

# 基于计算机技术的物联网数据处理与分析方法研究

张栩恒

中国矿业大学徐海学院 江苏徐州 221100

**摘要:** 物联网 (IoT) 技术的发展使得海量异构数据在感知层、网络层与应用层之间高速流动, 推动了智能城市、工业互联、智慧医疗等领域的快速发展。然而, 由于物联网数据具备高频率、强时序性、异构性等特征, 其在采集、传输、存储与分析过程中面临诸多挑战。本文基于计算机技术, 从边缘计算、云计算、大数据平台等角度出发, 系统研究了物联网数据的处理架构、分析方法与关键技术。提出了一种融合边缘预处理与云端智能分析的混合处理模型, 结合机器学习、流处理与可视化技术, 实现对物联网数据的高效处理与智能洞察。

**关键词:** 物联网; 数据处理; 计算机技术; 边缘计算; 智能分析

物联网通过感知设备、通信网络与智能平台的深度融合, 使物理世界与数字世界实现互联互通, 形成了复杂的数据生态系统。从智能制造到智慧交通, 物联网设备产生的实时数据流正以前所未有的规模涌入数据中心, 这对数据处理能力提出了更高要求。传统的数据处理架构已难以满足物联网对低延迟、高并发、异构兼容等需求。为应对这一挑战, 计算机技术不断演进, 边缘计算、云计算、大数据平台、人工智能等技术逐渐成为物联网数据处理的关键支撑。

## 一、物联网数据特征与处理挑战

### (一) 物联网数据的类型、结构与生成模式

物联网数据需从类型、结构、生成模式三方面梳理特征, 为处理方案设计提供依据<sup>[1]</sup>。数据类型按来源与用途可分为感知数据、设备状态数据、环境数据、交互数据, 不同类型数据对应差异化处理需求。结构层面呈现“结构化、半结构化、非结构化”混合特征: 结构化数据格式固定, 易直接存储与分析; 半结构化数据有基本结构但字段灵活; 非结构化数据无固定格式, 需特殊处理。生成模式具有“高频性、连续性、分布式”特点: 感知设备按固定或动态周期持续生成数据, 部分场景下采集频率达毫秒级; 数据生成点分散于不同区域的设备终端, 形成分布式数据源头, 且生成量随设备数量

增加呈指数级增长, 需针对性设计高效的数据接收与整合机制。

### (二) 数据处理面临的实时性与异构性难题

物联网数据处理在实时性与异构性方面面临双重难题, 制约处理效率与应用价值<sup>[2]</sup>。实时性难题源于两方面: 一是部分场景对数据处理延迟要求严苛, 如工业设备故障预警需在数据生成后毫秒级内完成分析并触发警报, 否则可能引发安全事故; 二是数据生成量波动大, 高峰时段瞬时数据量激增, 易导致处理系统拥堵, 延长响应时间, 例如智能制造场景中生产线满负荷运行时, 数据量较常态增长数倍, 传统处理架构难以应对。异构性难题体现在数据格式、协议、质量的差异: 格式上包含数值、文本、图像等多种类型, 需不同处理流程; 协议上设备采用LoRa、NB-IoT、WiFi等不同通信协议, 数据接入时需多协议适配; 质量上存在噪声干扰、数据缺失、异常值等问题, 如传感器受环境影响产生偏差数据, 或传输过程中因网络波动导致数据丢失, 需通过预处理提升数据质量, 但多样化的质量问题增加了处理复杂度。

### (三) 对计算架构与技术支持的需求分析

物联网数据处理对计算架构与技术支持提出“分布式、弹性化、一体化”需求, 以适配数据特征与处理挑战<sup>[3]</sup>。计算架构需求方面, 需构建“边缘-云端”协同的分布式架构: 边缘侧需具备数据预处理与实时分析能力, 就近处理高频实时数据, 减少数据传输带宽占用与延迟; 云端需承担海量数据存储、复杂分析与全局决策任务, 实现数据深度挖掘。弹性化需求体现在架构需随数据量与处理需求动态调整: 数据高峰时自动扩展计算

**作者简介:** 张栩恒 (2002.10.10-), 性别: 男, 民族: 汉, 籍贯: 江苏省苏州市, 学历: 本科, 单位: 中国矿业大学徐海学院, 职称: 无, 研究方向: 计算机系物联网专业。

与存储资源，避免系统过载；低谷时收缩资源，降低成本，需依托虚拟化、容器化技术实现资源弹性调度。技术支持需求涵盖多维度：数据接入层需多协议兼容技术，实现不同设备数据的统一接收；数据预处理层需数据清洗、格式转换、质量修复技术，解决异构性问题；数据存储层需支持多结构数据的混合存储技术，兼顾存储效率与查询性能；数据分析层需实时计算、离线分析、智能挖掘技术，满足不同场景的处理需求，同时需安全技术保障数据传输与存储过程中的隐私与完整性。

## 二、基于计算机技术的数据处理架构设计

### （一）边缘计算与边云协同模型设计

边缘计算与边云协同模型需围绕“就近处理、高效协同”构建，平衡实时性与全局优化需求。边缘计算节点部署在物联网设备附近，承担本地化数据处理任务：首先对采集的数据进行预处理，过滤噪声、补全缺失值，减少无效数据传输；其次完成实时分析任务，如设备异常检测、阈值判断，快速生成本地决策，例如工业边缘节点可实时分析设备振动数据，发现异常后立即控制设备停机；同时实现数据缓存，将高频实时数据暂存本地，按规则向云端传输关键数据，降低带宽消耗。

### （二）云计算平台在海量数据处理中的应用

云计算平台凭借“海量存储、高性能计算、灵活扩展”优势，成为物联网海量数据处理的核心支撑。在数据存储方面，云计算平台提供分布式存储服务，支持结构化、半结构化、非结构化数据的混合存储，通过数据分片与副本机制，确保海量数据存储的可靠性与可用性，同时支持弹性扩容，可根据数据增长需求动态增加存储节点，满足物联网数据长期积累的存储需求。在数据计算方面，云计算平台提供批处理与流处理结合的计算能力：批处理适用于海量历史数据的离线分析，如通过分析过去一年的设备运行数据挖掘故障规律；流处理适用于实时数据的持续分析，如实时计算设备运行参数的均值、峰值，及时发现异常，且计算资源可随数据量动态调整，避免资源浪费或不足。

### （三）分布式存储与流处理框架优化

分布式存储与流处理框架的优化需针对物联网数据特征，提升存储效率与处理性能。分布式存储优化聚焦三方面：一是采用“分层存储”策略，根据数据访问频率与重要性，将热数据（高频访问的近期数据）存储于高性能介质，冷数据存储于低成本介质，平衡性能与成本；二是优化数据分片策略，结合物联网数据的时间关

联性与设备关联性，按“时间+设备ID”分片存储，提升按设备或时间范围查询的效率；三是引入数据压缩与索引技术，针对不同类型数据采用适配的压缩算法，同时建立多维索引，加速数据检索。

## 三、物联网数据智能分析方法研究

### （一）基于机器学习的异常检测与预测模型

基于机器学习的异常检测与预测模型需结合物联网数据特征，实现“精准识别、提前预警”，提升数据应用价值<sup>[4]</sup>。异常检测模型针对物联网数据中的异常值与异常模式，通过三类方法实现：一是无监督学习方法，无需标注数据，直接从海量正常数据中识别偏离模式的异常数据，适用于缺乏异常样本的场景；二是监督学习方法，利用标注的正常与异常样本训练模型，实现对新数据的异常分类，适用于异常模式明确的场景；三是半监督学习方法，结合少量标注样本与大量未标注样本训练，平衡数据需求与检测精度。预测模型聚焦设备状态、数据趋势的提前预判，采用时间序列预测算法，通过分析历史时序数据，学习数据变化规律，预测未来一段时间内的参数值，如预测设备未来一周的温度变化趋势，或预测设备故障发生时间，为预防性维护提供依据，模型需具备自适应能力，可根据新数据动态调整参数，提升预测准确性。

### （二）时序数据建模与行为模式识别

物联网中大量时序数据的建模与行为模式识别，是挖掘数据关联规律、实现智能决策的关键。时序数据建模需突出时间维度特征，构建适配不同数据规律的模型：针对平稳时序数据，采用传统统计模型捕捉趋势与周期性；针对非平稳、非线性时序数据，采用深度学习模型，通过多层网络结构学习复杂的时间依赖关系，提升建模精度。建模过程中需注重数据预处理，通过数据归一化、平稳性处理、特征提取，将原始时序数据转化为适用于模型输入的格式，同时采用滑动窗口等技术，将时序数据分割为样本序列，支撑模型训练与推理。

### （三）数据可视化与决策辅助系统设计

数据可视化与决策辅助系统需实现“数据直观呈现、决策科学支撑”，推动物联网数据从“可看”向“可用”转化。数据可视化设计需兼顾多维度需求：一是实时监控可视化，采用仪表盘、动态图表展示实时数据，支持数据实时刷新与异常高亮提醒，帮助用户快速掌握系统状态；二是历史趋势可视化，通过折线图、热力图展示数据随时间的变化趋势，支持多数据对比与交互式探索，

便于用户分析规律；三是关联关系可视化，采用网络图、Sankey图展示数据间的关联，清晰呈现复杂关系。决策辅助系统在可视化基础上，整合数据分析结果，提供决策支持：一是自动生成分析报告，汇总关键数据、异常情况、趋势预测，减少人工分析工作量；二是提供决策建议，基于异常检测与预测结果，推荐应对措施；三是支持模拟分析，用户可调整参数模拟不同决策的效果，辅助选择最优方案，提升决策的科学性与效率。

#### 四、系统实现与应用实例验证

##### (一) 系统功能架构与模块化实现方案

物联网数据处理系统需采用“分层架构+模块化设计”，确保功能完整、可扩展、易维护。功能架构分为数据接入层、预处理层、存储层、分析层、应用层：数据接入层负责接收不同设备、不同协议的数据，提供多协议适配接口；预处理层完成数据清洗、格式转换、质量修复，输出高质量数据；存储层实现多结构数据的分布式存储，支持冷热数据分层；分析层包含实时分析、离线分析、智能挖掘模块，满足不同场景需求；应用层提供可视化界面、API接口、决策辅助功能，支撑具体应用。模块化实现方案将各层功能拆分为独立模块：数据接入模块、数据清洗模块、存储管理模块、实时计算模块、机器学习分析模块、可视化展示模块等，各模块通过标准化接口通信，可独立开发、测试、部署。同时采用容器化技术封装模块，实现模块的快速部署与弹性扩展，支持根据应用需求灵活增减模块，例如新增图像数据处理功能时，仅需开发并部署图像预处理与分析模块，无需修改现有架构，提升系统灵活性与可维护性。

##### (二) 智能制造场景下的数据处理案例分析

智能制造场景下的物联网数据处理系统需聚焦“设备运维、生产优化”核心需求，通过实际应用验证系统有效性<sup>[9]</sup>。案例中，系统部署于某汽车零部件生产线，接入数百台生产设备的运行数据，涵盖温度、振动、转速、电流等参数。数据处理流程如下：边缘节点实时采集设备数据，完成噪声过滤与异常初步判断，将正常数据按分钟级汇总后传输至云端，异常数据立即推送至云端并触发本地预警；云端存储历史数据，通过批处理分析设备运行规律，利用机器学习模型预测设备故障，同时通过流处理实时监控生产参数，确保生产过程稳定。

##### (三) 性能评估与优化策略研究

物联网数据处理系统的性能评估需从多维度开展，结合评估结果制定优化策略，持续提升系统性能<sup>[6]</sup>。性能评估维度包括数据处理延迟、吞吐量、存储效率、准确性、可靠性，通过模拟不同数据量、不同场景的测试环境，获取各项指标数据，对比设计目标与行业标准，识别性能瓶颈。优化策略针对评估发现的问题制定：若处理延迟超标，优化边缘侧预处理流程，减少数据传输量，同时升级流处理框架的计算资源；若存储效率低，优化数据压缩算法与分层存储策略，增加冷数据压缩比；若准确性不足，扩充训练数据集，优化机器学习模型参数。

#### 结语

随着物联网技术的不断发展，数据处理与分析能力已成为其核心竞争力之一。本文基于计算机技术，提出了面向物联网环境的混合型数据处理与智能分析方法，融合了边缘计算的实时性与云计算的强处理能力，结合机器学习与可视化技术，实现了对复杂物联网数据的高效管理与深度洞察。实践验证表明，该方法在多个实际应用场景中具备良好的性能与可扩展性。

#### 参考文献

- [1] 严杭. 基于计算机技术的物联网数据处理与分析方法研究[J]. 中国宽带, 2025, 21(12): 49-51.
- [2] 武琳琳. 面向智慧校园的物联网数据融合与异常检测机制[J]. 物联网技术, 2025, 15(22): 125-127+132.
- [3] 夏川. 物联网环境下分布式计量检测数据融合算法及检定结果溯源研究[J]. 实验室检测, 2025, 3(21): 45-47.
- [4] 沙芳宇. 基于物联网的电力工程数据识别系统研究[J]. 电气技术与经济, 2025, (10): 347-349.
- [5] 苏唱. 物联网与计算机大数据处理技术分析[J]. 电子技术, 2025, 54(06): 172-173.
- [6] 崔孟阳, 王会琳, 薛保星, 高晓丽. 泛在电力物联网的边缘节点感知自适应数据处理方法分析[J]. 通信电源技术, 2020, 37(03): 131-132.