

# 计算机大数据信息处理技术策略的相关研究

吴 翔

武昌职业学院 湖北武汉 430200

**摘要:**目前,计算机大数据信息处理技术在各行业展现出了巨大的应用潜力和价值。随着信息技术的不断发展,各种信息数据正以前所未有的速度和规模增长。如何高效存储、管理和利用这些海量数据已成为关注的焦点。大数据信息处理技术的创新与应用将直接决定企业获取洞见、提高竞争力的水平,以及政府部门做出科学决策的能力。同时,数据安全、处理效率和数据质量等问题也亟须解决。本文分析和探讨了计算机大数据信息处理技术的重要性、面临的问题以及关键技术,以期为该领域的可持续发展贡献力量。

**关键词:**计算机;大数据;信息处理技术;策略

大数据信息处理技术是指利用现代信息技术手段,对大规模、高增长率和多样化的数据进行有效获取、存储、管理、分析和应用的技术体系。它涵盖了数据采集、存储、处理、挖掘、传输、可视化等多个环节,是实现大数据价值的关键支撑。当前,随着云计算、物联网、移动互联网等新兴技术的迅猛发展,大数据信息处理技术在各行业广泛应用,已成为推动科技创新、商业决策和社会治理现代化的重要力量。

## 1. 计算机大数据信息处理技术的重要性

### 1.1 市场和用户需求分析

在商业领域,计算机大数据信息处理技术可以帮助企业更深入地洞察市场和用户需求。通过对海量消费行为数据、网络舆情数据、产品使用反馈数据等进行分析,企业可以及时发现市场热点和新兴趋势,快速捕捉消费者的喜好变化,预判潜在需求<sup>[1]</sup>。同时,企业还可以基于用户画像、购买路径等数据挖掘结果,进行精细化营销策略制定,实现营销投资的精准化和个性化,从而提升营销转化效率,增强企业的市场竞争力。

### 1.2 科学研究的推进

科学研究中常常需要处理大规模复杂的实验数据、观测数据等,借助大数据技术可以极大提高数据处理和分析的效率。海量数据下发现新问题、提出新假设、验证研究假设的过程将更加流畅高效。以天文研究为例,借助大数据处理技术可以快速整合多个天文望远镜、探测器产生的PB级数据,挖掘发现新的天体现象,验证理论假说。医学研究中也可以通过分析医疗大数据,发现疾病前兆和易感人群,推动

新药研发。

### 1.3 国家事务管理

在政府治理层面,大数据技术可以为科学决策注入新的动力。通过汇总分析人口普查、经济运行、社会调查等各类民生数据,政府可以更清晰掌握社会发展的客观状况,及时发现问题和隐患;基于对互联网舆情、民意反馈等的监测分析,政府可以及时感知民意民智民愿。将政策措施的实施效果数据进行分析比对,可以持续优化完善各项决策部署。所以,大数据视角下科学决策将更具前瞻性、精准性和有效性。

## 2. 计算机大数据信息处理遇到的问题和不足

### 2.1 数据安全性问题

数据安全性问题贯穿于整个大数据信息处理的生命周期。从数据采集阶段,物联网设备、移动终端等数据源的安全性就容易受到威胁,一旦遭受攻击或植入恶意代码,将影响源头数据的真实性和完整性<sup>[2]</sup>。在数据传输过程中,网络通信又面临中间人攻击、流量劫持等风险,导致数据被窃取或篡改。而在数据存储和处理环节,分布式系统的复杂性也给数据安全带来新的挑战,存在内外部攻击、数据泄露等风险。如果核心数据资产遭到破坏,将造成不可估量的损失。对于企业来说,机密数据外泄将严重影响竞争力。对于个人来说,隐私数据泄露将给生活带来困扰。

### 2.2 数据处理效率问题

随着数据规模的持续增长,传统的数据处理方式已经无法满足实时性和高效性的要求。以往基于单机或小规模集群的处理模式,在面对PB甚至EB级的数据时,处理效率

将大幅降低，很难实现快速响应和决策支持。数据处理效率低下，不仅会影响大数据分析的时效性，错失获取洞见的关键时机，还可能使得实时业务系统运行缓慢或瘫痪。比如在电商营销、社交推荐等时间敏感型场景，如果数据处理效率低下，无法及时处理最新用户行为、商品库存等动态数据，就无法即时捕捉用户兴趣偏好的变化趋势，错过了为用户推送个性化营销和内容推荐的最佳时机，必然会导致营销转化率和推荐体验大幅下降，影响业务收益。

### 2.3 数据质量问题

数据质量问题涉及数据的完整性、准确性、一致性、时效性等多个维度，并且贯穿于大数据处理的全生命周期。例如在数据采集环节，传感器故障、网络丢包等就可能造成数据缺失或数据错误。在数据存储过程中，系统故障、硬件老化等情况也可能造成数据损坏或丢失。此外，数据来源的异构性和复杂性也使得数据整合过程中容易出现不一致的情况。如果数据质量出现问题，将严重影响后续分析和决策的质量。在医疗健康领域，依赖于不完整、失真的数据进行病情诊断和用药指导，必然会给患者带来严重的健康风险。在金融风控领域，数据缺陷将无法有效识别异常行为和风险隐患，可能造成资金损失。在政府决策层面，基于质量不佳的数据制定的政策措施也将事与愿违、适得其反。

## 3. 计算机数据信息处理技术

### 3.1 信息存储技术

当前主流的信息存储技术包括关系数据库、NoSQL数据库和分布式文件系统等。关系数据库以结构化的方式存储数据，适用于处理高度结构化的数据，具有数据一致性和完整性保证、事务处理能力强等优势，但在处理非结构化和半结构化数据时受到限制<sup>[9]</sup>。为了应对大数据场景下多样化的数据类型和格式，NoSQL数据库应运而生，它打破了关系数据库的限制，支持灵活的数据模型，适合存储结构化、半结构化和非结构化数据，处理性能卓越，但缺乏关系数据库的事务一致性保证。而对于存储海量数据，尤其是非结构化数据，分布式文件系统可以提供高可靠、高吞吐的存储服务，具有良好的扩展性和容错能力。这三种存储技术各有特色，组合使用可以满足不同场景下的数据存储和管理需求，是大数据信息处理的重要基石。

### 3.2 信息安全技术

信息安全技术是保障大数据信息处理安全的关键，可

以防止数据遭受泄露、篡改和其他攻击，维护数据的完整性和保密性。信息安全技术在大数据处理中主要包括数据加密、访问控制和身份认证等多个层面。其中，数据加密技术通过使用密钥算法如AES、RSA等，将明文数据转换为无法被直接识别的密文形式，从而在存储和传输环节保护数据的机密性。即使密文被窃取，也无法获取原始数据内容，有效防止了数据泄露<sup>[4]</sup>。以医疗大数据为例，患者的电子病历等隐私数据在上传至云端前需要先进行加密，确保数据在传输和存储期间的安全性。访问控制技术则通过建立精细的权限控制策略，限制对系统资源的访问，只有经过明确授权的用户或应用才能访问相应的数据资源，有效防止未经授权的访问。在大数据处理过程中，需要根据不同岗位和业务需求，细化划分数据访问权限，从而实现最小权限原则。此外，身份认证技术要求系统的所有主体在访问资源前必须提供合法的身份凭证，经过身份验证后方可获取相应的访问权限，确保系统免受非授权主体的攻击和入侵。通过多重防护，这些安全技术共同构筑了大数据处理的安全防线。除此之外，防火墙、入侵检测、漏洞扫描、安全审计等一系列安全技术措施也为大数据处理系统的安全防护提供了有力支撑。

### 3.3 信息数据收集和传播技术

信息数据收集和传播技术在大数据信息处理中承担着获取和输送数据的重要职责。传感器技术通过各种传感器装置可以感知并采集自然界和工业领域的海量数据，如物联网环境下的温度、压力、位置等各类传感数据。在数据收集方面，还涉及网络爬虫、数据抓取等技术，用于获取互联网上的结构化和非结构化数据资源。而将采集到的数据高效传输到存储和处理环节，则离不开通信技术的支撑。当前，5G、物联网通信等新型通信技术的发展，使得数据传输更加便捷、实时和高效，为大数据信息的高速流动提供了通路。同时，内容分发网络、对等网络等技术也为大数据的快速分发奠定了基础。

### 3.4 信息处理技术的发展

信息处理技术的发展是推动计算机大数据应用的核心动力，数据分析技术通过挖掘海量数据中的统计规律和隐含知识，为决策提供数据支撑。其中，商业智能、大数据分析等技术手段可以帮助企业发现潜在的市场机遇和潜在风险。数据挖掘技术通过应用各种算法模型，可以自动从海量数据中发现有价值的知识模式和内在关联。例如关联规则挖掘可

用于发现购物篮分析中商品之间的关联规律,为个性化推荐和交叉销售提供依据;聚类分析算法则可以将具有相似特征的数据对象划分为同一类别,用于客户细分、图像识别等场景。另外,通过序列模式挖掘可发现时间序列数据中的周期性规律,适用于异常检测、流量预测等。而机器学习算法则赋予计算机模拟人类学习的能力,使之能够自主分析数据,持续优化分析模型,自动进化出新的数据处理方式,大大提高了数据处理的智能化水平。这些创新处理技术的不断突破和集成应用,为大数据信息处理打开了新的视野,释放出巨大的应用潜能。

### 3.5 信息采集与加工方面

信息采集与加工是大数据信息处理的基础环节,负责获取数据源和对数据进行预处理。高质量的数据源是分析的前提,如何发现、获取和整合来自不同领域、不同格式的海量数据,需要综合运用网络爬虫、数据抓取、传感器采集等多种技术手段。同时,由于大数据来源庞杂,数据质量参差不齐,需要对采集到的原始数据进行全面的加工处理,包括数据清洗、数据集成、数据转换等步骤。另外,数据清洗需要采用数据填充、数据平滑、异常值处理等方法解决数据缺失、噪音数据、不一致性等问题。数据集成则需要建立数据标准和规范,通过数据抽取、转换、加载等 ETL 流程将异构数据集中到数据仓库或数据湖中,实现统一的数据存储格式;数据转换可以利用编码、归一化等手段将加工后的数据转化为结构化、标准化的形式,以满足后续分析工具对数据

格式的要求。

### 4. 结束语

总的来说,计算机大数据信息处理技术为各行业领域的决策和创新注入了新的动力,其重要性日益凸显。然而,数据安全性、处理效率和数据质量等问题也亟待解决。当前,信息存储、安全、收集传播、处理分析等各项核心技术的不断突破,为大数据处理赋能,但仍需要持续创新。信息采集与加工作为基础环节,对获取高质量数据源至关重要。未来,随着新技术的涌现和融合应用,大数据处理必将实现质的飞跃,为挖掘数据价值、支撑科学决策提供有力支撑,推动数字经济、数字社会、数字政府的高质量发展。

### 参考文献

- [1] 李长挺. 大数据时代计算机信息处理技术及策略 [J]. 电子世界, 2021,(24):21-22.
- [2] 陈莹. 大数据时代背景下计算机信息处理技术的相关研究 [J]. 数字通信世界, 2019,(04):46.
- [3] 王姝蓓. 大数据时代背景下计算机信息处理技术的相关研究 [J]. 通讯世界, 2018,(11):3.
- [4] 孟秉能. “大数据”背景下高中计算机信息处理技术策略研究 [J]. 新课程教学 (电子版), 2018,(10):99.

### 作者简介:

吴翔 (1983.1.28 - ), 男, 汉, 湖北荆州, 大学本科, 研究方向: 计算机大数据。