

# CyberSoLIM: 基于知识驱动的在线数字土壤制图原型系统\*

江净超<sup>1,2</sup> 朱阿兴<sup>1,3†</sup> 秦承志<sup>1</sup> 杨琳<sup>1</sup> 刘峰<sup>4</sup> 张淑杰<sup>1,2</sup>

(1 资源与环境信息系统国家重点实验室(中国科学院地理科学与资源研究所),北京 100101)

(2 中国科学院大学,北京 100049)

(3 Department of Geography, University of Wisconsin-Madison, Madison, WI 53706, USA)

(4 土壤与农业可持续发展国家重点实验室(中国科学院南京土壤研究所),南京 210008)

## CyberSoLIM: KNOWLEDGE-DRIVEN ON-LINE DIGITAL SOIL MAPPING PROTOTYPE SYSTEM

Jiang Jingchao<sup>1,2</sup> Zhu A-Xing<sup>1,3†</sup> Qin Chengzhi<sup>1</sup> Yang Lin<sup>1</sup> Liu Feng<sup>4</sup> Zhang Shujie<sup>1,2</sup>

(1 State Key Lab of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China)

(2 University of Chinese Academy of Sciences, Beijing 100049, China)

(3 Department of Geography, University of Wisconsin-Madison, Madison, WI 53706, USA)

(4 State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, CAS, Nanjing 210008, China)

**关键词** 数字土壤制图;知识驱动;在线应用系统

**中图分类号** P209 **文献标识码** A

土壤属性空间分布信息是进行土壤侵蚀、水文、生态和养分循环等地质模拟研究的重要基础信息,数字土壤制图是以定量土壤—景观模型为理论基础、以空间分析和数学方法为技术手段的土壤调查与制图方法,为快速地获取高精度、高分辨率的土壤分布信息提供了有效的途径<sup>[1-2]</sup>。数字土壤制图涉及比较复杂的土壤推测模型及数据准备过程,用户实际应用时通常需要借助特定的工具软件<sup>[2-3]</sup>。根据其适用范围可将数字土壤制图软件分为两类:(1)通用的地统计软件工具,如 Gstat<sup>[4]</sup>、ArcGIS Geostatistical Analyst<sup>[5]</sup>、GS plus<sup>[6]</sup>等;(2)专门用于数字土壤制图的软件,如 FuzME<sup>[7]</sup>、SoLIM Solutions<sup>[8]</sup>、ArcSIE<sup>[9]</sup>、TAL<sup>[10]</sup>等。上述工具在应用时存在如下主要问题:建模过程主要依赖于用户专业知识,导致非专家用户难以完成;软

件的实现通常为单机版,用户安装配置较为繁琐,且算法库不易扩展,难以利用网络上越来越多封装为 Web Service 的模型和算法;所采用的串行算法运行效率低下。针对上述问题,迫切需要构建一种新型的数字土壤制图工具,以满足用户对简易和高效制图的需求。

本文综合知识驱动、开源 Web GIS、Web Service 等技术,设计并完成了基于知识驱动的在线数字土壤制图原型系统(CyberSoLIM)的构建。

## 1 CyberSoLIM 系统的设计与构建

### 1.1 在线式的数字土壤制图

现有数字土壤制图软件通常是单机版,要求用户在本地计算机安装后才能使用,一些大型平台型

\* 国家 863 计划项目(2011AA120305)、国家科技支撑计划课题任务(2013BAC08B03-4)、中国科学院地理资源研究所优秀青年人才基金项目(2011RC203)、科技部国际科技合作项目(2010DFB24140)共同资助

† 通讯作者, E-mail: axing@reis.ac.cn

作者简介:江净超(1986—),男,河北邢台人,博士研究生,主要从事 Web GIS 和地学智能建模研究。E-mail: jiangjc@reis.ac.cn

收稿日期:2013-03-04;收到修改稿日期:2013-05-22

的软件工具配置较为繁琐。如采用浏览器/服务器(B/S)模式,用户借助浏览器便可以完成制图操作,无需安装任何插件,最大限度地降低用户使用门槛。基于开源 Web GIS 地图服务器 MapServer<sup>[11]</sup>将土壤推测模型计算得到的土壤属性信息进行发布,并交由开源 Web GIS 客户端 Openlayers<sup>[12]</sup>在浏览器中显示,用户不必借助其他软件进行土壤属性信息的可视化。

通过将数字土壤制图的土壤推测模型和数据准备算法封装为 ASP.NET Web Service,可具有良好的跨平台、互操作和扩展性,同时还可以快速集成网络上封装为 Web Service 的其他模型和算法,便于支持更为复杂的应用模型的构建。

## 1.2 知识驱动下的应用模型的建立

数字土壤制图中的土壤推测模型最常用的输入数据是定量刻画地形环境要素的各种地形属

性<sup>[13]</sup>,这些地形属性通常是由栅格数字高程模型(DEM)通过各种数字地形分析算法计算得到<sup>[1,14]</sup>。数字土壤制图应用模型运行前重要的一部分工作是设定数字地形分析 workflow。数字地形分析 workflow 的设定可以看作是根据已有数据和目标数据确定所需数字地形分析算法和算法执行顺序的过程,对用户提出了较高的专业要求,为实际应用带来不便<sup>[15]</sup>。

为降低用户的建模难度,本文采用知识驱动<sup>[16]</sup>的方式辅助用户完成算法选择和算法执行顺序的确定。知识驱动建模的前提是建模知识的显式表达。工作流建模知识为算法与其输入输出的关系,即特定的算法具有特定的输入和输出,本文以资源描述框架(RDF)三元组<sup>[17]</sup>(即主体、谓词和客体)的方式表达此类知识,RDF 具体实例如图 1 所示。

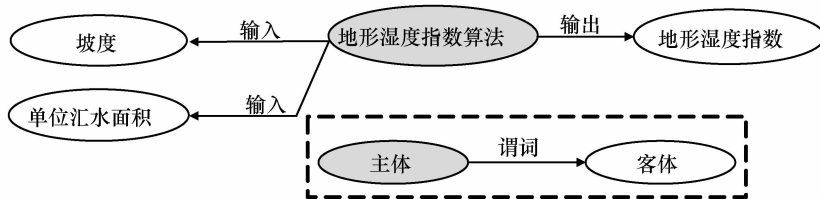


图 1 资源描述框架的“主体-谓词-客体”三元组示意图

知识驱动建模的推理过程如下列步骤:

步骤(1):用户指定目标数据  $T$ ;

步骤(2):在 RDF 知识库中查找输出为  $T$  的算法  $A_i (i=1, 2, \dots, n; \text{其中 } n \text{ 为算法的个数})$ ;

步骤(3):对于  $A_i$  的每个输入  $I_j (j=1, 2, \dots, m; \text{其中 } m \text{ 为每个算法的输入项的个数})$ ,提示用户做如下判断:已有数据是否与  $I_j$  相匹配,如果相匹配,则针对  $I_j$  的推理停止;否则,将  $I_j$  作为目标数据  $T$ ,回到步骤 2 继续推理,直至  $A_i$  的所有输入均与已有数据相匹配。

该建模过程不需要用户掌握专业知识,用户仅需要指定目标数据、判断系统生成的当前模型所需的各种输入是否均已有数据匹配。建模过程结束后得到算法序列  $A_n, A_{n-1} \dots A_2, A_1$ ,即为由已有数据得到目标数据  $T$  所需的算法及其执行顺序。

## 1.3 土壤推测模型的并行化

传统的串行计算技术已经不能满足用户对快速数字土壤制图的需求。计算机已经进入多核时代,并行计算技术可以使用多个处理器完成同一任务,有效加快问题的求解速度,为提高数字土壤制

图的生产效率提供了一条有效的途径。

土壤推测模型众多,如基于模糊数学<sup>[18-19]</sup>、专家知识<sup>[20]</sup>、决策树<sup>[21]</sup>、神经网络<sup>[22]</sup>等。其中,基于样点个体代表性的土壤属性推理模型突破了现有方法对样点集全局代表性的严格要求,能够利用任意样点集进行制图,并且可以度量推测的不确定性,实际应用表明该模型可取得很好的推测效果<sup>[23]</sup>。

本文以基于样点个体代表性的土壤属性推理为例,利用 .NET 4.0 Framework 提供的共享存储器模式的多线程任务并行编程库 TPL<sup>[24]</sup>对其进行了并行化实现。

## 2 CyberSoLIM 系统的特点与案例研究

### 2.1 CyberSoLIM 系统架构

综合上述的开源 Web GIS、Web Service、知识驱动、并行计算等技术,本文设计并初步实现了基于知识驱动的在线数字土壤制图原型系统——CyberSoLIM。

CyberSoLIM 采用 B/S 三层架构,如图 2 所示,整个架构分为表现层、应用层和数据层。

(1)表现层:基于浏览器的富客户端,不需要安装任何插件。土壤属性信息显示功能用以显示数字土壤制图的结果;可视化建模功能实现制图流程的图形化,并结合应用层中的知识驱动建模实现流程自动构建。

(2)应用层:响应浏览器的制图请求,并返回制图结果。知识驱动建模模块完成应用模型的构建,生成执行脚本;Web 服务匹配模块根据执行脚本匹配相应的 Web 服务,创建 Web 服务链;工作流引擎 Apache ODE 调用 Web 服务链,完成数字土壤制图;土壤属性信息发布模块负责将制图结果交由浏览器端显示。

(3)数据层:包括数据库、知识库和 Web 服务库三部分,数据库存储制图区边界、土壤环境要素和代表性样点数据;知识库存储知识驱动建模所用的算法知识;Web 服务库包含了土壤推测模型和数字地形分析算法的 WebService。

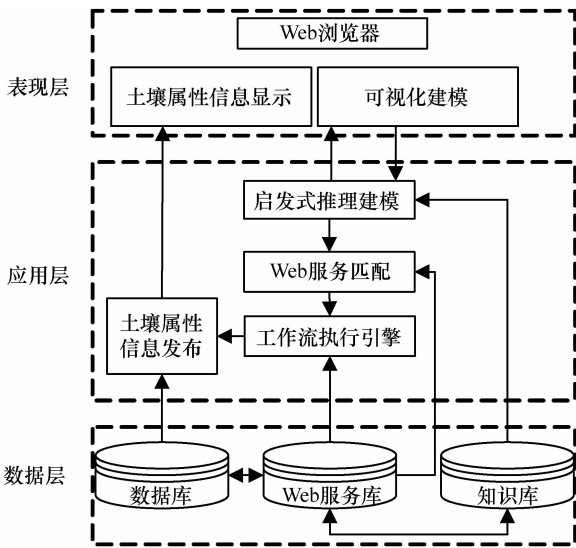


图 2 CyberSoLIM 系统架构

## 2.2 运行环境和主要功能

服务器端的硬件环境:4 路 Intel 6 核处理器,内存大小为 32G,硬盘空间为 2.7TB;软件环境:64 位 Windows Server 2008 操作系统、Tomcat6.0 应用服务器、MapServer 服务器、Apache ODE BPEL 执行引擎、IIS7.0 Web 服务器。客户端支持各种主流的浏览器,如 IE、Chrome、Opera 和 Firefox 等。

CyberSoLIM 提供了基于知识驱动的智能可视化数字土壤制图界面,用户以在线方式通过浏览器

进行土壤制图。系统的主要功能包括:兼容多种浏览器的在线制图、样点和环境要素数据的上传下载、启发式制图应用建模、推测模型参数设置、基于并行计算的高性能推测制图、制图成果发布与共享等。

## 2.3 CyberSoLIM 应用的案例研究

本文使用 CyberSoLIM 以安徽宣城为研究区进行数字土壤制图作业,推测 0~20cm 土壤黏粒含量的空间分布信息,以此案例来讨论 CyberSoLIM 的性能和实用性。

研究区位于安徽省东南部宣城的三个区县(宣州区、郎溪县、广德县),其地理位置为  $48^{\circ}53'24'' \sim 48^{\circ}59'24''N, 125^{\circ}8'24'' \sim 125^{\circ}16'12''E$ ,面积约  $5\ 900\text{ km}^2$ 。已有的环境要素数据包括降水 (Precipitation)、气温 (Air temperature) 和母质 (Parent material)。需要通过知识驱动建模得到的环境要素数据包括 4 个地形属性:地形湿度指数 (Topographic Wetness Index, TWI)、坡度 (Slope gradient)、沿剖面曲率 (Profile curvature) 和沿等高线曲率 (Plan curvature)。上述环境要素数据格式为 ASCII,分辨率为 90 m,栅格数目为  $1\ 242 \times 928$  (行数  $\times$  列数)。土壤样点个数为 64 个,记录为 CSV 文件格式。

用户使用 CyberSoLIM 对研究区 0~20cm 土壤黏粒含量进行在线数字土壤制图的基本过程如下:(1)通过浏览器访问 CyberSoLIM;(2)选择目标研究区并启动土壤属性制图指令(土壤属性选择 0~20cm 土壤黏粒含量,其他参数使用系统默认值);(3)系统自动在知识驱动下启发式地完成应用模型的构建,包括对模型所需输入的各种地形属性数据(如 TWI、Slope gradient 等)的计算过程建模(如图 3 所示),该过程中计算工作流的建立只要求用户指定目标数据和输入数据,不需其他额外的干预;(4)运行应用模型,得到 0~20cm 土壤黏粒含量的空间分布信息(如图 4 所示)。

上述应用实例证明了 CyberSoLIM 在制图实践中的可行性。此外,由于所用的土壤推测模型实现了并行化,在开启多个线程的情况下,与串行程序(单进程)相比,可缩短制图时间,提高效率。由表 1 可知,在 4 个线程时,该应用案例的土壤推测模型运行时间约为串行程序的 0.4 倍。在开启超过 4 个线程时,运行时间继续缩短,但不太明显,因此系统默认情况下开启 4 个线程。

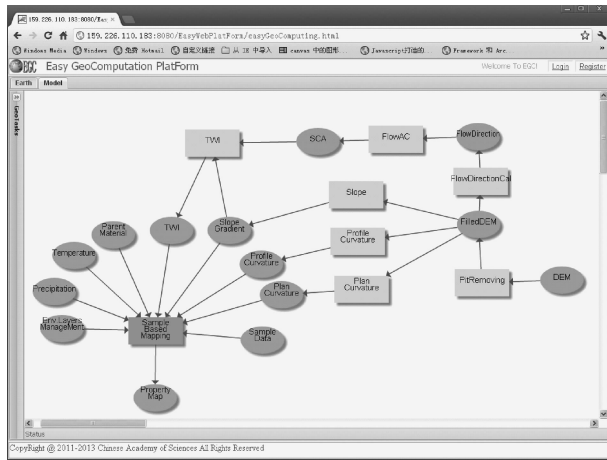


图3 数字土壤制图流程构建

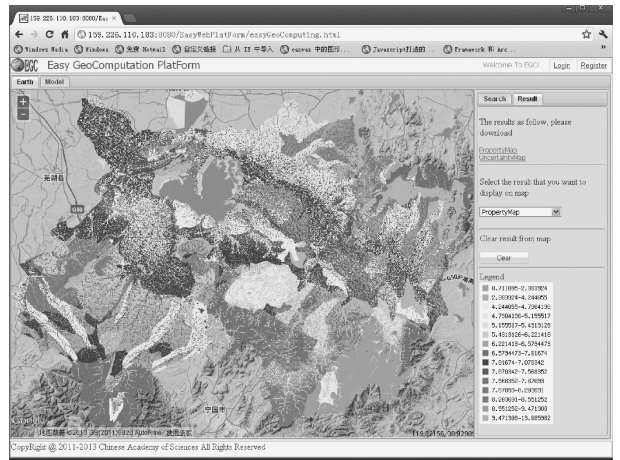


图4 土壤黏粒含量的空间分布

表1 制图时间随系统线程数目的变化情况

线程数目(个)	运行时间(秒)	线程数目(个)	运行时间(秒)
1	252	6	76
2	149	7	66
3	131	8	59
4	108	9	53
5	89	10	50

CyberSoLIM 基本实现了在线式的数字土壤制图功能,借助网络可使更多的用户以便捷、智能的方式制图,在一定程度上摆脱了对单机软件的依赖。CyberSoLIM 的服务器端由于采用了并行计算技术,具有较传统的串行程序更高的性能,使用户获得了更好的制图体验,与现有的数字土壤制图工具相比,具有在线制图、免安装配置、不依赖用户建模知识等优点(表2)。

表2 CyberSoLIM 与部分现有数字土壤制图工具的对比

土壤制图工具	支持在线制图	安装配置	依赖知识	性能
地统计软件(如 Gstat)	否	较复杂	强烈	一般(串行计算)
专用软件(如 FuzME)	否	易	强烈	一般(串行计算)
CyberSoLIM	是	否	很少	高(并行计算)

### 3 结论

本文针对现有数字土壤制图工具存在的问题,设计并初步实现了基于知识驱动在线数字土壤制图原型系统 CyberSoLIM。应用实例表明,通过采用 CyberSoLIM,用户无需安装数字土壤制图软件和

任何插件,借助浏览器便可完成制图,从而有效降低用户的使用门槛;基于知识驱动建模使得用户无需掌握地形环境要素数据准备方面的知识,从而降低用户制图的建模难度。应用门槛的降低也使得 CyberSoLIM 为非土壤专业的用户(如相关管理部门乃至公众)提供了一个使用和参与数字土壤制图的简易快捷方式。

致谢 感谢北京师范大学学卢岩君和张桂铭在系统开发中给予的无私帮助。

### 参考文献

- [1] 朱阿兴,李宝林,裴韬,等. 精细数字土壤普查模型与方法. 北京:科学出版社,2008
- [2] 孙福军,雷秋良,刘颖,等. 数字土壤制图技术研究进展与展望. 土壤通报,2011,42(6): 1502—1507
- [3] Rossiter D G. A compendium of online soil survey information. [http://www.itc.nl/~rossiter/research/rsrch\\_ss\\_sw.html](http://www.itc.nl/~rossiter/research/rsrch_ss_sw.html), 2013—02—04
- [4] Pebesma E J. Gstat. <http://www.gstat.org/>, 2013—02—04
- [5] ESRI. ArcGIS geostatistical analyst. <http://www.esri.com/software/arcgis/extensions/geostatistical>, 2013—02—04
- [6] Gamma Design Software. Geostatistics for the environmental sciences, GS + 7.0. Gamma Design Software LLC, Plainwell, Michigan, USA, 2007
- [7] Minasny B, McBratney A B. FuzME. <http://sydney.edu.au/agriculture/pal/software/fuzme.shtml>, 2013—02—04
- [8] Zhu A X. SoLIM solutions. <http://solim.geography.wisc.edu/software/index.htm>, 2013—02—04
- [9] Shi X. ArcSIE. <http://www.arcsie.com/index.htm>, 2013—02—04
- [10] Sung C T B. Texture autoLookup (TAL). <http://www.christopherherteh.com/tal/index.html>, 2013—02—04
- [11] 宋现锋,刘军志,吴建国,等. 开源代码技术的 FLASH 地图

- 实现方法. 地球信息科学, 2006, 8(4): 101—105
- [12] OpenLayers. OpenLayers: Free maps for the Web. <http://www.openlayers.org/>, 2013—02—04
- [13] McBratney A B, Mendoca Santos M L, Minasny B. On digital soil mapping. *Geoderma*, 2003, 117(1/2): 3—52
- [14] 秦承志, 卢岩君, 包黎莉, 等. 简化数字地形分析软件(Sim-DTA)及其应用——以嫩江流域鹤山农场的坡位模糊分类应用为例. 地球信息科学学报, 2009, 11(6): 737—743
- [15] Qin C Z, Lu Y J, Zhu A X, et al. Software prototyping of a heuristic and visualized modeling environment for digital terrain analysis// 11th international conference on GeoComputation. University College London, UK, 2011
- [16] Lu Y J, Qin C Z, Zhu A X, et al. Application-matching knowledge based engine for a modelling environment for digital terrain analysis// GeoInformatics 2012. The Chinese University of Hong Kong, 2012
- [17] Antoniou G, Harmelen F. 陈小平译. 语义网基础教程. 北京: 机械工业出版社, 2008
- [18] 杨琳, 朱阿兴, 李宝林, 等. 应用模糊  $c$  均值聚类获取土壤制图所需土壤—环境关系知识的方法研究. 土壤学报, 2007, 44(5): 784—791
- [19] Zhu A X, Yang L, Li B L, et al. Construction of membership functions for predictive soil mapping under fuzzy logic. *Geoderma*, 2010, 155(3/4): 166—174
- [20] 朱阿兴, 李宝林, 杨琳, 等. 基于 GIS 模糊逻辑和专家知识的土壤制图及其在中国应用前景. 土壤学报, 2005, 42(5): 844—851
- [21] McBratney A B, Odeh I O A, Bishop T F A, et al. An overview of pedometric techniques for use in soil survey. *Geoderma*, 2000, 97(3/4): 293—327
- [22] Minasny B, McBratney A B. The neuro-m method for fitting neural network parametric pedotransfer functions. *Soil Science Society of America Journal*, 2002, 66(2): 352—361
- [23] 刘京, 朱阿兴, 张淑杰, 等. 基于样点个体代表性的大尺度土壤属性制图方法. 土壤学报, 2013, 50(1): 12—20
- [24] Campbell C, Johnson R, Miller A, et al. 曹译文, 邹雪梅, 李岸译. 设计模式——. NET 并行编程. 北京: 清华大学出版社, 2012

(责任编辑: 檀满枝)