

企业会计信息 传递效率优化中的高效分布式计算算法设计研究

委芮 邹昱瑄

陕西铁路工程职业技术学院, 陕西 渭南 714000

摘要: 在现代企业管理中, 会计信息的高效传递对决策支持和资源配置具有重要意义。然而, 传统会计信息系统在面对海量数据处理、实时分析需求时, 易受传输延迟和计算效率限制。本文结合分布式计算技术, 设计了一种优化的高效分布式算法, 旨在提升会计信息的传递效率。研究通过构建数据分区模型与节点优化机制, 验证了算法在实际场景中的可行性和应用价值, 为企业会计信息管理提供了一种切实可行的解决方案, 同时, 为应对复杂业务场景下多维度数据的实时处理需求提供具有扩展性和鲁棒性的技术支持。

关键词: 企业会计信息; 传递效率; 分布式计算; 算法设计; 优化研究

Research on High-Efficiency Distributed Computing Algorithm Design for Optimizing Corporate Accounting Information Transmission

Wei,Rui Zou,Yuxuan

Shaanxi Railway Institute, Weinan, Shaanxi, 714000, China

Abstract: In modern enterprise management, efficient accounting information transmission is crucial for decision-making and resource allocation. However, traditional systems face limitations in massive data processing and real-time analysis due to transmission delays and computational inefficiencies. This study designs an optimized distributed computing algorithm to enhance transmission efficiency. By developing data partitioning models and node optimization mechanisms, the research validates the algorithm's feasibility and practical value, providing enterprises with viable accounting information management solutions while offering scalable and robust technical support for real-time multidimensional data processing in complex business scenarios.

Keywords: Corporate accounting information; Transmission efficiency; Distributed computing; Algorithm design; Optimization research

DOI: 10.62639/ssps20.20250103

企业会计信息系统不再仅仅扮演数据记录的角色, 而是逐步转向支持实时决策和资源配置的关键工具。然而, 传统集中式系统在处理大规模分布式会计数据时, 易受到计算能力瓶颈和数据传输延迟的限制, 难以满足现代企业高效运作的需求^[1]。

一、相关概念界定

(一) 分布式计算技术应用概述

分布式计算是一种将计算任务分散到多个节点上, 通过各节点协同工作来完成复杂计算任务的技术。在计算节点的部署和管理方面, 分布式计算提供了一种既灵活又高效的解决方案, 使得资源的利用效率得以最大化。复杂的会计信息结构、高度动态化的数据交互, 以及对实时性和准确性的严格要求, 使得传统分布式计算框架的应用遇到了不小的阻碍。尽管如此, 这种以节点协作为核心的技术为企业级应用带来的可能性依然是不可忽视的, 尤其是在数据存储和分析需求日益扩大的背景下, 其优势逐步显现。

(二) 企业会计信息传递的关键技术

企业会计信息传递是现代企业管理中不可或缺的环节。如何使大量复杂信息能够在多个部门和节点之间快速而可靠地流转, 一直是信息管理领域的核心议题。从学术研究的角度来看, 优化信息传递效率的本质在于对数据分区策略、节点通信性能和算法鲁棒性三个方面的持续改进。通过精细化的数据分区, 可以有效减少跨节点的通信成本。以 MapReduce 为代表的分布式计算框架, 已在大规模数据分析中展现出强大的处理能力^[2]。

(三) 算法优化方向

在探讨企业会计信息管理中分布式计算的优化路径时, 算法设计无疑是最为重要的突破口。近年来, 研究热点逐渐聚焦于动态任务调度与负载均衡机制的改进上。动态调度策略旨在通过更高效的资源分配方法提升系统整体性能, 而负载均衡则力图在资源使用效率和系统稳定性之间找到最佳平衡点。某些基于深度学习的预测模型已经被应用于负载分布的预估, 但其较高的计算复杂度使得直接迁移到企业会计信息管理中存在一定挑战。

(稿件编号: SE-25-3-1027)

作者简介: 委芮 (1994-), 女, 汉族, 籍贯: 甘肃省民勤县, 硕士, 助教, 研究方向: 会计实务、企业成本管理、现金流管理。
基金项目: 陕西铁路工程职业技术学院科研基金项目: “多元公共监督主体对会计信息质量提升有效性研究” (2024KYYB16)。

二、方法与模型

(一) 总体架构设计

本研究设计了一种分布式计算框架, 包含数据采集模块、分布式存储模块、计算任务调度模块和结果反馈模块, 各模块之间紧密协作, 形成了完整的功能闭环。通过任务拆分和节点分工, 提升了整体系统的计算效率。

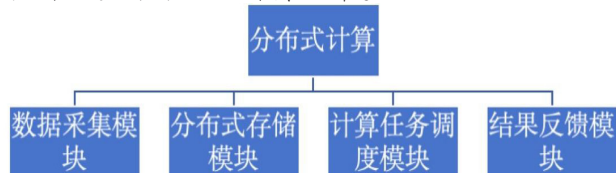


图1 总体架构图

1. 数据采集模块

为适应多样化的数据输入格式和动态更新的特点, 该模块采用了多线程采集和预处理的策略。原始数据在采集过程中根据其特征被归类标记, 并进行初步的清洗与校验, 避免错误数据对后续计算的干扰。

2. 分布式存储模块

采用分布式文件系统 (如 HDFS) 分片存储数据, 支持高并发访问和容错能力。设计分类存储方案, 将基础信息 (如客户档案) 以高一致性存储, 动态信息 (如交易记录) 则优先高并发访问需求。节点间动态迁移机制缓解了负载不均问题, 提升了读取效率和高负载场景下的稳定性。

3. 计算任务调度模块

采用基于优先级的动态任务分配模型, 以任务重要性和紧急程度计算优先级, 结合节点的实时计算能力动态分配资源。引入贪心算法缩短关键任务等待时间, 并合理分配低优先级任务资源, 优化了系统整体性能。

4. 结果反馈模块

整理并返回分布式计算结果, 设计分步反馈机制, 即对于实时性要求较高的任务, 优先返回核心结果; 而对于批处理任务, 则待所有计算完成后行再行汇总输出。满足实时性需求, 同时减少资源冗余^[3]。模块内置数据校验和异常处理功能, 通过多重验证和自动重试, 增强了反馈结果的可靠性。

(二) 算法核心设计

1. 数据分区策略

数据分区是分布式计算中提高效率的基础环节, 决定了数据在各节点间的分布方式。

设计原则: 将具有相似特征的数据分配至同一计算节点, 避免频繁的跨节点数据传输。通过分析会计信息数据的结构特征 (如日期、业务类型、金额范围等), 提取关键维度, 形成分区规则。

具体实现步骤如下: 第一步: 对数据进行初步聚类, 利用 k-means 算法确定数据的相似度分组; 第二步: 使用加权特征函数

$$D_{similarity}(x) = \sum_{i=1}^n w_i \cdot f_i(x)$$

其中, w_i 表示特征权重, $f_i(x)$ 表示特征函数值, 计算数据点的分区权重;

第三步: 将分区结果存储至节点分配表, 确

保相似数据分布在相邻节点或同一节点。

2. 任务调度优化

任务调度是提升分布式系统计算效率的重要手段。为提高会计信息处理的实时性和关键任务优先响应能力, 设计了一种基于优先级的动态任务分配算法。

设计原则: 以任务的重要性和紧急性为主要优先级指标, 动态调整任务分配顺序, 使得高优先级任务能在有限资源下优先完成。

算法流程如下:

第一步: 为每个任务设置优先级权重

$$W_i = \alpha \cdot U_i + \beta \cdot T_i$$

其中, U_i 表示任务紧急性, T_i 表示任务复杂度;

第二步: 基于贪心算法动态分配计算资源, 选择当前可用资源中能够最小化任务完成时间的节点;

第三步: 定期检查任务优先级表, 对低优先级任务进行延迟调度, 确保资源高效利用。

3. 负载均衡机制

在分布式计算环境中, 节点负载的不均衡会导致部分节点资源过载, 而其他节点资源闲置。本研究提出了一种实时监测和动态调整的负载均衡机制, 旨在提高系统整体资源利用率和稳定性。

设计目标: 避免单点过载, 最大化利用集群资源, 同时确保高并发环境下的系统稳定性。

实现方法如下:

第一步: 实时负载监测: 通过监测各节点的 CPU 使用率、内存占用率和任务队列长度, 构建负载状态矩阵 $L(t) = [l_1(t), l_2(t), \dots, l_n(t)]$, 其中 $l_i(t)$ 表示第 i 个节点在时间 t 的负载;

第二步: 动态任务迁移: 基于节点负载状态, 设计迁移策略 $\Delta T = \operatorname{argmin}_j \sum_{i=1}^n L_i(t)$, 将任务从高负载节点迁移至低负载节点;

第三步: 资源回收与优化: 在任务完成后, 自动回收资源并重新分配, 避免资源浪费。

(三) 公式与实现

1. 数据分区策略的数学模型

数据分区策略通过分析数据的特征相似性, 确保相似数据分配至同一计算节点, 从而降低跨节点通信开销, 提高计算效率, 数学模型如下:

$$P(x) = \frac{D_{similarity}(x)}{\sum_{i=1}^n D_{similarity}(i)}$$

其中:

$P(x)$: 表示数据 x 的分区权重;

$D_{similarity}(x)$: 表示数据 x 与其他数据点的相似度;

n : 为总数据量。

实现步骤:

第一步: 相似度计算: 利用欧几里得距离或余弦相似性计算 $D_{similarity}(x)$:

$$D_{similarity}(x, y) = \frac{\sum_{i=1}^m x_i \cdot y_i}{\sqrt{\sum_{i=1}^m x_i^2} \cdot \sqrt{\sum_{i=1}^m y_i^2}}$$

其中 m 表示数据维度。

第二步：分区权重归一化：将 $D_{\text{similarity}}(x)$ 归一化以生成分区权重 $P(x)$ ；

第三步：节点分配：根据 $P(x)$ 值将数据分配至对应计算节点。

2. 动态调度策略的优化模型

动态任务调度是分布式计算中的关键环节，通过合理分配计算资源，确保高优先级任务得到快速处理，优化目标：

$$\text{Maximize} \sum_{i=1}^n W_i \cdot S_i$$

其中：

W_i ：任务 i 的权重，反映任务的重要性；

S_i ：节点 i 的计算能力，通常以 CPU 使用率、内存占用等指标衡量；

n ：任务总数。

实现方法：

第一步：任务优先级计算： $W_i = \alpha \cdot U_i + \beta \cdot T_i$

其中 U_i 为任务紧急性， T_i 为任务复杂度， α ， β 为调节参数。

第二步：贪心算法调度：

按 W_i 值对任务排序；

依次选择计算能力 S_i 最大的节点进行任务分配。

第三步：实时调整：在每次调度后，更新节点的剩余计算能力，重新计算任务优先级。

三、实验与结果分析

本章节对设计的高效分布式计算算法进行了实验验证，从实验环境的搭建到结果的详细分析，全面评估算法在企业会计信息传递效率优化中的表现。

（一）实验环境

为了验证算法性能，本实验采用了下表所示的环境和数据集配置：

表 1 实验环境和数据集配置表

环境要素	配置描述
硬件	分布式服务器集群，包含 8 个节点，每节点配置为：CPU 16 核、内存 64GB、存储 2TB
软件	Hadoop 分布式框架，Python 开发环境（Python 3.8，pandas、NumPy 等库）
数据集	模拟企业会计信息数据集，包含 100 万条记录，数据字段包括日期、金额、科目编码、摘要等
评估指标	平均传递延迟、计算效率（数据处理速度）、系统鲁棒性（高负载下的稳定性）

（二）实验结果

实验通过对传递延迟、计算效率和系统鲁棒性三个维度进行测量，得到以下对比数据：

1. 传递延迟分析

新算法采用了基于数据特征的自适应分区方法，减少了跨节点通信开销，显著降低了数据传递延迟。

表 2 传递延迟分析结果

算法	平均传递延迟（ms）
传统分布式算法	150
新算法	105
优化幅度	-30%

如上表 2 所示，新算法在传递延迟上优化了

30%，证明了分区策略的高效性。

2. 计算效率分析

新算法通过优先级动态调度模型提升了关键任务的执行效率，在处理百万级数据集时表现出明显优势。

表 3 计算效率统计结果

算法	数据处理速度（条 / 秒）
传统分布式算法	7,000
新算法	9,800
提升幅度	+40%

如上表 3 所示，新算法在大规模数据处理中的计算效率提升了 40%，验证了任务调度优化的有效性。

3. 系统鲁棒性分析

在模拟高负载场景（任务队列超出系统容量的 1.5 倍）下，新算法通过动态负载均衡机制表现出更高的稳定性。

表 4 系统鲁棒性分析结果

算法	任务成功率（%）
传统分布式算法	78
新算法	92
提升幅度	+14%

如上表 4 所示，新算法在高负载情况下，任务成功率提升了 14%，避免了因单节点过载导致的系统性能下降。

4. 结果总结

表 5 为实验结果的综合对比总结。结果表明：基于高效分布式计算设计的新算法在企业会计信息传递效率优化中表现优异，不仅显著降低了传递延迟，还大幅提升了计算效率和系统鲁棒性，为企业级数据处理提供了可靠的技术支撑。

表 5 实验总结结果

评估指标	传统算法表现	新算法表现	优化幅度
平均传递延迟（ms）	150	105	-30%
数据处理速度（条 / 秒）	7,000	9,800	+40%
任务成功率（%）	78	92	+14%

四、结语

综上所述，本文设计的高效分布式计算算法，在企业会计信息传递效率优化中展现了卓越的性能，通过实验验证，其在降低传递延迟、提升计算效率和增强系统鲁棒性方面均取得了显著的改进。这一研究不仅为企业级会计信息系统的优化提供了实用的方法论支持，也为分布式算法在复杂业务场景中的应用开辟了新的可能性。然而，在实际部署中，算法的适应性和可扩展性仍需进一步考量，特别是在面对多变的数据结构和动态的业务需求时，如何实现更智能、更灵活的优化机制将成为未来研究的重点。

参考文献：

[1] 刘晓颖，陈卓．从方法论看企业信息不对称——基于企业会计信息视角[J]．商场现代化，2022，(08)：143-145.
[2] 汪頔．林业会计信息平台发展的困境及信息化策略分析[J]．湖北林业科技，2022，51（02）：86-88+47.
[3] 孙聪，贾晓霞．绿色企业会计信息披露质量内部影响因素的条件组态研究[J]．财会月刊，2021，(12)：91-100.