

# 基于多维数据挖掘的智能矿山安全隐患分析与管理\*

王晔<sup>1,2</sup>, 姚旭龙<sup>1,2</sup>, 于光远<sup>1,2</sup>, 张艳博<sup>1,2</sup>, 陶志刚<sup>3,4</sup>, 赵继忠<sup>5</sup>

1. 华北理工大学 矿业工程学院, 河北 唐山市 063210;
2. 河北省矿山绿色智能开采技术创新中心, 河北 唐山市 063210;
3. 隧道工程灾变防控与智能养建全国重点实验室, 北京 100083;
4. 中国矿业大学(北京) 力学与土木工程学院, 北京 100083;
5. 首钢滦南马城矿业有限责任公司, 河北 唐山市 063500)

**摘要:**为挖掘智能矿山建设过程中大量安全隐患数据的潜在价值,以山东某矿山为例,对其2014—2023年的安全隐患历史数据进行全面分析,构建了矿山安全管理的多维度分析模型。首先,采用多层感知器(MLP)构建隐患事故识别的人员、设备及环境分类模型;运用潜在狄利克雷分布(LDA)主题模型,将设备隐患归类为照明、运输、支护、电气、消防、爆破、通风及杂物八大主题。然后,基于Apriori算法原理,从非结构化隐患文本中提取关键信息,挖掘分析不同隐患特征及隐患主题之间的关系。最后,利用多维度分析与数据可视化技术相结合的手段,对数据挖掘结果进行了深入分析。结果表明,设备类隐患是该矿山安全管理中需特别关注的高风险领域,顶板欠缺支护、斜坡道路面坑洼、安装开关接地等是显著的隐患主题和关联规则, S16181、S18165等区域是该类型隐患的聚集区。研究构建的多维度分析模型可为矿山安全隐患分析提供参考。

**关键词:**智能矿山;安全隐患管理;数据挖掘;多维分析;关联规则

**中图分类号:**TD77 **文献标识码:**A **文章编号:**1005-2763(2025)10-0173-09

## Analysis and Management of Safety Hazards in Intelligent Mines Based on Multidimensional Data Mining

WANG Ye<sup>1,2</sup>, YAO Xulong<sup>1,2</sup>, YU Guangyuan<sup>1,2</sup>, ZHANG Yanbo<sup>1,2</sup>, TAO Zhigang<sup>3,4</sup>, ZHAO Jizhong<sup>5</sup>

1. College of Mining Engineering, North China University of Science and Technology, Tangshan, Hebei 063210, China;
2. Hebei Mining Green Intelligent Mining Technology Innovation Center, Tangshan, Hebei 063210, China;
3. State Key Laboratory for Tunnel Engineering, Beijing 100083, China;
4. School of Mechanics and Civil Engineering, China University of Mining and Technology-Beijing, Beijing 100083, China;
5. Shougang Luannan Macheng Mining Co., Ltd., Tangshan, Hebei 063500, China)

**Abstract:** In order to explore the potential value of a large amount of safety hazard data in the construction process of intelligent mines, taking a mine in Shandong as an example, comprehensive analysis of its historical safety hazard data from 2014 to 2023 was conducted, and a multidimensional analysis model for mine safety management was constructed. Firstly, a Multi-Layer Perceptron (MLP) was used to construct a personnel, equipment, and environmental classification model for identifying hazards and accidents. Using the Latent Dirichlet Allocation (LDA) topic model, equipment hazards were classified into eight topics of lighting, transportation, support, electrical, firefighting, blasting, ventilation, and miscellaneous. Then, based on the principle of Apriori algorithm, key information was extracted from unstructured hazard text, and the relationship between different hazard features and topics was explored and analyzed. Finally, deep analysis of the data mining results was conducted using a combination of multidimensional analysis and data visualization techniques. The results indicate that equipment related hazards are high-risk areas that require special attention in the safety management of the mine. The lack of support for the roof, potholes on sloping road surfaces, and installation of switch grounding are significant hazard topics and associated rules, and

\* 收稿日期:2024-11-09

基金项目:国家自然科学基金资助项目(52074123);河北省自然科学基金资助项目(E2022209143)

作者简介:王晔(2000—),女,安徽亳州人,硕士研究生,主要研究方向为矿山隐患数据挖掘。E-mail:wwwye2022@163.com

通信作者:姚旭龙(1983—),男,河北张家口人,博士,教授,主要研究方向为岩体工程灾害智能预警、智能开采理论与技术。E-mail:yaoxulong@ncst.edu.cn

the areas such as S16181 and S18165 are gathering areas for this type of hazard. The multidimensional analysis model constructed by the research can provide a basis for the analysis of mining safety hazards.

**Key words:** Intelligent mine, Safety hazard management, Data mining, Multidimensional analysis, Associated rules

## 0 引言

矿山安全管理是保障矿山安全生产的关键<sup>[1]</sup>。近年来,我国矿山行业安全形势总体稳定,但仍存在诸多隐患,矿山安全管理尚有诸多亟待解决的问题。矿山安全事故多与管理不善、操作失误等因素有关,通过对历史安全隐患数据的深入挖掘和分析,能有效预防和控制事故<sup>[2-3]</sup>。

随着智能矿山建设的推进,矿山安全管理信息化水平显著提升,并积累了海量的历史安全隐患数据<sup>[4]</sup>。然而,这些数据尚未得到充分利用,难以有效识别矿山安全管理中的薄弱环节。因此,深入挖掘和分析历史安全隐患数据,对提升矿山安全管理水平具有重要意义。

学者们将数据挖掘技术应用于矿山安全隐患分析,取得了一定的研究成果。WANG Ze 等<sup>[5]</sup>利用数据挖掘方法从大量文本中提取危险隐患信息,建立了煤矿危险源网络以捕捉危险源之间的相关性。QIU Zunxiang 等<sup>[6]</sup>通过文本挖掘来识别煤矿事故的关键致因。LI Shuang 等<sup>[7]</sup>利用数据挖掘技术,深度挖掘煤矿事故案例数据,探索风险因素的交互机制。CAO Xian'gang 等<sup>[8]</sup>提出了一种基于多源文本的煤矿设备维护大型语言模型(XCoalChat),以更好地管理和利用现有的海量煤矿设备维护知识。上述研究成果对矿山企业安全隐患排查、决策制定具有重要意义,但主要从事事故致因角度分析,侧重识别事故的直接原因,忽略了其他潜在的影响维度,研究结果存在一定的局限性。还有部分学者采用多维度视角分析矿山安全隐患。郭对明等<sup>[9]</sup>基于大数据技术构建隐患多维分析模型,分析隐患在时间和空间维度上的分布规律。李国清等<sup>[10]</sup>运用“六何分析方法”建立了多维数据集,综合分析矿山安全隐患。WU Bing 等<sup>[11]</sup>从多个角度分析了我国 70 年来煤矿安全事故的变化趋势,以深入了解事故的特点。LI Hongxia 等<sup>[12]</sup>创新性地将智能采矿与多维分析相结合,建立了瓦斯爆炸灾害预警指标体系。上述研究采用多维度分析方法,提升了对安全隐患的识别能力,增强了安全管理的针对性和有效性。总体来

看,现有研究在利用数据挖掘技术分析矿山安全隐患方面取得了一定进展,但大多从单一角度入手,忽视了隐患的多维特征。因此,有必要采用更全面的分析视角,深入挖掘隐患的时空多维度特征及其内在关联规律。

本文通过对矿山企业历史隐患数据进行深入分析,旨在揭示隐患发生的特征规律,识别矿山安全管理中的薄弱环节,为提升矿山安全管理水平提供决策支持。研究对原始隐患数据进行预处理,解决了文本数据非结构化和非标准化的问题。在此基础上,提出针对矿山安全管理的“四维分析模型”对数据进行多维分析,构建多层感知器分类模型识别三大类隐患(人员、设备、环境),并采用潜在狄利克雷(Latent Dirichlet Allocation, LDA)主题模型对设备类隐患进行深入挖掘,提取出 8 个重要主题。随后,使用 Apriori 算法分析不同主题之间的关联规律,为管控矿山安全隐患提供决策依据。本研究创新性地结合文本挖掘、数据分析与可视化技术,以多维视角系统分析了矿山安全隐患,不仅丰富了该领域的理论研究,也为实际应用提供了有效的分析框架和可操作建议,对保障矿山安全生产具有重要意义。

## 1 安全隐患数据分析架构

### 1.1 安全隐患数据的收集与预处理策略

由于中国矿山安全事故记录尚未形成统一的报告格式和措辞标准,因此事故记录数据具有非结构化和非标准化特征<sup>[13-14]</sup>。本研究以某矿山企业 2014—2023 年的安全隐患事故记录为对象,收集了 11 644 条有效数据,涵盖事故时间、地点等基本信息,为后续的文本挖掘分析奠定了基础。

在数据预处理环节,首先剔除重复值、空值和异常值,以确保数据的有效性和代表性。随后采用基于词典的分词方法,引入“采矿工程”“矿山安全隐患信息分析”等专业词库,对文本进行精准分词,同时使用正则表达式去除分词结果中的标点符号,并利用停用词表过滤无意义的虚词,进一步优化分词结果,为后续的文本分析提供准确且有价值的输入数据。部分安全隐患分词结果见表 1。

表 1 部分安全隐患分词结果

Table 1 Segmentation results for selected safety hazard texts

隐患记录文本	分词结果
西南探矿巷路面积水、淤泥较多。	“西南”“探矿巷”“路面”“积水”“淤泥”“较多”
K-08 <sup>#</sup> 自卸车 12:45 上行时,驾驶员未按规定佩戴安全帽。	“自卸车”“上行”“驾驶员”“未”“按规定”“佩戴”“安全帽”
斜坡道坑洼不平、两帮淤泥、石块、积水等杂物较多,复查没整改,给予区域责任人管理不到位罚款 200 元。	“斜坡道”“坑洼”“不平”“两帮”“淤泥”“石块”“积水”“杂物”“复查”“没”“整改”“给予”“区域”“责任人”“管理”“不到位”“罚款”
掘进作业面利用高压风通风,无照明设施。顶板两帮浮石多,路面文明生产差。风带距迎头较远,吊挂不规范,使用电水泵。耙装机电缆接头多,未整改。	“掘进”“作业面”“利用”“高压风”“通风”“无”“照明”“设施”“顶板”“两帮”“浮石”“多”“路面”“文明”“生产”“差”“风带”“距”“迎头”“远”“吊挂”“不”“规范”“使用”“电水泵”“耙装机”“电缆”“接头”“多”“未”“整改”

## 1.2 矿山安全管理多维度分析模型

为全面理解矿山安全管理隐患的潜在风险及其内在联系,学者们构建了多维度安全管理分析模型<sup>[15-16]</sup>。在此研究基础上,本文提出了“四维分析模型”,围绕矿山安全管理,从隐患类型、发生时间、空间分布和致因分析 4 个维度进行综合分析,提供一种简洁而全面的研究框架。

(1) 隐患类型维度:描述矿山安全隐患的具体类型,包括人员操作不当、设备故障、环境因素等。

(2) 发生时间维度:描述隐患发生的时间特点,如年度、季度或月度变化规律。

(3) 空间分布维度:描述隐患在矿山空间中的分布情况,如主斜坡道、辅助斜坡道、生产工作面等。

(4) 致因分析维度:分析导致隐患发生的根源,如支护不力、照明设备故障、通风系统异常等。

本文提出的“四维分析模型”聚焦矿山安全管理核心要素,后续研究将基于此框架对实际案例进行深入分析,为矿山安全管理实践提供切实可行的建议。

## 2 研究方法

本研究针对智能化矿山系统的数据离散化特点,基于多层感知器对矿山安全隐患进行分类识别,并结合隐患主题挖掘、分类分析及关联规则挖掘,构建了一套高效的数据汇聚与关联分析机制,为矿山安全管理提供决策支持。

### 2.1 基于多层感知器的矿山安全隐患分类

矿山安全管理涉及多种类型的隐患,各类隐患具有不同的特征和影响因素<sup>[17-18]</sup>。为了提高分类效率和效果,本研究将“人-机-环”要素融入矿山隐患分类框架,将隐患分为人员隐患、设备隐患和环境隐患三大类,并采用多层感知器作为分类模型。

多层感知器(Multi-layer Perception, MLP)是一种前馈神经网络模型,能够学习和表示复杂的非线性关系,具有强大的特征映射能力<sup>[19]</sup>。MLP 结

构示意图 1,其每个神经元与前一层所有神经元相连,并通过权重进行连接。

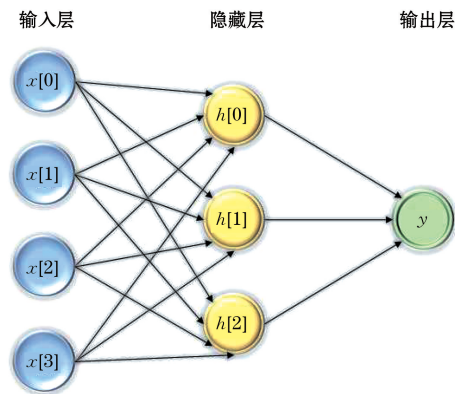


图 1 MLP 结构示意图

Fig.1 Schematic of the MLP structure

为进一步提升分类性能,将文本数据转换为词袋向量作为 MLP 的输入,并采用网格搜索算法优化模型参数。最佳参数组合为[“hidden\_layer\_sizes”: (200, ), “activation”: “tanh”, “solver”: “adam”]。

### 2.2 安全隐患主题挖掘与分类

将安全隐患粗略分为人员、环境和设备三大类,为进一步细化各大类的具体隐患类型,本研究采用潜在狄利克雷分布主题挖掘模型<sup>[20]</sup>,对该矿山企业安全管理隐患文本进行潜在主题挖掘。LDA 通过概率图模型发现文本数据中的主题结构,其原理如图 2 所示。通过建立文档-主题和主题-词语潜在关系的方式,LDA 可有效挖掘文本语料中的隐藏主题结构。

由于矿山安全事故记录尚未标准化,单个事故记录可能涉及多个隐患主题,LDA 主题模型无法准确处理该情况。因此,本研究结合 LDA 主题挖掘结果与点互信息(PMI)方法,设计了一种多主题分类器,通过提取高相关性特征进一步提升分类准确性。具体流程如图 3 所示。

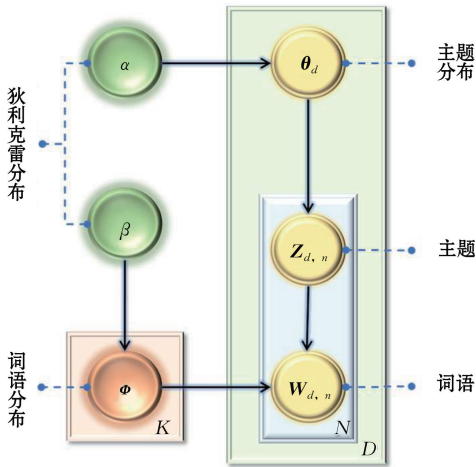


图2 LDA主题模型结构

Fig.2 Structure of the LDA topic model

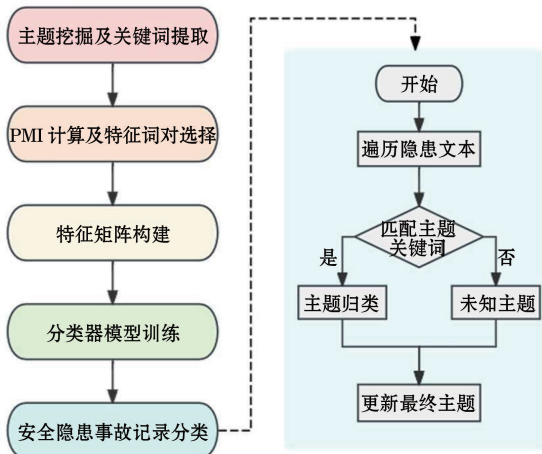


图3 多主题分类流程

Fig.3 Flow for multi-topic classification

### 2.3 基于 Apriori 算法的隐患关联规则挖掘

在矿山安全事故分析中,揭示安全风险因素之间的潜在关联规则对理解事故原因至关重要。Apriori(先验)算法是一种经典的关联规则挖掘算法,被广泛应用于探索导致事故的不确定因素之间的关联特征<sup>[21-23]</sup>,能够为决策提供依据。Apriori 算法的基本原理是通过发现频繁项集和关联规则来揭示数据集中的潜在关联性<sup>[24]</sup>。支持度和置信度是关联规则挖掘中常用的评估指标。

(1) 支持度(S):用于衡量项集或规则在数据集中出现的频率。支持度越高,表明该项集或规则在数据集中越普遍,越能代表整体特征。

$$S(X \rightarrow Y) = P(X \cap Y) \quad (1)$$

(2) 置信度(C):用于衡量规则的强度和可靠性,描述前提条件成立时结论发生的可能性。置信度越高,表明当前提 X 成立时,结论 Y 发生的可能性越大,规则越可靠。

$$C(X \rightarrow Y) = P(Y | X) \quad (2)$$

通过设定的支持度和置信度阈值,可筛选出满足条件的关联规则,帮助矿山管理者制定精准有效的安全管理策略,降低事故风险,提高作业安全性和可持续性。

## 3 试验结果

### 3.1 安全隐患主题 MLP 分类结果

对 11 644 条安全隐患事故数据进行随机采样和训练-测试集划分(测试集占 25%),采用多层感知器模型分类隐患事故,得到人员、设备和环境三种主要隐患类型占比和词云分布,如图 4 所示。



图4 隐患类型占比及词云分布

Fig.4 Proportion of hazard types and word cloud distribution

由图 4 可知,该企业设备类隐患较为突出,涉及照明、支护、通风、爆破等多个主题;相比之下,环境和人员引起的隐患类型较为单一且易于识别。针对设备类隐患,可深入挖掘其潜在主题并进一步分析导致因素,以明确设备管理和维护的重点,减少设备类隐患的发生。人员和环境类隐患相对明确,可采

取针对性的措施进行管控。

### 3.2 安全隐患主题的挖掘与分类分析结果

利用 LDA 主题模型对数据进行隐患主题挖掘分析。困惑度用于评估语言模型对给定数据的预测能力,较低的困惑度表示模型能够更好地对数据进行建模,并产生更准确的预测结果。设置主题数目

范围为 2~30,步长为 5,进行 1 000 次迭代,得到了困惑度与主题数目的关系,如图 5 所示。

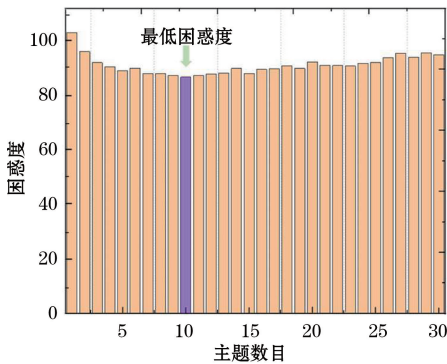


图 5 困惑度与主题数目的关系

Fig.5 The relationship between perplexity and the number of topics

结果表明,当主题数目为 10 时,模型达到最低困惑度,即最佳主题数目为 10。在此基础上,将相近主题合并,最终提取出 8 个主要隐患主题:照明、运输、支护、电气、消防、爆破、通风和杂物。

表 3 为 8 个主要隐患主题的具体信息。其中,支护主题包含的隐患数目最多,其关键词包括锚杆、支护、浮石、裂隙、危石等,表明企业需要重视矿井的支护工作,确保采矿过程的安全性和稳定性。通风主题的关键词包括风带、风机、风管等,反映了企业通风系统在各关键部件存在问题。矿山企业需要建立健全的通风系统管理制度,并落实日常检查维护工作,确保各个环节的通风设备处于最佳状态,从而有效预防和控制通风隐患。

表 3 主题对应自定义隐患主题词

Table 3 Customized hazard topic words corresponding to the identified topics

序号	隐患主题	隐患数目	隐患关键词
1	照明	1 757	“照明”“不亮”
2	运输	372	“卡车”“错车”“驾驶员”“停车”“刹车”“指示灯”“铲运机”
3	支护	4 216	“锚杆”“支护”“浮石”“裂隙”“危石”
4	电气	1 574	“配电箱”“接地”“变压器”“电缆”“电源”“配电柜”“漏电”
5	消防	472	“灭火器”“水管”“消防箱”“无水”“阀门”“管路”“水泵”
6	爆破	834	“导爆管”“炸药箱”“爆炸”“残药”“残炮”“爆破”“盲炮”“裸放”
7	通风	1 842	“风带”“风机”“通风”“风管”“风筒”
8	杂物	206	“标志牌”“防护栏”“警示牌”“护栏”“电话”“通讯”“信号”

### 3.3 关联规则挖掘结果

基于 Apriori 算法原理,对设备类型中各主题的隐患特征进行挖掘。经过多组测试,设置最小置信度为 0.4,最小支持度为 0.005,共获得 532 条关联规则。表 4 为 10 条典型的关联规则。

表 4 典型关联规则

Table 4 Typical association rules

规则	支持度	置信度
顶板,欠缺→支护	0.030 831 3	0.994 459 8
斜坡道,坑洼→路面	0.006 956 3	0.920 454 5
安装,开关→接地	0.005 410 5	0.887 323 9
吉普车→超载	0.007 643 4	0.872 549 0
未戴→安全帽	0.006 355 2	0.860 465 1
裸放→炸药	0.006 183 4	0.857 142 9
消防→水管	0.007 385 8	0.693 548 4
配电箱→接地	0.041 223 0	0.562 719 8
采联→危石	0.015 973 9	0.500 269 0
上锁→导爆管	0.008 502 2	0.487 684 7

“顶板,欠缺→支护”规则表明,当存在顶板问题时,往往伴随有支护设备缺失或不足的问题,这意味

着支护设备的缺陷是导致顶板隐患的重要原因之一。因此,在矿山安全管理中,需要加强对顶板支护设备的检查和维护,确保支护设备能够及时有效地发挥作用。

“配电箱→接地”规则反映出配电系统的接地问题是一个容易被忽视但非常重要的隐患。企业在矿山电气设备管理中,应高度重视配电系统的接地问题,并采取必要的修复和维护措施。

“裸放→炸药”规则表明,炸药存放不善易导致爆炸隐患。在矿山爆破作业中,企业应加强对爆破作业全过程的规范管控,确保炸药的安全使用,从而杜绝各类爆破作业安全隐患。

## 4 安全隐患多维度分析

遵循“What-You-Find-Is-What-You-Fix”原则<sup>[25]</sup>,采用“四维分析模型”选取最具有代表性的三个交叉维度进行分析讨论。

### 4.1 “类型-致因”维度分布规律

矿山隐患通常由多种因素的相互作用产生,具

有复杂的内在联系。为深入分析不同类型隐患(人员、设备、环境)的主要成因,对隐患类型与成因之间的分布规律进行分析,以期管理者制定针对性的分类管控措施提供依据。

(1) 设备类隐患。设备类隐患主要涉及照明、运输、支护、电气、消防、爆破、通风等多个主题,主题间的相关性热力图如图 6 所示。可以看出,“支护-照明”“照明-通风”“支护-通风”“电气-通风”等主题间存在强关联性。企业在优化设备管理时,应综合考虑上述主题特征及其相互关联性,加强设备日常维护和检修,确保设备稳定运行。

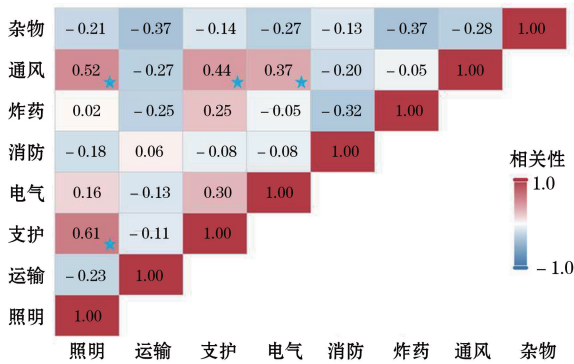


图 6 隐患主题相关性热力图

Fig.6 Heat map of hazard topic correlations

(2) 人员类隐患。主要涉及违规行为,如驾驶员超速超载、未佩戴安全帽、人货混载等,其根源在于个体安全意识薄弱、管理制度不完善、监管机制缺失等。针对这些隐患,企业需要加强培训和监督,健全安全管理体系。

(3) 环境类隐患。主要包括路面积水、道路不平整、遗留杂物等,易引发交通事故和工地伤害,对员工和设备安全构成潜在威胁,其根源在于环境整治不到位、维护不及时等。因此,企业需加强环境整治和日常维护,确保生产环境安全。

在综合分析三类隐患的主要成因和相互关联的基础上,可为矿山管理者制定以下分类管控措施:

(1) 定期检查设备运行状态,加强设备维护保养,提高设备使用效率和安全性,重点管控支护、照明、通风等关联主题,预防设备类隐患发生;

(2) 加强员工安全教育培训,提高员工安全意识和技能水平,健全安全管理制度和流程,强化监督检查机制,有效控制人员类隐患;

(3) 加强环境整治和安全巡查,及时清理环境

隐患,修复破损设施,保持工作区域整洁有序,保持通风通畅、路面平整,减少环境隐患对生产安全的影响。

通过综合管控不同类型隐患间的关联,采取上述针对性措施,企业可有效降低安全事故风险,提升安全管理水平。

#### 4.2 “时间-空间”维度分布规律

为深入研究矿山安全事故发生的时空特征,选择月份和班次两个不同粒度的时间维度,分析事故频发地点的时空分布规律,以期有针对性地采取预防和应对措施提供依据。

(1) 月份维度。图 7 为月份-隐患地点热力图,展示了前 6 个隐患易发地每个月份的隐患分布情况。可以看出,主斜坡道和辅助斜坡道全年存在较高的安全隐患,表明这两个区域的安全管理长期存在问题,需要进行重点整治。生产作业区域如 S16181、S18156 和 S19156 的隐患呈现出明显的时间波动特征,需要根据不同时期的具体生产需求,采取针对性的管控措施。此外,斜井口在 3 月和 7 月出现了安全隐患高峰,这可能与斜井口作为矿山重点物流通道,在生产高峰期人员和车辆流量增大有关。当前该区域的安全管理仍存在薄弱环节,需要矿山制定专项整治方案,如增加监控覆盖、优化人员调配、加强现场管理等。

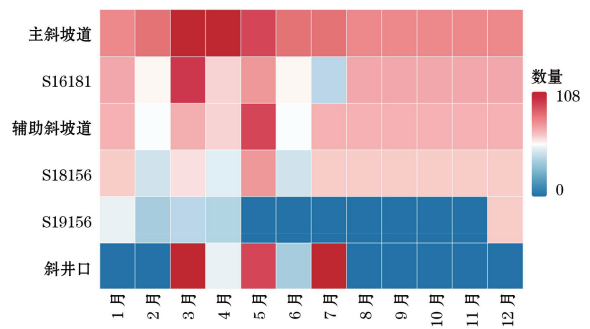


图 7 月份-隐患地点热力图

Fig.7 Heat map of month-hazard location

(2) 班次维度。图 8 为班次-隐患地点桑基图,展示了不同班次发生隐患的空间分布情况。整体来看,8:00 班的隐患发现频率(24:00 至次日 8:00)明显高于 16:00 班和 24:00 班,表明夜班时段隐患发生整体活跃度较高。主斜坡道和辅助斜坡道作为关键通道,在各个时段都保持较高的使用频率,突出了其在矿山生产中的重要性。



图 8 班次-隐患地点桑基图

Fig.8 Sankey diagram of shift-hazard location

### 4.3 “空间-类型”维度分布规律

不同区域、不同类型隐患的形成机理和管控难点存在差异,为深入探究矿山安全隐患分布规律,对隐患类型与空间分布进行综合分析。

图 9 展示了三种隐患类型在前 10 个隐患频发地上的分布特征。可以看出,人员类型隐患集中在主斜坡道、辅助斜坡道和斜井口;S16181、S18165 等地段则是设备类型隐患的聚集区。

为更直观地展示隐患的空间分布特点,绘制了前 6 个隐患频发地的隐患词云分布,如图 10 所示。可以看出,主斜坡道主要存在吉普车超载、路面不平等问题;辅助斜坡道路面状况不佳,影响车辆行驶安全;S16181、S18156 和 S19156 等区域存在顶板支护不稳等隐患;斜井口则容易存在违章和超载等问题。上述分析结果可为矿山针对性安全管理提供参考,根

据不同区域的隐患特点,需采取针对性的管控措施,如加强道路养护、规范车辆运行、优化顶板支护等。

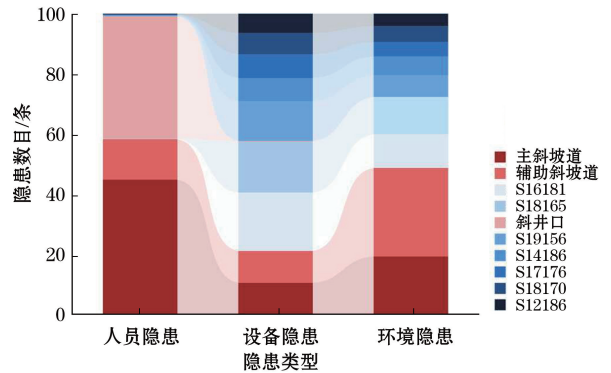


图 9 三种隐患类型的空间分布特征

Fig.9 The spatial distribution characteristics of three types of hazards



图 10 前 6 个隐患频发地的隐患词云分布

Fig.10 Word cloud distribution for the top 6 high-frequency locations of safety hazards

## 5 结论

本研究创新性提出了一种矿山安全隐患识别与评估框架,成功融合了数据挖掘技术和多维度分析方法。利用 LDA 主题模型和关联规则挖掘,从矿山安全隐患数据中识别出 8 个主要隐患主题,并深入分析了不同隐患类型之间的相互关系。研究为矿山安全管理领域提供了新的视角,得到以下主要结论。

(1) 设备类隐患是该矿山安全管理中需特别关注的高风险领域。对设备类型中各主题的隐患特征进行挖掘,获得了 532 条关联规则。顶板欠缺支护、斜坡道路面坑洼、安装开关接地等强关联风险需要重点进行防范。

(2) “支护-照明”“照明-通风”“支护-通风”“电气-通风”等主题间存在强关联性;人员类隐患主要涉及违规运输、未佩戴安全帽、人货混载等行为;环境类隐患主要包括路面积水、道路不平坦、遗留杂物等问题。

(3) 安全隐患的时空分布规律表明,主斜坡道和辅助斜坡道全年存在较高的安全隐患,斜井口在 3 月和 7 月出现安全隐患高峰,需要根据具体生产需求采取针对性的管控措施;夜班时段隐患发生整体活跃度较高,需要加强夜间监管。

(4) 人员类型隐患集中在主斜坡道、辅助斜坡道和斜井口,设备类型隐患聚集于 S16181、S18165 等地段。其中,主斜坡道主要存在吉普车超载、路面不平等问题,辅助斜坡道路面状况不佳,S16181、S18156 和 S19156 等区域存在顶板支护不稳等隐患,斜井口易出现违章和超载等问题。针对不同区域的隐患特点,需采取针对性的管控措施,如加强道路养护、规范车辆运行、优化顶板支护等,以确保矿山安全生产。

### 参考文献(References):

[1] 白润才,付恩三,马力,等.露天煤矿安全-绿色-高效-低碳协同开采技术体系[J].煤炭学报,2024,49(1):298-308.  
BAI Runcai, FU Ensan, MA Li, et al. Collaborative mining technological system of safety-green-high efficiency-low carbon for open pit coal mine[J]. Journal of China Coal Society,2024,49(1):298-308.

[2] 付净,聂方超,荆德吉,等.数据挖掘在预测组织事故防控效果中的应用[J].中国安全生产科学技术,2020,16(10):152-157.  
FU Jing, NIE Fangchao, JING Deji, et al. Application of data mining in predicting prevention and control effect of organizational accidents[J]. Journal of Safety Science and Technology,2020,16(10):152-157.

[3] JISKANI Izhar Mithal, HAN Shuai, REHMAN Atta Ur, et al. An integrated entropy weight and grey clustering method-based evaluation to improve safety in mines[J]. Mining Metallurgy & Exploration,2021,38(4):1773-1787.

[4] 牛莉霞,赵蕊.大数据时代煤矿安全风险治理模式研究[J].煤矿安全,2022,53(7):241-245.  
NIU Lixia, ZHAO Rui. Research on coal mine safety risk management model in the era of big data[J]. Safety in Coal Mines,2022,53(7):241-245.

[5] WANG Ze, LI Huajiao, TANG Renwu. Network analysis of coal mine hazards based on text mining and link prediction[J]. International Journal of Modern Physics C,2019,30(7):1940009.

[6] QIU Zunxiang, LIU Quanlong, LI Xinchang, et al. Construction and analysis of a coal mine accident causation network based on text mining[J]. Process Safety and Environmental Protection,2021,153:320-328.

[7] LI Shuang, YOU Mengjie, LI Dingwei, et al. Identifying coal mine safety production risk factors by employing text mining and Bayesian network techniques[J]. Process Safety and Environmental Protection,2022,162:1067-1081.

[8] CAO Xian'gang, XU Wangtao, ZHAO Jiangbin, et al. Research on large language model for coal mine equipment maintenance based on multi-source text[J]. Applied Sciences-Basel,2024,14(7):2946.

[9] 郭对明,李国清,胡乃联,等.基于文本挖掘的矿山安全隐患大数据分析可视化[J].工程科学学报,2022,44(3):328-338.  
GUO Duiming, LI Guoqing, HU Nailian, et al. Big data analysis and visualization of potential hazardous risks of the mine based on text mining[J]. Chinese Journal of Engineering,2022,44(3):328-338.

[10] 李国清,李学玉,侯杰,等.矿山安全隐患辨识与预警大数据分析系统研发[J].金属矿山,2022(6):129-137.  
LI Guoqing, LI Xueyu, HOU Jie, et al. Development of big data analysis system for mine safety hazard identification and early warning[J]. Metal Mine,2022(6):129-137.

[11] WU Bing, WANG Jingxin, ZHONG Mingyu, et al. Multidimensional analysis of coal mine safety accidents in China: 70 years review[J]. Mining, Metallurgy & Exploration,2023,40(1):253-262.

[12] LI Hongxia, ZHANG Yiru, YANG Wanli. Gas explosion early warning method in coal mines by intelligent mining system and multivariate data analysis[J]. Plos One,2023,18(11):e0293814.

[13] 王国法.煤矿智能化最新技术进展与问题探讨[J].煤炭科学技术,2022,50(1):1-27.  
WANG Guofa. New technological progress of coal mine intelligence and its problems[J]. Coal Science and Technology,2022,50(1):1-27.

- [14] 王国法,庞义辉,任怀伟,等.智慧矿山系统工程及关键技术研究与实践[J].煤炭学报,2024,49(1):181-202.  
WANG Guofa, PANG Yihui, REN Huaiwei, et al. System engineering and key technologies research and practice of smart mine[J]. Journal of China Coal Society,2024,49(1):181-202.
- [15] TONG Ruipeng, ZHANG Yanwei, CUI Pengcheng, et al. Characteristic analysis of unsafe behavior by coal miners: multi-dimensional description of the pan-scene data [J]. International Journal of Environmental Research and Public Health,2018,15(8):1068.
- [16] CHEN Hong, QI Hui, LONG Ruyin, et al. Research on 10-year tendency of China coal mine accidents and the characteristics of human factors[J]. Safety Science,2012,50(4):745-750.
- [17] MENG Qingfeng, TAN Shaohong, LI Zhen, et al. A review of game theory application research in safety management [J]. IEEE Access,2020,8:107301-107313.
- [18] XI Jianfeng, GUO Hongyu, JIAN Tian, et al. A classification and recognition model for the severity of road traffic accident [J]. Advances in Mechanical Engineering, 2019,11(5):1-8.
- [19] LIN Ruiyuan, ZHOU Zhiruo, YOU Suya, et al. Geometrical interpretation and design of multilayer perceptrons[J]. IEEE Transactions on Neural Networks and Learning Systems,2024,35(2):2545-2559.
- [20] BLEI David M, NG Andrew Y, JORDAN Michael I. Latent dirichlet allocation [J]. Journal of Machine Learning Research,2003,3(4/5):993-1022.
- [21] CAI Qiuru. Cause analysis of traffic accidents on urban roads based on an improved association rule mining algorithm[J]. IEEE Access,2020,8:75607-75615.
- [22] GUO D G. Correlation analysis of coal mine safety accidents based on Apriori algorithm [J]. Fresenius Environmental Bulletin,2022,31(9):9355-9361.
- [23] LIU Ying, LIANG Yunpei, LI Quanguai. Cause analysis of coal mine gas accidents in China based on association rules [J]. Applied Sciences,2023,13(16):9266.
- [24] WU Xindong, KUMAR Vipin, QUINLAN J Ross, et al. Top 10 algorithms in data mining [J]. Knowledge and Information Systems,2008,14(1):1-37.
- [25] LUNDBERG Jonas, ROLLENHAGEN Carl, HOLLNAGEL Erik. What-You-Look-For-Is-What-You-Find—The consequences of underlying accident models in eight accident investigation manuals[J]. Safety Science,2009,47(10):1297-1311.