

# Integrating Multi-modal Seismic Damage Defense Basic Data Based on Spatiotemporal Big Data Technology

Xin Wu

Yunnan Province Map Institute, Kunming, Yunnan, 650000, China

## Abstract

This paper aims to discuss the method and application of multi-mode integration of basic data based on spatial-temporal big data platform, integrate multi-source data analysis, and propose a set of system methods of multi-modal prevention basic data integration. The research content includes data preprocessing, data fusion and key technical links, such as data mining and visualization, etc. The research results indicate that this method is of great significance in improving the scientific and effective nature of earthquake damage prevention. It can effectively enhance the accuracy and response efficiency of earthquake damage prevention predictions, and provide new technical concepts and practical references for the research paper in the field of earthquake damage prevention.

## Keywords

spatial otemporal big data technology; earthquake damage defense; multi-modal data integration; data preprocessing

## 基于时空大数据技术实现多模态震害防御基础数据整合

吴鑫

云南省地图院, 中国·云南昆明 650000

## 摘要

论文旨在通过构建时空大数据平台、整合多源数据、深度时空数据分析等方式, 探讨基于时空大数据技术实现多模态震害防御基础数据集成的方法与应用, 提出一套多模态震害防御基础数据集成的系统方法。研究内容包括通过实际应用案例进行效果评估的数据预处理、数据融合和关键技术环节, 如数据挖掘和可视化等。研究结果表明, 该法对提高震害防御的科学性和有效性具有十分重要的意义, 可以有效提高震害防御预测准确率和响应效率, 并为论文研究提供了新的技术理念和在震害防御领域的实践借鉴。

## 关键词

时空大数据技术; 震害防御; 多模态数据整合; 数据预处理

## 1 引言

地震灾害因其突发性和破坏性, 对人类社会构成严重威胁, 有效的震害防御策略对于减轻其影响至关重要。时空大数据技术的发展为震害防御提供了新的视角和工具, 通过整合和分析多源数据, 能够增强震害预测的准确性和应急响应的效率。论文探讨如何利用时空大数据技术实现多模态震害防御基础数据的整合, 以期震害防御提供更为科学和系统的方法论支持。

## 2 时空大数据技术在震害防御中的应用

### 2.1 时空大数据平台构建

时空大数据平台构建旨在优化震害防御中多源异构数

据整合与分析, 设计遵循模块化原则, 划分数据采集、存储、处理及分析四个核心层级。数据采集层配备预配置 API 与数据适配器, 负责实时捕获来自不同监测设备和数据库的多种数据格式, 确保数据完整与时效。数据存储层部署分布式存储解决方案, 采用如 Cassandra 的 NoSQL 数据库, 提供高吞吐量数据读写能力和灵活扩展性, 支持跨数据中心数据复制, 增强数据持久性与可用性。数据处理层由自动化工作流构成, 执行数据清洗、标准化与质量控制任务, 利用预定义规则和算法识别并纠正数据异常值与不一致性, 提供高质量数据集供后续分析<sup>[1]</sup>。数据分析层集成多种先进工具, 包括机器学习与统计分析软件包, 对处理后数据深入分析, 识别地震活动模式与趋势, 预测地震风险, 并提供基于数据的洞察给决策者。平台提供交互式用户界面, 允许用户通过图形化工具和仪表盘访问数据与分析结果, 支持自定义查询与报告生成, 使非技术用户轻松利用平台功能。时空大数据平台通过精心设计架构实现震害防御数据高效管理和分析, 为地震预测与应急响应提供强大技术支持系统。

【作者简介】吴鑫(1984-), 男, 中国云南红河人, 本科, 工程师, 从事测绘地理信息在行业应用及新型测绘技术研究。

## 2.2 多源数据整合

涉及地震监测数据、地质调查数据、遥感影像等多种数据源的整合，多源数据整合在震害防御中发挥着关键作用。这一过程在增强数据全面性、准确性的同时，也提供了丰富的震害防御资料。涵盖资料清洗、格式标准化、差错校正等内容，在资料整合前需进行资料预处理，以确保资料质量。利用特征级或决策级的整合技术，将不同来源的资料组合起来，将彼此之间的关系以及潜在的模型展现出来。遥感影像提供的是地表宏观视图，而详细的时间序列信息则由地震监测数据提供，二者相互融合，可以帮助震前模式异常的准确识别和预警精确度的提高。数据挖掘和机器学习算法可以从数据中提取特征和模式，帮助理解地震发生的原因和传播机制。在实际应用中，整合地震监测和地质调查资料，包括抗震设计、制定应急预案、开展公众教育等，可以对地震风险进行更准确的评估和制定防御措施。

## 2.3 时空数据分析

时空数据分析在震害防御中发挥关键作用，为预测地震可能性和影响范围提供依据，通过地震监测、地质调查和遥感影像等多源资料的分析，揭示地震活动的时空分布特点。利用地震监测资料对地震波传播特性进行分析，对震源位置、烈度等进行测定。地质勘测资料，提供地震发生的原因和潜在风险区域资料，遥感影像图显示地震发生前后地表异常的地表宏观变化情况。技术上，从海量数据中提取特征和模式，构建地震预测模型，评估发生概率，采用决策树、随机森林、支持向量机等数据挖掘和机器学习算法。而数据可视化技术则有助于直观地呈现复杂的信息，便于决策。实践中，通过历史地震和地质构造资料分析时空资料，增强地震风险评估在特定区域的准确性，支持地震设计、应急预案的制定和民众教育等方面的措施。时空数据分析将在震害防御中展现更广阔的应用前景和显著成效，因为技术进步和数据获取能力的提高。时空大数据平台架构图如图1所示。

## 3 多模态震害防御基础数据整合方法

### 3.1 数据预处理

数据预处理是数据分析和机器学习过程中至关重要的一步，它涉及对原始数据进行清洗、转换和规范化，以便于后续的建模和分析。在多模态震害防御基础数据整合中，数据预处理包括处理缺失值、异常值、数据标准化和归一化等关键步骤<sup>[2]</sup>。缺失值的处理通常采用插值填充或使用均值、中位数等统计量进行替代，如使用公式填充值  $=\mu$  来表示使用均值填充缺失值的情况。对于异常值，采用箱线图检测或基于绝对离差中位数 (MAD) 的方法，通过计算各观测值与平均值的距离总和来识别和处理异常值，公式用于定义  $3\sigma$  原则，识别偏离均值  $3\sigma$  之外的离群点。

数据标准化和归一化是确保数据在统一尺度上进行比较的重要手段，标准化通常采用 z-score 方法，将数据转换

为均值为 0、标准差为 1 的分布，公式描述了这一过程。归一化则将数据缩放到  $[0, 1]$  区间内，使用公式进行转换。这些预处理步骤对于提高数据分析的准确性和模型的预测性能至关重要，通过这些方法，可以确保数据的质量和一致性，为后续的数据分析和建模打下坚实的基础。



图1 时空大数据平台架构图

### 3.2 数据融合

在多模态震害防御基础数据整合中，数据融合涉及将不同传感器和数据源的信息合并成统一视图。特征级融合和决策级融合是两种主要方法。特征级融合从各数据源提取特征，整合成综合特征集，可采用加权平均法或卡尔曼滤波法等技术。加权平均法适用于数据分布均匀且无异常值的情况，而卡尔曼滤波法则在符合特定统计模型时提供最优估计。决策级融合则在数据源已做出决策后进行，整合各决策结果，得出最终决策，具有容错性强、开放性好、处理时间短、数据要求低、分析能力强等特点，但预处理和特征提取要求较高。数据融合旨在消除冗余和矛盾，提高信息提取的及时性和可靠性，为震害防御提供更丰富细致的信息。

### 3.3 数据挖掘与可视化

数据挖掘与可视化作为多模态震害防御基础数据整合中的关键环节，使得复杂数据集变得易于理解和操作。常用算法如决策树、支持向量机和随机森林等能够从大量数据中提取有用模式和关系<sup>[3]</sup>。例如，决策树算法通过递归地将数据集分割成越来越小的子集来构建模型，每个分割基于某个特征的阈值，如公式  $D_i = \{x \in D | x_j \leq \theta\}$  所示，其中  $D$  为数据集； $D_i$  为分割后的数据子集； $x_j$  为特征，而是分割阈值。

数据可视化则将挖掘出的信息以图形或图表形式展现，

使非专业人士也能直观理解数据背后含义。在震害防御中,可视化技术帮助专家与决策者识别地震活动热点区域、潜在地震风险及应急响应优先级。例如,通过地理信息系统(GIS)映射技术,地震数据与地理信息结合生成地震风险地图,提供直观决策支持。实际应用中,数据挖掘与可视化结合已证明对提升震害防御效率和准确性具有重要作用,对历史地震

数据挖掘,研究人员能够识别地震活动周期性与空间分布特征,而可视化工具使这些复杂信息传递与交流更高效。随着数据挖掘与可视化技术进步,其在震害防御中应用将更广泛,为地震预测与防灾减灾提供更有力的支持。多模态震害防御基础数据整合流程如图2所示。



图2 多模态震害防御基础数据整合流程图

## 4 应用案例与效果评估

### 4.1 应用案例

在加州实施的综合地震预警体系中,系统实时地对地震的活动进行监视、分析,通过综合地震的实时监测资料、地质资料以及卫星的遥感影像等进行综合分析。通过支持向量机(SVM)、神经网络等机器学习算法,该系统可以对地震发生的概率进行预判,并对发生的震害的范围加以评估和测量的能力。通过对历史的地震资料以及地质构造特征的分析 and 地震活动的模式的学习来提高预警的准确度,在这些算法中得到了很好的体现。例如,用来划分不同类别的数据,用SVM的算法来分析超平面上的位置,其中 $w$ 和 $b$ 为模型参数; $x$ 为输入特征向量。通过优化超平面上的位置来划分不同类的数据,从而达到帮助决策人和大众更好地理解地震隐患的目的,从而起到行之有效的防范措施,这一应用的案例表明,在震害防御中,多模态数据的集成是非常有效的,也是很实用的,为其他区域的震害预警系统提供一个借鉴。

### 4.2 效果评估

效果评估是衡量数据整合方法有效性的关键环节,在多模态震害防御基础数据整合的应用案例中。通过统计指标如精度、召回率、F1积分等的运用,对实际震害防御中数据融合、挖掘技术的表现进行量化分析。例如,评估分类模型性能的两个重要指标——精确度(Precision)和召回率(Recall),分别体现了模型预测的精确性和完整性<sup>[4]</sup>。F1积分则是将精确度和召回率进行调和平均,用来综合考虑这两个指标,计算公式为F1。如地震预警系统性能评估指标表1所示。

实际情况下,如美国加州综合地震预警系统在地震发生前通过集成多种资料来源和应用先进的资料挖掘技术,系统可提供预警争取到宝贵的应急响应时间,疏散公共场所的时间。效果评估表明,该系统通过GIS映射、直观地展示地震风险区域,为决策者提供强有力的支持,在多次地震事件

中,系统对地震活动进行了成功的预测,并通过数据可视化技术得到了很好的发挥。这些考核结果表明,在震害防御方面,多模态数据集成技术的应用价值十分显著,应用前景十分广阔。

表1 地震预警系统性能评估指标表

指标	描述	数值
精确度	预测为地震的事件中实际发生地震的比例	0.85
召回率	实际发生地震的事件中被正确预测为地震的比例	0.90
F1分数	精确度和召回率的调和平均	0.875
预警时间	系统发出预警到地震实际发生之间的时间	30分钟
误报率	预测为地震但实际未发生的事件比例	0.10
漏报率	实际发生地震但未被预测的事件比例	0.05

## 5 结语

论文深入探讨基于时空大数据技术实现多模态震害防御基础数据整合的方法与应用,涵盖数据预处理、融合、挖掘及可视化,展示数据整合技术在震害防御中重要作用。研究表明,整合多源数据并应用先进分析技术,可显著提高地震预测准确性与应急响应有效性。伴随技术进步,未来震害防御将更加依赖大数据与人工智能技术,实现高效精准灾害管理和减灾目标。

### 参考文献

- [1] GB/T 42528—2023 时空大数据技术规范[S].
- [2] 马昭辉,张璐.地理编码技术在智慧六盘水时空大数据云平台中的应用[J].长春工程学院学报(自然科学版),2022,23(3):64-67.
- [3] 李甜甜,王志涛,马东辉.区域抗震防灾综合防御体系框架与对策研究[C]//中国城市规划学会,成都市人民政府,面向高质量发展的空间治理——2020中国城市规划年会论文集(01城市安全与防灾规划),北京工业大学建筑与城市规划学院,2021.
- [4] 彭文娟.基于多模态数据的地点语义建模算法研究[D].天津:天津大学,2018.