

# 面向CIM平台的多源异构数据融合关键技术

张泽林 王凯强 安海娜

中国市政工程西北设计研究院有限公司 甘肃兰州 730030

**摘要:** 本文梳理了国内外多源异构数据融合研究现状,探讨CIM在集成多源异构数据、推动城市数字化治理、优化公共服务等方面的重要意义。文章重点关注CIM作为数字孪生城市核心技术的应用路径,结合BIM、GIS、物联网等技术手段,研究多源数据整合及三维可视化方案,提出建设基于CIM的智慧城市管理平台的可行性分析,旨在为未来CIM技术研究与实践提供理论依据和应用参考。

**关键词:** 城市信息模型;智慧城市;数据整合;数字孪生;平台建设

随着城市化进程加快,城市治理面临管理复杂度提升、资源分配不均和环境压力增大的挑战,传统的城市管理方式已难以满足现代化城市发展的需求,智能化、数字化转型成为必然趋势。在新型城建与新基建融合发展的背景下,信息技术的深度应用不仅关系到城市运行的效率,还直接影响公共服务质量与社会治理水平。城市信息模型(City Information Modeling, CIM)作为一种新兴的技术平台,能够整合多源异构数据,全面感知并动态分析城市运行发展状态,从而提升城市综合管理能力。

## 一、多源异构数据融合的研究背景

随着智慧城市建设的不断推进,城市管理和服呈现数据驱动、智能化的特征。然而,城市运行过程中产生的数据具有来源多样、类型复杂、格式异构等特点,包括BIM模型、GIS地理空间数据、物联网感知数据以及交通、环保、能源等行业业务数据。这些数据不仅在结构和语义上存在差异,而且在时间、空间尺度上也表现出高度分散和非标准化特征,导致数据共享和交互难度加大。如何实现多源异构数据的高效融合,构建统一的数据框架,成为提升城市智能化管理水平的核心任务。多源异构数据融合不仅是CIM平台建设的技术基础,也是实现数字孪生城市 and 智能决策支持的关键环节,具有重要的理论研究价值和应用实践意义<sup>[1]</sup>。

**课题项目:** 甘肃省住房和城乡建设厅建设科技项目《基于CIM多源异构数据的三维整合平台开发及应用》(编号:JK2024-24)

**作者简介:** 张泽林(1990.01——),汉族,男,陕西西安人,高级工程师,主要从事BIM/信息化研究。

## 二、面向CIM平台的多源异构数据融合需求分析

CIM平台作为城市信息基础设施,其核心目标是整合并管理城市全生命周期的数据资源,满足规划、建设、运营及服务的智能化发展要求。由于平台需要同时处理来自BIM、GIS、IoT及大数据分析等多维度的信息,必须具备强大的数据接入、语义映射与模型转换能力,实现不同数据源间的互操作性和一致性。此外,CIM平台需满足实时性和高并发处理要求,保证海量动态数据能够被快速采集、传输和可视化呈现,支持空间分析、三维建模及动态仿真。针对异构数据的复杂性,平台还需建立标准化的数据分类体系与统一编码规则,并结合云计算与分布式存储,实现高效的数据管理和安全共享目标。只有打造完善的数据融合机制,才能构建兼具精细化表达和高可扩展性的CIM平台,为智慧城市的精准治理和科学决策提供坚实的数据基础。

## 三、面向CIM平台的多源异构数据融合关键技术

### (一)数据标准化与预处理

在CIM平台中,多源异构数据包括BIM模型、GIS空间数据、物联网实时感知信息以及视频、传感器、业务系统产生的结构化与非结构化数据。这些数据在格式、语义、坐标系及精度方面存在显著差异,若缺乏统一标准,将直接导致数据在存储、交互和分析过程中出现不一致性。数据标准化与预处理成为CIM平台建设的首要环节,其核心目标是借助统一的数据分类体系、编码规则及语义模型,保障数据在属性、结构和空间基准上的一致性。标准化过程中,需要引入国际标准,并结合国家BIM、GIS规范,构建兼容CIM应用的混合标准体系<sup>[2]</sup>。同时,为适应不同采集来源,需进行数据清洗与异常检测,

剔除冗余和冲突信息，保证数据完整性与准确性。在空间数据处理上，需统一坐标系统，完成二维与三维数据的融合投影，确保地理精度；在属性信息上，通过语义解析和本体映射消除命名歧义，实现跨领域数据的语义互通。进行标准化与预处理工作，可以为后续的数据融合、模型构建与动态仿真提供高质量的数据基础，降低系统集成难度，提升平台的可扩展性和智能化水平。

### （二）融合架构与平台设计

面向CIM平台的数据融合架构须具备高并发处理能力、跨领域兼容性和可扩展性，适应城市级别的超大规模数据集成需求。平台设计应遵循分层架构理念，通常包括数据采集层、融合处理层、存储管理层和应用服务层。数据采集层负责对接各类数据源，借助API接口、流数据接入和ETL工具实时采集异构数据；融合处理层利用中间件和语义引擎完成格式转换、语义匹配和空间关联，采用多模态融合算法实现二维GIS与三维BIM的深度耦合，形成可视化的城市信息模型。在存储管理层，需引入分布式数据库和云计算架构，支持大规模数据的高效存储、索引和检索，结合时空数据库实现动态更新与快速查询。应用服务层构建开放API和微服务架构，支持上层应用如规划模拟、应急调度、设施管理等多场景需求。此外，平台需集成可扩展的数据安全与访问控制机制，保证敏感信息在共享过程中的合规性<sup>[3]</sup>。设计科学的融合架构与平台，CIM系统能够实现多源异构数据的高效整合、统一管理和智能服务，全面支撑智慧城市建设和数字孪生应用的落地实施。

### （三）融合方法与算法

CIM平台的多源异构数据融合不仅涉及结构化、半结构化和非结构化数据的整合，还需实现空间、语义及时间维度的统一表达。因此，融合方法需以多模态融合为核心，结合规则驱动、模型驱动和数据驱动的多种策略。其中，基于语义的融合方法是实现跨领域数据一致性的关键，构建本体模型，将BIM、GIS及物联网数据的属性和语义进行映射，解决数据描述不一致的问题。与此同时，空间融合方法需采用坐标转换和空间索引技术，精确叠加不同坐标系数据。需借助数据流融合技术实时处理动态传感器数据和视频流，结合滑动窗口算法与时间序列分析保证融合结果的时效性。在算法实现方面，采用基于图论的多源数据关联算法，可以借助节点—边结构表征异构数据关系，优化跨域数据的匹配效率；而针对大规模数据，需结合分布式计算框架与

MapReduce模型，实现融合过程的并行化与负载均衡。此外，为应对数据冗余和冲突问题，可引入基于置信度加权的贝叶斯推理方法，对多源数据进行可靠性评估与融合，从而在保证一致性的基础上提升融合精度。这种融合方法不仅保障了CIM平台数据的完整性与一致性，还为后续的三维可视化、数字孪生及智能决策提供了坚实的数据基础。

### （四）实时融合与动态更新

CIM平台在智慧城市运行过程中，需要处理大量来自物联网传感器、视频监控和实时监测系统的数据，这要求平台具备高效的实时融合与动态更新能力。为实现这一目标，平台需构建基于流式计算的实时处理框架，采用分布式流处理引擎（如Apache Flink或Spark Streaming）对多源异构数据进行增量式融合，促使数据在毫秒级响应的同时保持一致性。实时融合过程中，需引入事件驱动机制，以消息队列作为数据传输通道，结合时空索引实现数据的精准定位与快速匹配，进而保障复杂城市场景下的实时可视化与动态分析。在动态更新方面，CIM平台需支持增量更新和局部刷新策略，避免全量重建带来的性能瓶颈，结合缓存机制与多级索引优化查询效率。针对高频数据流，需借助滑动窗口聚合与数据压缩算法，降低网络传输和存储压力，保障数据时效性和可靠性<sup>[4]</sup>。为了实现与城市运行场景的深度联动，还需在平台中引入基于微服务的动态部署架构，借助API接口实现模型数据与外部系统的实时交互，支持应急调度、交通优化及能源管理等应用场景，提升CIM平台对复杂城市运行状态的响应能力，为智慧城市实现“感知—分析—决策—反馈”的闭环提供关键技术保障。

## 四、面向CIM平台的多源异构数据融合系统构建

### （一）系统整体架构设计

面向CIM平台的多源异构数据融合系统整体架构需遵循分层、模块化和可扩展的原则，保障海量数据处理的高效性和灵活性。系统设计采用“采集—传输—融合—存储—应用”五层结构。采集层负责多源数据的获取，包括来自BIM模型、GIS地理信息、IoT传感器、视频监控及业务系统的数据，利用多协议适配器兼容各类硬件和软件接口。传输层基于高带宽网络与消息队列机制，采用MQTT、Kafka等分布式消息中间件，实现异步传输与流数据缓冲，提高数据实时性和稳定性。融合层是系统的核心，构建基于语义映射的融合引擎，结合坐标转换、几何校正和语义对齐技术，统一表达空间、属

性和时间维度，并借助多模态融合算法实现BIM与GIS深度耦合目标。存储层采用分布式数据库与时空数据库的混合架构，结合NoSQL与关系型数据库，满足大规模结构化和非结构化数据的存储需求，利用云计算平台完成弹性伸缩任务<sup>[5]</sup>。应用层面向智慧城市的业务场景，提供规划模拟、运维管理、应急调度等服务，并凭借三维可视化技术进行全域动态展示。整体架构设计强调系统的开放性和标准化，采用微服务架构和容器化部署，支持跨平台运行与多终端访问。

## （二）技术路线与实现方案

系统的技术路线以“标准先行、融合驱动、分布式协同”为原则，围绕数据采集、语义融合、模型构建和应用开发展开。在数据采集阶段，采用多通道接入和协议转换机制，保证各类异构数据实时上传；借助ETL流程完成数据抽取、清洗和转换，为融合提供高质量输入。在融合实现方面，基于本体模型构建统一的语义框架，实现跨领域属性映射和语义对齐，在空间数据处理环节，采用分层投影和多坐标系统转换技术，使BIM和GIS数据能够在同一三维空间中无缝集成。针对实时感知数据，部署流式计算引擎，采用窗口聚合和事件驱动模型，实现高频数据的增量式融合目标。在模型构建阶段，采用多维索引和三维建模算法，将静态模型与动态监测信息融合，形成支持可视化和预测分析的数字孪生体。应用开发方面，基于微服务架构和开放API，支持智慧交通、应急响应和基础设施运维等场景的业务接入。整个技术路线以分布式计算框架（如Hadoop、Spark）为支撑，结合云平台实现弹性扩展，并利用GPU加速和并行处理技术提升大规模三维模型渲染性能。该方案在保证系统高性能的同时，兼顾可维护性和扩展性，能够适应智慧城市多样化的业务需求。

## （三）系统性能优化与安全保障

针对CIM平台对高性能和高安全性的双重要求，系统性能优化主要从数据处理、计算调度和渲染加速三个维度入手。在数据处理方面，采用分布式文件系统和内存计算框架，结合列式存储和数据压缩技术，显著降低I/O开销，提高海量数据查询和写入速度。在计算调度上，引入动态负载均衡和容器编排技术，根据任务负载

实现计算资源的自动分配与迁移，避免系统瓶颈并提升服务的并发处理能力。三维可视化模块采用GPU加速渲染和LOD（Level of Detail）技术，保证在大规模模型加载时实现流畅的交互体验。借助CDN加速和边缘计算优化用户端访问性能，满足多终端实时交互需求。在安全保障方面，系统采用多层防护架构，包括数据加密传输、存储加密及基于角色的访问控制（RBAC），防止数据泄露与非法访问；引入区块链技术对关键操作进行溯源，实现数据可信流转。在运行安全方面，部署入侵检测与异常行为监控系统，借助AI算法识别潜在威胁并实现实时预警，保证平台在复杂网络环境下的稳定运行。

## 结语

系统研究面向CIM平台的多源异构数据融合技术，本文提出了以数据标准化、语义对齐、融合算法及实时更新为核心的技术体系，构建了面向智慧城市的融合架构与实现方案。研究表明，语义映射结合空间投影的多模态融合方法能够有效解决跨领域数据一致性问题，流式处理与分布式计算架构显著提升了实时性与并发性能。未来研究应聚焦融合算法的智能化演进、动态更新的全生命周期管理以及面向数字孪生的应用拓展，促使CIM平台在城市治理中发挥更高的实践价值。

## 参考文献

- [1] 黄凌锋. 基于多源异构数据融合的道路信息采集系统设计研究[J]. 信息记录材料, 2025, 26(07): 176-178+187.
- [2] 张娴. 基于多源异构数据融合的政务云数据治理关键技术研究[J]. 网络安全技术与应用, 2024, (12): 60-62.
- [3] 邱新忠, 汤赛, 徐晓红, 等. 多源异构三维空间数据融合关键技术研究及应用[J]. 浙江国土资源, 2024, (09): 38-40.
- [4] 夏莹杰. 多源异构海量交通数据与交通模型融合关键技术及应用. 浙江省, 杭州远眺科技有限公司, 2022-05-11.
- [5] 李仕峰. 多源异构数据时空融合关键技术研究与应用[J]. 地理空间信息, 2021, 19(10): 19-21+25+149.