

引文格式:侯东阳.地表覆盖网络化信息发现方法研究[J].测绘学报,2017,46(1):133. DOI:10.11947/j.AGCS.2017.20160480.
HOU Dongyang.Method of Land Cover Web Information Discovery[J]. Acta Geodaetica et Cartographica Sinica,2017,46(1):
133. DOI:10.11947/j.AGCS.2017.20160480.

地表覆盖网络化信息发现方法研究

侯东阳^{1,2,3}

1. 山东师范大学地理与环境学院,山东 济南 250014; 2. 国家基础地理信息中心,北京 100830; 3. 中国矿业大学环境与测绘学院,江苏 徐州 221116

Method of Land Cover Web Information Discovery

HOU Dongyang^{1,2,3}

1. College of Geography and Environment, Shandong Normal University, Jinan 250014, China; 2. National Geomatics Center of China, Beijing 100830, China; 3. School of Environment Science and Spatial Informatics, China University of Mining and Technology, Xuzhou 221116, China

地表覆盖及其变化是全球变化研究、生态环境评估以及地理国情监测的重要基础信息。近年来,通过网络分发和共享,形成了由地表覆盖数据、相关辅助数据及与地表覆盖变化相关的事件信息组成的地表覆盖网络化信息。地表覆盖网络化信息的发现是开展相关领域科学研究和研制协作服务平台的重要工作。然而,日益增多的地表覆盖网络化信息存在分散、互不联通的特点,致使其发现和获取较为困难。尽管其他领域已有许多网络化信息发现方法,但是在地表覆盖领域存在着地名主题词普通化、时间分布平等化及信息搜索表层化的问题。为此,本文从其地名主题词特殊化、时间意图明确化和信息搜索双层化的特殊要求出发,提出了地表覆盖网络化信息发现的研究方法。本文的主要研究内容如下:

(1) 提出了基于地名空间特征的双向量主题表示及相关度计算方法,解决了地名主题词普通化问题;从地名及其空间关系的作用出发,将地名从传统单向量中分离出来,形成由地名和普通关键词独立组成的双向量;然后,借鉴地理信息检索中分别判断查询词关键词相关度和查询空间范围相关度的思想,分步计算和判断普通主题相关度、地名相关度和综合相关度;其中,在计算地名主题相关度时引入了相等、相交、包含、被包含和相离5种空间关系因子,用于传递关联地名的权重。基于此,发展了一种基于地名空间特征的地表覆盖网络化信息发现方法,实现了地名主题词特殊化的要求,其查准率和 F-Score 值分别提高了约 10%和 7%。

(2) 提出了顾及时间意图的主题相关度计算和 URL 优先级分配方法,解决了时间分布平等化问题;首先,提出了基于 Google 趋势数据的时间意图识别方法,用于估算主题的起始时间和量化其时间分布;然后,分步计算时间相关度和普通主题相关度,用于强化起始时间的作用;接着,将时间分布量化值以指数函数的形式融入到传统 URL 优先级分配方法中,实现了时间分布优先化的要求;

并基于此发展了一种顾及时间意图的地表覆盖网络化信息搜索方法,其查准率平均提高了约 10%。

(3) 提出了面向深层网络的地表覆盖网络化信息获取方法,解决了信息搜索表层化的问题;其是在传统主题爬行的基础上,利用关键词或规则匹配来识别和理解脚本信息源及其内容,用于获取隐藏在脚本内容中的深层地表覆盖网络化信息;基于此分别提出了面向深层网络的地理标记文本搜索方法和地表覆盖 Web 地图服务发现方法,满足了地表覆盖验证样本主动采集和地表覆盖 Web 地图服务发现的需要;结果表明,只需修改脚本识别和脚本内容理解的关键词或规则即可主动发现不同类型的深层网络信息。

(4) 研制了地表覆盖网络化信息发现系统。根据 GlobeLand30 在线信息服务平台和全球地理信息协作平台的需要,基于论文提出的方法研制了地表覆盖文本信息搜索子系统和地表覆盖 Web 地图服务发现子系统,实现了地表覆盖网络化信息的发现、检索及地图可视化的功能。

中图分类号:P228 文献标识码:D
文章编号:1001-1595(2017)01-0133-01
基金项目:国家自然科学基金(41231172)

收稿日期:2016-09-28

作者简介:侯东阳(1986—),男,2016年6月毕业于中国矿业大学,获工学博士学位(指导教师:陈军教授),研究方向为时空网络数据搜索、挖掘与分析。

Author: Hou Dongyang (1986—), male, received his doctoral degree from China University of Mining and Technology on June 2016, majors in spatio-temporal Web data discovery, mining and analysis.

E-mail: houdongyang1986@163.com