

中文引用格式:曾大林,肖方正,姜志超,等.数据驱动下建筑施工安全隐患关联分析与预警策略[J].中国安全科学学报,2025,35(5):56-63.

英文引用格式:ZENG Dalin, XIAO Fangzheng, JIANG Zhichao, et al. Data-driven correlation analysis and early warning strategy for potential safety hazards in construction[J]. China Safety Science Journal, 2025, 35(5): 56-63.

数据驱动下建筑施工安全隐患关联分析与预警策略*

曾大林 教授,肖方正**,姜志超,张海洋
(山东建筑大学 管理工程学院,山东 济南 250101)

中图分类号:X928.03

文献标志码:A

DOI: 10.16265/j.cnki.issn1003-3033.2025.05.0883

资助项目:山东省重点研发计划项目(2021CXGC011204);山东省住房城乡建设科技计划立项项目(2024KYKF-JZGYH129)。

【摘要】 为深入挖掘施工安全隐患数据的潜在预警价值,提高隐患排查治理效率,利用文本挖掘技术、关联规则和复杂网络理论,研究数据驱动下建筑施工安全隐患关联分析与预警策略。首先,借助文本挖掘技术标准化降维处理某地2023年度的1405条建筑施工安全巡检记录,并提取出67个安全隐患特征;其次,通过Apriori关联规则算法得到70个频繁项集,125条强关联规则,并确定隐患的关联类型;然后,结合复杂网络理论构建施工安全隐患特征网络,基于网络结构指标和个体指标分析,并结合特征模块化分析,辨识关键安全隐患特征;最后,提出一种基于特征数据驱动的安全隐患预警策略。结果表明:文本挖掘技术能有效降维非标准化隐患数据;基于关联规则的隐患特征网络能充分挖掘隐患数据间的潜在关联,加强隐患预警信息的可靠性;预警策略有助于解决传统隐患排查中无序、低效的问题。

【关键词】 数据驱动; 建筑施工; 安全隐患; 关联规则; 预警策略; 复杂网络

Data-driven correlation analysis and early warning strategy for potential safety hazards in construction

ZENG Dalin, XIAO Fangzheng, JIANG Zhichao, ZHANG Haiyang

(School of Management and Engineering, Shandong Jianzhu University,
Ji'nan Shandong 250101, China)

Abstract: To explore the potential early warning value of construction safety hazard data and improve the efficiency of hazard identification and control, this study investigated a data-driven approach for the association analysis and early warning strategy of construction safety hazards by integrating text mining, association rule mining, and complex network theory. First, 1405 construction safety inspection records in 2023 were standardized and dimensionally reduced using text mining techniques, resulting in the extraction of 67 safety hazard features. Then, 70 frequent itemsets and 125 strong association rules were obtained using the Apriori algorithm, and the types of hazard associations were identified. Subsequently, a hazard feature network was constructed based on complex network theory. Key hazard features were identified

* 文章编号:1003-3033(2025)05-0056-08; 收稿日期:2025-01-12; 修稿日期:2025-03-21

** 通信作者:肖方正(2000—),女,山东肥城人,硕士研究生,主要研究方向为施工现场人机交互安全。E-mail:18553847056@163.com。

through structural and node-level indicators, combined with feature modularity analysis. Finally, a feature-driven early warning strategy was proposed. The results show that text mining effectively reduces the dimensionality of unstructured hazard data. The hazard feature network based on association rules successfully reveals hidden associations within the data and enhances the reliability of early warning information, providing clear direction and reference for on-site hazard detection. The early warning strategy helps address the issues of disorder and inefficiency in traditional hazard inspections.

Keywords: data-driven; construction; safety hazards; association rules; warning strategy; complex networks

0 引言

事故源于隐患的理念已被广泛认可,加强安全隐患排查治理、实现隐患预警对降低建筑施工现场事故率有着积极作用^[1]。然而,安全隐患种类繁多且关联复杂,加之其高度不确定性和频发特性,给施工现场的隐患排查治理工作带来挑战^[2]。随着建筑业安全管理向智能化与信息化转型,隐患排查治理模式逐渐由经验主导转向数据驱动^[3-5]。目前,安全隐患数据的采集效率与规模正逐步提升,但由于隐患数据采集未标准化,施工现场积累了大量非结构化的隐患数据,仅靠人力难以充分发掘和发挥数据的潜在价值。因此,如何高效提取文本记录中隐患特征,挖掘隐患间关联并创造预警价值,亟需展开深入研究。

在施工安全隐患的数据降维和特征提取方面,现有研究多依赖经验或规范标准进行识别与分析,存在片面性和主观性^[6]。为此,一些研究尝试从数据驱动视角展开研究,基于事故调查报告、安全巡检记录和施工文档等文本数据^[2,7],结合文本挖掘和自然语言处理等方法逐渐实现隐患文本中关于隐患名称、类型、整改情况等特征的提取^[8]。但多数研究将隐患特征视为孤立存在,忽视了隐患特征间的潜在联系。因此,需要进一步研究安全隐患间的关联,并从中挖掘出更多潜在预警信息,以指导隐患排查治理的实践。研究施工安全隐患的复杂关联机制尚处于起步阶段。有关影响因素耦合关系的研究多采用决策实验室分析法(Decision-Making Trial and Evaluation Laboratory, DEMATEL)和解释结构模型(Interpretative Structural Modeling, ISM)^[9],以及贝叶斯网络^[10]等方法。其中,DEMATEL-ISM法难以应对隐患数据的动态变化,贝叶斯网络则对数据质量要求较高,部分先验概率确定需依赖专家知识。相比之下,复杂网络作为近年来备受关注的方

法,不仅能有效降低主观干扰,还能适应隐患数据的动态变化^[4]。在复杂网络模型的构建过程中,边的定义

直接决定模型能否准确反映安全隐患的核心特征与规律,其构建方法主要有事故链^[4]、共现率^[5]和关联规则挖掘^[11]3种。事故链通过分析事故报告确定隐患优先级来构建网络,耗费人力且易受主观因素影响。共现率是通过计算隐患同时出现的频率来确定边,避免了主观因素干扰,但难以确定隐患间是否存在实质性联系。关联规则挖掘引入提升度,提高了关联分析的有效性,且数据无需满足特定的分布或假设检验条件,剔除了主观干扰。

鉴于此,笔者拟运用文本挖掘技术进行隐患数据的标准化降维和特征提取,基于Apriori关联规则和复杂网络理论构建符合工程实际的隐患特征网络,充分挖掘隐患特征间的潜在关联关系,明确隐患关联分类和隐患特征分类,辨识关键特征,进而提出基于数据驱动的施工安全隐患预警策略,以提高现场安全隐患排查治理效率。

1 建筑施工安全隐患特征提取

1.1 数据来源

收集某地2023年度对53个建设项目的施工安全巡检记录作为原始数据,共计1405条。每条隐患数据包括项目编号、巡检日期、隐患问题表述及巡检重点等信息。部分隐患记录见表1。

表1 安全隐患记录(部分)

Table 1 Safety hazard records (partial)			
项目编号	巡检日期	隐患问题表述	巡检重点
2350101	2023-05-31	施工单位技术负责人未到岗履责	安全行为
2350501	2023-06-08	钢网架预埋件底部混凝土存在孔洞	工程实体
2360601	2023-06-02	脚手架作业层未挂设随层平网	现场施工
⋮	⋮	⋮	⋮
2360901	2023-06-07	塔机附着装置附墙支座与建筑物连接螺栓松动	起重机械

- 1) 扫描初始数据库并计数,选择满足支持度阈值的候选项集,得到频繁项集 T_1 。
- 2) 通过连接和剪枝产生候选项集。
- 3) 重复步骤 1) — 2),直至无法发现频繁项集 T_k 。
- 4) 产生每个频繁项集 T 的所有非空子集 S 。
- 5) 对于每个非空子集 S ,如果满足置信度阈值,则输出关联规则。

2.2 复杂网络构建

复杂网络是一种将系统中元素抽象为节点,元素间关系抽象为边的网络模型^[3]。为实现隐患特征与网络节点的一一对应,基于二阶强关联规则确定网络中关键节点和边,节点表示关键隐患特征,边表示规则前后项的关联关系,置信度作为边的权重。

特征网络结构整体特征由图密度、网络直径、平均路径长度和模块度等指标描述^[5]。其中,图密度用于衡量节点间连边的紧密程度,网络直径为任意 2 个节点之间的最长距离。平均路径长度为网络中任意 2 个节点之间的平均距离,用于衡量网络的信息传播效率,值越低表示信息传播效率越高。模块度用于衡量网络中节点聚类或社区结构的程度,数学表达式如下:

$$Q = \frac{1}{2m} \sum_{ij} \left(M_{ij} - \frac{k_i k_j}{2m} \right) \delta(i, j)$$

(3)

式中: m 为边的个数; M_{ij} 为节点 i 和 j 间的连边数; k_i 、 k_j 为节点 i 和 j 的度; $\delta(i, j)$ 为 1 时 i 和 j 在同一社区, $\delta(i, j)$ 为 0 时 i 和 j 在不同社区。

特征网络结构个体特征由出度中心性 G 、入度中心性 I 、介数中心性 B 、特征向量中心性 E 衡量^[3,5,11]。其中, G 为节点 i 指向其他节点的边的数量,表示隐患的主动转移能力。 I 为其他节点指向节点 i 的边的数量,表示隐患特征受其他特征的影响程度。 B 为通过节点 i 的最短路径条数,表示某些隐患在网络中的“桥梁”作用,数学表达式如下:

$$B_i = \sum_{s \neq i \neq t} \frac{n_{st}^i}{g_{st}}$$

(4)

式中: n_{st}^i 为经过节点 i 且为最短路径的路径数量; g_{st} 为连接 s 和 t 的最短路径数量。 E 为节点 i 的重要程度,侧重于评估节点相邻节点的数量和质量,衡量节点的长期影响力,数学表达式如下:

$$\lambda E_i = A^T E_i$$

(5)

式中: E_i 为邻接矩阵 A^T 的特征向量; λ 为对应特征向量。

3 建筑施工安全隐患特征关联分析

3.1 隐患特征关联规则挖掘

合适的支持度和置信度阈值对于生成符合要求的关联规则至关重要^[13]。鉴于建筑施工现场环境复杂、潜在隐患众多,为避免遗漏一些关键特征的关联规则,将支持度和置信度阈值分别设置为 0.01 和 0.2。最终共挖掘出 70 个频繁项集和 125 条强关联规则,提升度均大于 1,说明这些规则的关联性明显且具有意义,部分典型关联规则见表 2。

表 2 关联规则(部分)

Table 2 Association rules (partial)

序号	关联规则	支持度	置信度	提升度
1	符合要求/符合规范/符合标准,作业层⇒脚手架	0.012 1	1	7.025 0
6	洞口/井口⇒防护措施/防护设施	0.023 5	0.916 7	13.701 2
12	支架⇒模板	0.014 2	0.833 3	11.478 8
14	连墙件⇒脚手架	0.014 2	0.800 0	5.620 0
17	指导性⇒专项施工方案	0.012 8	0.782 6	7.281 9
25	监理单位,签章/盖章/章⇒验收	0.012 1	0.708 3	5.888 8
33	后浇带⇒模板	0.012 1	0.653 8	9.006 4
37	签章/盖章/章⇒记录	0.023 5	0.600 0	5.403 8
48	基础⇒塔机/塔式起重机/起重机械/塔吊	0.022 8	0.516 1	3.699 8
64	脚手架⇒符合要求/符合规范/符合标准	0.065 5	0.460 0	2.051 7

1) 由关联规则 1 和 64 可知:脚手架的安全使用与规范要求密切相关。作业层若符合规范要求,脚手架配套使用几乎是必然的。

2) 由关联规则 6 可知:洞口/井口出现时,有约 90% 的概率伴随出现防护措施/防护设施。因此,此类高风险区域应重点落实防护措施设置情况。

3) 关联规则 25 和 37 表明:监理单位签章行为与验收、记录间存在明显关联。监理单位在验收和记录过程中发挥着关键作用,通过其签章或盖章,可认为相关施工环节或设备已通过必要的验收程序并得到准确记录。

3.2 隐患特征关联分类

不同安全隐患特征下的隐患存在重叠部分。如“楼梯、楼体临边防护未设置挡脚板,防护立杆间距大于 2 m 规定”同时表征防护措施/防护设施、挡脚板、立杆和楼梯等特征。基于关联规则,根据作业空间、作业时间、同一诱因和潜在对象将隐患特征关联

续表 4

序号	模块命名	模块含义	隐患特征数量	隐患数量	主要特征关联类型
4	施工机械和设备设置	涉及施工机械设备的操作规范和设备设置的隐患特征 例如:塔机/塔式起重机/起重机械/塔吊、说明书、人员、设施/设备/设置	10	596	空间接触、时间交叉、相同潜在对象
5	安全管理行为	涉及现场安全管理行为规范的隐患特征 例如:记录、验收、签字/签名、监理单位、签章/盖章/章、桩基	9	378	时间交叉、相同潜在对象
6	安全防护和工程实体	涉及工程实体质量和关键位置安全防护的隐患特征 例如:符合要求/符合规范/符合标准、防护措施/防护设施、洞口/井口、平台、楼梯、结构、混凝土、质量	12	599	空间接触、同一诱因、相同潜在对象

注:文中所用数据中,与桩基关联的特征均属于安全管理行为,故桩基被聚类到安全管理行为这一模块。

1) 各模块的安全隐患特征数量分布在 3~12 的范围。其中,模块 6 和模块 4 的安全隐患数量最多,分别为 599 和 596 项;而模块 2 包含的安全隐患数量最少,仅有 87 项。

2) 各模块特征的关联类型也不相同,模块 5 和模块 2 隐患多涉及相同潜在对象,如记录和监理/总监,一旦发生,需重点排查与这 2 个隐患潜在对象相关的隐患特征。模块 1 包含的关联类型更为复杂,主要有空间接触、时间交叉和同一诱因。当施工现场有隶属于该模块的隐患发生时,应重点排查作业空间内与该隐患时间交叉和有共同诱因的隐患,采用整体管控和分层排查相结合的策略。

4.2 安全隐患特征网络个体指标分析

以出度中心性为例进行排序,安全隐患特征网络的个体指标分析结果见表 5。

表 5 个体指标分析结果(部分)

Table 5 Individual indicators analysis results(partial)				
隐患特征	<i>G</i>	<i>I</i>	<i>B</i>	<i>E</i>
符合要求/符合规范/符合标准	14	1	0.004	0.567
验收	11	2	0.101	1.000
脚手架	10	1	0.011	0.533
资料	6	3	0.102	0.577
记录	6	3	0.042	0.771
交底	6	0	0.080	0.286
模板	5	3	0.039	0.128
混凝土	3	5	0.010	0.082
防护措施/防护设施	3	4	0.035	0.136
签章/盖章/章	3	4	0.015	0.483
作业层	3	1	0.000	0.014
力矩	2	4	0.002	0.040
塔机/塔式起重机/起重机械/塔吊	2	4	0.000	0.257
平台	2	2	0.047	0.176

1) 除少数关键隐患特征外,大部分隐患特征的

出度中心性差异不大。说明网络由少数核心特征主导,整体为多中心结构。出入度差异较大的节点有符合要求/符合规范/符合标准、脚手架、验收、交底等,对施工现场安全管理具有直接且重要的影响,应作为现场安全管理工作的重点关注对象。相对而言,混凝土、塔机/塔式起重机/起重机械/塔吊等隐患特征の出度中心性低而入度中心性高,表明它们更易受到网络中其他隐患的影响,通常是在其他隐患产生后才被触发。在资源受限的情况下,可适度降低这些隐患的监控优先级。

2) 介数中心性的排名较高的是资料、验收、交底、平台和记录,表明这些特征在隐患传播中起枢纽作用,大部分隐患通过它们来触发其他隐患,应优先管控以降低连锁反应风险。

3) 特征向量中心性排名较高的是验收、记录、资料、符合要求/符合规范/符合标准和脚手架,说明它们与大量重要节点紧密相连,具有较大的综合影响力,是需要重点防范的核心隐患。

5 基于数据驱动的安全隐患预警策略

在建筑施工安全隐患特征网络模型的基础上,进一步提出一种数据驱动的安全隐患预警策略,即依托建筑施工现场积累的隐患文本数据驱动,构建面向现场应用的隐患预警信息平台,如图 4 所示。该平台包括 4 个功能模块:

1) 数据积累模块。人工导入施工现场的隐患记录台账,持续迭代平台数据,逐步提高数据驱动预警的准确性。

2) 隐患输入模块。实时记录施工现场所发现的隐患,如巡检员在检查现场施工时发现隐患并输入到平台,经平台标准化处理后生成标准化描述,并同步传输到关联运算模块和数据积累模块。

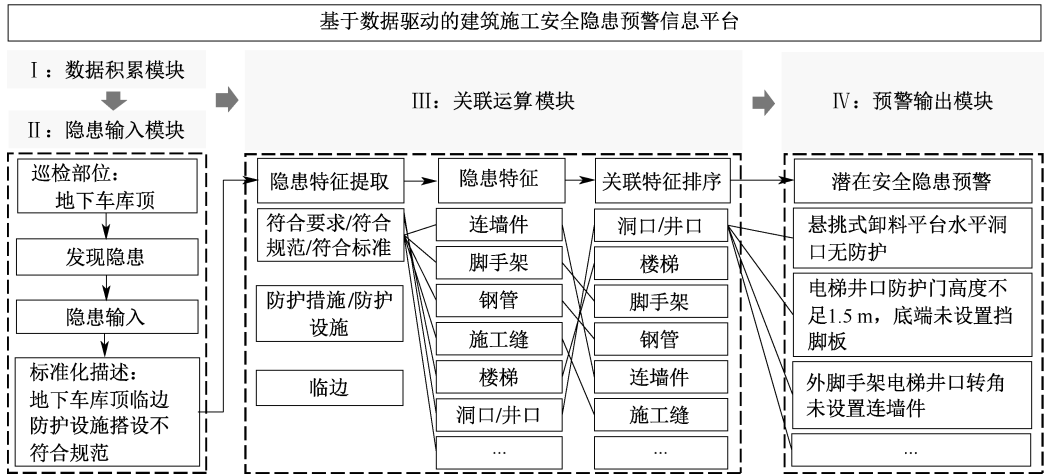


图 4 隐患预警信息平台示例

Fig. 4 Example of hazard early warning information platform

3) 关联运算模块。基于隐患特征网络模型,提取标准化隐患描述的隐患特征,并进行关联分析等处理,得到关联特征排序。

4) 预警输出模块。确定关键关联特征后,从平台检索关联隐患,输出预警信息,为确定下一步的巡检重点提供依据。

上述预警策略通过及时反馈潜在隐患信息,能够为巡检员提供隐患排查方向,解决因无针对性排查造成的低效巡检问题。

6 结 论

1) 从数据驱动视角,利用文本挖掘,建立隐患数据的标准化及特征提取方法,实现 1405 条非标准

化隐患记录到 67 个隐患特征的数据降维,为现场隐患排查治理提供关键信息支撑。

2) 结合 Apriori 算法和复杂网络理论构建隐患特征网络,实现安全隐患间关联分析,基于隐患特征网络的隐患预警策略,辅助优化隐患预警流程并升级隐患预警工具,帮助建筑施工现场精准开展隐患排查治理,快速、准确地实现隐患预警。

3) 文中仅关注分词结果中具有实际意义且频率较高的名词,但其他词性的分词结果也可能蕴含重要信息,后续研究将考虑更多词性的潜在特征及其关联性。此外,文中侧重于描绘基于众多巡检项目数据关联预警的情况,未来会针对某个具体项目积累数据,验证该关联预警方法的有效性。

参 考 文 献

[1] 赵东风, 刘尚志, 王浩水, 等. 基于风险的隐患定义及管控研究[J]. 中国安全科学学报, 2023, 33(2): 1-8.
ZHAO Dongfeng, LIU Shangzhi, WANG Haoshui, et al. Unsafe-condition definition and control based on risk[J]. China Safety Science Journal, 2023, 33(2): 1-8.

[2] TIAN Dan, LI Mingchao, SHEN Yang, et al. Intelligent mining of safety hazard information from construction documents using semantic similarity and information entropy[J]. Engineering Applications of Artificial Intelligence, 2023, 119: DOI: 10.1016/j.engappai.2022.105742.

[3] 陈述, 朱雨萍, 陈云, 等. 基于复杂网络的水电工程施工安全隐患时序特性[J]. 中国安全科学学报, 2022, 32(8): 61-66.
CHEN Shu, ZHU Liping, CHEN Yun, et al. Sequential characteristics of safety hazards in hydropower project construction based on complex networks[J]. China Safety Science Journal, 2022, 32(8): 61-66.

[4] 尼颖升, 张卫强, 司涵, 等. 基于复杂网络的桥梁施工事故成因识别及安全管理对策[J]. 公路交通科技, 2023, 40(增2): 164-173.
NI Yingsheng, ZHANG Weiqiang, SI Han, et al. Bridge construction accident causation identification and safety management strategy based on complex network[J]. Journal of Highway and Transportation Research and Development, 2023, 40(S2): 164-173.

- [5] 刘梅, 许林宇, 廖彬超, 等. 基于数据驱动的施工安全隐患特征网络分析与预警策略[J]. 清华大学学报:自然科学版, 2023, 63(2): 191–200.
- LIU Mei, XU Linyu, LIAO Binchao, et al. Data-driven network analysis of construction hazard characteristics and warning strategy[J]. Journal of Tsinghua University: Science and Technology, 2023, 63(2): 191–200.
- [6] MENG Qingfeng, PENG Qiyuan, LI Zhen, et al. Big data technology in construction safety management: application status, trend and challenge[J]. Buildings, 2022, 12(5): DOI: 10.3390/buildings12050533.
- [7] XU Na, MA Ling, LIU Qing, et al. An improved text mining approach to extract safety risk factors from construction accident reports[J]. Safety Science, 2021, 138: DOI: 10.1016/j.ssci.2021.105216.
- [8] ZHU Yuqian, LIAO Huimin, HUANG Dengchi. Using text mining and multilevel association rules to process and analyze incident reports in China[J]. Accident Analysis and Prevention, 2023, 191: DOI: 10.1016/j.aap.2023.107224.
- [9] 陈赓, 刘湘慧, 陈玉斌. 基于决策实验室分析法及解释结构模型的交通隧道施工事故致因链构建[J]. 科学技术与工程, 2021, 21(35): 15 222–15 229.
- CHEN Yun, LIU Xianghui, CHEN Yubin. Construction of the cause chain of traffic tunnel construction accidents based on DEMATEL-ISM [J]. Science Technology and Engineering, 2021, 21(35): 15 222–15 229.
- [10] 杨文安, 李佳欣. 基于动态贝叶斯网络的装配式建筑吊装施工安全风险[J]. 安全与环境学报, 2024, 24(4): 1 328–1 336.
- YANG Wenan, LI Jiaxin. Safety risk analysis of hoisting construction of prefabricated buildings based on dynamic Bayesian network[J]. Journal of Safety and Environment, 2024, 24(4): 1 328–1 336.
- [11] 冯胤伟, 刘正江, 蒋子怡, 等. 基于关联规则挖掘和复杂网络理论的船舶碰撞事故影响因素分析[J]. 大连海事大学学报, 2023, 49(3): 31–44.
- FENG Yinwei, LIU Zhengjiang, JIANG Ziyi, et al. Analysis of factors affecting ship collisions based on association rule mining and complex network theory[J]. Journal of Dalian Maritime University, 2023, 49(3): 31–44.
- [12] 郑彬彬, 冯婷婷, 王佳贺, 等. 基于文本挖掘的城镇燃气事故致因及关联分析[J]. 中国安全科学学报, 2023, 33(7): 190–195.
- ZHENG Binbin, FENG Tingting, WANG Jiahe, et al. Causes and correlation analysis of urban gas accidents based on text mining [J]. China Safety Science Journal, 2023, 33(7): 190–195.
- [13] 陈述, 习俊博, 王建平, 等. 水电工程施工安全隐患关联规则挖掘[J]. 中国安全科学学报, 2021, 31(8): 75–82.
- CHEN Shu, XI Junbo, WANG Jianping, et al. Mining association rules of near-misses of hydropower projects construction [J]. China Safety Science Journal, 2021, 31(8): 75–82.



作者简介: 曾大林 (1981—),男,山东临朐人,博士,教授,主要从事复杂项目管理、项目治理、建筑经济、智能建造、人机交互等方面的研究。E-mail: zengdalin@sdjzu.edu.cn。