机化搜索算法, 虽然不能得到最优解, 但是可以通过迭代不断得到更优解, 从而得到符合用户要求的满意解, 而且遗传算法计算量也较小^[10-11], 所以本文选择遗传算法来实现基于QoS的服务选择。

4.1 染色体的树型编码

组合服务模型用结构化流程表示, 它可以分解为较细粒度的子结构, 子结构之间遵循顺序、概率、并发和循环4种组合关系进行组合。同样地, 分解可以递归地施加于子结构, 直至所有的子结构都为原子结构为止。通过这种递归, 结构化流程可以表示为一棵树^[12]。本文对染色体采用树型编码, 采用上述编码方式, QoS的计算可以直接映射到表2中的计算方法, 而且中间节点可以保存服务组合QoS计算的中间结果, 减少计算量。

4.2 适应度函数设计

本文采用标准化后各QoS属性的加权和作为工作流的适应度函数, 即:

$$f = \sum_{i=1}^n w_i V_i \quad (1)$$

式中 w_i 为QoS属性的权重, 表示用户对QoS属性的关注程度, $0 \leq w_i \leq 1$, $\sum_{i=1}^k w_i = 1$; V_i 为标准化后的工作流的QoS属性。要把工作流的各个QoS属性标准化以后再进行加权, 是因为QoS各个属性量纲不一致, 如价格和可用性会相差好几个数量级; 目标方向不一致, 如价格越低越好, 而可用性越高越好。如果不进行标准化, 这些属性不具有可比性。标准化的具体方法如下:

对于成本型目标, 如费用和执行时间, 有:

$$V_i = \begin{cases} \frac{Q_i^{\max} - Q_i}{Q_i^{\max} - Q_i^{\min}} & Q_i^{\max} - Q_i^{\min} \neq 0 \\ 1 & Q_i^{\max} - Q_i^{\min} = 0 \end{cases} \quad (2)$$

对于效益型目标, 如成功率和可用性, 有:

$$V_i = \begin{cases} \frac{Q_i - Q_i^{\min}}{Q_i^{\max} - Q_i^{\min}} & Q_i^{\max} - Q_i^{\min} \neq 0 \\ 1 & Q_i^{\max} - Q_i^{\min} = 0 \end{cases} \quad (3)$$

式中 Q_i^{\max} 和 Q_i^{\min} 分别表示所有组合方案中第*i*个QoS属性的最大值和最小值。在利用表2中公式计算出 Q_i 后, 就可以利用式(2)或(3)计算标准化的属性值, 标准化的属性值在0~1之间。在计算 Q_i^{\max} 和 Q_i^{\min} 时, 不必遍历所有可能的组合方案。如为了计算价格属性的最大值 $Q_{i\text{Price}}^{\max}$, 只需要对每个任务选择价格最高的服务, 然后求和即可。

对于带限制条件的问题, 用罚函数法将限制条件与目标函数综合在一起构成适应度函数^[13]:

$$\text{Fitness} = f - \lambda \sum_{j=1}^n \left(\frac{\Delta P_j}{R_{j\max} - R_{j\min}} \right)^2 \quad (4)$$

式中 $R_{j\max}$ 和 $R_{j\min}$ 分别为第*j*个约束条件中的最大值与最小值; n 为约束条件个数; λ 为惩罚系数, 属于经验值, 在本文中取值为0.1。 ΔP 定义如下:

$$\Delta P_j = \begin{cases} P_j - R_{j\max} & P_j > R_{j\max} \\ 0 & R_{j\min} \leq P_j \leq R_{j\max} \\ R_{j\min} - P_j & P_j < R_{j\min} \end{cases} \quad (5)$$

5 实验分析

通过实验对本文提出的方法的有效性进行验证。实验环境为Pentium IV 2.0G处理器、1 024 MB内存和Windows XP操作系统, 编程语言为Python 2.4。工作流中各任务的候选服务的QoS属性值在一定的取值范围内随机产生。采用精英遗传算法, 交叉率为0.7, 变异率为0.1, 种群大小默认为50, 当候选服务数超过50时, 种群大小取值与候选服务数相等, 以保证解的质量。

考虑到在线性约束、线性目标函数情况下, 整数规划可以求得最优解, 可以用整数规划得到的最优解验证遗传算法解的质量。设计目标函数和约束条件为:

$$\begin{cases} \max & f = 0.6V_{\text{duration}} + 0.4V_{\text{price}} \\ \text{constraints} & Q_{\text{duration}} < 20, \quad Q_{\text{price}} < 600 \end{cases} \quad (6)$$

图1给出了整数规划得到的最优解、遗传算法得出的服务选择方案和随机生成的符合约束的选择方案的适应度对比情况, 图中曲线是在候选服务个数为10的条件下、重复进行20次运算的结果。代数为100的遗传算法在20次实验中有11次达到最优解, 未达到最优解时与最优解差别也较小。

图2给出了任务数为8、最大代数为100的情况下, 计算时间随服务候选数变化的曲线。当服务数超过50时, 计算时间随任务数呈线性增加。因为当候选服务数超过50时, 种群大小取值为候选服务数。

这种做法在保证解的质量的同时,也使算法具有可以接受的时间复杂度。

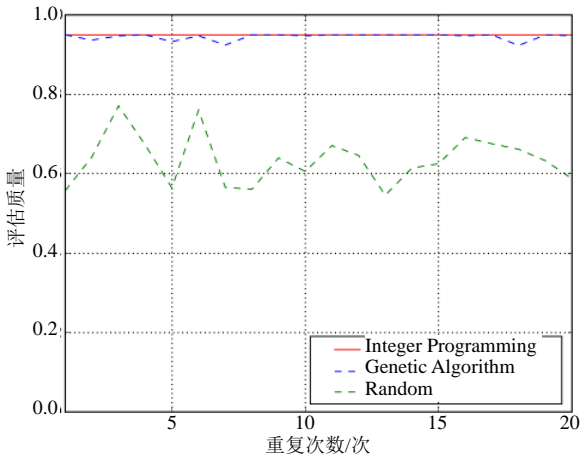


图1 遗传算法、整数规划与随机方案的适应度对比

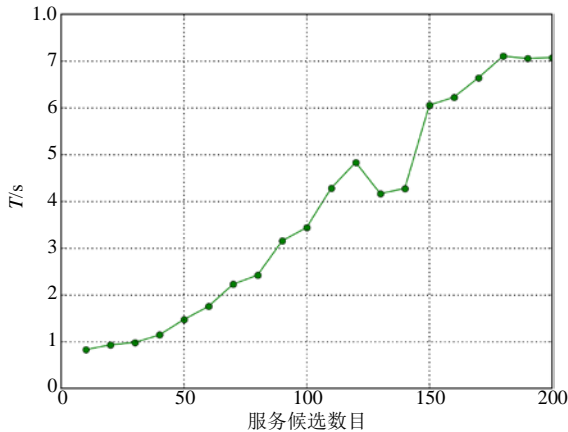


图2 计算时间随服务候选数目变化情况

图3给出了候选服务数为10、最大代数为100的情况下,计算时间随任务数变化的曲线。计算时间随任务数呈线性增加,在实际应用中这种时间复杂度是可以接受的。

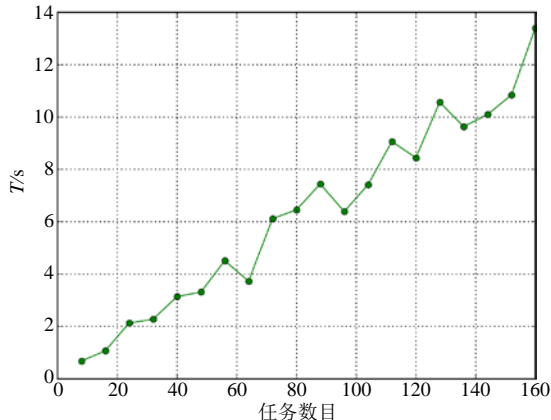


图3 计算时间随任务数目变化情况

6 结束语

本文对OGC服务标准提出了QoS扩展,并在此基础上对GIS工作流的QoS进行优化,实现了一种基

于遗传算法,对QoS进行全局优化的服务选择方法,并通过实验验证了该方法的可行性。

与商业Web服务相比, GIS服务传输和处理的数据要大得多,而数据量的大小直接影响到GIS服务的QoS属性,如执行时间、成功率和价格等。如何在考虑数据量对QoS参数影响的情况下,对GIS工作进行优化是下一步需要研究的内容。

参 考 文 献

[1] SONG Xian-feng, RUI Xiao-ping, HOU Wei, et al. An OGC standard-oriented architecture for distributed coal mine map services[J]. Journal of China University of Mining and Technology, 2008, 18(3): 381-385.

[2] KEENS S. Discussions, findings, and use of WPS in OWS-4[R]. Open Geospatial Consortium, 2006.

[3] 刘军志, 宋现锋, 汪超亮, 等. 基于OGC WPS的遥感图像分布式检索系统研究[J]. 地理与地理信息科学, 2008, 24(4): 1-5.

LIU Jun-zhi, SONG Xian-feng, WANG Chao-liang, et al. A distributed retrieval system for remote sensing images based on OGC WPS[J]. Geography and Geo-Information Science, 2008, 24(4): 1-5.

[4] 贾文珏, 李 斌, 龚健雅. 基于工作流技术的动态GIS服务链研究[J]. 武汉大学学报: 信息科学版, 2005, 30(11): 982-985.

JIA Wen-jue, LI Bing, GONG Jian-ya. Research on dynamic GIS chain based on workflow technology[J]. Geomatics and Information Science of Wuhan University, 2005, 30(11): 982-985.

[5] The opengis abstract specification topic 12: opengis service architecture[S/OL]. [2002-3-21]. <http://www.opengis.org/docs/022112>.

[6] 郑向宏, 李院春, 李增智. 面向语用Web服务的QoS评价模型研究[J]. 电子科技大学学报, 2007, 36(6): 1477-1484.

ZHENG Xiang-hong, LI Yuan-chun, LI Zeng-zhi. Research on pragmatic Web oriented QoS evaluation model[J]. Journal of University of Electronic Science and Technology of China, 2007, 36(6): 1477-1484.

[7] 刘书雷. 基于工作流的空间信息服务聚合技术研究[D]. 长沙: 国防科技大学, 2005.

LIU Shu-lei. Research on the composition techniques of geo-spatial Web services based on workflow[D]. Changsha: National University of Defense Technology, 2005.

[8] 张晓光, 李 琰, 王海洋. 一种基于遗传算法QoS敏感的Web服务组合方法[J]. 山东大学学报(理学版), 2007, 42(9): 56-61.

ZHANG Xiao-guang, LI Yan, WANG Hai-yang. A kind of QoS-sensitive web service composition method based on genetic algorithm[J]. Journal of Shandong University (Natural science), 2007, 42(9): 56-61.

(下转第315页)