

基于大数据技术的传动轴磨削加工过程中碳足迹的核算

殷俊, 李仁旺

浙江理工大学机械工程学院, 浙江 杭州

收稿日期: 2023年10月16日; 录用日期: 2023年11月23日; 发布日期: 2023年11月30日

摘要

本文通过分析传动轴磨削加工过程中碳排放量的影响因素, 建立了较为准确的碳足迹核算模型。并且鉴于传统核算方法存在的一些问题决定将大数据技术运用于碳足迹的收集与核算。传统的碳足迹收集及核算方法在面临海量与无界数据时存在着诸如算力不足、收集困难等一系列问题, 而本研究利用大数据技术, 能够收集和分析大量的生产数据和能耗数据, 为碳足迹核算模型提供实时准确的元数据。

关键词

碳足迹, 大数据, 传动轴, 磨削

Calculation of Carbon Footprint in Grinding Process of Drive Shaft Based on Big Data Technology

Jun Yin, Renwang Li

School of Mechanical Engineering, Zhejiang Sci-Tech University, Hangzhou Zhejiang

Received: Oct. 16th, 2023; accepted: Nov. 23rd, 2023; published: Nov. 30th, 2023

Abstract

By analyzing the influencing factors of carbon emission in the grinding process of drive shaft, a more accurate carbon footprint calculation model is established in this paper. In view of some problems existing in traditional calculation methods, it is decided to apply big data technology to the collection and calculation of carbon footprint in the grinding process of drive shaft. Traditional

methods of carbon footprint collection and accounting have a series of problems such as insufficient computing power and collection difficulties when faced with massive and unbounded data. However, this study uses big data technology to collect and analyze a large number of production data and energy consumption data to provide real-time and accurate metadata for carbon footprint accounting models.

Keywords

Carbon Footprint, Big Data, Drive Shafts, Grinding

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着全球环境问题日益突出, 碳足迹核算成为了评估企业和制造业对环境影响的重要指标之一。传动轴磨削加工作为制造业中的重要工序, 对环境的影响也不可忽视。德国和美国等国在机床磨削加工过程的能耗及碳排放研究领域取得了较大的成果, 他们的研究主要集中在磨削过程的能耗模型建立, 以及通过优化加工工艺、改进机床设计和使用新型切削液等方式来降低能耗和碳排放。国内对磨削加工过程中碳排放的研究较晚, 但近年来发展迅速, 目前的研究主要集中在磨削过程的能耗分析, 通过建立能耗模型, 分析磨削过程中的能耗组成和影响因素。

然而, 由于大多数研究采用的是传统的碳足迹核算方法, 因此可能存在数据获取不准确、分析困难等问题。本文旨在基于大数据技术, 提出一种有效的传动轴磨削加工过程碳足迹核算模型, 以解决传统方法存在的问题。本文将根据传动轴磨削加工过程碳排放量的影响因素建立传动轴磨削加工过程碳足迹核算模型, 并基于此模型利用大数据技术进行相关核算工作。

2. 传动轴磨削加工过程碳足迹核算模型

2.1. 传动轴磨削加工碳排放量的影响因素

2.1.1. 磨削用量

磨削过程中废屑产生的碳排放通常与废屑的体积成比例。而磨削用量则与工件磨削转速以及磨削的深度呈正比。一般而言, 磨削比 e_s 可用于代表磨床的节能水平[1], 其公式如下(1)所示[2]。

$$e_s = \frac{E_m}{V} = \frac{P_m}{V_w a_p b} \quad (1)$$

其中 E_m 为磨床加工过程中产生的总能耗, V 为磨除体积, V_w 为工件线速度, b 为工件宽度, a_p 为磨削深度。零件在磨削加工过程中, 磨削用量一般可以表达为式(2) [2]:

$$\Delta V_w = \pi r(t) l d_w \quad (2)$$

根据上述公式可推断, 零件的碳排放量与磨削深度密切相关, 这表明零件磨削过程中的加工参数对碳排放量起着重要作用。

2.1.2. 砂轮磨损

砂轮的磨损量通常与其加工总量成正相关。当砂轮在待加工毛坯上高速加工时, 会产生氧化、热应

力等磨损, 从而影响零件的碳排放的总量。用“磨削比”可以很好地衡量砂轮磨损带来的碳排放, 其指的是砂轮加工过程中砂轮的磨削量与加工总磨削材料量的比值, 在稳定磨削过程中通常为常数, 用符号表示, 表达式为(3) [2]:

$$G = \frac{\Delta V_w}{\Delta V_s} \quad (3)$$

式中 ΔV_w 表示以时间为单位的砂轮磨削量, ΔV_s 表示以时间为单位的材料磨削量, 磨削比可以进一步表示为式(4) [2]:

$$G = \frac{K(V_w/V_s)^{0.158}}{F'_n} \quad (4)$$

式中 K 为磨削过程相关系数, F'_n 为法向磨削力。从而结合上述结论可得砂轮在加工过程中所损耗的体积公式可表达为(5) [2]:

$$\Delta V_s = \frac{\pi l d_w (st/2) F'_n}{K(V_w/V_s)^{0.158}} \quad (5)$$

式中 V_s 为砂轮线速度, st 为磨削余量。通过上式可以得到砂轮磨损具体的理论计算公式, 从而为下文建立零件磨削加工碳足迹模型提供理论基础。

2.1.3. 磨削液

磨削液在磨削加工中扮演着重要角色。它不仅能够提供砂轮和被加工工件的润滑, 还能降低温度和排放磨削废屑。由于磨削加工的高速运动, 会导致被加工零件的热变形和表面烧伤。而磨削液的使用可以显著减少这些损耗对加工过程的影响。一般来说, 磨削液需要定期更换, 并且添加的量会根据磨床的磨削液箱池而有所不同。因此, 本文将根据时间来计算磨削液的消耗, 以评估其对传动轴磨削加工过程的碳排放影响。

2.1.4. 磨床能耗

磨削加工过程中产生的碳排放量与物料消耗以及所使用的磨床能耗密切相关。同一批次的毛坯在不同型号的磨床上加工, 最终的碳排放量会有所不同。这种差异通常可以追溯到数控磨床的各个子系统, 包括加工系统和辅助系统。加工系统负责控制磨床的运转和材料加工过程, 而辅助系统提供支持和辅助功能, 同时也会对能耗产生一定的影响, 如照明系统、排屑系统等。因此, 选择合适的磨床型号和优化磨床的能耗是减少碳排放的重要措施。

2.2. 建立传动轴磨削加工碳足迹核算模型

对于常见的传动轴来说, 一般使用数控磨床对其进行批量加工, 因此本文引用磨削加工通用碳排放核算公式进行计算[3]。

数控磨床总体碳排放的核算主要包括数控磨床驱动电能的碳足迹核算、磨削加工过程中资源消耗的碳足迹核算两个方面。如下是对这两个方面的具体核算过程。

2.2.1. 数控磨床驱动电能的碳足迹核算

数控磨床磨削加工过程中电能的消耗主要有以下几个方面:

1) 磨床工作的基本耗能

磨床的基本能耗 E_b 的核算公式为(6):

$$E_b = \int_0^t P_b dt = P_b \cdot t \quad (6)$$

式中 E_b 为磨床的基本能耗, 磨床的基本功率为 P_b , 工作时长为 t 。

2) 主轴频率转换的功耗

主轴频率转换能耗 E_T 主要包括砂轮和导轮电机的耗能。记砂轮电机和导轮电机的功率分别为 P_w 和, 则主轴频率转换能耗 E_T 的核算公式为(7):

$$E_T = \int_0^t (P_w + P_B) dt \quad (7)$$

3) 工件材料磨削去除能耗

工件材料磨削去除能耗功率公式为(8):

$$P(t) = \mu b \tau^{-2\epsilon} \cdot V_F^{1-\epsilon} V_S^\epsilon \quad (8)$$

式中 $P(t)$ 为数控磨床工件材料磨削去除能耗, μ 为材料去除系数, b 为磨削加工的宽度, τ 为磨粒的有效间隔, ϵ 为常数, 取值范围为 0.2~0.5 之间, V_F 为磨削件的线速度, V_S 为砂轮加工的线速度。

因此工件材料磨削去除能耗 E_C 的核算公式为(9):

$$E_C = \int_0^t P(t) dt \quad (9)$$

4) 机器响应的能耗

传动轴磨削过程中会对磨床产生响应能耗 E_r , 其具体核算公式如下(10):

$$E_r = \int_0^t F_i \cdot V_n dt \quad (10)$$

其中 F_i 为传动轴和导轮之间的静摩擦力, V_n 为速度。

综上所述传动轴磨削加工过程中电能的总耗能 E 的核算公式为(11):

$$E = E_b + E_T + E_C / \gamma + E_r \quad (11)$$

式中 γ 磨削效率, 根据传动轴磨削加工的实际情况一般取 0.7~0.85。

2.2.2. 磨削加工过程中资源消耗的碳足迹核算

磨削加工过程中资源的消耗主要有以下几个方面:

1) 磨削液的消耗

传动轴磨削加工过程中需要使用磨削液来提供对砂轮和被加工工件的润滑, 同时降低加工温度, 磨削液的消耗公式如下(12):

$$R_G = \int_0^t T(t) dt \quad (12)$$

式中 R_G 为磨削液的消耗量, $\Delta T(t)$ 为第 t 种材料随时间变化的消耗量。

2) 砂轮磨损消耗

砂轮在传动轴毛坯上高速加工时, 会产生氧化、热应力等一系列磨损现象。砂轮磨损消耗的公式为(13):

$$R_w = \int_0^t \frac{\Delta V_s(t)}{V_l - V_B} dt \quad (13)$$

式中 R_w 为砂轮的磨损总量, $\Delta V_s(t)$ 为砂轮随加工时间而变化的磨损量, V_l 为砂轮的理生命周期内可供使用的体积, V_B 为砂轮的修复体积。

3) 磨屑排放消耗

磨屑排放消耗是指在传动轴磨削加工过程中, 由于材料的去除, 产生的磨屑需要进行处理和处置, 这其中就会产生一定的能源消耗[4]。磨屑排放的消耗公式为(14):

$$R_p = \int_0^t \Delta V(t) \cdot \rho_m dt \tag{14}$$

式中 R_p 为传动轴磨削加工过程中的磨屑排放的总消耗, $\Delta V(t)$ 为磨屑的排放量, ρ_m 为切屑的密度。

因此传动轴磨削加工过程中资源的消耗 R 的核算公式为(15):

$$R = R_G + R_w + R_p = \int_0^t \left(T(t) + \frac{\Delta V_s(t)}{V_I - V_B} + \Delta V(t) \cdot \rho_m \right) dt \tag{15}$$

综上所述, 传动轴磨削加工的碳足迹核算模型为公式(16):

$$CE_M = C_{Ener} \cdot E + C_{Mat} \cdot R \tag{16}$$

式中 CE_M 为传动轴磨削加工过程的总碳排放, C_{Ener} 为电能碳排放系数, C_{Mat} 为材料碳排放系数。根据《省级温室气体清单编制指南》2005 年我国区域电网单位供电平均二氧化碳排放表的数据, 本文电能碳排放系数的取值参考表 1:

Table 1. Carbon emission coefficient of electricity in various regions of China
表 1. 中国各区域电能碳排放系数

电网名称	包含省份	碳排放系数(kgCO ₂ /kWh)
华北电网	北京、天津、河北、山西、山东、内蒙古	1.246
东北电网	辽宁、黑龙江、吉林	1.096
华东电网	上海、江苏、浙江、安徽、福建	0.928
华中电网	河南、江西、四川、重庆、湖北、湖南	0.801
西北电网	陕西、青海、甘肃、新疆、宁夏	0.977
南方电网	广西、广东、云南、贵州	0.714
海南电网	海南	0.917

本文选用的材料碳排放系数主要是根据文献[5], 具体数值见表 2:

Table 2. Carbon emission factor of commonly used materials
表 2. 常用物料碳排放因子

常用物料	载能量(MJ/kg)	标煤当量(kg)	标煤碳排放因子(kgCO ₂ /kg)	物料碳排放因子(kgCO ₂ /kWh)
钢	32.0	1.09		2.69
铁	26.2	0.90	2.47	2.22
铝	191.0	6.53		16.13

3. 传动轴磨削过程中碳足迹的收集和核算

上文通过对传动轴磨削过程中碳排放特性的分析建立了具体的碳足迹核算模型, 由分析得到的碳排放源以及建立的碳足迹核算模型可以知道, 进行碳足迹核算所需的各个系统的数据量较大、数据项较多, 因此我们需要考虑采用何种技术对碳排放数据进行收集及计算。Hadoop 是一种大数据处理技术, 它采用 HDFS 这种分布式文件系统, 存储大量数据的存储, 而相较于 Hadoop 原生的大数据处理框架 MapReduce

而言, 使用 Spark 来处理传动轴磨削过程中的碳排放数据可以带来更多的优势, 首先, Spark 提供了比 MapReduce 更快的计算速度, 因为它利用了内存计算和基于内存的数据存储。这使得处理大规模数据变得更加高效。其次, Spark 具有更灵活的数据处理能力, 支持多种数据源和数据格式, 可以更方便地进行数据转换和分析。此外, Spark 还提供了丰富的 API 和功能, 如 Spark SQL、Spark Streaming 和 Spark MLlib, 可以满足不同的数据处理和分析需求, 因此本文采用 Spark 代替 MapReduce 的处理。

3.1. 碳足迹核算所需基础数据

根据碳排放源的分析 and 碳足迹核算模型的建立, 我们需要大量实时加工数据来支持该模型。为了获取传动轴磨削加工过程中的碳排放结果, 我们需要基础数据, 包括磨床系统中加工子系统和辅助子系统产生的碳排放数据等。目前有很多研究针对生产车间机床运行状态采集和生产信息采集进行了探索。这些研究表明车间数据采集已经较为成熟。本章将重点讨论和总结碳足迹核算所需的基础数据项, 为后续的数据存储和碳足迹核算做好准备。根据对碳排放源的分析, 我们将基础数据分为加工系统产生的碳排放数据和辅助系统产生的碳排放数据两大类, 并进一步细分为若干小类。图 1 列出了碳足迹核算所需基础数据项的信息。

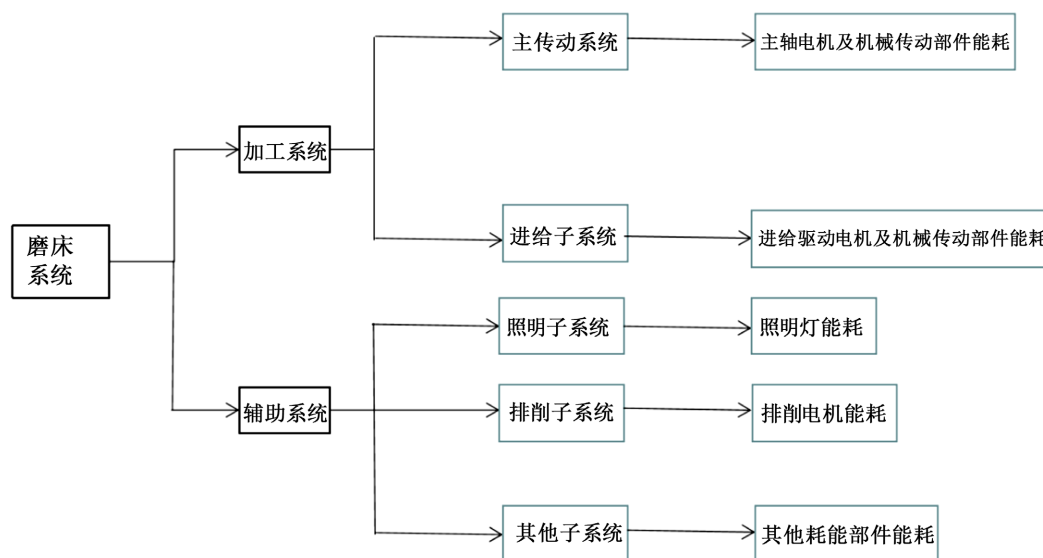


Figure 1. Basic data items on carbon emissions generated by each system

图 1. 各系统产生的碳排放基础数据项

图中主要相关数据项说明见表 3。

Table 3. An explanation of the underlying data item

表 3. 基础数据项解释说明

	说明
主传动系统	磨床运行过程中主轴电机以及机械传动部件的能耗
进给子系统	进给驱动电机及机械传动部件的能耗
照明子系统	照明灯能耗
排屑子系统	排屑电机能耗
其他子系统	其他耗能部件能耗

3.2. 基础数据存储与处理需求分析

为精确估算传动轴磨削加工过程中的碳排放, 必须密切追踪机床各系统在生产过程中的使用情况。但是, 鉴于整个加工过程中机床产生的实时数据量之大, 我们需寻求一种高效的数据存储方式。

研究发现, 仅记录一台数控机床所有状态就会产生 TB 级别的数据量, 而由于技术瓶颈, 传统关系型数据库无法妥善处理 TB 级别的数据存储。因此, 我们需要寻找一种能够妥善管理大数据的存储框架。在存储实时数据的同时, 我们还需要不断的实时处理这些数据以获取碳排放结果。鉴于数据量巨大, 我们需要借助计算机进行计算分析, 因此也需要一个大数据处理框架。

3.3. Hadoop 技术框架

3.2.1. Hadoop 功能特点

Hadoop 是一个开源软件框架, 专门为处理大量数据和大规模数据计算任务而设计。它的优势不仅在于开源性质, 更重要的是它在数据存储和计算方面的高度扩展性。Hadoop 由 HDFS 和 MapReduce 两个主要组件构成。HDFS 是 Hadoop 的分布式文件系统, 它拥有强大的存储能力和高存储效率。MapReduce 是一款处理大规模无界数据流的分布式计算框架, 与传统的计算框架相比, 其在计算规模和效率上具有显著的优势, 而 spark 在其基础上计算能力和性能更进一步[6]。

1. 可扩展性: Hadoop 可以方便地扩展集群的规模, 通过增加更多的节点来增加存储容量和计算能力。它可以适应不断增长的数据量和计算需求, 保持良好的性能和可靠性;

2. 高容错性: Hadoop 通过数据的冗余备份和自动故障恢复机制, 确保数据的安全性和可靠性。它将数据划分为多个块, 并在集群中的不同节点上存储多个副本, 当某个节点发生故障时, 可以快速恢复数据并继续进行计算任务;

3. 高性能: Hadoop 通过将数据和计算任务分布在多个节点上, 并采用并行处理和本地化计算等技术, 以实现高效的数据处理和计算。它能够充分利用集群中的计算资源, 提高数据处理的速度和效率;

4. 适应多种数据类型和工作负载: Hadoop 不仅可以处理结构化数据, 还可以处理半结构化和非结构化的大数据。它可以适应不同类型的工作负载, 包括批处理、交互式查询、流处理等多种数据处理场景。

由以上对 Hadoop 功能特点的介绍可知, Hadoop 主要就是应对大数据量的存储以及对海量数据的处理, 因此 Hadoop 完全可以满足碳足迹核算基础数据的存储与处理的需求, 可将碳足迹核算所需基础数据存储到 HDFS 中, 并依据碳足迹核算模型, 采用 spark 计算框架对这些基础数据进行处理, 从而得到碳足迹核算结果。

3.2.2. HDFS 分布式文件系统

面对传动轴磨削加工过程中的实时数据繁杂且数据量较大的问题, 通过网络通信将这些数据存储到 HDFS 中。HDFS 是一种主从模式的存储系统结构[7], 可以由单台服务器扩展到上千台服务器, 其系统结构如图 2 所示。

NameNode 是主节点, 负责存储文件系统的元数据, 包括文件和目录结构、权限信息、文件位置等。它是 Hadoop 集群的核心, 控制所有的数据存储和数据分析过程。

DataNode 是工作节点, 主要负责存储和检索数据。当客户端需要读取或写入数据时, 它们会直接与 DataNode 进行通信。每个 DataNode 管理自己节点上的数据存储, 并定期向 NameNode 报告其存储的块信息, 以便 NameNode 能够进行全局的文件系统管理。

Client 是一个客户端程序, 主要负责与 HDFS 进行交互, 实现文件的读写操作。当客户端需要读取文件时, 它首先会连接到 NameNode 并获取文件的元数据信息, 包括文件的块信息和所在的 DataNode 的信息。然后, 客户端直接从 DataNode 中读取数据。当客户端需要写入文件时, 它会向 NameNode 发出

请求, 获取可以写入的 DataNode 的信息。然后, 客户端将数据直接写入到 DataNode, 并且同时也会将数据的副本写入其他的 DataNode, 以实现数据的冗余和容错。写入完成后, 客户端会向 NameNode 报告文件的元数据信息。

Block 是数据存储和处理的基本单位。在 HDFS 中, 文件被切分成一系列的块, 然后这些块被存储在集群的不同节点上。每个块都有一个唯一的块 ID, 以及一份元数据信息, 包括块的大小、创建时间、副本数量等。

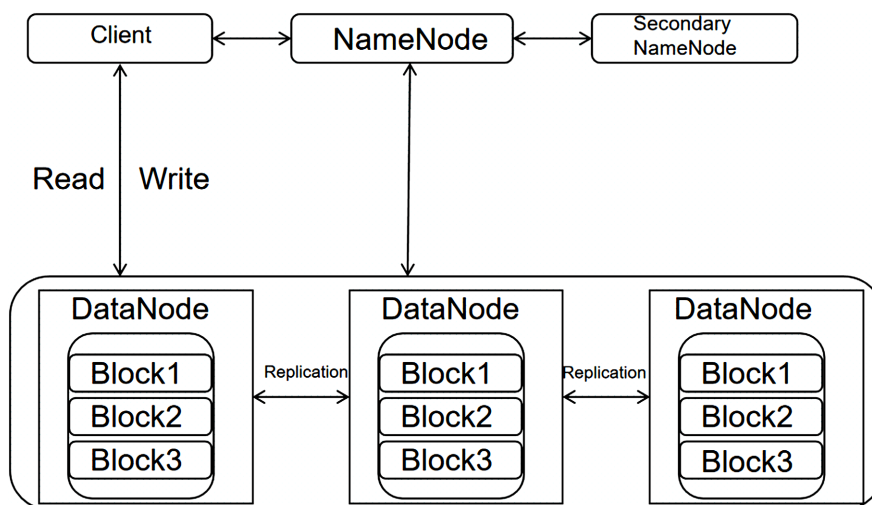


Figure 2. HDFS distributed storage system
图 2. HDFS 分布式存储系统

3.2.3. Spark 分布式计算框架

为提高计算效率, 在碳足迹计算模型的基础上, 可以采用 Spark 分布式计算框架来处理基础数据。与 MapReduce 不同, Spark 采用了基于内存的计算模型, 能够在处理大规模数据时提供更快计算速度。

Spark 的主要思想是将大规模的计算任务分解为一系列的操作, 通过构建一系列的转换操作和动作操作来处理数据。Spark 的运行过程主要由 Driver、Executor 和 Cluster Manager 三部分组成。Driver 是 Spark 应用程序的主控制器, 负责解析应用程序代码并将任务分发给 Executor 进行执行。Executor 是 Spark 集群中的工作节点, 负责实际的任务执行。Cluster Manager 则负责协调和管理 Executor 的分配和调度。在处理计算任务时, Spark 采用了弹性分布式数据集(Resilient Distributed Datasets, 简称 RDD)来表示数据集。RDD 是一个可分区、可并行计算的数据集合, 可以在计算过程中被缓存到内存中, 提高计算速度。Spark 的处理过程也可以分为两个阶段: 转换阶段和动作阶段。转换操作是对 RDD 进行转换, 生成新的 RDD, 而动作操作是对 RDD 进行计算, 并返回结果。在实际开发过程中, 可以通过编写一系列的转换操作和动作操作来处理碳足迹计算模型的基础数据。Spark 提供了丰富的 API 和函数库, 可以方便地进行数据处理和计算。通过使用 Spark 分布式计算框架, 可以充分利用集群中的计算资源, 并通过并行化计算提高计算效率[8]。同时, Spark 的内存计算模型也能够加速数据处理过程, 从而更快地得出最终的计算结果。

Spark 处理传动轴碳排放数据具体流程图如图 3 所示。

图中, 每个 Map 阶段的函数会产生键值对形式的结果, 为了更好的给 Reduce 分配计算任务, spark 框架还会对 Map 阶段的结果进行整理合并, 然后再作为输入数据传递给 reduce 函数进行处理[9]。

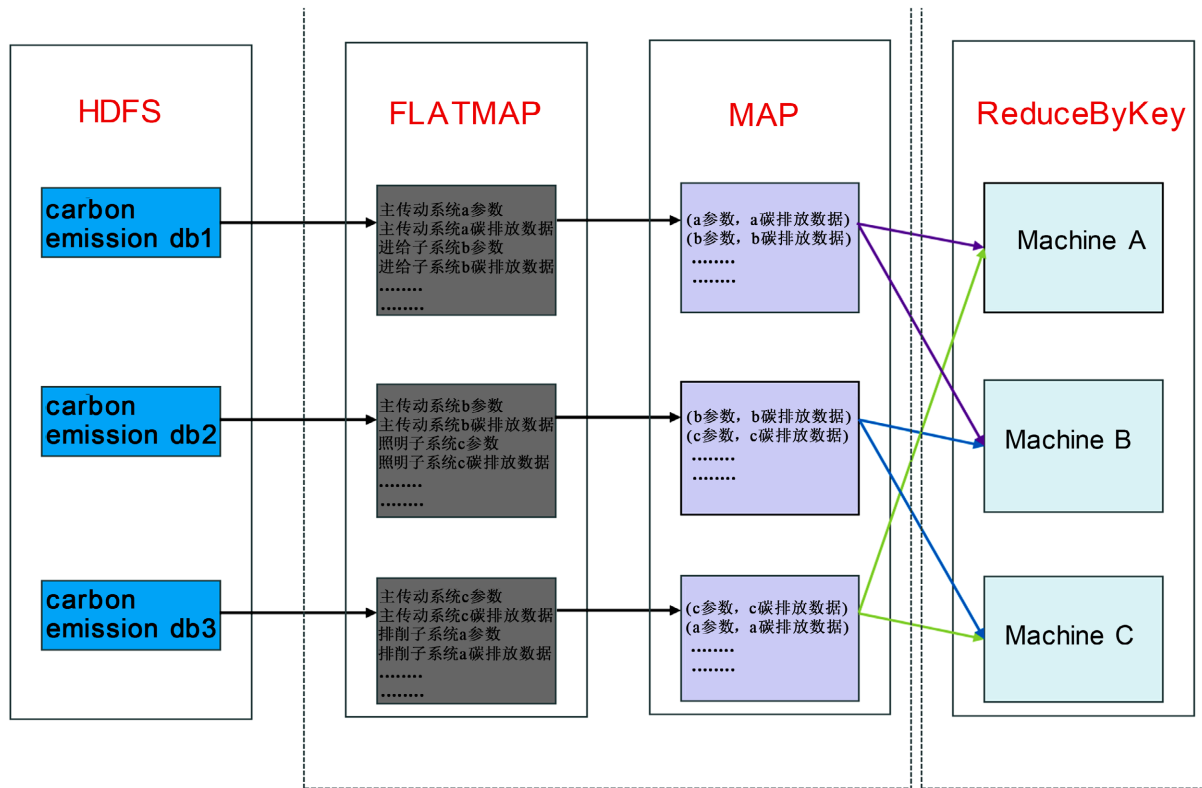


Figure 3. Spark processes the propeller shaft carbon emission data flow chart
图 3. Spark 处理传动轴碳排放数据流程图

3.4. 基于 HDFS 的基础数据存储

根据上文探讨的传动轴磨削加工过程碳足迹核算所需的各系统的基础数据可知, 这些数据是在不间断实时产生的, 即所谓的无界数据流, 因此数据量十分庞大, 实时性要求较高。在 Spark 对这些数据处理之前, 我们需要制定一套行之有效的存储策略, 将基础数据按照一定的格式存储到 HDFS 中, 以便于后续 Spark 的数据处理操作。

3.4.1. 数据采集存储流程

在抽取磨削过程产生的实时数据之前, 我们需要将工件信息, 机床信息等一系列静态数据优先存储到 postgresql 数据库中, 以便于后续与实时数据的汇总使用。因此本文制定了如图 4 所示的数据采集存储流程。

碳排放的实时数据来自于磨床的各个子系统, 因此我们采用常见的软件采集器和硬件采集器并用的方式, 实时监测机床各个系统的状态信息。WEB SERVICE 除了提供配置信息外, 还提供数据访问接口 Sop 和数据展示 web 模块, 将数据采集系统采集到的数据提供给外部系统访问。同时我们通过 etl 抽取之前预先保存到 postgresql 中机床以及工件的静态数据到 kafka 中, 然后利用 kafka-connect 将数据按照数据存储策略加载到 HDFS 中, 完成数据的汇总。

3.4.2. 基于问题分析树型图的数据存储策略

为了整理出有条理的存储策略, 借助问题分析树型图来梳理各种碳排放因素。问题分析树型图是一种以树状结构来展示问题分析和解决方案的方法。它以一个问题作为根节点, 通过不断细分和关联, 将问题分解成更具体的子问题, 并找出解决每个子问题的具体方案。这种图形化的表示方法有助于组织和展示问题的分析过程, 使人们更清晰地理解问题的本质和各个方面的关系将其运用到传动轴磨削加工数

据的研究上可以进一步的细化分析每一个系统产生碳排放数据的具体原因, 也有助于决定本文的存储策略。传动轴磨削加工过程碳排放的问题分析树型图如图 5 所示。

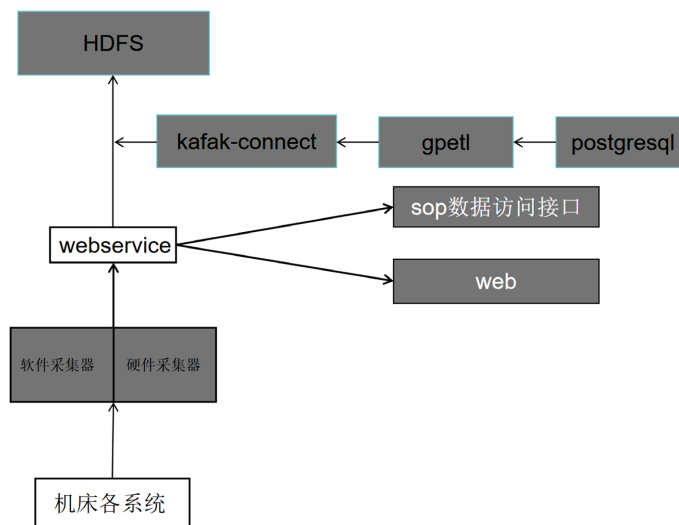


Figure 4. Data acquisition and storage process
图 4. 数据采集存储流程

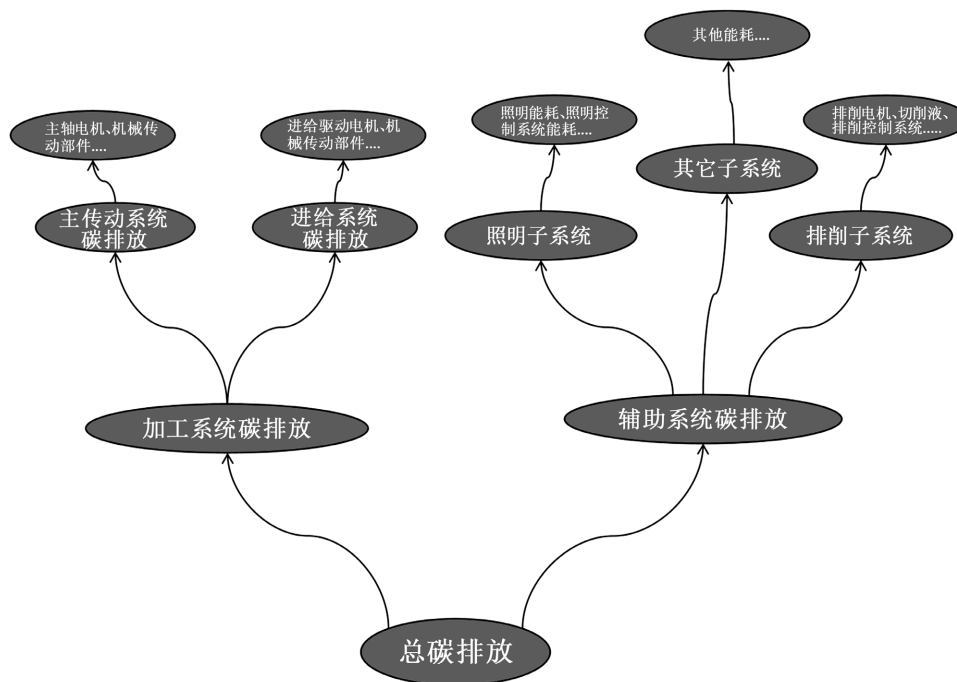


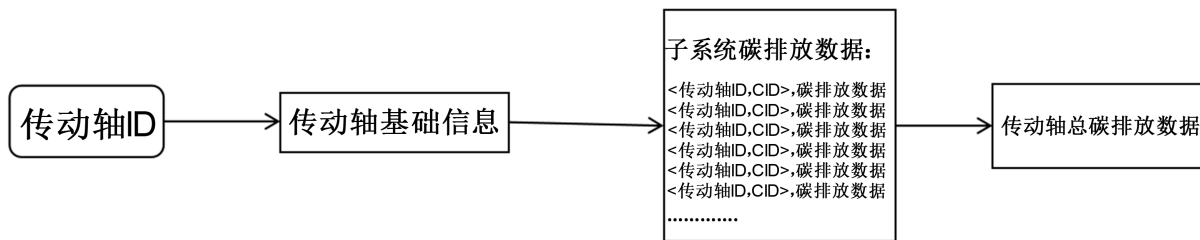
Figure 5. Tree diagram of the problem analysis of carbon emissions during the grinding process of the drive shaft
图 5. 传动轴磨削加工过程碳排放的问题分析树型图

根据图 5 所示, 由于每一个主系统下的子系统完全不相同, 因此可以为每一个子系统设置一个唯一的系统 ID, 并且由于子系统产生碳排放的因素可能有多, 因此可以为每一个因素也设置一个因素 ID, 这样系统 ID 和因素 ID 就可以唯一的表示一个碳排放数据的出处, 我们将这个唯一 ID 记为 CID; 这样的话, 我们可以制定如下的数据存储策略:

1. 传动轴信息存储：传动轴 ID - 传动轴基础信息。
2. 传动轴加工过程信息存储：传动轴 ID-CID - 碳排放数据。

3.5. 基于 Spark 的碳足迹核算流程

要将 Spark 分布式计算应用到零件加工过程的碳足迹核算中，首先需要根据碳足迹核算模型确定数据的处理流程，然后定义 Map 和 Reduce 两个阶段的数据处理函数。如图 6 所示为传动轴磨削加工过程核足迹核算的数据处理流程。首先根据传动轴 ID 从 HDFS 中获取传动轴的基础信息数据，然后根据 CID 查询出各个子系统的碳排放量，最后将各系统的碳排放叠加得出整个传动轴磨削加工过程的碳排放量。



注：CID 代表子系统 ID 和子系统碳排放因素 ID 的联合唯一 ID。

Figure 6. Data processing flow chart

图 6. 数据处理流程图

根据数据处理流程图建立 Map 和 Reduce 阶段的计算任务,如图 7 所示为运用 Spark 进行碳足迹核算两个阶段的数据流图。

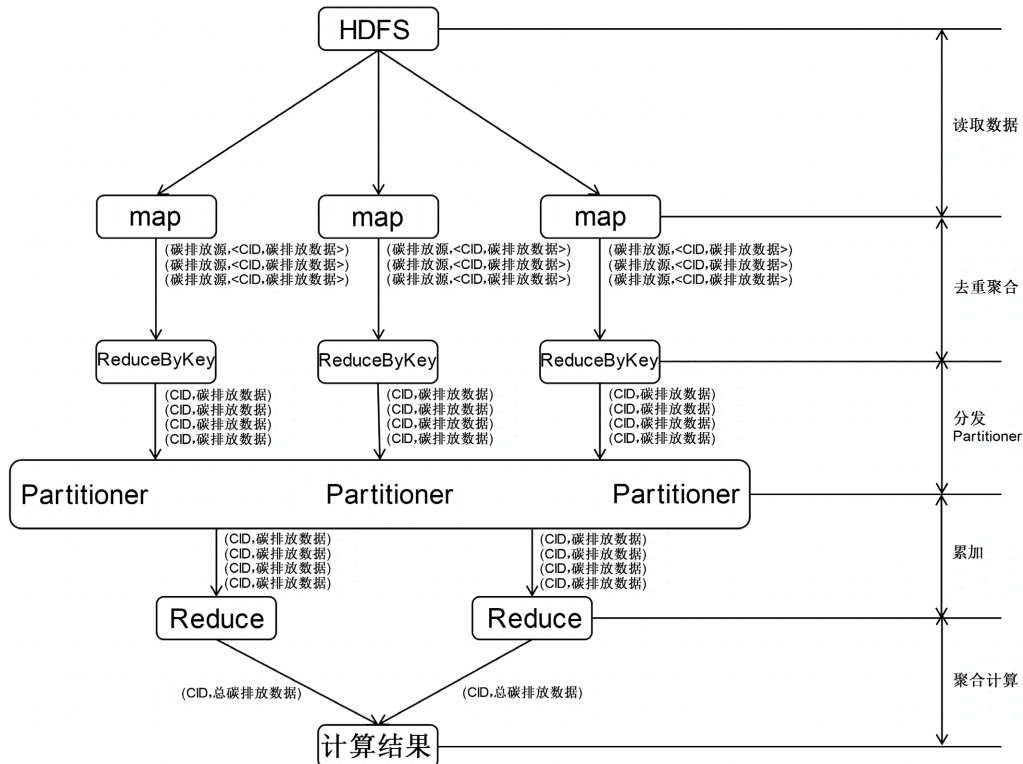


Figure 7. A data flow diagram of the two phases of Spark's carbon footprint accounting

图 7. Spark 进行碳足迹核算的两个阶段的数据流图

我们使用 `map` 函数来对获得的数据块中的各个子系统的各个因素的碳排放进行计算。输出数据的结构可以使用键值对的形式表示, 其中键是碳排放源的编号, 值是一个包含 `CID` 和碳排放量的对象。因此, `Map` 阶段的数据输出格式为<碳排放源编号, (`CID`, 碳排放量)>。

为了减少输出文件数据的冗余并提高数据传输效率, 在 `Spark` 中可以使用 `ReduceByKey` 操作对 `map` 函数的结果进行合并处理。该操作会将具有相同键的数据进行聚合, 将键相同的数据的碳排放量进行相加。这样, 可以将多个数据项合并为一个, 并将 `CID` 作为新的键, 碳排放量作为值。此时的数据格式为<`CID`, 碳排放量>。

接下来, 使用 `Partitioner` 接口来决定将数据交给哪个 `reduce` 函数进行处理。`Partitioner` 根据键或值以及 `reduce` 的数量来划分数据, 确保具有相同 `CID` 的数据被分配到同一个 `reduce` 任务中。在 `Spark` 中, `reduce` 函数的数量不是根据不同的 `CID` 确定的, 而是根据配置文件或集群资源进行调整。

在 `Reduce` 阶段, `reduce` 函数接收 `Map` 阶段的输出结果作为输入, 并根据 `CID` 对碳排放量进行累加。`Reduce` 阶段的数据输出格式仍为<`CID`, 碳排放量>。与 `Map` 阶段不同的是, `Reduce` 阶段的目标是对工序碳排放量进行累加, 而不需要进行其他计算[10]。

4. 总结

本文主要针对传动轴磨削加工过程中的碳足迹核算进行了研究。鉴于目前碳足迹核算研究中的一些问题和不足, 从碳足迹核算模型的建立、碳排放实时数据的收集、碳足迹核算方面的具体数据的处理等方面开展了深入研究。提出借助大数据技术进行加工过程中碳排放数据的实时收集和计算。文章的研究理论和成功有利于企业实现对传动轴磨削加工过程中碳足迹的实时监控, 企业可根据该模型对传动轴磨削加工的加工工艺进行相应的优化, 最终实现低碳制造。

基金项目

浙江省 2023 年度“尖兵”“领雁”研发攻关计划(2022C01SA111123), 国家自然科学基金资助项目(51475434)。

参考文献

- [1] 赵旭东. 基于 OPC 的数控加工单元数据采集技术研究[D]: [硕士学位论文]. 成都: 电子科技大学, 2016.
- [2] 王杰. 面向低碳的定位销生产过程中碳足迹核算及其加工参数优化[D]: [硕士学位论文]. 杭州: 浙江理工大学, 2023. <https://doi.org/10.27786/d.cnki.gzjlg.2022.000747>
- [3] 刘文英. 磨削加工中磨削液的选择[J]. 工业技术创新, 2016, 3(6): 1120-1123. <https://doi.org/10.14103/j.issn.2095-8412.2016.06.017>
- [4] 迟玉伦, 芦华, 徐亮亮, 等. 基于磨削去除率建立无心磨削 ERWC 碳排放模型的方法[P]. 中国专利, CN115577512A, 2023-01-06.
- [5] 刘朝晖, 张为民, 金希, 等. 制造过程中的广义碳排放计量方法研究[J]. 机械科学与技术, 2019, 38(4): 544-550. <https://doi.org/10.13433/j.cnki.1003-8728.20180219>
- [6] 朱毓. 基于 Hadoop 与 Spark 的大数据处理平台的构建研[J]. 互联网周刊, 2023(7): 87-89.
- [7] 丁浩. 基于 SpringBoot 项目融合查询 HDFS 分布式文件系统[P]. 中国专利, CN116126793A. 2023-05-16.
- [8] 曹海平. 依托于 Spark 平台的大数据挖掘技术分析[J]. 软件, 2022, 43(7): 84-86.
- [9] 梁芷梧. 云计算中 MapReduce 分布式并行处理框架的研究[D]: [硕士学位论文]. 武汉: 湖北工业大学, 2017.
- [10] Zhang, J.H. and Wang, H.L. (2023) Analysis of CLARANS Algorithm for Weather Data Based on Spark. *Computers, Materials & Continua*, **76**, 2427-2441. <https://doi.org/10.32604/cmc.2023.038462>