

可视化建模技术在 GIS 中的研究与应用

崔 雪, 石伟伟, 李永超

(北京超图软件股份有限公司 北京 100015)

摘要: 可视化建模是利用围绕现实想法组织模型的一种思考问题的方法。该技术以其简化需求、易于理解问题、沟通问题等特点被广泛应用于软件开发的各个环节中。本文分析了空间信息处理的特点,结合可视化建模和工作流的技术,提出了空间处理建模的概念,并探讨了空间处理建模的建模流程及应用方式。

关键词: 可视化建模; GIS; 工作流; 空间信息处理; 空间处理建模

中图分类号: P208 文献标识码: B 文章编号: 1672 - 5867(2013) S0 - 0015 - 04

Research and Application of Visual Modeling in GIS

CUI Xue, SHI Wei - wei, LI Yong - chao

(SuperMap Software Co., Ltd., Beijing 100015, China)

Abstract: Visual modeling is an approach of modeling based on thinking around a realistic idea. The technology is widely adopted in software developments for its advantages in simplifying the requirements and facilitating comprehension and communication. This paper analyzes the characteristics of spatial information processing, and by combining visual modeling and workflow technology, proposes the concept of spatial processing modeling, and eventually explores its modeling process and application methods.

Key words: visual modeling; GIS, workflow; spatial information processing; spatial processing modeling

0 引言

GIS (Geographic Information System 地理信息系统) 的重点在空间信息处理,优势在空间分析。目前, GIS 已广泛应用于各个行业,且应用日益深入,涉及越来越多的复杂分析处理过程,这些过程由一系列相互联系又具有一定次序关系的空间操作(如数据转换等)或空间分析模型(如缓冲区分析、叠加分析等)组成,而它们之间又存在着复杂的关系,串联、并联、分支、或者层次结构。这种空间分析操作流程称为空间信息处理过程(Spatial Process)。

GIS 本身提供了基本的空间操作功能和空间分析模型,但却缺乏对其所构成的复杂空间信息处理过程的有效管理,不能有效的构建空间分析模型及每个模型间的关系,使得原本复杂的空间信息处理过程变得更加复杂,再加上有大量人工参与,导致过程和结果可控性差。因此,对空间信息处理过程进行过程建模,辅助大型 GIS 应用的规划、管理和实施就显得尤其重要。

本文分析了空间信息处理过程的基本特性,结合可视化建模和工作流的技术,提出了空间处理建模的概念,实现 GIS 应用的过程可视化建模。

1 空间信息处理

空间信息处理过程包括 3 种基本要素:空间活动、过程数据和过程关系流。空间活动是空间信息处理过程的基本活动,通常对应一组空间操作或空间分析模型,由人或计算机软件执行。空间活动的逻辑顺序关系构成空间信息处理过程的控制流和操作流。过程数据是空间活动处理的对象,通常对应一定格式的空间数据,它是前驱空间活动的输出,也是后继空间活动的输入。过程数据的处理流程构成空间信息处理过程的数据流。过程关系流是空间活动和过程数据之间的连接弧,它具有一定的方向,标志着数据和控制的流向。

空间信息处理过程为一个有向图, $SP = (A, D, F)$, 是一个三元组。其中 A 为有向空间活动集; D 为有向过程数据集; F 为表示为 A 和 D 有向关系集。并且满足:

- 1) 空间活动和过程数据是两个不同的元素,即 $A \cap D = \Phi$ 。
- 2) SP 中至少有一个元素,即 $A \cup D \neq \Phi$
- 3) $F \subseteq (A \times D) \cup (D \times A)$ 是关系流集, F 是从 A 和 D 构造出来的。

收稿日期: 2013 - 06 - 20

作者简介: 崔雪(1982 -)女,山东青岛人,硕士,2008年毕业于广西师范学院地图学与地理信息系统专业,主要从事平台软件技术支持工作。

4) $dom(F) \cup cod(F) = A \cup D$ 其中 $dom(F) = \{x|y: (x, y) \in F\}$ $cod(F) = \{y|x: (x, y) \in F\}$ 分别为 F 所含的有序偶的第一个和第二个元素所组成的集合,代表 F 的定义域和值域。

5) 空间活动 A 、过程数据 D 和关系流 F 都是空间相关,受空间语义的约束。空间活动只能与空间数据有直接的关系流,孤立的数据和孤立的的活动不具有现实意义。

在 GIS 中,对于空间信息的处理过程一般采用两种方式:一种是用 GIS 桌面软件处理的方式;另一种是在 GIS 开发平台上写代码开发的方式;桌面软件处理,需要按步骤操作,前一步操作完成后进行下一步操作,每一步操作都需要人工监控,人工输入参数。而代码方式可以将整个空间信息处理过程用代码串起来,使整个操作连续执行,直到最后输出结果,执行过程中无需人工监控和干预,但缺点是代码需要专业的软件开发工程师去编写,这样无论是开发难度还是时间上相比桌面操作都要困难很多,且复用性差,需求的每一次改变,都要开发人员改代码。

2 可视化建模在空间信息处理中的应用

可视化建模 (Visual Modeling) 是利用围绕现实想法组织模型的一种思考问题的方法。模型对于了解问题、沟通问题、模仿企业流程、准备文档、设计程序和数据库来说都是有用的。建模促进了对需求的更好的理解。可视化建模就是以图形的方式描述所开发的系统的过程,它提供了一种从不同的视角观察被开发系统的机制。

好的建模能够帮助我们:

- 1) 鉴别需求和沟通信息;
- 2) 着眼于系统的组件如何相互作用,而不是陷于具体的细节;
- 3) 使你能够了解设计组件的相互关系;
- 4) 通过使用一个共同的图形语言,改进跨团队的沟通。

可视化建模在修改和复用方面同样很有效,例如客户有了一个新的需求,分析后发现与已有的一个需求很类似,只是某几个环节需要更改,那么可以直接拿前一个建模后的成果将其中几个元素稍加改动即可,不用重新建模,无论从时间上还是人力上都能做到非常好的节约。

鉴于以上谈到的可视化建模优点,笔者将可视化建模技术应用到 GIS 中,目的是使 GIS 空间信息处理的构建和操作流程能像可视化建模一样简单、易于理解,且复用性高。GIS 的空间信息处理的每一个空间活动都可以将其构建成为一个可视化的模型,如何将 GIS 的操作功能按照空间活动的特征封装成为有输入输出的模型是首先需要考虑的问题。为此需要仔细分析空间活动的元模型。空间活动元模型如图 1 所示。空间操作和分析模型是空间活动的核心要素。任何复杂的空间操作和分析模型都可以由最基本的空间原子活动组合产生,空间原子活动对应到 GIS 上是一个空间基本操作。空间活动同时还包括用于过程逻辑的属性信息和状态信息,这些信息在过程

演进中具有重要的作用。

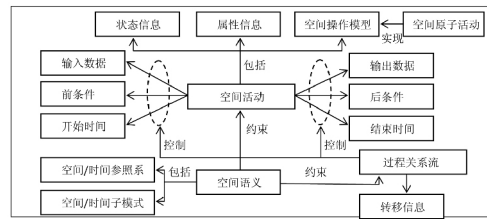


图 1 空间活动元模型图

Fig. 1 Meta - model diagram of space activities

接下来需要解决的问题是如何定义、组织和规范过程数据和过程关系流,如何控制信息处理操作过程。为此,我们结合了 workflow 的技术。Workflow 就是业务过程的部分或整体在计算机应用环境下的自动化,它主要解决的是使在多个参与者之间按照某种预定义的规则传递文档、信息或任务的过程自动进行,从而实现某个预期的业务目标,或者促使此目标的实现。Workflow 是工作流程的计算模型,将工作流程中的工作如何前后组织在一起的逻辑和规则在计算机中以恰当的模式进行表示并对其进行计算。

结合可视化建模和 workflow 技术,本文提出了空间处理建模的概念。空间处理建模首先是获取可以解决一项问题的若干个空间处理;其次,编排这些空间处理的执行逻辑顺序;最后,正确设置空间处理的输入参数以及空间处理间的数据传递关系。空间处理建模所构建的数据处理流程不同于普通可视化建模所构建的模型流程图,空间处理建模所构建的流程是可执行的,通过执行可以实现空间信息处理任务,获得预期结果数据。

结合空间处理建模这一概念,超图软件研发了 SuperMap GeoProcessor 6R 产品,该产品是进行 SuperMap 空间处理建模的可视化操作场所。为了更好地服务于应用层的用户,空间处理建模工具隐藏了空间处理建模复杂的代码实现和逻辑概念,用户无需进行二次开发,只需通过图形化的绘制流程图的方式构建空间处理模型,形成一个数据处理、加工的流程,然后,正确设定各个空间处理所需的数据和参数信息,最终执行空间处理模型,执行空间信息的处理。

在设计 SuperMap GeoProcessor 6R 时,考虑到其适应性,跨平台问题是首先要解决的关键问题,不仅要支持 Windows 操作系统,还要能够在各种 Linux、Unix 系统上运行。至下而上分析,首先,SuperMap GIS 底层的 UGC 类库是跨平台的;接下来是在 UGC (Universal GIS Core) 的类库之上封装模型,因为 java 有良好的跨平台特性,为此 java 依据空间活动元模型的原理将 UGC 类库提供的空间操作的接口和类进行二次封装,封装成模型元素需要的类和接口;然后,在封装模型界面元素时选择了 flash 语言,在 Flash Builder 工具里,将 java 封装好的接口和类包装成用于可视化建模的模型元素,该模型所具有的信息和原理与空间活动元模型的信息和原理相同。无论是在用微软 .NET 框架下的 SuperMap 桌面 GIS 产品还是在支

持跨平台的服务器产品里,只需将 SuperMap GeoProcessor 6R 安装在指定目录下,便可以在这两个产品中进行可视化建模。SuperMap GeoProcessor 6R 的技术架构如图 2 所示。

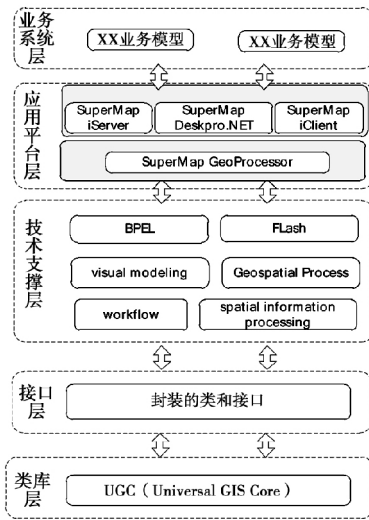


图 2 SuperMap GeoProcessor 技术架构图
Fig.2 SuperMap GeoProcessor technical architecture diagram

根据图 2 所示的技术架构,我们来分析该工具的工作原理。当在桌面或服务器产品提供的建模工具里建模完成后,点击“执行”操作时首先 Flash 会将模型翻译成一个 xml 文件,该 xml 文件是按照 BPEL(Business Process Execution Language, 业务流程执行语言)格式记录的,文件里记录了每一个可执行的操作,部分内容如图 3 所示。BPEL 是一种使用 XML 编写的编程语言,用于自动化业务流程,这些流程能够在任何一个符合 BPEL 规范的平台或产品上执行,也曾经被称作 WSBPEL 和 BPEL4WS,优点为具有可移植性。在生成该 xml 后,桌面产品或服务器产品会调用其自己开发语言 C#或 java 调用 java 进程在后台执行,使其按 xml 中的内容依次执行每一步操作,最终输出结果,完成一系列操作。



图 3 SuperMap GeoProcessor 可执行 XML 文件部分内容截图
Fig.3 SuperMap GeoProcessor's screenshot part of the executable XML file

另外,建模后的界面也可以保存成一个 xml 文件,但这个 xml 中记录的内容与图 3 所示的内容不同,界面的 XML 文件记录的是每个模型的名称和位置,如果需求变化时,可以在已有的基础上微作调整,达到复用的目的。

3 空间处理建模流程及应用

以育苗选址的一个例子来说明 SuperMap GeoProcessor 的建模流程及应用方式。现有某区域高程数据、水系数据、主要道路数据和分园分布数据,按以下要求和标准寻找适宜的位置。

1) 要选择平坦的地区,要求将坡度值 X 进行划分,坡度等级:平坡 $x \leq 5^\circ$;缓坡 $5^\circ < X \leq 15^\circ$;斜坡 $15^\circ < x \leq 25^\circ$;陡坡 $25^\circ < x \leq 35^\circ$;急坡 $35^\circ < x \leq 45^\circ$;险坡 $x > 45^\circ$ 。影响因子按照坡度等级依次赋值 5,4,3,2,1,0。该需求需要用到两个操作,坡度分析于用对原 DEM 数据进行坡度分析,栅格代数运算用于将坡度分析后的栅格数据根据此条件赋予 5,4,3,2,1,0 不同的值。

2) 要选择较好的朝向,本题认为较好的朝向为东南、南和西南,朝向值在区间 $[90, 270]$ 或者值为 -1 时,朝向影响因子为 3,其他取值为 0。该需求需要用到两个操作,坡向分析于用对原 DEM 数据进行坡向分析,栅格代数运算用于将坡向分析后的栅格数据根据条件赋予 3 或 0 值。

3) 要选择适当的高程,当高程值在区间 $[1, 350, 1, 700]$ 、 $[1, 800, 2, 150]$ 时,影响因子为 1;当高程值在区间 $(1, 700, 1, 800)$ 时,影响因子为 3,其他取值为 0。该需求需要进行三次栅格代数运算,分别按条件赋 1,3,0 值。

4) 要选择靠近湖泊的地方,要求生成湖泊的缓冲区,缓冲区半径 $D \leq 1,000$ m 的影响因子赋值 5, $1,000 < D \leq 1,500$ m 的影响因子赋值 3, $1,500 < D \leq 2,000$ m 的影响因子赋值 1,其他取值为 0。该需求需要将湖泊线数据集做多重缓冲区分析,然后将分析后的矢量数据转成栅格数据,再将区域外的数据裁剪掉,最后用栅格代数运算对栅格数据赋 5,3,1,0 值。

5) 不要靠近主要街道上,要求生成主要街道的缓冲区,缓冲区半径 $D \leq 300$ m 的影响因子赋值 0, $300 < D \leq 800$ m 的影响因子赋值 1, $800 < D \leq 1,300$ m 的影响因子赋值 2,其他取值为 3。该需求与需求 4 相同,操作步骤同上。

6) 不要选择在公园内。将以上几个操作分别产生的栅格数据集做栅格代数相加的运算,使其合并成一个栅格数据集,然后用公园的面数据与相加后产生的数据集做裁剪操作去除公园部分,最后得出的栅格数据集就可以展示出各项分析后的结果值,供育苗选址使用。

在 SuperMap GeoProcessor 里,将每个分析模型在如图 4 所示的右上角模型窗口中选择合适的模型(如坡度分析模型)拖到窗口中,然后用有方向的剪刀将其正确连接起来,在右下角窗口中设置每个模型的输入输出参数,这里要注意箭头的方向,表示处理的先后和数据流转的方向;同时也要考虑并行的问题,要确定各个功能是否能并行处理,并行固然提高效率,但要注意顺序,从本例的分析上看,前 5 个需求互相并无联系和先后次序,因此可以并行操作。建模后的界面如图 4 所示。

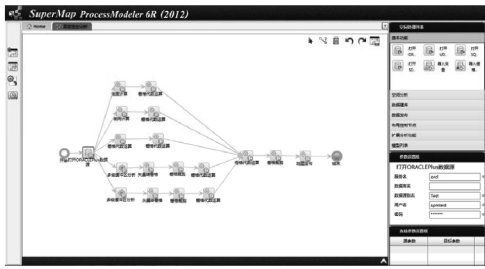


图 4 育苗选址建模图

Fig. 4 Nursery site modeling diagrams

4 结束语

空间信息处理过程是大型 GIS 应用的关键,将空间信息处理的过程做到简单、易用、高效,就解决了大型 GIS 应用的一个重要难题。研究空间信息处理过程的内在联系和规律,将可视化建模技术应用其中,再结合作流技术,使用建模工具,提供对大型 GIS 应用的分析、管理、执行和控制的支持,从而可以将 GIS 应用的分析层面、处理

(上接第 6 页)

程为导入、拼接、入库、建立索引等操作。TM 处理单幅用时 30 min,串行预计时间 440 h(约 18 d),并行化后处理时间为 3 d 左右。

从上述结果可以看出,进行并行化后,数据处理的效率得到大幅度的提升。通过硬件的满负荷运转,充分利用了系统资源,不仅降低了数据处理过程中的等待时间,成本和能源消耗都得以大幅度降低。

6 结束语

云计算将带动 GIS 技术、应用模式、部署方式都发生重大变化,高性能的分析服务和高效能的数据处理是其重要的组成部分。并行算法和架构对 GIS 运算提升效果非常明显,能够更好地满足大规模服务、海量数据处理、实时遥感与位置服务分析等应用需要。

目前,采用并行来提升 GIS 效能技术正在发展之中,尚存在需要突破的技术包括:并行算法、一致性、容错与恢复、处理模型、普适性等。预计未来几年,并行 GIS 技术将会获得大的发展,包括:采用 GPU 来提升空间分析算法效率,采用多核并行提升单机处理速度,采用异步并行来改善用户交互能力,采用分布式并行来进行高效数据处理,采用 Hadoop 进行空间数据分析与挖掘应用,采用高性能计算集群解决 3 维建模、网络分析、时空数据挖掘、传感器实时分析等复杂应用的需要。

地理空间数据处理和分析是地理信息系统建设和运行的重要环节,并行处理技术复杂、涉及层面多、不同数据的要求差异大,在此谨将个人的理解以及可能对 GIS 的影响的一些思考和实践与大家分享,希望与各位同仁继续探讨。

层面和管理层面的工作有机的结合起来,更有效规划管理 GIS 应用工作的实施,提高工作质量和效率。

参考文献:

- [1] 丁华祥,唐力明.空间处理建模技术的概念和应用[M].测绘通报,2009(1):64-67.
- [2] 黄柳青,王满红.构件中国:面向构件的方法与实践[M].北京:清华大学出版社,2006.
- [3] 何凯涛,唐宇,刘书雷.基于工作流的 GIS 服务动态聚合方法、技术体系和参考模型研究[J].国防科技大学学报,2008,30(1):47-52.
- [4] 贾文珏,李斌,龚健雅.基于 workflow 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.
- [5] 高勇,鄢伦,刘瑜.空间信息处理过程建模研究[J].北京大学学报(自然科学版),2004,40(6):914-921.

[编辑:胡雪]

参考文献:

- [1] 黄骞.解放地理信息[J].新经济导刊,2012(12):34-36.
- [2] Li, L., Li, M. F., Goodchild, and L. Anselin (2013) A geospatial cyberinfrastructure for urban economic analysis and spatial decision - making [J]. ISPRS International Journal of Geo - Information 2: 413 - 431. DOI: 10.3390/ijgi2020413. [549]
- [3] S. Wang, L. Anselin, B. Bhaduri, C. Crosby, M. F. Goodchild, Y. Liu, and T. L. Nyerges (2013) CyberGIS software: a synthetic review and integration roadmap [J]. International Journal of Geographical Information Science DOI: 10.1080/13658816.2013.776049. [546]
- [4] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and Matei Zaharia (2009) Above the Clouds: A Berkeley View of Cloud Computing [J]. Technical Report No. UCB/EECS-2009-28
- [5] 黄骞,王尔琪.基于云计算的空间关联信息管理与服务技术研究[J].测绘与空间地理信息,2011,34(S0):68-72.
- [6] Yan Liu, Kaichao Wu, Shaowen Wang, etc. A MapReduce approach to $G_i^*(d)$ spatial statistic [J]. HPDGIS '10: Proceedings of the ACM SIGSPATIAL International Workshop on High Performance and Distributed Geographic Information Systems 2010.
- [7] 王尔琪,黄骞,滕寿威.云计算模式与 GIS 技术创新[J].测绘与空间地理信息,2011,34(S0):5-13.

[编辑:王悦辉]