

基于改进的自适应遗传算法的 GIS 服务组合研究

柳 玲 冉 唯 许超超

(重庆大学软件学院 重庆 400044)

摘 要 针对当前 GIS 系统中集成应用模型时存在的数据和功能冗余、模型难以复用等问题,将 GIS 应用模型分解为子模型,与 GIS 功能一起,以服务的方式提供给用户,并对这些 GIS 服务定义服务质量(QoS);最后,提出了一种基于改进的自适应遗传算法(MAGA)的服务组合方法,以实现 GIS 服务组合的全局优化。

关键词 GIS 服务, QoS, 改进的自适应遗传算法(MAGA)

中图分类号 TP311.5 **文献标识码** A

Research on GIS Services Composition Based on Application Model

LIU Ling RAN Wei XU Chao-chao

(School of Software Engineering, Chongqing University, Chongqing 400044, China)

Abstract Recently, there are some problems in integrating application model into GIS system, such as redundancy, difficulty to reuse the model and so on. We proposed a method to solve these problems. At first, the GIS application model was decomposed into sub-models, and the sub-models and GIS functions were defined as services which are provided to users. Then it defined the QoS (quality of service) of the GIS services. At last, the paper proposed a service composition method based on modified adaptive genetic algorithm (MAGA). The method can achieve the goal of global optimization of GIS services composition.

Keywords GIS services, QoS(Quality of Service), Modified adaptive genetic algorithm(MAGA)

随着计算机技术和网络技术的发展, GIS(geographic information systems, 地理信息系统)技术在特定领域中的应用价值逐渐提升, GIS 系统中应用模型的集成已成为了研究热点。当前 GIS 应用模型在集成过程中存在很多问题, 如数据和功能冗余、模型难以复用等^[1]。文献[2]进行了基于 QoS 的 GIS 服务组合研究, 提出了对 OGC 进行 QoS 扩展和基于简单遗传算法(SGA)的服务组合方法。然而, 在简单遗传算法(SGA)中, 由于交叉概率 P_c 和变异概率 P_m 取为恒定值, 对基于 QoS 全局优化的进行服务选择时, 效率不高且存在“早熟”的可能^[3]。本文基于以上背景, 将 GIS 应用模型分解为子模型, 与 GIS 基本功能一起, 以服务的方式提供给用户, 并对 GIS 服务定义 QoS 属性; 最后, 提出了一种基于改进的自适应遗传算法(MAGA, Modified Adaptive Genetic Algorithm)的服务组合方法, 以实现 GIS 服务组合的全局优化。

1 GIS 服务研究

GIS 系统会有一些常见的公共功能, 如地图绘制、鹰眼、空间位置计算、地图打印等, 称其为 GIS 基本功能。通常实现一个 GIS 系统, 都需要实现一遍这些 GIS 基本功能, 从而造成了资源的重复浪费。GIS 系统还具有应用分析功能, 如量算与统计、预测与监测、规划与管理、辅助决策等, 这些应用分析功能是通过各种应用模型来实现的, 称其为 GIS 应用模

型。一个 GIS 应用模型往往将 GIS 功能和数据的部分或者全部与应用模型绑定在一起, 从而造成了功能或数据的冗余, 应用模型的难以复用。

针对这些问题, 本文将 GIS 应用模型从数据和过程的角度按照自上而下的方法进行分解, 得到若干子模型; 并将各个 GIS 基本功能与子模型封装为 GIS 服务。GIS 应用模型的分解, 能有效提高 GIS 服务查找的准确率和 GIS 服务的重用率。GIS 应用模型的分解方法需遵守以下原则: 第一, 空间信息优先原则; 第二, 完整性与独立性原则。GIS 应用模型中所使用的数据信息可以通过 OGC 提供的 WMS、WFS、WCS 等 GIS 服务得到^[4]。

2 GIS 服务的 QoS 参数

QoS 是影响服务组合方式选择的关键。文献[2]对 OGC 服务标准提出了 QoS 扩展, 扩展后的 GIS 服务包括 6 种标准的 QoS 参数, 即执行时间 D (Duration)、执行费用 P (Price)、成功执行率 S (SuccessfulRate)、信誉等级 R (reputation)、空间数据完整性 DI (Integrity) 和空间数据真实性 DA (Authenticity)。考虑到本文对 GIS 服务的定义, 并非所有的 GIS 服务都直接涉及到空间数据。用遗传算法解决优化问题时, 复杂的多变量会增加算法的计算时间及复杂度。为此, 本文定义 4 种 QoS 参数: 执行时间 T (time)、执行费用 C (cost)、

本文受中央高校基本科研业务费(CDJZR10090001)资助。

柳 玲(1970-), 女, 博士, 副教授, 主要研究方向为服务计算与面向服务的软件工程; 冉 唯(1986-), 女, 硕士生, 主要研究方向为面向服务的软件工程, E-mail: 625351618@qq.com; 许超超(1986-), 女, 硕士生, 主要研究方向为 Web 服务发现。

信誉等级 Rep (reputation)和可靠性 R (reliability)。空间数据 GIS 服务的信誉等级 Rep 和可靠性 R 完全可以表示文献[2] 中空间数据完整性 DI (Integrity)和空间数据真实性 DA (Authenticity),因为不完整、不真实的空间数据不会得到较大数值的信誉等级和可靠性。GIS 服务组合 SC 的 QoS 属性值是按照组合流程结构和各个单一服务的 QoS 属性值计算得到的。组合基本模型的流程结构分为串联、并联、选择、循环 4 类。

3 基于改进的自适应遗传算法的服务组合方法

针对基于 QoS 的服务组合问题,遗传算法是比较好的解决方案,虽然不能得到最优解,但可以通过不断迭代得到满足用户要求的满意解。

3.1 基于用户偏好的适应度函数设计

本文考虑到 GIS 应用模型的分解方法,采用文献[6]中提出的树型编码对染色体进行编码。因为 4 个 QoS 属性值的大小不一,可能会相差太大,没有可比性。本文采用文献[2]中的量化方法将 $Q(S)=\{T,C,Rep,R\}$ 中的 4 个属性进行标准化。量化后得到执行时间 T 、执行费用 C 、信誉等级 Rep 和可靠性 R 分别为 $L_{T_i}, L_{C_i}, L_{Rep_i}, L_{R_i}$ 。

基于上述定义,可以设置符合用户偏好的适应度函数:

$$f_i = \lambda_1 L_{T_i} + \lambda_2 L_{C_i} + \lambda_3 L_{Rep_i} + \lambda_4 L_{R_i} \quad (1)$$

式中, $\sum_{i=1}^4 \lambda_i = 1$, 用户可以根据自己的侧重点选择 λ_i 的系数值。例如:某用户看重服务的信誉和可靠性,可以将 λ_i 依次设置为 0.1, 0.1, 0.4, 0.4。这样得到的适应度函数即为: $f_i = 0.1L_{T_i} + 0.1L_{C_i} + 0.4L_{Rep_i} + 0.4L_{R_i}$ 。

3.2 遗传操作

3.2.1 选择

本文采用轮盘赌法和最佳个体保存法进行选择。最佳个体保存法是指对群体中适应度最高的个体不进行交叉和变异,直接复制到下一代中,这样可以确保在算法终止时得到的最后结果一定是历代出现的最高适应度的个体。

3.2.2 交叉和变异

影响遗传算法行为和性能的关键在于交叉概率 P_c 和变异概率 P_m 的选择。为了选择到适用的 P_c 和 P_m , 本文采用文献[3]提出的 MAGA 算法(改进的自适应遗传算法)对 P_c 和 P_m 进行选择。 P_c 和 P_m 公式如下:

$$P_c = \begin{cases} P_{c1} - \frac{P_{c1} - P_{c2}}{\text{Max}(f) - \text{Avg}(f)} (f^* - \text{Avg}(f)) & f^* \geq \text{Avg}(f) \\ P_{c1}, & f^* < \text{Avg}(f) \end{cases} \quad (2)$$

$$P_m = \begin{cases} P_{m1} - \frac{P_{m1} - P_{m2}}{\text{Max}(f) - \text{Avg}(f)} (\text{Max}(f) - f), & f \geq \text{Avg}(f) \\ P_{m1}, & f < \text{Avg}(f) \end{cases} \quad (3)$$

式(2)、式(3)中, P_{c2} 表示群体中最大适应度值个体的交叉率, P_{m2} 表示群体中最大适应度值个体的变异率, f^* 表示需要交叉的一对个体中较大的适应度值, f 表示需要变异的个体适

应度值, $\text{Max}(f)$ 表示群体中最大的适应度值, $\text{Avg}(f)$ 表示群体的平均适应度值。 $P_{c1}, P_{c2}, P_{m1}, P_{m2}$ 4 个参数取值范围均在 0~1 之间。

4 实验结果分析

实验环境为 CPU 为 Intel Pentium(R) Dual-Core, 内存大小为 3.49GB, 操作系统为 Window XP, 编程语言采用 C#。设定算法的群体规模为 400, 染色体交叉率 $P_{c1} = 0.9, P_{c2} = 0.6$, 染色体变异率 $P_{m1} = 0.1, P_{m2} = 0.001$ 。各个候选服务的 QoS 属性值在规定范围内随机产生, 如 $T \in (0, 20], C \in [0, 50], Rep \in [0, 5], R \in [0, 1]$ 。本文以校园地图服务为例, 采用图 1 所示的 GIS 服务组合流程和图 2 所示的 GIS 应用模型分解, 以 3.1 节所举例子中的适应度函数为目标函数, 设定组合服务的最高执行时间和最高价格分别为 30s 和 200 元。

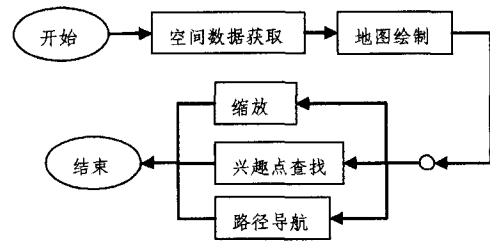


图 1 校园地图服务组合流程

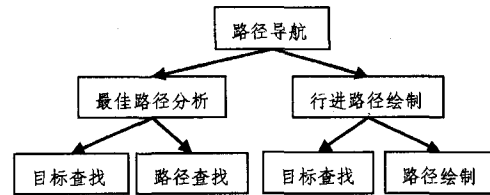


图 2 路径导航的模型分解

图 3 表明简单遗传算法(SGA)与改进的自适应遗传算法(MAGA)在服务群规模为 30 的实验中, 得到目标函数值的对比。可以看出, SGA 效率不高且容易早熟, MAGA 的组合方法明显优于 SGA 的组合方法。

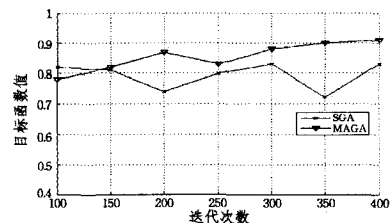


图 3 SGA 与 MAGA 各目标函数值的对比

结束语 本文就当前 GIS 系统中集成应用模型遇到各种问题的现状, 提出了 GIS 服务的构成方法, 并对 GIS 服务定义服务质量; 设计了一种基于改进的自适应遗传算法的服务组合方法, 并通过实验验证了其可行性和优越性。下一步的工作是研究如何将拥有海量空间数据的服务进行优化, 以提高 GIS 组合服务的全局 QoS。

参考文献

[1] 杨富平, 李林, 丰江帆, 等. 基于服务组合的 GIS 应用模型集成

- [2] 宋现锋, 刘军志. QoS 支持下的 GIS 服务链最优化问题研究[J]. 电子科技大学学报, 2010, 39(2): 298-301
- [3] 段玉倩, 贺家李. 遗传算法及其改进[J]. 电力系统及其自动化学报, 1998, 10(1): 39-52
- [4] Ma Sai, Li Min-ruo, Du Wei-chang. Service Composition for GIS

- [5] 刘书雷, 刘云翔, 张帆, 等. 一种服务聚合中 QoS 全局最优服务动态选择算法[J]. 软件学报, 2007, 18(3)
- [6] 蔡美玲, 高春鸣. 基于树型编码的遗传算法在 Web 服务选择中的应用[J]. 计算机工程与应用, 2007, 43(31): 214-218

(上接第 415 页)

就将其反馈到需求分析中, 重新开发消除了缺陷的安全性需求, 以降低系统进入危险状态的可能性。

归纳分析方法和演绎分析方法都可以用于输出安全性需求的脆弱性分析。较高的分析等级一般比较适合采用演绎分析方法, 此时安全性需求信息比较通用、宏观、不涉及细节。而归纳分析方法则更适合对较低分析等级内的分析领域进行分析, 此时有关安全性需求组成要素及其联系的信息比较具体、特殊。PHA 和 HAZOP 是两种可用于项目需求规约阶段的方法, 并且这两种方法都可用于归纳分析或演绎分析。PHA 能够识别关键的系统功能和严重的系统危险; 而 HAZOP 使用一些引导词表明规约如何被违背以及违背是否会导致危险等, 是一种更为详细、系统的分析方法, 可认为是一种重量级的 PHA。考虑到脆弱性分析活动是在项目早期的需求开发阶段开展的, 此处给出如图 5 所示的支持 PHA、HAZOP 方法应用的脆弱性分析过程。

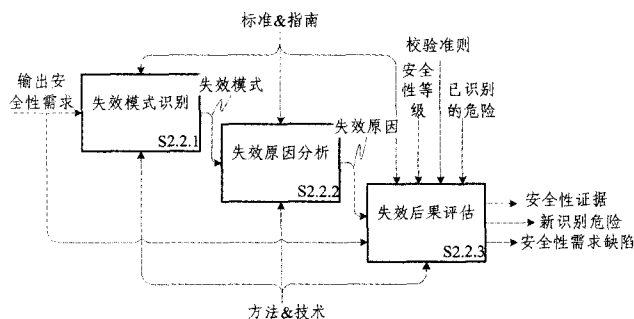


图 5 支持 PHA、HAZOP 的脆弱性分析活动

一旦脆弱性分析识别出具有能够导致危险后果的缺陷或者能够引起新的危险, 就会将其提交给需求分析, 以修正安全性需求。若脆弱性分析结果表明, 输出安全性需求既不具有缺陷, 也不会引入新的危险状态, 那么将给出证据, 表明输出的安全性需求已经满足校验准则。

结束语 本文采用系统建模与系统分析相结合的方式, 基于系统思维建构了一个软件安全性需求开发框架, 提出了一种集成了安全性分析的需求开发过程。该方法能够最大程度地约束安全性需求缺陷, 防止其向同一分析等级内的其它领域或下一分析等级传播, 同时有助于尽早地考虑重新生成安全性需求, 减弱后续安全性分析的复杂性。此外, 这种集成需求分析和安全性分析的策略能够不断生成证据, 支持安全性论据和案例的构建, 有助于减少后续认证代价。

尽管本文期望在需求开发阶段, 采用需求分析和安全性分析相结合的方法来确保安全性需求开发的完整性、正确性、一致性, 但毋庸置疑的是, 该目标能否实现还受到下述 3 个约束的影响: ①脆弱性分析并不能保证安全性需求依赖的所有

假设都能得到严格遵守; ②即使对安全性需求进行验证和确认, 仍然会驻留一些未被发现的缺陷; ③已有经验表明, 验证工作本身也可能存在缺陷。可见, 若可能, 还应尽可能早地开展安全性评估活动, 度量安全性需求的风险性, 预测软件对系统风险的贡献。

参 考 文 献

- [1] 褚文奎, 张凤鸣, 樊晓光. 综合模块化航空电子系统软件体系结构综述[J]. 航空学报, 2009, 30(10): 1912-1917
- [2] 422nd Test and Evaluation Squadron. Executive summary: aircraft accident investigation, F/A-22 S/N 00-4014[EB/OL]. http://www.f-22raptor.com/pdf/af_exsum_f22crash.pdf
- [3] 樊晓光, 褚文奎, 张凤鸣. 软件安全性综述[J]. 计算机科学, 2011, 30(5): 812-818
- [4] Lutz R R. Analyzing software requirements errors in safety-critical, embedded systems[C]//Proceedings of the International Conference on Software Requirements. IEEE, 1992: 53-65
- [5] McDermid J A. Software Safety: Where's the evidence? [C]//6th Australian Workshop on Industrial Experience with Safety Critical Systems and Software(SCS 2001). Brisbane: Australian Computer Society, 2001, CRPIT 3: 1-6
- [6] RTCA. DO-178B—1992 Software Considerations in Airborne Systems and Equipment Certifications[S]. Washington DC: Radio Technical Commission for Aeronautics, Inc, 1992
- [7] MoD. DEF-STAN 00-56 issue 4—2007 Safety Management Requirements for Defence Systems[S]. London: Ministry of Defence, 2007
- [8] 王拥军. 需求工程中的不确定性研究[D]. 西安: 西北工业大学, 2001
- [9] Wu W. Architectural reasoning for safety-critical software applications[D]. Heslington: University of York, 2007
- [10] Leveson N G. An approach to designing safe embedded software [M]. London: Springer Verlag, 2002, LNCS 2491: 15-29
- [11] Murray D P, Hardy T L. Developing safety-critical software requirements for commercial reusable launch vehicles[EB/OL]. http://www.faa.gov/about/office_org/headquarters_offices/ast/reports_studies/media/DMurray_SW%20REQTS_IAASS07_FINAL.pdf/
- [12] Gharajedaghi J. System thinking: managing chaos and complexity: a platform for designing business architecture[M]. Boston: Elsevier, 1999
- [13] Leveson N G. A Systems-Theoretic Approach to Safety in Software-Intensive Systems[J]. IEEE Transactions on Dependable and Secure Computing, 2004, 1(1): 66-86