

学校编码: 10384

密级_____

学号: 21220051302252

厦 门 大 学

硕 士 学 位 论 文

数据挖掘技术在环境规划管理的应用研究

Research and Applications on Data Mining Technology in

Environmental Planning and Management

郑振敏

指导教师姓名: 姜青山 教授

专 业 名 称: 环境管理

论文提交日期: 2009 年 11 月

论文答辩日期: 2009 年 12 月

2009 年 12 月

厦门大学博硕士学位论文摘要库

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下，独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果，均在文中以适当方式明确标明，并符合法律规范和《厦门大学研究生学术活动规范（试行）》。

另外，该学位论文为（）课题（组）的研究成果，获得（）课题（组）经费或实验室的资助，在（）实验室完成。（请在以上括号内填写课题或课题组负责人或实验室名称，未有此项声明内容的，可以不作特别声明。）

声明人（签名）：

年 月 日

厦门大学博硕士学位论文摘要库

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

1. 经厦门大学保密委员会审查核定的保密学位论文，
于 年 月 日解密，解密后适用上述授权。

2. 不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

厦门大学博硕士学位论文摘要库

摘要

环境系统是一个非常庞大而开放的复杂系统，而要对这个系统进行规划管理必然是一项艰巨和困难的任务。有效的环境规划管理是在对环境数据进行科学分析、合理评价、准确预测的基础上产生的，由于环境规划管理的整体性、综合性、区域性、动态性、信息密集等特点以及环境数据的不断膨胀使其需要有新的、有效的分析方法的支持。

数据挖掘是目前最有效的数据分析手段，用于发现大量数据所隐含的各种规律。利用数据挖掘算法（如关联、分类、回归、聚类等）建立环境规划管理的数据挖掘模型，对经过预处理（如筛选、补缺、标准化等）的环境数据进行分析 and 结果验证，并解释已验证结果在环境规划管理上的意义，可以有效解决环境规划管理中存在的数据丰富但可用信息缺乏的问题。

本文是在充分理解数据挖掘与环境规划管理各自的基本理论的基础上，探讨数据挖掘在环境规划管理中的应用。本文的主要研究内容包括：

- 1、系统总结了环境规划管理的基本理论，并论述环境规划管理的技术方法；完整地阐述数据挖掘的体系理论，探讨数据挖掘与环境规划管理结合的必要性和可能性；
- 2、在聚类分析方法基础上建立竞争性评价模型，该评价模型在指标集的取值范围内选取最优值建立一个虚拟最优城市，以虚拟最优城市为参照，引入优越指数概念，得出待评价城市的优越指数值，然后进行排序和评价；
- 3、采用自回归、指数平滑和ARIMA三种时间序列分析方法建立闽江福州段白岩潭断面水质动态预测模型，并对三个模型的模拟与预测效果进行比较和评价。

通过数据挖掘在环境规划管理中的实例应用，证明了两者结合的可行性和有效性。数据挖掘将为环境规划管理提供决策支持和技术保障。

关键词：数据挖掘；环境规划管理；聚类分析；时间序列

厦门大学博硕士学位论文摘要库

Abstract

Environmental system is an open system which is very large and complex. It should be an arduous and difficult task to design and manage this system. Environmental planning and management which is based on the scientific analysis, reasonable evaluation and accurate prediction from environmental data will be effective. Due to the expansion of environmental data and its characteristic including holistic, synthetical, regional, comprehensive, dynamic and information-dense, environmental planning and management needs a new effective analysis method to support itself.

In current conditions, data mining is the most effective means of data analysis which could extract various rules hidden in the mass data. The model of environmental planning and management is established by the data mining algorithm. Environmental data which has been preprocessed is analyzed by the model, and the results of analysis will be tested. The validated results are interpreted from the viewpoint of environmental planning and management. It should be an effective solution for the problems on environmental planning and management which have rich data, but have no sufficient available informations.

In this thesis, we discuss on applications of the data mining in the environmental planning and management. The main contents are as follow:

1. The basic theory and methods on environmental planning and management were systematically summarized. The integrated theory on data mining system was completely elaborated, and the necessity and possibility on the combination of Data mining and the environmental planning and management were discussed;
2. Competitive Assessment Model of City Environment Quality (CAMCEQ) was established. CAMCEQ is a Cluster-Based model and sets up a Virtual Best City (VBC) whose value of evaluating indicator is most excellent in the cities which will be assessed. 18 cities were compared to VBC, assessed and sorted by their value of Superiority Index (SI);
3. A dynamic prediction model was established and related to water quality in

the Baiyantian section of “Min” river, Fuzhou. The dynamic prediction model adopted three kinds of methods on time series analysis including regression, exponential smoothing and ARIMA. The results of simulation and forecast about the three methods were analyzed and assessed.

The examples of the application illustrates that the combination of both methods is feasible and effective. Data mining will provide decision support and technical support for environmental planning and management.

Keywords: Data Mining; Environmental Planning and Management; Clustering Analysis; Time Series

目 录

第一章 绪论.....	1
1.1 研究背景及选题意义.....	1
1.2 研究现状及存在问题.....	2
1.3 主要研究内容及特色.....	4
1.4 本文结构安排.....	6
第二章 环境规划管理研究概述.....	7
2.1 环境管理.....	7
2.1.1 环境管理的研究内容.....	7
2.1.2 环境管理的技术方法.....	8
2.2 环境规划.....	13
2.2.1 环境规划的研究内容.....	13
2.2.3 环境规划的技术方法.....	14
2.3 环境规划管理过程中存在的问题及解决方法.....	17
2.4 小结.....	20
第三章 数据挖掘在环境规划管理中的应用研究.....	21
3.1 数据挖掘技术.....	21
3.1.1 数据挖掘的概念和功能.....	21
3.1.2 数据挖掘的一般过程.....	23
3.2 数据挖掘在环境规划管理中的应用方法与实例.....	24
3.3 环境质量评价预测中的数据挖掘方法.....	28
3.4 小结.....	37
第四章 基于聚类分析的竞争性环境质量评价模型.....	38
4.1 引言.....	38
4.2 竞争性环境质量评价模型.....	39
4.2.1 环境质量评价指标集.....	40
4.2.2 虚拟最优城市.....	41
4.2.3 归一化.....	41

4.2.4 计算各城市的优越指数 S.....	42
4.3 竞争性环境质量评价模型的结果分析	43
4.3.1 选取指标集.....	43
4.3.2 虚拟最优城市.....	44
4.3.3 归一化.....	44
4.3.4 城市环境质量优越指数 S.....	45
4.3.5 “十五”期间城市环境综合整治定量考核办法的评价结果.....	46
4.3.6 评价结果对比分析.....	55
4.4 竞争性环境质量评价模型在环境规划管理中的应用	58
4.5 小结	59
第五章 基于时间序列分析的水质预测模型	60
5.1 引言	60
5.2 水质预测模型	61
5.2.1 数据采集.....	61
5.2.2 时间序列预测模型.....	62
5.3 水质预测模型结果分析	64
5.3.1 自回归 (AR) 模型.....	64
5.3.2 指数平滑模型.....	68
5.3.3 自回归求和滑动平均 (ARIMA) 模型.....	72
5.3.4 三种预测方法的预测结果对比分析.....	75
5.4 水质预测模型在环境规划管理中的应用	76
5.5 小结	77
第六章 总结与展望	78
6.1 总结	78
6.2 展望	79
参考文献.....	80
攻读硕士期间的研究成果	82
致谢	83

CONTENTS

Chapter 1 Introduction	1
1.1 Background and Significance	1
1.2 Research Status and Problems	2
1.3 Main Research and Innovations	4
1.4 Outline of Thesis	6
Chapter 2 Overview of Research on EP and EM	7
2.1 Environmental Management(EM)	7
2.1.1 Research Content	7
2.1.2 Methods of Technical Support	8
2.2 Environmental Planning (EP)	13
2.2.1 Research Content	13
2.2.3 Methods of Technical Support	14
2.3 Problems and Solutions in EP & EM	17
2.4 Summary	20
Chapter 3 Technology Application of DM in EM and EP	21
3.1 Technology of Data Mining(DM)	21
3.1.1 Conception and Function	21
3.1.2 General Process	23
3.2 Examples and Methods of DM Application in EM & EP	24
3.3 Methods of DM in Env. Quality Assessment and Prediction	28
3.4 Summary	37
Chapter 4 CEQAM Based on Clustering Analysis	38
4.1 Introduction	38
4.2 Competitive Env. Quality Assessment Model(CEQAM)	39
4.2.1 Evaluation Indexes	40
4.2.2 Virtual Best City(VBC)	41
4.2.3 Normalization	41

4.2.4 Calculation of Superiority Index(SI)	42
4.3 Analysis of Results for CEQAM	43
4.3.1 Evaluation Indexes.....	43
4.3.2 Virtual Best City.....	44
4.3.3 Normalization	44
4.3.4 Superiority Index of Env. Quality	45
4.3.5 Assessment Results by 10th-five period CCCE.....	46
4.3.6 Comparison and Analysis of results.....	55
4.4 Application of CEQAM in EM and EP.....	58
4.5 Summary.....	59
Chapter 5 WQPM Based on Time Series Analysis.....	60
5.1 Introduction.....	60
5.2 Water Quality Prediction Model(WQPM).....	61
5.2.1 Data Collection	61
5.2.2 Prediction Model of Time Series	62
5.3 Analysis of results in WQPM	64
5.3.1 Model of Autoregression.....	64
5.3.2 Model of Exponential Smoothing.....	68
5.3.3 Model of ARIMA.....	72
5.3.4 Comparison and Analysis of results.....	75
5.4 Application of WQPM in EP and EM.....	76
5.5 Summary.....	77
Chapter 6 Conclusions and Futrure Work.....	78
6.1 Conclutions	78
6.2 Future Work	79
References	80
Publications	82
Acknowledgements	83

第一章 绪论

进入20世纪以来,环境问题呈现出地域上扩张和程度上恶化的趋势。随着污染程度的加深和污染影响范围的扩大,各种污染之间交叉复合以及愈演愈烈的过度开发,环境问题已逐渐由区域性演变为全球性的问题^[1]。近40多年来的历史说明,工业化和城市化进程加快给人类带来一系列环境污染、生态破坏、资源枯竭,甚至出现一系列全球性环境问题。人类从这一系列严重的环境问题中,逐渐认识到环境污染和生态破坏,归根到底来自于人类过度的和盲目的社会经济活动^[2]。为了尽可能快的速度逐步恢复被损害了的环境,并减少甚至消除新的发展活动对环境的结构、状态、功能造成新的损害,保证人类与环境能够持久地、和谐地协同发展下去^[1],人类需要有效的环境规划和管理。

1.1 研究背景及选题意义

自工业革命以来,由于科学技术的迅猛发展,经济水平的快速提高和人口的急剧增长,人类从自然环境中获取的资源数量越来越多,已经大大超过其补给和再生增殖能力,造成了资源的枯竭和环境的退化;另一方面,人类活动排入环境的废弃物也与日俱增,严重干扰了自然的正常循环,引起了生态系统平衡的失调。

从1978到2005年,我国年均经济增长率高达9.4%,但这是一种高投入、高消耗、高污染的粗放型经济增长方式。这导致我国经济社会发展与资源环境的矛盾日益突出,在解决环境问题的同时,新的环境问题也不断凸现。发达国家上百年工业化过程中分阶段出现的环境问题,在我国20多年里集中出现,呈现结构型、复合型、压缩型特点。当前,我国正进入大范围生态退化和复合型环境污染的阶段。环境污染和生态破坏造成了巨大的经济损失,不仅危害群众健康,影响社会稳定和环境安全,而且对亚洲邻国造成了不同程度的影响,导致区域冲突、贸易争端与外交制掣,从而危及国家安全^[3]。这给我国的环境规划和管理带来了很大的挑战和紧迫性。

有效的环境规划和管理是在对环境科学实验数据和社会经济统计数据进行

科学分析、合理评价、准确预测的基础上产生的。环境系统是一个非常庞大而开放的复杂系统^[4]，利用传统的技术和手段分析这些海量的环境数据，难以进行综合研究，不能为环境规划管理提供充分的决策支持。因此，寻求新型的、高效的现代化方法和技术就显得尤为重要。

数据挖掘是20世纪80年代末崛起的一门新兴学科，是一个多专业交叉的研究和应用领域。它的研究内容主要是从数据库中发现深层次的知识，是当前国际上数据库和信息决策领域的最前沿研究方向之一。随着计算机硬件和软件的飞速发展，尤其是数据库技术与应用的日益普及，人们面临着快速扩张的数据海洋，如何有效利用这一丰富数据海洋的宝藏为人类服务，业已成为各行各业的决策支持研究工作者所重点关注的焦点之一；而计算机性能的提高和先进的体系结构的发展使数据挖掘技术的研究和应用成为可能。经过几十年的研究和实践，数据挖掘技术已经吸收了许多学科的最新研究成果而形成独具特色的研究分支。从目前的现状看，大部分学者认为数据挖掘的研究仍然处于广泛研究和探索阶段。一方面，数据挖掘的概念已经被广泛接受。在理论上，一批具有挑战性和前瞻性的问题被提出，吸引越来越多的研究者。另一方面，数据挖掘的大面积广泛应用还有待时日，需要深入的研究积累和丰富的工程实践^[5]。

由于数据挖掘可以有效地解决大容量数据的利用问题并从海量数据中获取有用的信息和知识，因而环境规划管理所面临的问题在一定程度上可以通过数据挖掘来缓解甚至彻底解决。由此可见，环境规划管理与数据挖掘的结合有其必要性和可行性。

1.2 研究现状及存在问题

当前，我国环境保护事业进入新的发展阶段，环境信息化建设已经提上议事日程并且呈加速推进状态。国家环境信息化建设未来的发展将建立在“三体系、四平台、五应用”框架基础上，主要内容包括环境信息标准规范体系、环境信息基础设施平台、环境信息应用支撑平台、环境信息资源共享平台、环境信息资源服务平台、环境信息安全保障体系和环境信息运行管理体系等，并在应用支撑平台上建立包含饮用水源管理、环境质量监测管理、污染源监督执法管理、污染总量控制管理、建设项目管理、自然保护区管理、危险废物管理、核设施监控管理、

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士论文摘要库