

基于空间数据挖掘的区域危险废物决策支持设计研究

杜鹏^{1,2}, 杨蕾^{1,2}, 夏斌¹, 陈红顺^{1,2}

(1.中国科学院广州地球化学研究所, 广东 广州 510640; 2.中国科学院研究生院, 北京 100086)

摘要 区域危险废物环境管理与风险控制是涉及到危险废物特性、环境、经济和社会等多因素的复杂过程,对资源环境的可持续发展和人体健康都至关重要。空间数据挖掘技术在传统数据挖掘的基础上增加空间维尺度,能够实现从空间数据库和非空间数据库中提取危险废物产生、转移、处置和管理过程中隐含的空间模式与特征。文章在传统的危险废物环境管理和风险控制方法的基础上,构建了区域危险废物空间数据挖掘的指标体系,并设计了一种基于空间数据挖掘技术的区域危险废物决策支持系统框架。该框架由数据层、数据挖掘层和显示层组成,为区域危险废物管理和决策提供了直观的、科学的和高效的支持,提供了一种新颖可行的思路和方法提高危险废物的环境管理和风险控制水平,具有一定的理论和现实意义。

关键词 危险废物; 空间数据挖掘; 决策支持

中图分类号: X32 文献标志码: A doi: 10.3969/j.issn.1003-6504.2010.04.046 文章编号: 1003-6504(2010)04-0196-04

Design of Regional Hazardous Waste Decision Support System Based on Spatial Datamining

DU Peng^{1,2}, YANG Lei^{1,2}, XIA Bin¹, CHEN Hong-shun^{1,2}

(1.Guangzhou Institute of Geochemistry, Chinese Academy of Sciences, Guangzhou 510640, China;

2.Graduate School, Chinese Academy of Sciences, Beijing 100086, China)

Abstract: Regional hazardous waste management and environmental risk management is a complicated process which relate to the characteristics of hazardous waste as well as environmental, economic, social and many other factors. Based on traditional methods of hazardous waste management and environmental risk management, combined with techniques of spatial datamining (SDM), a framework of regional hazardous waste decision support system was designed. Some applications of SDM on decision of regional hazardous waste were also proposed. Suggestions are made for the Province's mid and long-term (2007~2020) planning of nature reserves.

Key words: hazardous waste; spatial datamining (SDM); decision support; knowledge management

危险废物是指具有毒性、易燃性、爆炸性、腐蚀性、化学反应性、传染性等的,会对生态环境和人类健康构成危害或潜在危害的废物。危险废物在自然环境下不能被降解或具有很高的稳定性,能被生物富集,同时能致命或因累积引起有害的影响^[1]。危险废物造成的环境污染已成为当今世界环境保护工作面临的重要问题之一,联合国环境规划署已将危险废物污染控制列为全球重大环境问题。世界发达国家根据危险废物的全过程管理原则,已经建立了一整套的危险废物从产生、收集、运输、贮存、处理、处置等方面的法律法规、政策、制度、管理机构等综合管理体系^[2]。随着经济的迅速发展,我国危险废物的产生量越来越大、种类越来越繁杂、性质越来越复杂,且产生源数量分布

广泛,管理难度较大。截至2006年,我国危险废物产生量已经超过1084万吨,但同时我国危险废物管理水平、处理技术、处置设施等都与国际先进水平具有一定差距。我国如何利用先进的信息化技术和手段,提高危险废物的环境管理和风险控制的水平和效率,以减少危险废物对人群健康的威胁和生态环境的破坏,是一个重大的环境问题。

危险废物的环境管理和风险控制是一个多层次、多目标、多因素互相影响的复杂过程,与危险废物产生、转移、处置处理的空间位置和自然环境密切相关。探索如何有效的利用已有的危险废物相关资料和数据,结合危险废物特性和区域环境条件,通过空间数据挖掘(Spatial Datamining, SDM)技术,从中提取隐

收稿日期: 2009-05-02, 修回: 2009-07-01

基金项目: 国家自然科学基金重点项目(40534019)

作者简介: 杜鹏(1982-)男,博士研究生,主要研究领域为遥感与GIS应用、环境地理信息系统集成应用(手机)13450410252(电子信箱)dpng1982@gmail.com。

含的、危险废物相关空间的和非空间的模式和普遍特征,对危险废物的环境管理提供决策支持,降低危险废物对生态环境和人类健康的危害和威胁具有重要意义。本文结合以往的危险废物决策支持系统设计的研究成果,应用空间数据挖掘技术,提出对区域危险废物环境风险管理空间决策支持系统的框架设计,重点分析了区域层面危险废物决策支持问题。

1 空间数据挖掘概述

空间数据挖掘,也称基于空间数据库的数据挖掘和知识发现(Spatial Datamining and Knowledge Discovery,SDMKD),它可定义为在空间数据库和数据仓库的基础上,综合利用统计学方法、模式识别技术、人工智能方法、神经网络技术、粗集、模糊数学、机器学习、专家系统、可视化技术等,从大量的空间数据、管理数据、经营数据或遥感数据中析取出可信的、新颖的、感兴趣的、隐藏的、事先未知的、潜在有用的和最终可理解的知识,从而揭示出蕴含在空间数据背后客观世界的本质规律、内在联系和发展趋势,实现知识的自动或半自动获取,为管理和经营决策提供依据。SDM 的特点是在常规的事务性数据挖掘基础上增加了空间尺度维,使其能够从空间数据库中提取用户感兴趣的空间模式与特征、空间与非空间数据的普遍关系及一些隐含在数据库中的普遍数据特征。目前 SDM 技术的主要方法包括空间分析、统计分析、归纳学习、空间关联规则、挖掘方法、空间聚类、神经网络、决策树、模糊集理论、空间特征和趋势探测、云理论、证据理论、遗传算法等。

随着信息技术的不断发展和相关理论的不断完善,SDM 技术在各个领域应用的广度和深度上都在不断宽展和加强。在环境保护领域,SDM 在大气、水体和土壤污染防治等方面都已有所应用。作为重要污染物之一,每天会有大量不同来源的空间和非空间危险废物相关数据产生。通过 SDM 技术,结合危险废物的特征属性和专家知识,能够科学有效的将这些数据与其空间特征相结合,利用一种或多种空间数据挖掘方法,获得这些数据之间的深层的关联关系和隐含信息,为危险废物环境管理和决策提供科学依据,从而提高危险废物管理的水平和效率。

2 区域危险废物决策指标体系

危险废物具有所含毒物性质多样、污染源分布广泛、受暴露人群不确定、暴露途径多样、有毒有害物质在环境迁移过程中十分复杂等特性。根据可比式区域危险废物优先控制名录确定方法的概念模型,危险废

物的危害性源于危险废物本身所具有的有毒有害物质,但是其对环境和人类的危害又不能完全通过该种废物的毒性判断,而是应该根据将该废物的有毒有害成分释放到环境中,并被人体摄入后所产生的各种生物毒性效应的叠加效应进行判断。

在对危险废物管理进行数据挖掘和决策支持时,也应当将危险废物产生源、运输转移路线和集中处理地点所处区域空间因素考虑在内,从全局角度分析问题,指标体系除了危险特性和处置处理方式之外,还应当包含区域社会、经济、资源、环境等基本要素,这样才能立体全面的分析各种因素对危险废物决策支持的影响。

根据可比式区域危险废物优先控制名录确定方法的概念模型(图 1),危险废物空间决策指标体系可分为以下几类:

- (1)危险废物物化特性:有毒有害物质种类,危险废物浓度,溶解度,挥发度,饱和蒸气压,空气扩散系数,降解系数等物化特性参数等^[9]。
- (2)区域危险废物产生和处理处置情况:危险废物类别,数量,累计贮存数量,贮存、处置、利用方式和空间位置信息等。
- (3)区域典型环境条件及迁移参数:区域土壤特性、地表水和地下水信息、气候信息、动植物分布信息等。
- (4)区域经济社会环境参数:区域人口密度、区域产业分布数据等。
- (5)各种生物慢性生物毒性效应数据:急慢性毒性效应、致癌性、非致癌性等。

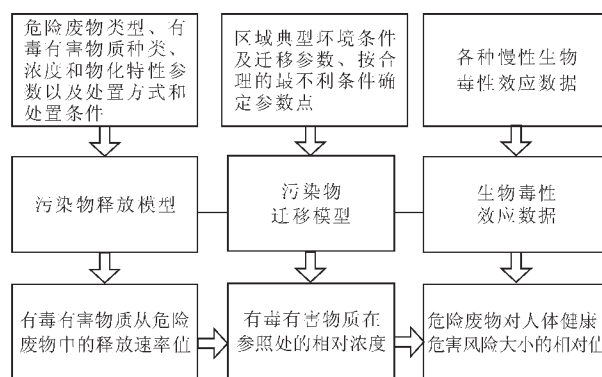


图1 可比式区域危险废物优先控制名录确定方法的概念模型^[11]
Fig.1 Conceptual model for determining the priority control directory of comparable regional hazardous waste^[11]

根据这些指标,根据危险废物相应的知识和规则赋予相应的权重,对区域危险废物管理进行决策支持,确定优先管理的目标,这些指标体系将构成危险废物基础信息数据库,为危险废物空间数据挖掘提供数据源。在设计危险废物决策支持系统指标体系时应

当具有一定的开放性和前瞻性,在使用过程中允许新的指标加入到决策支持中,是决策支持的可靠性能够不断的提高。

3 基于SDM的危险废物DSS框架设计

基于SDM的危险废物决策支持框架分为数据层、数据挖掘层和用户界面层三个层次。该框架通过对区域危险废物特征属性数据、危险废物申报登记数据、区域空间环境数据提炼、转化和空间变换等操作,构成区域危险废物空间数据仓库和空间数据立方体,为危险废物空间数据挖掘做数据准备。危险废物空间数据仓库中的多维数据通过空间在线分析处理(Online Analytical Processing, OLAP)和空间数据挖掘模块发现相应的危险废物管理知识和规则。发现知识和规则通过评价分析模块进行分析和评价,去除冗余信息,最终通过人机交互系统直观的展现给用户,并对发现的知识进行有效的管理,如图2所示。

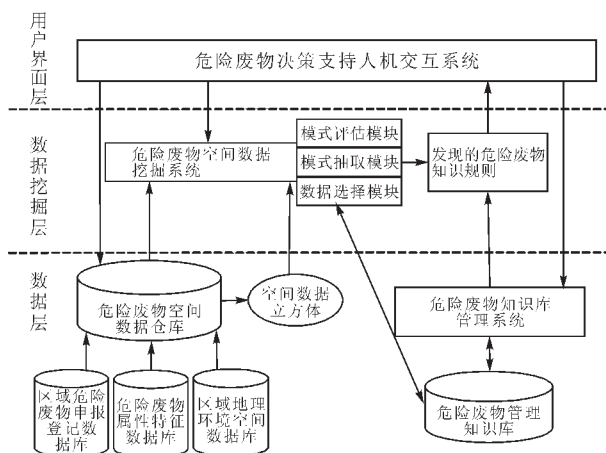


图2 空间数据挖掘的危险废物决策支持系统框架
Fig.2 The framework of hazardous waste decision support system based on SDM

3.1 数据层

数据层包括区域危险废物申报登记数据库,危险废物属性特征数据库,区域地理环境空间数据库,危险废物空间数据仓库和危险废物空间数据立方体几部分。通过空间数据引擎(Special Data Engine, SDE),实现危险废物属性数据、申报登记数据与空间数据关联,并对关联后的数据集成化、规范化后形成危险废物数据仓库体系和空间数据立方体,为数据挖掘做数据准备以提高挖掘的效率和精度。

3.1.1 危险废物数据库属性特征数据库

危险废物属性特征数据库中包括废物毒性、物理化学性质、有毒有害物质种类、溶解度、挥发度、饱和蒸气压、空气扩散系数、降解系数等物化特性数据以及危险废物急慢性毒性效应、致癌性、非致癌性等数据。

3.1.2 危险废物数据库属性特征数据库

区域危险废物申报登记数据库中包括危险废物产生单位基本经营数据、空间位置数据,产生危险废物数量、种类、形态、浓度、来源、含水率数据,主要有害成分含量及含量数据,废物贮存、处置、利用方式和数量等数据。

3.1.3 区域地理环境空间数据库

区域地理环境空间数据库包含区域基本地理要素空间数据、气象数据、地表水地下水数据、土壤类型数据、交通、土地利用数据等,人口密度及居民区数据,产业分布数据,动植物分布数据等。

3.1.4 区域危险废物空间数据仓库和数据立方体

区域危险废物空间数据仓库(Regional Hazardous Waste Spatial Data Warehouse, RHWSDW)是为危险废物属性特征数据库、区域申报登记数据库、地理环境空间数据库进行关联、提取、变化、转化后的危险废物多维信息集合,是空间危险废物空间数据挖掘的数据源。RHWSDW采用时间维、空间维、非时空维的多维组织结构,以提取各种维度和尺度上的区域危险废物相关信息,从而为区域危险废物空间决策管理提供数据支持。RHWSDW中的数据经过在线分析处理和数据挖掘处理后,将发现的知识 and 规则经过人机交互系统展现给用户。

本框架在危险废物空间数据挖掘过程中,采用构建数据立方体(Cube)的方法提高OLAP对数据仓库的访问速度。由于OLAP的多维数据模型和数据聚合技术可以组织并汇总大量的数据,因此,对数据立方体中的数据进行挖掘,可以利用已有的聚合信息和计算结果,提高数据挖掘算法的执行效率。

3.2 数据挖掘层

数据挖掘层通过危险废物空间数据挖掘系统选择空间数据挖掘库中的适当的挖掘方法和危险废物知识库中的已有规则和模式进行数据挖掘,发现新的、隐含的危险废物空间规则和模式,并将通过评估确认的规则和模式记录在危险废物知识库中。在本层中还将会把发现的知识通过评价分析模块进行评价,并通过人机交互系统转化为用户能够理解的方式呈现给用户。本框架中危险废物空间数据挖掘系统主要数据选择模块、模式抽取模块和评价分析模块组成。

3.2.1 数据选择模块

数据选择模块负责在数据仓库中选择与危险废物决策分析管理相关数据,并对选择的数据进行对象和属性的抽取,得到对不同决策有用的对象和属性。

3.2.2 模式抽取模块

模式抽取模块在数据选择模块的基础上,通过空

间数据引擎实现属性数据与空间数据的关联,实现空间数据的快速存取与分析,并利用机器学习、神经网络、决策树等合适的方法,对经过关联的空间数据和危险废物属性数据进行空间数据挖掘,从中发现一些隐含的空间规则与模式。

3.2.3 模式评估模块

危险废物空间数据挖掘过程中产生的规则和模式,需要采用相应的标准进行度量,从中去除冗余信息或已有的规则模式,发现具有实际意思的模式。如采用支持度、置信度、可信度等统计属性作为评估标准对预测型模式进行判断。在此过程中,为了取得更为有效的知识,可能会返回前面处理步骤中的某些步骤以反复提取,从而取得更为有效的知识。除此之外还要通过知识库管理系统对发现的知识进行一致性检测,使挖掘出的规则模式不与已有的规则和模式产生冲突。通过多轮反复的数据挖掘过程和调整,最终得到可靠的规则,并保存在知识库中进行管理。

3.3 显示层

显示层中包括人机交互系统。人机交互系统将发现的危险废物知识以地图、表格等多种直观形象的方式展现给用户,用户能够采取多种方式对发现的知识进行管理,并通过人机交互操作获得满意的数据挖掘结果。由于受人工智能和空间数据挖掘本身等技术的限制,危险废物空间数据挖掘和知识发现的过程需要反复的人机操作才能获得较为满意的结果。人机交互系统能够通过自然语言处理和语义查询在用户和系统之间提供相互联系的集成界面,最终实现危险废物空间管理的决策支持。

4 结语

区域危险废物决策支持受到区域社会经济环境、自然环境、危险废物特性等多方面多因素的影响,空间数据挖掘技术能够有效的整合影响危险废物环境管理决策的各类因素,为危险废物环境管理决策提供了新的思路和手段。例如利用空间聚类方法,将危险废物产生源空间数据与各种废物产生量数据进行空间聚类分析,实现危险废物产生源位置聚类,然后将危险废物产生行业与危险废物产生点位信息进行空间聚类分析,揭示出区域危险废物产生情况与行业之间空间分布规律和联系,结合区域居民区数据、人口密度数据,通过危险废物专家知识库确定区域优先管理和控制的危险废物行业和地区,降低危险废物的危险性,为确定区域优先控制的危险行业提供决策支持。

本文设计的基于SDM的危险废物决策支持系统

框架能够将包括区域空间要素在内的影响危险废物环境决策相关的参数进行抽取、净化、转换后,形成危险废物空间数据仓库,并采用不同的空间数据挖掘方法进行数据挖掘,通过危险废物知识库对数据挖掘结果进行评价、分析和管理的。本框架为危险废物环境管理和风险控制提供了一种新颖、开放和科学的信息化决策支持方式,在理论和实践都具有一定的研究价值。

【参考文献】

- [1] 李金惠,王琪,王洪涛,等.危险废物管理与处置处理技术[M].北京:化学工业出版社,2003.
Li Jin-hui, Wang Qi, Wang Hong-tao, et al. Management and Treatment Technology of Hazardous Waste[M]. Beijing: Chemical Industry Publishing House, 2003.(in Chinese)
- [2] 钟斌,孙绍锋.美国危险废物管理实践[J].环境保护,2008,48(22):74-76.
Zhong Bin, Sun Shao-feng. Hazardous waste management practice of the United States[J]. Environment Protection, 2008, 48(22): 74-76.(in Chinese)
- [3] 李金惠,聂永丰,刘鹏,等.危险废物管理区域决策支持系统的研究与开发[J].城市环境与城市生态,2000,13(6):15-17.
Li Jin-hui, Nie Yong-feng, Liu Peng, et al. Research and development of regional decision support system on hazardous waste management[J]. Urban Environment & Urban Ecology, 2000, 13(6): 15-17.(in Chinese)
- [4] 况代智.基于GIS技术的危险废物资源管理系统的研究[J].环境科学与技术,2008,31(10):151-153.
Kuang Dai-zhi. Hazardous waste management information system based on spatial information technology[J]. Environmental Science & Technology, 2008, 31(10): 151-153.(in Chinese)
- [5] 梅新,牛振国,吴孟泉,等.空间数据挖掘的区域循环经济决策支持框架设计[J].地球信息科学,2007,9(1):88-92.
Mei Xin, Niu Zhen-guo, Wu Meng-quan, et al. Design of the regional circular economy decision support system based on spatial datamining[J]. Geo-Information Science, 2007, 9(1): 88-92.(in Chinese)
- [6] 金萍.危险废物风险评估体系及管理模式的研究[D].合肥:合肥工业大学,2006.
Jin Ping. Risk Assessment System of Dangerous Waste and the Research of Management[D]. Hefei: Hefei University of Technology, 2006.(in Chinese)
- [7] 孙成忠,赵润怀,陈士林,等.基于聚类的空间数据挖掘技术在中药资源分析中的应用[J].测绘通报,2008,54(9):46-49.
Sun Cheng-zhong, Zhao Run-huai, Chen Shi-lin, et al. Application of spatial datamining technology in analysis of Chinese medicine resources based on clustering[J]. Bulletin of Surveying and Mapping, 2008, 54(9): 46-49.(in Chinese)