

基于 Web 文本的灾害信息挖掘研究进展*

李卫江, 温家洪

(上海师范大学 地理系, 上海 200234)

摘要: Web 文本中蕴含着丰富的以自然语言描述的非结构灾害信息和知识。基于 Web 文本自动提取和构造结构化、综合性灾害信息, 是灾害信息领域研究的前沿问题。目前国内外利用 Web 文本挖掘技术在灾时与灾后的应急响应与救援, 灾害的早期预警和风险分析方面进行示范应用; 同时在文本灾害信息的语义理解与抽取、文本灾害信息的时空匹配、以及文本灾害信息的不确定性和可靠性评价等关键技术领域迅速展开研究。我国应加强以 Web 文本为信息源的中文灾害信息挖掘关键技术、软件, 以及管理体系的研究, 以有效弥补灾害研究与管理过程中灾害数据共享困难, 以及可利用的动态实时、综合性灾害数据缺乏的薄弱环节, 提升灾害信息服务水平。

关键词: Web 文本; 灾害信息; 空间信息; 挖掘技术

中图分类号: X4 **文献标识码:** A **文章编号:** 1000-811X(2010)02-0119-05

0 引言

充足、准确、及时的灾害信息, 对于减少灾害风险具有重要意义^[1]。目前, 世界上很多国家、城市、研究和管理部门都在根据自己的需要致力于灾害专业数据库的建设, 著名的如全球历史灾害数据库 DesInventar(<http://www.desinventar.org/>)、全球级别的紧急灾害数据库 EM-DAT(<http://www.emdat.be/>)、慕尼黑再保险公司灾害数据库 NatCat(<http://www.munichre.com/>)、瑞士再保险公司数据库 Sigma(<http://www.swissre.com/>)、国内建设的中国自然灾害数据库等^[2]。但是, 由于灾害数据标准、数据来源的可靠性与广泛性、收录数据标准界定、数据管理范式(包括灾害特征类、字段名称、对应数据类型等规范的确定的)等的差异性, 很难实现灾害信息的应用层面共享^[3]。此外, 这些数据库在区域尺度、时间尺度、信息的精度、信息的时效性、信息条目的全面性等方面, 与具体的灾害应用需求(特别是非常规突发事件)往往存在较大的偏差。灾害信息收集的迟缓并由此造成的决策迟缓是一系列灾害管理中存在的主要问题。

海量的历史文本资料和 Web 文本资料成为灾害信息的重要来源。近年来, 随着语义 Web、Web2.0 等互联网新技术体系的出现, 一些新的网络文本如政府报告、新闻、博客、论坛、Google Earth、Wiki 等, 使 Web 日益成为灾害信息发布、交互的平台。以大量 Web 文本为灾害数据源, 自动抽取并整理包含灾害事件主题、时间、空间位置以及范围、直接损失及次生影响、致灾原因等内容的结构化的涉灾信息, 对传统结构化灾害数据库具有重要的补充意义^[4-5]。从灾害发展阶段看, 灾前、灾中与灾后均需要足够的规划、评价、应急与恢复和重建的现势与历史涉灾信息^[6]。一方面, 通过对区域范围内大量积累的, 文本语言记录的灾害历史资料的分析 and 挖掘, 有利于获取灾害事件发生发展的系统记录, 发现灾害事件存在的联系, 进而对灾害发生的类型、频度、时空分布特征进行分析, 为不同地域空间尺度的单灾种风险区划与综合灾害风险区划提供系统、准确的数据支持。另一方面, 在突发灾害事件下, 需要针对事件可能前兆和演化过程中的海量、异构、实时数据, 进行快速收集获取、整理、可视化分析和共享等处理, 需要把分布在不同系统、不同部门的涉灾信息全面、准确地挖掘出来, 并整合

* 收稿日期: 2009-10-15

基金项目: 上海市教委重点学科《地理学与城市环境》(J50402); 上海师范大学科研项目(SK200846); 上海师范大学重点培育学科项目(DZL801); 上海市科委科技攻关项目(08240514000)

作者简介: 李卫江(1976-), 男, 河南灵宝市人, 博士, 副教授, 主要研究方向: 空间信息挖掘与共享, 灾害信息学。

E-mail: Lwj@shnu.edu.cn

成可供决策参考的综合信息。文本挖掘提供了应急管理从不确定与半/非结构化资源中抽取现势的灾害信息的方法。

目前,传统的灾害数据库主要依赖于各级政府部门借助移动技术、网络技术,通过人工方式逐级上报和汇总信息,体系庞大,信息采集成本较高,信息的现势性也相对较差。文本挖掘技术则以庞大的各社会个体为灾害信息采集的“传感器”,以海量网络文本为信息源,以自动化挖掘为主要技术手段,实现灾害数据的抽取、筛选、鉴别和长期积淀。作为一种新兴的灾害信息获取手段,文本挖掘技术能够有效弥补传统灾害数据库建设方式的不足。

1 国内外发展现状与研究进展

1.1 应用现状

近年来,以 Web 文本为数据源,进行灾害信息挖掘的相关研究和示范应用已逐步展开。归纳起来,应用的重点集中在两个方面:灾时与灾后的应急响应与救援,以及灾害的早期预警。

1.1.1 灾时与灾后的应急响应与救援

灾害发生后,精确、及时、可靠的灾情信息,以及可利用的赈灾资源信息对救援决策起到关键性的作用。如何利用自动化文本信息提取技术和系统,把海量的、异构的信息转换为可供决策的行动信息,显得尤为重要。

如美国国土安全部的 DisasterHelp (<https://www.disasterhelp.gov/>) 一站式网站,在灾害发生后,以美国政府各部门最新的灾害形势报告为信息源,通过文本挖掘、过滤和分析,为各机构和团体提供及时有效的形势判断信息和各类赈灾资源的空间位置信息,以辅助各部门协同应急救援。目前该系统能够分析和处理的主要灾害种类有:自然灾害(洪涝、飓风、雷暴、龙卷风、暴风雪和极度低温、极端高温、地震、火山爆发、山体崩塌与泥石流、海啸、林火)、技术灾害(危险化学品泄露、日用化学品紧急事件、核电站紧急事件)和恐怖事件(爆炸、生物威胁、化学威胁)等。此外,美国计算机专家组建立的 AlertPedia (<http://www.alertpedia.com/>) 与 DisasterHelp 具有类似的功能。

由美国国家科学基金会(NSF)资助的 RESCUE 计划 (<http://www.itr-rescue.org/>),主要目标是在应对各类突发自然和人为灾害时,从根本上提升相关组织机构在收集、管理、分析和分发涉灾

信息方面的能力。该系统在灾害发生后,利用 SA (Situational Awareness) 技术,能够从灾时和灾后的各种不同的数据模型中(语音、文本、视频等),快速获取受灾人口信息(位置、统计信息)、救灾物资信息(食物、饮用水、避难场所),以及灾害事件的进程,然后根据获得的灾害信息进行灾害损失评估^[7]。以文本数据为例,该系统能够从各种网站搜索和下载相关网页信息,数据来源包括新闻、博客、卫星影像、政府部门的专业数据库等,下载的网页文本存在一个 IBM 的 DB2 数据库中,然后利用 XAR 软件对网页数据进行信息抽取和信息鉴别,并通过各种空间定位程序进行灾害信息的可视化。

英国工程和自然科学研究委员会(EPSRC)资助的 Armadillo e-Response 研究计划 (<http://e-response.org/>),致力于利用文本分析技术,研发一系列灾害信息挖掘工具,为灾害管理提供及时的、精确的涉灾信息。以伦敦地区应用为例,当紧急事件发生后,伦敦联合紧急事务控制中心(JESCC)就会启动该系统的 EmergencyTrigger Service、TripleStorageService、PostcodeDataService、LocationCentredSearchService、External Web Search Services、URIFinderService、DocumentCacheService、URICrawlerService、URIClassifierService、FastName Extractor、FastTelephoneExtractor 等功能模块,通过对灾害事件的空间定位及受灾区域范围界定、区域 Web 资源搜索、Web 文本分类、名字及电话号码关键信息抽取,以及涉灾信息的深度挖掘等任务处理,得到一个内容广泛的结构化灾害知识库,以辅助紧急响应决策^[8]。

为了落实 2005 年联合国世界减灾大会(WCDR)提出的《2005-2015 年兵库行动框架》,日本科学家 Hiroyuki Kameda 教授发起开展 DRH (<http://drh.edm.bosai.go.jp/>) 项目,提出基于 Web 建设“减灾技术与知识共享平台”的设想。在该平台中,关注了从非结构化资源中获取实时灾害信息的技术^[9],并极力倡导志愿者通过 Web 向共享平台提供各类灾害信息。

IBM 公司负责开发的开源 SAHANA 赈灾管理系统 (<http://www.sahana.lk/>),涉及失踪人员登记和管理、遇难人员登记和管理、避难所管理、志愿人员管理、援助需求管理、基于地图信息的急救方案信息、物资库存管理,以及信息报告管理等功能模块。SAHANA 具备 Web2.0 的特性,采用开放的用户注册和灾害信息采集手段。凭借其具备的文

本分析、灾害信息追踪、灾难态势地理信息可视化、同步及决策支持功能, SAHANA 在 2005 年斯里兰卡海啸、2005 年美国飓风、2006 年菲律宾南莱特岛泥石流、2006 年印度尼西亚 Yogyakarta 地震, 以及 2008 年汶川地震等重大灾害中发挥了重要作用, 有效地提高了救灾和援助的效率。

尹章才^[10]针对目前地震应急响应中灾情信息收集慢、费用大等问题, 提出基于 PPGIS 的社会化震害信息实时获取模型。以实现快速验证和修正基于灾害经验给出的地震影响和灾情评估结果, 为地震的应急指挥决策提供更可靠的基础信息。

1.1.2 灾害的早期预警和风险分析

目前, 以 Web 文本为信息源和利用文本挖掘工具, 进行流行病疫情的监控和预警方面应用较多。例如, 世界卫生组织(WHO)为了早期发现全球流行疾病的爆发, 开发了 GPHIN (<http://www.phac-aspc.gc.ca/>), 它能综合利用网络搜索、数据挖掘、自动翻译、自动化过滤等多项技术, 用于发现和跟踪重大公共卫生事件, 实现基于互联网的实时、早期风险预警。目前, 世界卫生组织、美国疾病防治中心、美国陆军医学情报中心都订购 GPHIN 数据库的服务, 联合国和世界卫生组织获得的大约 75% 的传染性病情报都来自 GPHIN^[11]。由国际传染病协会维护的 ProMED-PLUS 系统, 能够针对世界性的传染性流行病灾情的爆发, 从纯文本的医学病例报告中自动抽取病情信息, 并更新到数据库^[12]。美国 MITRE 公司开发 MiTAP 利用 Web 文本挖掘技术和自然语言处理技术, 进行 SARS 疫情的探测、监测和分析^[13]。美国 Clark Freifeld 和 John Brownstein 等开发的 HealthMap (<http://www.healthmap.org/>) 是一个自动查询、过滤、图像化文本报告的系统, 它能够把散布在世界各地的数据源, 通过网络集合在一起, 形成一个统一、综合的、关于目前全球传染病及其造成的人口和动物死亡的实时的地理分布视图。HealthMap 的数据库来源包括新闻资源(如谷歌新闻)、信息采集员, 以及官方预警信息(如世界卫生组织)。开源 Epi SPIDER (<http://www.epispider.org/>) 是一个基于网络的互动的流行病信息处理系统, 它通过邮件处理、新闻文本处理、GIS 空间定位, 直观展示全球流行病疫情信息, 有效提高对可威胁全球健康的新发传染病的监测能力。

此外, 国内相关灾害研究机构, 如北京师范大学在“十一五”国家科技支撑计划重点项目“综合风险防范(IRG)关键技术研究及示范”的支持下,

启动了中文智能化综合灾害信息检索平台的研发。该平台已先后试用于南方雪灾、四川汶川地震的灾后信息分析中。

1.2 关键技术研究进展

Web 文本中灾害信息的描述以非结构化的自然语言为主, 直接从文本中提取满足需求的有用灾害信息和知识, 是目前研究的前沿和热点问题。在文本灾害信息挖掘应用快速开展的同时, 仍有若干关键技术问题需要进一步研究。

1.2.1 文本灾害信息的语义理解和抽取

文本灾害信息的语义理解和抽取, 重点是解决文本语言信息的形式化问题, 即建立模糊的、定性的语言或认知与定量化的计算机模型之间的联系。因此, 需要综合利用 Web 挖掘技术、自然语言处理等技术, 按照一定的规则, 从文本中抽取有关灾害事件的命名实体, 如灾害事件、时间、空间位置、损失及影响、致灾原因等关键信息, 然后根据命名实体之间语义关系的推理和模拟, 来确定实体之间的关系, 进而达到对整个灾害事件信息的理解。命名实体识别、实体关系识别的难点在于中文分词和语义理解规则二个关键技术。

中文自然语言分词方法主要有机械匹配法、语义分词法和人工智能法 3 种类型^[14]。其中, 机械匹配法, 又称为字符串匹配法, 易于实现, 实际应用也最为广泛, 但难以处理未登录词, 无法有效克服歧义切分。语义分词法和人工智能法是近年发展起来的较为理想的分词方法, 它们在分词中引入了语义理解、人工智能等技术, 切分精度高, 但是消耗时间和资源较大, 实现的难度较高, 还处于初步的试验阶段。此外, 通用的分词算法很难有效满足具体应用要求, 在不同的应用领域, 需要根据应用特点设计有效的分词方法。目前直接针对灾害信息中文描述特点的分词研究还十分匮乏。

命名实体关系的理解主要通过模式匹配^[15]和机器学习^[16]相结合的方法实现。首先需要通过人工经验的方式或基于语料库统计的方式, 编制并建立实体关系模式库(语义理解规则), 然后再通过模式匹配的方法进行实体关系理解。由于灾害种类较多, 在灾害事件描述中, 包含了事件、时间、空间位置、损失, 以及影响方面(如人、财产、经济、社会)的实体, 并且实体之间存在时空关系、空间关系及因果关系, 使得在命名实体关系识别以及语义理解方面, 需要通过综合的分析推理机制实现。目前, 针对每一类灾害事件的文本描述特点, 进行实体之间的关系推理和提取是

一个研究的热点。

1.2.2 灾害事件时空位置描述及其时空匹配

时间和空间是灾害信息的重要属性。从 Web 文本抽取的是自然语言表达的、非结构化的时空信息,而灾害信息的时空分析是基于结构化的 GIS 数据库和计算模型。因此,必须建立自然语言描述的危害信息语义时空关系与 GIS 空间关系之间的转换模型,消除语义障碍,实现文本-地图的自动转换。美国 NCGIA 最早提出自然语言空间关系(natural language spatial relation)的概念,并展开自然语言与空间关系的研究。近年来,国内外相关研究主要集中在以下基本问题。自然语言中的空间词汇。例如,Mark^[17]和 Egenhofer^[18]对英语中反映线面拓扑空间关系的词汇的总结;乐小虬、杨崇俊等^[19]构建空间语义词典的研究。自然语言中空间关系描述的句法模式研究。例如,张雪英、闫国年^[20]针对 GIS 中自然语言空间关系查询请求表达的句法模式及其解析方法的研究。自然语言空间关系与 GIS 中计算模型的语义转换。例如,Mark 关于拓扑空间关系与自然语言之间的关系的研究;Shariff 等^[21]关于线面之间拓扑与距离自然语言空间关系的研究;Frank^[22]用东、西等 8 个方位和“同一”方位来描述空间对象的方位关系的研究;杜世宏等^[23]关于线、面之间的方位关系与自然语言之间的关系的研究;许珺^[24]关于线状地理特征空间关系的自然语言描述的形式化表达研究。

针对灾害事件,重点关注灾害发生的空间位置及其影响的范围。对于大空间尺度的、空间定位精度要求不高的灾害信息,可以通过不同级别的行政区地名库的模糊匹配实现空间匹配。而对于发生在空间结构相对复杂的城市区域的突发事件,其空间定位精度则要求较高。在城市地区,因灾害的类型、发生位置不同,对于灾害发生位置的描述表现出多元化。如,火灾事件多使用规则地址或路段定位描述;交通事故多使用路段、单一路口、动态路口定位描述。在描述灾害事件空间影响范围时,则更为模糊,往往使用地标名和衡量方位(如以南、往东等)、距离(如附近、旁边等)的动词、介词构成的短语。如何根据这些不规则的相对位置描述短语,进行灾害信息的空间定位和空间匹配,进而实现与 GIS 数据融合,是需要解决的一个难点问题。虽然,目前已经有成熟的地址匹配、地理编码技术可以实现从语义信息到二维空间的映射,但是这种定位往往是基于地名库的关键字匹配技术实现,需要位置描述非常规范和精确,而对于包含模糊空间

关系的位置表达方法,显然是不能满足的。针对灾害位置信息多元化而且复杂的描述特点,需要借鉴地名地标及其空间语义关系的模糊推理等技术,综合实现从文本信息到二维空间的映射。目前,针对多类型、异构的位置描述和空间参考方法,综合运用各种空间推理和空间转换方法实现灾害信息的精确定位,是一个复杂的知识密集型过程。此外,对于灾害空间影响范围的推理方法研究还比较少。

1.2.3 文本灾害信息的不确定和可靠性评价

准确、可靠的灾害信息是正确进行灾害分析与决策的重要基础。由于 Web 文本中灾害信息本身的错误、灾害信息描述的不确定性、灾害信息表达的不完全,或者信息提取和挖掘方法的不适应性,从 Web 文本中挖掘的灾害信息会存在一定的误差。因此,必须对各个环节的误差来源及不确定性、挖掘过程中灾害信息的误差传递机制进行研究,并建立综合性的灾害信息可靠性评价模型,以保证信息的可用性。

目前,主要通过定量化的指标和方法对挖掘结果进行可信度评估。然后根据评估结果,对抽取方法进行比较和优选、对系统参数进行调整。在评价信息的可靠性和系统的效果时,通常采用查准率(P)、查全率(R)和 F 值(F)进行检验。假设 Web 文本中所有通过人工识别应当被抽取出的信息个数为 T ,系统自动抽取的信息个数为 S , S 中符合人工识别的正确信息个数为 C ,则 $P = C/S$, $R = C/T$ 。 P 用于评价系统提取出的信息中正确的比例, R 用于评价被正确提取的信息的比例,以上指标值越大表示提取精度越高。 $F = P \cdot R \cdot (\alpha^2 + 1) / (R + \alpha^2 \cdot P) \cdot 100\%$,其中 α 表示 R 和 P 相对重要性的加权系数,如果 $\alpha = 1$,表示 R 和 P 同等重要。 F 用于综合评价系统的抽取精度,值越接近于 1 表示提取效果越好。

在评价过程中,首先利用 R 、 P 、 F 值,对灾害事件关键词、时间、位置、损失与影响等单项命名实体信息进行精度评价,然后设置各命名实体的相对重要性权重,进行命名实体识别精度的综合评价。根据精度评估的结果,进一步调整系统参数和优化抽取方法。

2 结论及展望

针对我国在灾害研究与管理过程中,灾害数据共享困难,以及可利用的动态实时、综合性灾害数据缺乏的薄弱环节,应尽快开展文本

灾害信息挖掘的关键技术、软件产品、管理体系的综合研究, 为灾害研究和管理提供及时、准确的数据获取技术保障, 进一步提升灾害信息服务水平。

2.1 文本灾害信息挖掘技术以及软件产品研发

在关键技术方面。①应针对不同的灾种, 重点研究和解决利用中文自然语言处理方法从 Web 文本中抽取灾害事件主题、时间、空间位置、损失及影响、致灾原因等 5 类关键信息的方法, 并通过定量模型对灾害信息的可靠性和精确性进行评价。②需要根据灾害位置描述的特点, 研究和建立灾害信息的多形式时空参考方法体系和描述规则, 根据不同的位置描述规则、灾害信息空间定位和空间匹配的算法模型和程序, 实现灾害信息的文本 - 地图转换。③根据灾害事件的紧急性、衍生性、关联性等特点, 通过关联规则挖掘算法, 对潜在的次生灾害信息以及风险进行识别, 为灾害的早期预警与风险防范提供技术支持。

在软件产品研发方面。可以充分借鉴国内外的一些开源 Web 文本挖掘工具, 如 GATE、WHISK、RAPIER、SRV 等, 进行本土化改造, 并增加对中文灾害信息处理功能。

2.2 文本灾害信息挖掘管理体系的完善

文本灾害信息挖掘软件系统的运行, 需要一个完善的管理体系作为保障。①建立灾害信息以及风险的监测机制。需要设计一个敏锐、开放的灾害信息以及风险感应机制, 对 Web 环境中所出现的异常现象做出及时反应, 进而迅速捕获社会环境中的灾害风险信息。②建立基于 Web 的灾害信息分析、加工机制。对获取的灾害信息进行分析、过滤、加工, 使之成为有质量的预报、预警。③Web 灾害信息鉴别机制。建立计算机程序自动化和专家队伍知识经验相结合的模式, 通过完善的灾害信息鉴别、审查机制, 使正确的灾害信息得以及时传播, 而错误的、误传的灾害信息提前过滤。

参考文献:

[1] International Federation of Red Cross and Red Crescent Societies. World Disasters Report 2005: Focus on information in disasters [EB/OL]. [2007-10-28]. <http://www.ifrc.org/publicat/wdr2005/>.

[2] 王静爱, 史培军, 朱骊, 等. 中国自然灾害数据库的建立与应用[J]. 北京师范大学学报: 自然科学版, 1995, 31(1): 121-26.

[3] 刘耀龙, 许世远, 王军, 等. 国内外灾害数据信息共享现状

研究[J]. 灾害学, 2008, 23(3): 109-13, 18.

[4] Paula K, Dunbar. Increasing public awareness of natural hazards via the Internet [J]. Natural Hazards, 2007, 42(3): 529-536.

[5] Peduzzi P, Dao H, Herold C. Mapping Disastrous Natural Hazards Using Global Datasets [J]. Natural Hazards, 2005 (35): 265-289.

[6] 史培军, 杜鹏, 叶涛, 等. 加强综合灾害风险研究, 提高应对灾害风险能力[J]. 自然灾害学报, 2006, 15(5): 1-6.

[7] Ma Yiming, Dmitri V Kalashnikov, Sharad Mehrotra, et al. On - Demand Information Portals for Disaster Situations [C]//Proceedings of IEEE International Conference on Intelligence and Security Informatics Location, New Brunswick, NJ, USA, May 23-24, 2007.

[8] Chapman S, Ciravegna F. Focused Data Mining for Decision Support in Emergency Response Scenarios [C]//Proceedings of ISWC, Athens, GA, USA, November 5-9, 2006.

[9] Hiroyuki Kameda. Keynote Presentation, Information sharing for technology and knowledge based on implementation strategies - Disaster Reduction Hyperbase (DRH) project [C]//Proceedings of Sixth DPRI - IIASA Forum on Disaster Risk Management - Risk and Challenges for Business and Industry, Istanbul, Turkey, August 13-17, 2006.

[10] 尹章才, 章光, 李井岗, 等. 基于 PPGIS 的社会化震害信息获取模型研究[J]. 灾害学, 2008, 23(3): 135-139.

[11] 缪其浩, 江世亮. 非官方信息源担当应急预警器[N]. 文汇报, 2008-6-29(7).

[12] Roman Yangarber, Lauri Jokipii, Antti Rauramo, et al. Information Extraction from Epidemiological Reports [C]//Proceedings of HLT/EMNLP, Vancouver, B. C, Canada, October 6-8, 2005.

[13] Damianos L E, Bayer S, Michael A Chisholm, et al. MiTAP for SARS Detection [C]//Proceedings of HLT - NAACL, Boston, MA, USA, May 2-7, 2004.

[14] 邱均平, 文庭孝, 周黎明. 汉语自动分词与内容分析法研究[J]. 情报学报, 2005, 24(3): 9-17.

[15] Aone C, Ramos - Santacruz M. Rees; A large - scale relation and event extraction system [C]//Proceedings of the 6th Applied Natural Language Processing Conference, Seattle, Washington, USA, April 29 - May 4, 2000.

[16] Soderland S. Learning information extraction rules from semi - structured and free text [J]. Machine Learning, 1999, (34): 233-272.

[17] Mark D, Gould M. Interaction with geographic information: a commentary [J]. Photogrammetric Engineering & Remote Sensing, 1991(57): 1427-1430.

[18] Egenhofer M J. Multi - modal spatial querying [C]//Proceedings of Seventh International Symposium on Spatial Data Handling, London, UK, August 12-16, 1996.

[19] 乐小虬, 杨崇俊, 于文洋. 基于空间语义角色的自然语言空间概念提取[J]. 武汉大学学报: 信息科学版, 2005, 30(12): 1100-1103.

[20] 张雪英, 闻国年. 自然语言空间关系及其在 GIS 中的应用研究[J]. 地球信息科学, 2007, 9(6): 77-81.

[21] Shariff A R B M, Egenhofer M J, Mark D M. Natural language spatial relations between linear and areal objects: The topology and metric of English - language terms [J]. International Journal of Geographical Information Science, 1998, 12(3): 215-246.

[22] Frank A U. Qualitative spatial reasoning about distances and directions in geographic space [J]. Journal of Visual Languages and Computing, 1992, 3(4): 343-371.

[23] 杜世宏, 王桥, 李顺. GIS 中自然语言空间关系定义[J]. 武汉大学学报: 信息科学版, 2005, 30(6): 533-38.

[24] 许珺. 关于线状地理特征空间关系的自然语言描述的形式化表达[J]. 遥感学报, 2007, 11(2): 152-58.