

KUANGSHAN KONGJIAN
XINXI JICHENG YU
SHUJU WAJUE

矿山空间信息集成 与数据挖掘

陈应显 王志宏 著

矿山空间信息集成与数据挖掘

陈应显 王志宏 著



煤炭工业出版社

· 北 京 ·

图书在版编目 (CIP) 数据

矿山空间信息集成与数据挖掘/陈应显等著. --北京:煤炭工业出版社, 2017

ISBN 978 - 7 - 5020 - 6211 - 8

I. ①矿… II. ①陈… III. ①空间信息技术—应用—矿业工程
②数据处理—应用—矿业工程 IV. ①TD

中国版本图书馆 CIP 数据核字 (2017) 第 256620 号

矿山空间信息集成与数据挖掘

著 者 陈应显 王志宏
责任编辑 刘永兴 尹燕华
责任校对 孔青青
封面设计 于春颖

出版发行 煤炭工业出版社 (北京市朝阳区芍药居 35 号 100029)
电 话 010 - 84657898 (总编室)
010 - 64018321 (发行部) 010 - 84657880 (读者服务部)
电子信箱 cciph612@126.com
网 址 www.cciph.com.cn
印 刷 北京建宏印刷有限公司
经 销 全国新华书店

开 本 787mm × 1092mm¹/₁₆ 印张 7¹/₄ 字数 164 千字
版 次 2017 年 11 月第 1 版 2017 年 11 月第 1 次印刷
社内编号 9091 定价 35.00 元

版权所有 违者必究

本书如有缺页、倒页、脱页等质量问题,本社负责调换,电话:010 - 84657880

前 言

进入 21 世纪，信息化浪潮席卷全球，数字化方兴未艾，数据挖掘和知识发现（Data Mining and Knowledge Discovery）成为最活跃的研究领域之一。

为了在信息时代获得竞争优势，各国竞相加大信息产业投资，加速信息技术应用，以尽快提高国家信息化水平。在我国，信息化建设已成为一项重要的战略国策。2016 年 12 月的《煤炭工业发展“十三五”规划》明确提出了“能源科技创新日新月异，以信息化、智能化为特征的新一轮能源科技革命蓄势待发”。

随着我国矿山行业数字化和信息化进程的加快，矿山企业产生了大量的多源、异构和分布的矿山三维空间信息，同时也出现了“数据越丰富，知识越贫乏”的局面，亟须对矿山空间三维信息的集成管理和深入挖掘进行研究。

本书结合管理科学、信息科学、矿业工程、数据库理论、数据挖掘理论、计算机图形学和 CAD 技术等诸多理论和方法，以理论分析和实际应用验证相结合的方法，对矿山空间信息集成与数据挖掘进行研究。

全书共分 7 章，在分析矿山三维空间信息来源的基础上，将矿山三维空间信息分为原始信息、成果信息和生产信息三大类，为矿山三维空间信息的集成和管理提供信息基础；通过对各种空间数据挖掘方法的讨论，根据矿山空间信息的特点及功能的要求，选择适合的数据挖掘方法，并设计相应的数据挖掘算法实现对矿山三维空间信息的数据挖掘；针对主要空间信息聚类方法没有考虑空间数据的方向变化的这一特点，提出了基于方向的空间信息聚类方法和利用基于梯度的空间信息聚类方法；结合矿山空间信息的特点，分别提出了基于约束三角剖分建立矿山空间信息表面模型的方法和基于不规则四面体建立矿山空间信息实体模型的方法；实现了对矿山三维空间信息的集成以及对各类矿山空间信息模型的集成管理；在对矿山系统需求分析的基础上，

提出了以矿山空间数据库为中心，对系统进行方案设计，提出了矿山信息系统集成整体逻辑结构，并对系统进行开发，建立了基于模型集成的矿山系统并已在实际矿山中得到应用。

由于作者水平所限，书中若有错误或不当之处，真诚欢迎同行专家不吝赐教。

作者

2017年7月

于辽宁工程技术大学

目 录

1 绪论	1
1.1 研究的背景与意义	1
1.2 矿山三维空间数据挖掘与集成的研究进展	2
1.3 存在的问题与不足	7
1.4 研究思路及主要研究内容	8
2 矿山三维空间的信息分析	10
2.1 矿山空间对象	10
2.2 矿山三维空间信息	13
2.3 小结	19
3 矿山三维空间信息的数据挖掘	20
3.1 空间数据挖掘的定义与特点	20
3.2 空间数据挖掘的理论方法	20
3.3 钻孔信息的挖掘	29
3.4 基于人工神经网络的矿床界面预测	32
3.5 矿山测量信息的空间聚类	36
3.6 小结	37
4 基于方向的空间信息聚类	38
4.1 聚类方法分析	38
4.2 基于方向的空间信息聚类算法	43
4.3 基于梯度的空间信息聚类	46
4.4 基于梯度聚类的矿山断层信息挖掘	48
4.5 小结	50
5 矿山三维空间信息模型	51
5.1 基于表面模型的方法	51
5.2 基于约束三角剖分的矿山表面信息模型	54
5.3 基于实体建模的方法	57
5.4 基于不规则四面体的矿山实体信息模型	60

5.5 小结	62
6 矿山三维空间信息集成	63
6.1 矿山三维空间的信息流分析	63
6.2 等高线信息模型	66
6.3 表面信息模型	68
6.4 实体信息模型	74
6.5 小结	79
7 矿山三维空间信息集成系统开发与应用	80
7.1 矿山信息集成系统的需求分析	80
7.2 矿山信息集成系统的整体设计	81
7.3 系统开发环境选择及接口实现	83
7.4 矿山空间数据库的实现及空间信息管理	87
7.5 空间数据库与图形系统集成的实现	95
7.6 系统的功能模块及应用	95
7.7 小结	98
参考文献	99

1 绪 论

矿山三维空间信息的采集、处理、表达和服务等技术设备的迅速发展,使得矿山三维空间信息的数量急剧膨胀。为了充分利用矿山空间信息资源,对空间信息的集成化和数据挖掘受到了广泛的关注和重视。

1.1 研究的背景与意义

进入 21 世纪,信息化浪潮席卷全球,数字化方兴未艾,数据挖掘和知识发现(Data Mining and Knowledge Discovery)成为最活跃的研究领域之一。

为了在信息时代获得竞争优势,各国竞相加大信息产业投资,加速信息技术应用,以尽快提高国家信息化水平^[1,2]。在我国,信息化建设已成为一项重要的战略国策。2016 年 12 月的《煤炭工业发展“十三五”规划》明确提出了“能源科技创新日新月异,以信息化、智能化为特征的新一轮能源科技革命蓄势待发。”

自 1999 年首届“国际数字地球”大会上提出了“数字矿山”(Digital Mine, DM)概念以来,数字矿山的思想已开始深入人心,数字矿山科学研究与技术攻关已悄然兴起^[3-7]。数字矿山是对真实矿山整体及相关现象的统一认识与数字化再现,即将矿山生产、安全、矿山地理、地质、矿山建设等综合信息全面数字化,其目的是利用信息技术及现代控制理论与自动化技术去动态详尽地描述与控制矿山安全生产与运营的全过程。以高效、安全、绿色开采为目标,保证矿山经济的可持续增长,保证矿山自然环境的生态稳定。近年来,已取得了很多研究成果^[8-12]。

随着我国矿山行业数字化和信息化进程的加快,矿山企业产生了大量的多源、异构和分布的矿山三维空间数据,同时也出现了“数据越丰富,知识越贫乏”的局面,亟须对矿山空间三维数据深入挖掘进行研究。如何在浩瀚的空间数据和人们的知识渴求之间建立一个桥梁,的确是一个巨大的挑战,而以聚类、分类等人工智能技术为基础的空间数据挖掘为迎接这个挑战提供了新的技术支撑。

空间数据挖掘(Spatial Data Mining, SDM)是数据挖掘(Data Mining, DM)的分支学科,但 SDM 不同于一般的 DM,有别于常规的事务型数据库的数据挖掘,比一般数据挖掘的发现状态空间理论增加了空间尺度维(Scale)。简单地讲,空间数据挖掘是指从空间数据库中提取隐含的、用户感兴趣的空间和非空间的模式、普遍特征、规则和知识的过程。对矿山三维空间信息的数据挖掘是为了得出对矿山分析、计划和生产有用的知识。通过对矿山三维空间信息的数据挖掘,将极大地提高数据处理效率和质量,对加快我国的“数字矿山”建设有积极意义,具有广泛的社会应用前景。

对空间对象的分析是对空间信息的深入挖掘。空间数据挖掘和知识发现是空间数据库技术、空间数据获取技术、计算机技术、网络通信技术和管理决策支持技术等发展到一定

阶段的产物，是多学科相互交融和相互促进的新兴边缘学科，汇集了人工智能、模式识别、数据库、空间信息学、统计学等各学科技术的成果。

矿业是我国基础工业的支柱，是国民经济建设的基础，矿山信息化、集成化建设是我国信息化建设的一个重要组成部分。其建设目标是：充分利用现代信息技术，开发利用矿山信息资源，优化矿山活动全过程，用信息化改造我国传统矿山，提升我国矿山的管理水平和技术水平，提高我国矿山企业的竞争能力，实现我国矿业的现代化。

矿山空间信息集成是在计算机软、硬件支持下，运用系统工程和信息科学的理论与方法，综合、动态地获取、存储、管理、分析和描述整个或部分矿山表面与地下矿体模型，并服务于矿山的生产和设计，是空间信息采集、管理和应用的综合体。不仅有与矿山对象空间分布有关的空间数据，还有描述各矿山空间对象特征的属性数据和复杂的空间关系。为此，选择“矿山空间信息集成与数据挖掘”作为研究对象。鉴于当前我国矿山信息化建设仍然处于一个较低的阶段和水平上，对此进行深入研究既有较高的学术价值，又有很大的应用价值。

1.2 矿山三维空间数据挖掘与集成的研究进展

1.2.1 相关技术的进展

空间数据挖掘是综合利用统计学方法、模式识别技术、人工智能方法、人工神经网络技术、模糊数学、机器学习、专家系统和相关信息技术等按照一定的度量值和临界值抽取空间知识及与之相应的预处理、空间抽样和数据变换的一个多步骤相互链接、反复进行的人机交互过程。可以归纳为数据准备（了解应用领域的先验知识、生成目标数据集、数据清理、数据简化与投影）、数据挖掘和知识发现（数据挖掘功能和算法的选取，在空间的关联、特征、分类、回归、聚类、函数依赖等特定的规则中搜索感兴趣的知识）以及数据挖掘后处理（知识的解释、评价和应用）三部分。

空间数据挖掘的研究比一般的关系数据库和事务数据库的研究要晚，但近几年已经引起广泛的兴趣。目前，国内外已经开展了空间数据挖掘与知识发现方面的研究。

美国明尼苏达大学、加拿大西蒙弗雷泽大学、德国慕尼黑大学、芬兰赫尔辛基大学以及澳大利亚等国家的许多大学和研究所，都有空间数据挖掘的成果报道。这些研究者大多具有计算机科学背景，他们一般把空间数据挖掘作为数据挖掘的一个应用领域，研究的重点是提高原有数据挖掘算法在空间数据库的执行效率。测绘遥感界的学者们在特征提取、模式识别等研究中已经做了许多空间数据挖掘的工作。目前，在空间数据挖掘系统的开发方面，国际上有代表性的通用 SDM 系统有：GeoMiner, Descartes 和 ArcView GIS 的 S-PLUS 接口。加拿大 Simon Fraser 大学计算机科学系的数据挖掘研究小组，在 MapInfo 平台上建立了空间数据挖掘的原型系统 GeoMiner，实现了空间数据特征描述、空间区分、空间关联、空间聚类和空间分类等空间数据挖掘方法。ArcView GIS 的 S-PLUS 接口是著名的 ESRI 公司开发的，它提供工具分析空间数据中指定的类。

我国许多科研院所和高校先后展开了对空间数据挖掘和知识发现的理论和应用研究，国家对空间数据挖掘和知识发现也给予了极大的重视。国家自然科学基金、国家高技术研究发展计划（863 计划）、国家重大基础研究发展计划（973 计划）和教育部博士点基金，

以及军事国防等纵向、横向的科研或应用基金,都相继把空间数据挖掘和知识发现列为资助项目范围。

矿山三维空间信息集成系统是对不同时间、不同空间和不同部门的空间数据的集合,是对空间数据进行表达、分析、处理和对矿山活动进行模拟的过程,以空间数据管理、数据建模、数据挖掘及多维可视化表达技术为基础。

因此,矿山空间信息集成系统的研究和数据库技术、网络技术、数据挖掘技术、计算机图形学及科学可视化等技术密切相关,矿山空间信息集成系统的研究以这些技术的发展与应用为基础。本节将简要概述这些相关技术的发展与应用。

1. 数据库与网络技术的发展与应用

数据库(Data Base)是以一定的组织方式存储在一起的相互关联的数据集合,能最大限度地减少数据冗余,增强数据间的联系,实现对数据的合理组织和灵活存取^[13]。从1963年C. W. Bachman设计开发综合数据存储系统(Integated Data Store, IDS),开创了数据库技术的先河。数据库技术的发展经历了层次、网络数据库系统,发展到关系型数据库系统和以面向对象为主要特征的数据库系统后,数据库技术与分布式网络通信技术、面向对象技术、多媒体技术、人工智能技术以及可视化技术相互渗透,使数据库技术与网络技术成为信息化建设领域的基础技术。当今市面上大型的数据库系统包括 Oracle、SQL Server、Sybase 和 DB2 等。

目前,数据库管理系统(Data Base Management System, DBMS)已进入人们的日常生活。DBMS 为用户访问数据库提供了事务处理、操作进程的并发和共享数据的访问、数据的备份与恢复、数据的存储管理与容错机制等服务功能,数据库管理软件的开发使得人们在使用数据库时不需要了解数据的具体存取与管理方式。计算机网络是利用通信线路和交换、传输等通信设备,将地理上分散在不同地点,并具有独立功能的多套计算机系统互连接起来,按照网络协议进行数据通信,实现资源共享的计算机系统的集合。计算机网络为用户间提供资源共享、信息的快速传输与集中处理等信息服务功能。尤其是 Internet/Intranet 的迅猛发展,为信息的交流和共享、部门间的协同运作提供了技术的保证,同时也预示着网络化办公时代来临,涌现了大量的基于 C/S (Client/Server) 模式和基于 B/S (Browse/Server) 模式的各种数据库管理系统。

随着计算机通信设备和网络操作系统与应用软件的发展,数据集成能够通过应用间的数据交换而达到集成目的,解决数据的分布性和异构性问题。数据集成包括数据访问网关和数据复制两种方式。数据访问网关通过建立全局数据模式和进行全局查询处理来统一、透明地访问各种异构的数据资源,而数据复制技术通过对数据进行提取、转换、传输和加载操作来实现将数据从一个地方复制到另一个地方,完成数据的集中与汇总。

矿山空间信息集成系统正是通过一系列应用软件,利用计算机网络技术与数据库技术来对各种分布、异构的数据资源进行全局、统一、高效的访问和管理,为矿山的各种应用和决策支持提供一个良好的数据基础。

2. 数据挖掘技术的发展

数据挖掘(Data Mining, DM)源于数据库的知识发现(Knowledge Discovery in Database, KDD)技术,是通过从数据库中抽取隐含的、未知的、具有潜在使用价值信息的过

程^[14]。在实际的工程应用中,由于数据挖掘是 KDD 过程中最关键的步骤,许多学者认为数据挖掘就是知识发现^[14-16]。数据挖掘技术涉及机器学习、模式识别、统计学、人工智能、专家系统以及科学可视化等技术领域,随着这些技术的发展,无论在理论研究还是实际应用中,数据挖掘技术均获得了很大的进步^[17-20]。

空间数据挖掘 (Spatial Data Mining) 是指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其他一些隐含在数据库中的普遍的数据特征^[21]。近年来,空间数据挖掘技术从理论和应用上都获得了迅速的发展,在空间数据结构、多源数据集成、空间数据的管理、空间决策支持模型以及地学数据挖掘模型与 GIS 集成等方面均取得了不少的成就^[22-31]。

将数据挖掘方法和技术引入矿山空间信息分析领域将有助于提高矿山空间信息分析智能化的水平,对于为矿山的生产和开发提供有力的分析工具具有重要意义。

3. 可视化技术的发展

可视化技术是随着计算机图形学的发展、科学可视化技术、虚拟现实 (Virtual Reality, VR) 等技术与地球科学相结合而发展起来的综合性技术。

计算机图形学 (Computer Graphics) 是研究怎样利用计算机来显示、生成和处理图形的原理、方法和技术的一门学科^[21]。计算机图形学自 20 世纪 50 年代形成以后,随着计算机图形设备和图形处理软件的发展,在地球科学、医药、工业、商业、艺术及广告娱乐业等方面获得了广泛的应用。计算机辅助设计与计算机辅助制造 (CAD/CAM) 更是在各种应用领域中获得了广泛的发展和发挥着巨大的作用,包括各类 CAD 软件的开发和计算机几何造型的研究与应用,如 SDRC 公司于 1979 年发布了世界上第一个完全基于实体造型技术的大型 CAD/CAE 软件 I-DEAS; 美国 Autodesk 公司开发的 AutoCAD 软件,其应用已遍及全球各行各业。国内外许多著名大学、实验室从六七十年代就相继开始了计算机辅助几何造型的研究,在理论和技术上均取得了不少成果^[35-37]。

科学可视化 (Scientific Visualization) 又名科学计算可视化 (Visualization in Scientific Computing), 是发达国家 20 世纪 80 年代后期提出并发展起来的一个新的研究领域,指的是运用计算机图形学和图像处理技术,将科学计算过程中及计算结果的数据转换为图形及图像在屏幕上显示出来并进行交互处理的理论、方法和技术,已经应用于医学及医疗、地震勘探、气象预报、分子结构、流体力学、有限元分析、天体物理、海洋观察、地理信息、大气物理、环境保护等领域^[38,39]。

虚拟现实技术研究交互式实时三维图形在计算机环境模拟方面的应用,开始于 20 世纪 60 年代,是一种高度逼真地模拟人在自然环境中视、听、动等行为的人机界面技术。它通过头盔式的三维立体显示器、数据手套、三维鼠标、数据衣 (Data Suit)、立体声耳机等使人能完全沉浸在计算机生成的一种特殊三维图形环境,并且人可以操作控制三维图形环境,实现特殊目的。许多著名大学、大公司的研究中心都在研究和推广应用可视化与虚拟现实技术,如: NASA 的 AMES 研究中心, LAWRENCE LIVERMORE 国家实验室, LOS ALAMOS 国家实验室, IBM 研究中心, UIUC 的 NCSA, 海军研究实验室等。虚拟现实系统能以实时的速度生成有逼真感的景物图形,即三维全彩色、有明暗、纹理和阴影的图像。近年来,虚拟现实技术在我国无论是理论技术研究还是工程应用方面也均获得了很

大的发展^[40-46]。

科学可视化技术与地球科学相结合,在地学领域内有着广阔的应用前景。地学工程数据的可视化、地质体几何造型与矿化空间的可视化、时空过程的模拟等均要用到可视化技术。许多公司也正从事三维可视化软件技术在矿床开发评价中的应用,还有许多学者在开展三维地面景观模型、地质体的三维表达与可视化、地震发生及油藏模拟等可视化方面的研究。VR 技术在地学过程模拟中也获得了广泛的应用^[46],虚拟地质与虚拟矿山成为 VR 技术在工程应用方面的一大主题。以图形、图像的方式逼真再现三维地质实体以便真实地重建地下目标的结构、描述资源分布的状况,为模拟实验矿山开采提供决策的凭据,可视化分析地质现象的特征与空间分布,已经成为地学现代化研究领域的新分支。

1.2.2 国外矿山空间信息集成系统的进展

20 世纪 70 年代后期,计算机开始在矿山企业应用。国外一些矿山开始建立信息系统,如英国的 MINOS 系统,加拿大的 TMMS 系统,澳大利亚帕拉布多矿的 MIS 系统,苏联的 ACY 系统,等等,这些系统的功能不是很强,但标志着矿山应用信息技术的开始^[47]。

20 世纪 90 年代,国外矿山信息系统建设发展较快。在国外很多矿山,信息系统得到了广泛应用(如 AMSKAN 矿山信息系统、AQUILA 采矿系统公司的实时信息管理系统、ENDAKO 铝矿信息系统,等),信息系统已能够将矿山所需的绝大部分图纸与数据存储在磁盘上,几乎全部地测、采掘计划、爆破设计等方面的图件以及储量管理、生产计划、生产统计、设备、库存、人事、工资、销售、财会等方面的报表,均可由计算机辅助作业、管理、存贮及输出,大大减轻了工作量,使工程技术人员、管理人员摆脱了繁重的事务作业,效率大为提高^[48-51]。

尽管在 20 世纪 90 年代国外矿山信息系统建设取得了巨大的进展,但客观地分析:这种进展更多地体现在过程自动化方面,在整体信息应用与管理决策支持上进步并不明显。对此,加拿大著名的矿业专家 Z. LUKACS 曾撰文指出:“虽然矿业公司已经意识到先进信息技术的经济效益,但是在管理信息系统领域中,矿业落后于其他行业,在矿业利用新的信息技术之前,应迅速找出信息管理问题,矿山生产、运营所需要的信息也同样应该构成有效经营管理决策的基础^[52]。”

进入 21 世纪,由于市场竞争加剧,国外的矿山企业和专家学者对信息系统建设更为重视,期望通过矿山信息系统建设获得竞争优势。加拿大露天采矿研究和技术协会与艾伯塔大学采矿与石油工程学院于 2000 年 6 月共同主办了一个学术研讨会议,对有关矿山信息管理问题进行了深入的探讨。在这次会议上, M. SCOLBLE、LMOTTLA 等诸多知名矿业专家对矿山企业、原设备制造厂和软件销售商的典型实例进行了深入分析与探讨,一致认为:全采矿系统(主要包括:①在线机器监测、控制和定位系统;②一体化数据库系统,可使作业、维护和管理群体快速应用的、灵活的软件系统;③开放型结构、定向目标的矿山模型和 GIS 系统;④全面的矿山规划系统;⑤具有足够带宽的双向移动通信系统)和有效的管理决策支持系统将是矿山信息系统建设的趋势,系统规划与集成、标准化问题、信息系统管理机构建设问题将成为矿山信息系统建设中的重点与难点^[53-56]。

针对以上重点与难点,国外一些矿山和研究机构已经投入巨大的人力和物力进行研

究^[57-65]。例如,芬兰赫尔辛基技术大学岩体工程实验室、欧托昆普矿业公司、TOMROCK 公司等数家机构预算 800 万美元,针对凯米矿山进行了一项名为“智能矿山实施”的研究计划,主要从以下 3 个方面进行了研究:如何实施先进技术、如何有效利用信息、如何对矿山人员进行针对新系统的培训^[66];加拿大 GEMCOM 公司投入巨资进行了 Gemeom Enterprise Mining System 的开发,其目标是形成一个矿山企业大型集成信息系统^[67];智利天主教大学采矿中心进行了采矿工程计算的开放系统标准研究,等等^[68]。

1.2.3 国内矿山空间信息集成系统的现状

我国矿山应用信息技术始于 20 世纪 80 年代。一些矿山、矿冶设计院、大学借鉴国外的经验,开始探索计算机技术在矿山设计和生产中的应用,并开发出了一些矿山应用软件和系统。如 20 世纪 80 年代初德兴铜矿首次应用计算机进行了克里格法储量计算,之后在梅山铁矿、驾鹿金矿、刁泉银矿、金渠沟金矿、银洞坡金矿、玉龙铜矿等进行了克里格法储量计算,这标志着我国矿山开始应用计算机技术。这在推动我国矿山信息系统建设方面起到了积极的作用。随后,矿山信息系统建设在我国进一步发展,例如铜陵凤凰山选矿厂建立了选矿自动控制系统。

20 世纪 90 年代以来,我国矿山信息系统建设发展较快。先后有一批矿山进行了信息系统的开发(或引进应用),并取得了较好的效果。如德兴铜矿、永平铜矿、山西孝义铝矿等一些矿山利用符合国际矿业标准的系统进行了采矿设计。德兴铜矿于 1998 年引进美国模块公司用 GPS 定位的 DISNTCH 系统,实现了露天采矿生产的计算机调度,提高了设备利用率和矿山生产管理水平^[69]。德兴铜矿大山选矿厂引进了选矿自动控制系统,该系统以 Windows NT 为平台,采用了现场总线系统,并可通过 Internet 实现生产过程监控^[70]。云南铜业公司进行了其信息系统建设的一期工程,建立了 NOVELL 网,应用了 ORACLE 数据库,并先后开发了一些应用系统,如人事管理系统,这些系统的应用对于企业领导及时掌握生产情况起到了积极的支持作用^[71]。高峰矿业有限责任公司进行了网络 MIS 系统的开发^[72]。福建紫金矿业股份有限公司开发了生产管理信息系统,提高了其生产效率^[73]。神东公司开发了网络信息系统,使企业的管理更加规范化、更加高效^[74],等等。

虽然我国矿山企业的信息系统建设取得了很大的进展,但是与国外先进矿山以及其他信息化程度较高的行业(如金融、电信、保险、精细化工等行业)比较起来,我国矿山企业信息系统建设仍然处于一个较低的阶段和水平上^[75-81]。这具体表现在:

(1) 我国是一个矿业大国,矿山企业在 2001 年就已经达到 153068 个,其中中型以上企业有近 3000 个^[82]。除了少部分矿山企业有一定的信息系统基础外,大部分矿山还处于信息系统建设的初始状态,虽然已经开始在某些职能部门引入单系统(如财务管理系统),但没有明确的建设思路,缺乏对信息系统建设的整体规划。

(2) 矿山信息系统建设对于迅速提升我国矿山生产水平和管理水平的重要性,还没有成为我国大多数矿业主管部门和大部分矿山企业领导的共识。我国矿山信息系统建设落后于整个国家信息化建设的步伐是一个不容争议的事实。在全国其他行业进行信息系统建设的大潮中,矿山企业无论在投资上,还是在项目上都居于后列。其中一个主要原因就是矿业主管部门和矿山主要管理者的重视程度不够。从这个意义上来说,矿山信息系统建设不仅仅是矿山技术部门的工作,而更应该是矿业主管部门和矿山主要领导者的战略决策问

题。

(3) 通过对我国部分矿山的现有信息系统进行实地调研和深入分析,可以看出:我国现有的矿山信息系统主要是针对生产中的某个环节或者管理中的部分结构化问题而建设的,这些信息系统确实提升了局部工作效率,但是难以从整体上提升矿山的生产与管理水平,难以对矿山高层领导进行战略管理提供有力的支持,这也是出现上述第二种情况的一个重要原因。

(4) 我国国内自主进行矿山应用软件开发力度不足,引进先进矿业国家的应用软件的力度也不够。因此,在进行资源的优化利用、采矿系统的优化设计、经济效益的优化评估、生产运营的综合分析等诸多方面缺乏优秀的软件进行支持。这影响了我国矿山信息系统建设的速度和水平,也在一定程度上造成了我国矿山信息系统建设中的“硬硬软软”现象,即硬件较好,而软件的素质却一般。

(5) 我国矿山信息系统建设方面的技术队伍明显不足。近10年来,各行各业都在吸引信息化人才,不失时机地建立其行业、部门、企业的信息化队伍。实践证明:人才是产业信息化重要保证。但在矿业领域,信息人才一直是一个薄弱环节。特别是既熟悉矿山实际情况又精通企业信息系统建设的复合型人才更是非常缺乏。

美国著名的信息系统专家诺兰(R. L. Nolan)曾提出过著名的企业信息系统建设阶段模型(诺兰模型)^[83],诺兰模型的6个阶段分别是:初始阶段、传播阶段、控制阶段、集成阶段、数据管理阶段和成熟阶段。根据诺兰模型,对我国矿山信息系统建设现状的分析,可以得到这样的结论:我国矿山信息系统建设仍然处于企业信息系统建设的低级阶段,即“初装、传播、控制”阶段。这突出体现在:矿山对企业信息系统建设缺乏整体规划;现有矿山信息系统难以对矿山海量数据进行有效管理和深层次的综合应用,不能为矿山高层领导进行战略管理提供有力支持。

1.3 存在的问题与不足

经过矿山科研工作者的大量工作,矿山空间信息系统的研究取得了显著的成果,但是由于矿山空间信息的多样性、复杂性和差异性,矿山建模方法的适应性,这一领域仍然存在很多问题亟待解决。

(1) 矿山空间信息管理的研究不够。由于矿山空间信息量巨大,数据来源不同,数据格式差异大,缺乏较为统一的标准,因此矿山空间信息管理往往具有强烈的针对性,移植性很差,阻碍了矿山数字化和信息化的发展。

(2) 矿山空间信息模型集成的要求过低。不同的建模方法在特定的适应条件下,往往能够很好地满足某些方面的需求,但在另外一些方面却很难实现。由于矿山各专业、各层次需求的多样性,单一的矿山空间模型难以满足矿山各方面的业务需求,这就对矿山空间信息模型集成提出了更高的要求。

(3) 矿山空间信息集成系统的规划和设计常常缺乏矿山远景规划的指导,信息系统的开发忽略了与矿山组织结构、人员、知识结构等的关系,并且开发出的信息系统缺乏弹性,难以适应业务过程的需要。

(4) 矿山空间信息建模与工程应用脱节。矿山空间信息建模是要求具有非常专业的

计算机知识、地质学知识、测绘学知识和采矿知识，往往只有多年从事矿山理论和实践的研究人员才能完成地质模型的建立和维护。这就限制了矿山空间信息模型的使用和发展，并进而限制建立在矿山空间信息模型基础上的工程应用。

(5) 矿山空间信息缺乏深入的挖掘。从最初的地质勘查开始，积累了大量的空间信息，通过人工从中获取的知识只是其中很少的部分，如何用数据挖掘的方法代替人工从空间信息中获取知识，或者从大量的空间信息中获得更多的知识是对矿山空间信息集成系统的更高的要求。

1.4 研究思路及主要研究内容

1.4.1 研究目标及研究思路

为了对矿山三维空间信息进行更为有效的管理和深入的应用，在矿山三维空间信息集成及数据挖掘、空间信息模型的理论和方法，以及矿山三维空间信息集成系统建立和应用等方面仍有许多尚待进一步研究的问题。

因此，研究目标是：对矿山三维空间信息进行分析和数据挖掘，以矿山三维空间信息集成为基础，建立矿山三维空间信息集成系统，并对其应用进行研究。

针对这一目标，本书如图 1-1 所示思路开展研究，即：

(1) 对矿山空间对象的信息及其特征进行深入分析，将矿山空间信息划分为 3 类，即原始信息、成果信息和生产信息，为矿山三维空间信息集成系统提供信息基础。

(2) 对矿山三维空间信息进行深入挖掘，分别实现了对煤岩参数、钻孔信息和测量信息的数据挖掘，并提出了基于方向的空间信息聚类方法和基于梯度的空间信息聚类的新方法，为矿山三维空间信息的进一步应用提供了理论基础。

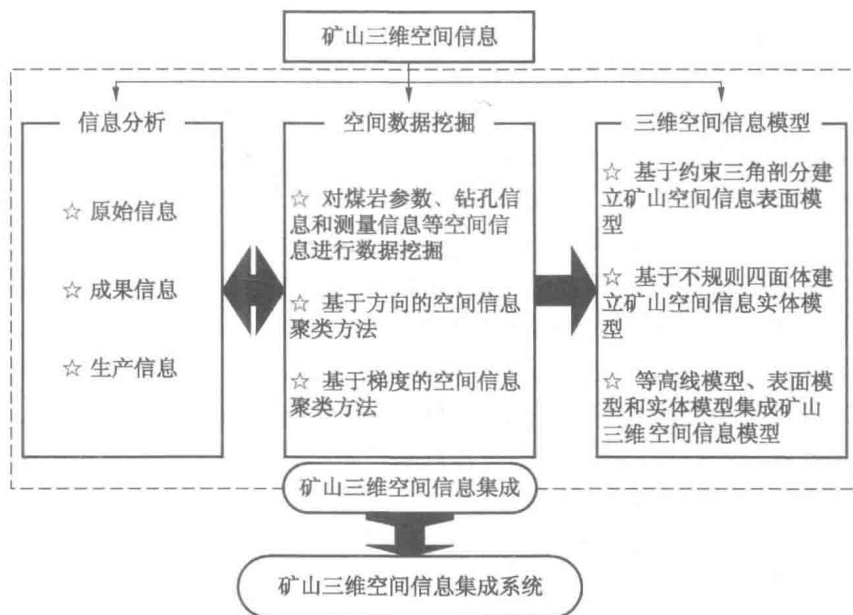


图 1-1 研究思路及主要研究内容

(3) 在对矿山三维空间信息模型进行理论研究的基础上, 提出由等高线信息模型、表面信息模型和实体信息模型集成的矿山三维空间信息模型, 这是矿山系统应用的关键。

(4) 对原始信息、成果信息、生产信息和矿山三维空间信息模型分别进行信息流分析, 得出各类矿山空间信息间的内在联系, 实现了矿山三维空间信息的集成。

(5) 建立矿山三维空间信息集成系统, 并对系统进行应用。

1.4.2 主要内容

根据以上研究目标和研究思路, 结合管理科学、信息科学、矿业工程、数据库理论、数据挖掘理论、计算机图形学和 CAD 技术等诸多理论和方法, 以理论分析和实际应用验证相结合的方法, 对矿山三维空间信息集成系统及其应用进行研究。主要内容如下:

(1) 介绍矿山三维空间信息集成系统的研究意义、研究现状及发展趋势, 提出了研究目标, 并给出了研究思路和主要内容。

(2) 在分析矿山三维空间信息来源的基础上, 将矿山三维空间信息分为原始信息、成果信息和生产信息三大类。为矿山三维空间信息的集成和管理提供信息基础。

(3) 通过对各种空间数据挖掘方法的讨论, 根据矿山空间信息的特点及功能的要求, 选择适合的数据挖掘方法, 并设计相应的数据挖掘算法实现对矿山三维空间信息的数据挖掘。

(4) 针对主要空间信息聚类方法没有考虑空间数据的方向变化的这一特点, 提出了基于方向的空间信息聚类方法, 并设计了方向聚类算法, 用实验数据对算法进行了验证; 利用基于梯度的空间信息聚类方法, 实现对矿山断层信息的聚类, 为断层的识别提供了新方法。

(5) 从基于面表示的模型、基于体表示的模型、基于混合表示的模型三大类角度详细地讨论了各种常见空间信息模型的特点。结合矿山空间信息的特点, 分别提出了基于约束三角剖分建立矿山空间信息表面模型的方法和基于不规则四面体建立矿山空间信息实体模型的方法。

(6) 针对单一矿山空间信息模型的不足, 对由等高线模型、基于约束三角剖分的表面模型和基于不规则四面体的实体模型进行集成, 进而实现了对矿山空间信息模型的集成管理。对原始信息、成果信息、生产信息和矿山空间信息模型三者相互间的信息流进行分析, 得到各类矿山空间信息间的内在联系, 实现对矿山三维空间信息的集成。

(7) 在对矿山系统需求分析的基础上, 提出了以矿山空间数据库为中心, 包括矿山数据管理、矿山三维模型和矿山应用三层结构的概念模型。对系统进行方案设计, 提出了由数据层、数据库平台和应用层组成的矿山信息系统集成的整体逻辑结构。选择开发工具和平台对系统进行开发, 建立了基于模型集成的矿山系统并已在矿山中得到正式的应用。

2 矿山三维空间的信息分析

矿山三维空间可以看作一个广义的地质空间,包括狭义地质体和人工工程。矿山空间对象从成因方面可分为自然地质对象和人工工程对象,从矿山空间的几何特征和空间三维特征上可分为点对象、线对象、面对象、体对象。矿山三维空间信息是对矿山空间对象特征的描述,是建立矿山空间信息集成系统的数据基础。

2.1 矿山空间对象

2.1.1 空间对象概念及特征

将空间现象或事物进行抽象得到空间对象或称空间实体、空间目标^[89]。将矿山空间中的空间现象或事物进行抽象后得到矿山的空间对象。

空间对象包含空间数据、非空间数据和空间关系。从复杂性角度划分,空间对象分为简单对象和复杂对象。简单对象指最基本的、不可再分的对象,或者说没有通过其他对象集成得到;复杂对象是指通过简单对象集成得到,而且复杂对象可以分解为一系列的简单对象。从几何角度划分,空间对象可以用点(0维)、线(1维)、面(2维)、体(3维)来表示。

空间对象是地理空间中有实际意义的最小地理单元,是能独立地反映空间共同定义的实体。它具有空间信息的基本特征,即空间位置特征、属性特征和空间关系特征等。

1. 空间位置特征

空间对象的位置是空间对象和空间基准坐标系之间的关系,可以通过测量并且转换为平面基准坐标。常用的平面基准坐标是用经纬度、空间直角坐标、平面直角坐标、极坐标等方式来表示。空间位置特征也称为几何特征,具体包括空间实体的位置、大小、形状、分布状况等。按照几何定位特征和空间维数,地理空间实体分为点、线、面、表面和体5类地理空间实体。

2. 属性特征

属性是人们通过对周围空间实体的认识、了解和解释,并在头脑中形成相应的对空间对象的定义、描述和说明^[90]。属性特征(又称非空间特征)是与地理空间实体相联系、具有地理意义的数据,用于表达事物本质特征和对实体的语义定义,以区别于其他实体。

3. 空间关系特征

空间关系与人类认识、传输与改造现实世界的活动息息相关,是人类对于空间认知结果的高度概括,是人类所形成的空间概念的最重要的基本组成部分,形成了人类进行空间描述、推理与分析的基础^[91,92]。在通常的情况下,描述与记忆一个空间实体的位置时,不是以几何坐标的形式给出,而是以它与周围物体的关系的形式给出。这些语义属性比几何描述更基本、更重要。如一个学校在哪两条路之间,靠近哪个道路交叉口;一块农田距