文章编号:1671-251X(2021)12-0121-07 DOI:10.13272/j.issn.1671-251x.2021040004

基于 Hadoop 生态圈的选煤数据中台设计

赵鑫, 王然风, 付翔

(太原理工大学 矿业工程学院, 山西 太原 030024)

扫码移动阅读

摘要:针对现有选煤厂信息管理系统采用的接口不规范,导致数据重复采集,且各系统相互独立,对多源异构数据处理能力弱等问题,基于 Hadoop 生态圈大数据技术,提出了一种基于 Hadoop 生态圈的选煤数据中台设计方案。通过主数据管理系统、企业服务总线定义数据标准实现系统集成;设计归一化、相关系数矩阵和噪声异常点检测程序实现数据处理;结合 D-S(Dempster-Shafer)证据理论、Hadoop 与 Hive 数据仓库设计多源异构数据融合子系统,实现数据融合;利用 Highcharts 数据可视化组件实现数据交互式的可视化展示。实际应用结果表明,该数据中台实现了主数据定义标准与系统集成接口规范化,提高了选煤数据处理能力,实现了多源异构选煤数据融合共享、数据实时交互式的可视化展示。

关键词:选煤厂智能化; Hadoop 生态圈; 数据中台; 大数据; 多源异构数据; 数据融合; D-S 证据理论

中图分类号:TD948 文献标志码:A

Design of coal preparation data center platform based on Hadoop ecosystem

ZHAO Xin, WANG Ranfeng, FU Xiang

(College of Mining Engineering, Taiyuan University of Technology, Taiyuan 030024, China)

Abstract: The existing coal preparation plant information management system uses nonstandard interface, which leads to repeated data collection, and each system is independent of each other, and the capability to process multi-source heterogeneous data is weak. In order to solve above problems, based on big data technology of Hadoop ecosystem, a coal preparation data center platform design scheme based on Hadoop ecosystem is proposed. The system integration is realized by defining data standards through master data management system and enterprise service bus. Normalization, correlation coefficient matrix and noise abnormal point detection programs are designed to realize data processing. DS (Dempster-Shafer) evidence theory, Hadoop and Hive data warehouse are combined to design multi-source heterogeneous data fusion subsystem to realize data fusion. Highcharts data visualization components are used to achieve interactive visualization of data. The practical application results show that the data center platform realizes the standardization of master data definition standard and system integration interface, improves the processing capability of coal preparation data, realizes the fusion and sharing of multi-source heterogeneous coal preparation data, and realizes the real-time interactive visualization of data.

Key words: intelligent coal preparation plant; Hadoop ecosystem; data cente platform; big data; multi-source heterogeneous data; data fusion; D-S evidence theory

---

收稿日期:2021-04-02;**修回日期**:2021-12-04;**责任编辑**:张强,郑海霞。

基金项目:山西省应用基础研究计划重点自然基金项目(201901D111007);山西省关键核心技术和共性技术研发攻关专项项目(2020XXX004)。

作者简介:赵鑫(1996-),男,山西太原人,硕士研究生,研究方向为选煤智能化与大数据开发,E-mail:zhaoxin3158@163.com。

引用格式:赵鑫,王然风,付翔. 基于 Hadoop 生态圈的选煤数据中台设计[J]. 工矿自动化,2021,47(12):121-127.

ZHAO Xin, WANG Ranfeng, FU Xiang. Design of coal preparation data center platform based on Hadoop ecosystem[J]. Industry and Mine Automation, 2021, 47(12): 121-127.

0 引言

目前智能化煤矿建设已成为国家战略[1]，在智能化选煤厂建设过程中，主要关注设备及其控制的智能化分析，而数据智能化是选煤厂智能化的有机组成部分，对实现选煤数据精细化管理具有重要意义[2-5]。

在物联网、大数据、云计算、人工智能为核心的新一代信息技术支撑下，许多学者对选煤数据智能化进行了研究。匡亚莉[6]针对智能化选煤厂建设，提出选煤大数据系统研究的目标是全面的数据融合与集成。张磊[7]针对选煤厂数据共享实时性差和信息传输准确性低等问题，提出利用 PLC 二次回路技术提高现场设备运行数据采集的精确度，并通过对系统集成接口进行深度开发，建立了全面的信息管理系统。高晓茜等[8]为提高选煤厂调度员数据记录效率，设计了基于 B/S 架构的信息化调度管理系统，实现了选煤厂生产数据自动记录、集成和共享。孙小路等[9]研究了选煤实时数据、关系数据库中离散数据的存储规则，并通过 Web API（Application Programming Interface，应用程序接口）实现数据共享。选煤厂信息管理系统的应用，提高了选煤厂生产数据的管理效率，但不同系统采用的接口不规范，导致数据重复采集，使得信息严重冗余，且各系统相互独立，对多源异构数据处理能力弱，数据融合缺乏大数据技术支撑。

针对上述问题，本文提出了基于 Hadoop 生态圈的选煤数据中台设计方案：通过 MDM（Master Data Management，主数据管理）系统实现主数据标准规范化；通过 ESB（Enterprise Service Bus，企业服务总线）实现系统集成接口规范化；设计归一化、相关系数矩阵和噪声异常点检测程序实现数据处理；结合 D-S（Dempster-Shafer）证据理论与 Hadoop 生态圈技术设计多源异构数据融合子系统，实现数据融合共享；运用 Highcharts 数据可视化组件实现数据实时交互式的可视化展示。

1 选煤数据中台架构

选煤数据中台由数据感知层、基础设施层、中台实现层和业务服务层组成，其总体架构如图 1 所示。

（1）数据感知层负责数据的自动采集与感知。通过 PLC、传感器等智能终端，采用 Modbus、OPC 协议自动感知洗选、人员定位等数据；使用 WebService、IEC104 协议实现 SCADA（Supervisory Control and Data Acquisition，数据采集与监视控制）、ERP（Enterprise Resource Planning，企业资源计划）、OA（Office Automation，办公自动化）等系统的关系型数据库中计划、生产、能源、环保、安全、设备等管控数据的自动采集；通过录像机、视频环网使用 RTSP（Real Time Streaming Protocol，实时流传输协议）接入监控视频数据。

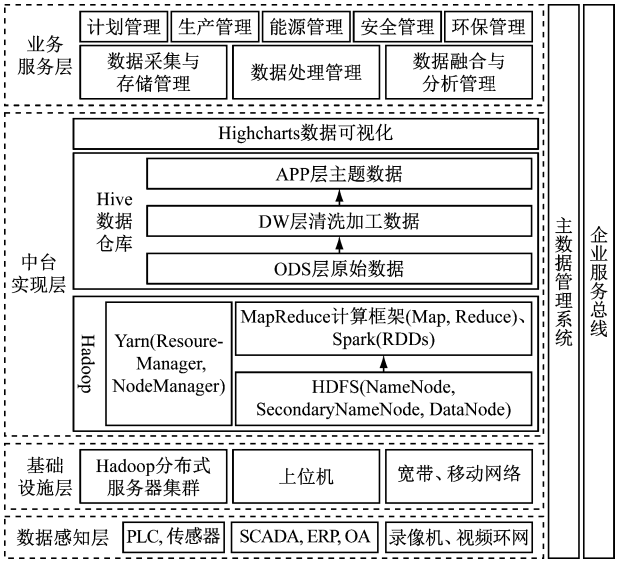

图 1 选煤数据中台总体架构

Fig. 1 Overall structure of coal preparation data center

（2）基础设施层由选煤厂互联网数据中心部署分配的 Hadoop 分布式服务器[10-11]集群、上位机及宽带、移动网络组成。Hadoop 分布式服务器集群通过分布式架构解决了单台服务器多线程并行计算耗时多，大规模数据存储量大、网络带宽等问题。通过部署服务器集群[12-13]将大数据技术的相关模块部署在不同服务器并形成备份，以解决高并发访问量情况下的负载不均衡问题，避免了因某一台服务器单点故障而影响整个数据中台运行。Hadoop 分布式服务器集群由 5 台服务器组成，每台服务器彼此独立，相互之间仅通过消息队列通信。

（3）中台实现层由 Hadoop 框架、Hive 数据仓库和 Highcharts 数据可视化组件组成。

Hadoop 框架包括 HDFS（Hadoop Distributed File System，Hadoop 分布式文件系统）、Yarn（Yet Another Resource Negotiator，另一种资源协调者）、MapReduce 计算框架和 Spark 4 个组件。HDFS 组件中 NameNode 作为主节点，负责获取数据位置信息；SecondaryNameNode 作为辅助节点，管理元数据信息；DataNode 作为从节点，负责读写、存储相关数据。Yarn 组件中 ResourceManager 执行用户的计算请求，并负责合理分配资源；NodeManager 负责管理 Hadoop 集群中的单个计算节点。MapReduce 计算框架在 Map 阶段将大规

模的计算任务拆解成多个子任务,在 Reduce 阶段将子任务计算结果合并为最终的计算结果。该组件需要处理的原始数据及计算结果通过 HDFS 存储,计算的同时需要 Yarn 并行提供资源调度。Spark 通过 RDDs(Resilient Distributed Datasets,弹性分布式数据集)执行引擎将中间数据在服务器内存中迭代存储,克服了 MapReduce 计算框架在磁盘中运算的缺陷。

Hive 数据仓库面向特定主题,关注联机分析处理与管理决策。Hive 数据仓库包括 ODS(Operational Data Store,数据运营)层、DW(Data Warehouse,数据仓库)层、APP(Application,数据应用)层。ODS 层实现与其他数据中台的关系数据库数据表、文档、图片、点击流日志的同步;DW 层对数据进行清洗加工、分析与挖掘,并形成特定业务主题数据存入 APP 层;APP 层提供针对特定场景的个性化业务数据,供后续数据分析与挖掘使用。

Highcharts 数据可视化组件用于实现数据交互式的可视化展示。

(4) 业务服务层依据选煤厂及煤炭集团业务需求,将数据中台功能进行模块化封装,根据用户操作权限展示与操作不同功能页面。

2　选煤数据中台关键技术

面对 TB 级别、彼此孤立存在的海量多源异构数据,数据中台通过 Hadoop 生态圈技术实现数据感知、集成、处理、融合、分析、可视化展示的全生命周期管理。

2.1　多源异构数据集成技术

(1) MDM 系统。数据中台的主数据包括通用基础类(地区、民族、资金单位)数据、单位类(内部单位、外部单位)数据和产品类(原煤、中煤、精煤、矸石)数据等。针对主数据分类、编码不一致,关键信息录入空值、不规范等问题,数据中台设计 MDM 系统,根据 DB14/T 2245—2020《煤炭洗选企业标准化管理规范》,明确定义增量数据中间表的相关字段、数据类型与映射关系。以产品类主数据中间表字段为例,通过 MDM 系统实现各生产部门按需选择主数据及相应的属性字段,为 ESB 提供集成的主数据字段映射关系,相关信息见表 1。

表 1　主数据煤炭类中间表字段及相关信息

Table 1　Master data coal intermediate table fields and related information

序号	数据项	类型	限制条件	说明	数据长度/字符	显示列
1	状态代码	字符	必填	中台自动生成	6	state_code
2	状态描述	字符	必填	中台自动生成	80	state_desc
3	产品代码	字符	必填	中台自动生成	6	pro_code
4	产品名称	字符	必填	手动输入	80	pro_name
5	产品描述	字符	必填	中台自动生成	80	pro_desc
6	产品分类代码	字符	必填	中台自动生成	6	sort_code
7	产品分类描述	字符	必填	中台自动生成	80	sort_desc
8	硫分	浮点	非必填	手动输入	32	sulfur
9	灰分	浮点	非必填	手动输入	32	ad
10	挥发分	浮点	非必填	手动输入	32	volatile
11	发热量	浮点	非必填	手动输入	32	qnet,ar
12	全水分	浮点	非必填	手动输入	32	mt

(2) ESB。针对现有数据处理软件采用不同的编程语言、数据库及系统集成接口不规范等问题,数据中台采用 SOA(Service Oriented Architecture,面向服务)架构的 ESB 模式实现系统集成接口规范化。SOA 架构将不同的应用功能单元视为一个服务,对服务定义标准化的接口与契约,实现松耦合的总线型拓扑结构。

ESB 模式按照数据流的传输方向,将访问服务分为生产者(数据提供方)与消费者(数据接收方),如图 2 所示。生产者在数据发生变化时调用 ESB 接口传入 Json 文件,并发起流程;ESB 调用数据接收接口,使用单条同步的方式进行状态同步;ESB

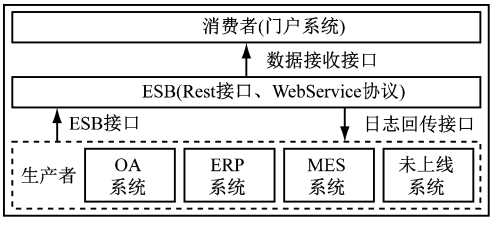

图 2　ESB 架构

Fig. 2　ESB architecture

通过日志回传接口向生产者反馈消费者的状态同步结果。当未来新上线的系统需要通过 ESB 与数据中台集成时，统一注册并拓展 Rest 服务接口，使用标准 WebService 协议实现消息通信、服务检测等功能。

2.2 多源异构数据处理技术

针对数据采集、集成过程中数据原始量纲不统一，多随机变量间相关关系无法判断，数据集包含不一致、不完整的“脏数据”等问题，数据中台借助 Hadoop 生态圈的 ETL（Extract－Transform－Load，数据仓库技术），研究并设计归一化程序、相关系数矩阵程序和噪声异常点检测程序进行数据处理，为后续数据融合与分析提供数据支撑。

（1）归一化（去量纲）。将有量纲的数据集通过比例缩放转换为无单位的小范围标量。通过线性函数离差归一化方法（式（1）），将数据集映射到[0，1]区间。

$$x^* = \frac{x - \min X}{\max X - \min X} \tag{1}$$

式中：x^* 为归一化数据；x 为原始数据；X 为原始数据集。

（2）相关系数矩阵。根据 MapReduce 计算框架的键值对思想，利用 Pearson 公式对数据集进行映射与规约，计算每个特征值之间的相关系数。MapReduce 计算框架在 Map 阶段通过多线程并行循环读取行数据与列数据并进行分布式计算，在 Reduce 阶段将分组算出的系数汇总并输出。以商品煤质量鉴定单中煤泥化验结果为例，将全水分、灰分和发热量 3 个字段作为数据处理对象，选取 500 组数据作为数据集，见表 2。

表 2 商品煤质量鉴定单

Table 2 Commercial coal quality appraisal sheet

序号	全水分/%	灰分/%	发热量/(J·g^{-1})
1	10.8	34.76	17.572
2	10.6	32.29	20.481
3	10.7	33.58	19.891
4	10.6	35.51	17.928
5	10.5	33.49	18.208
⋮	⋮	⋮	⋮
495	10.7	34.73	19.409
496	10.6	33.52	18.367
497	10.7	33.19	19.095
498	10.5	35.55	18.108
499	10.8	35.58	17.179
500	10.4	36.12	17.789

采用 Pearson 公式对数据集进行运算，运算结果的特征关系系数值域为[－1，1]，通过 Highcharts 数据可视化组件中的矩阵表示随机变量间的相关性关系，如图 3 所示。

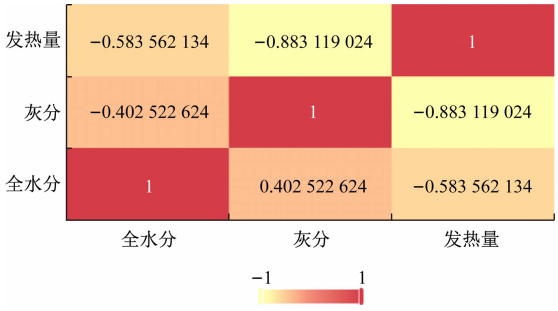

图 3 相关系数矩阵

Fig. 3 Correlation coefficient matrix

由 Pearson 公式定义可知，相关系数＞0 表示二者之间呈正相关，相关系数＜0 表示二者之间呈负相关。由图 3 可知，发热量与全水分的相关系数为－0.583 562 134，绝对值位于 0.40～0.69 之间，表示二者之间呈中度负相关；发热量与灰分的相关系数为－0.883 119 024，绝对值位于 0.70～0.89 之间，表示二者之间呈高度负相关。通过相关系数矩阵程序处理，为下一步数据挖掘发热量与全水分、灰分的线性关系提供支持。

（3）噪声异常点检测。数据中台引入 Spark 框架下的 DataFrame 对象，借助四分位距原理检测一组测试数据中是否存在噪声异常数据。DataFrame 对象将 Hive 数据仓库、Json 文件、关系型数据库的数据表作为数据源，调用内置 approxQuantile()方法传入列名及计算分位点等参数，并求出每个字段的边界；调用 select()、filter()方法循环遍历数据集，最终筛选出噪声异常数据。

2.3 多源异构数据融合技术

针对选煤厂海量数据融合的需求，结合经典 D-S 证据理论[14-15]、Hadoop 和 Hive 数据仓库，设计了多源异构数据融合子系统。D-S 证据理论运用 Dempster 合成准则和信度函数描述按时序获得的要素间相互关系。数据融合流程如图 4 所示，其中 $F_N(\cdot)$，$B_N(\cdot)$，$P_N(\cdot)$ 分别为 MapReduce 将大规模计算任务拆分成 N 个子任务的正交和组合函数、信任函数、似然函数。

（1）调用 API（Application Programming Interface，应用程序接口）中的 showTables()方法判断数据源表是否存在，若存在则调用 selectTable()方法选择目标数据源表、调用 map()方法选择表中的具体字段，否则转至数据集成与预处理。

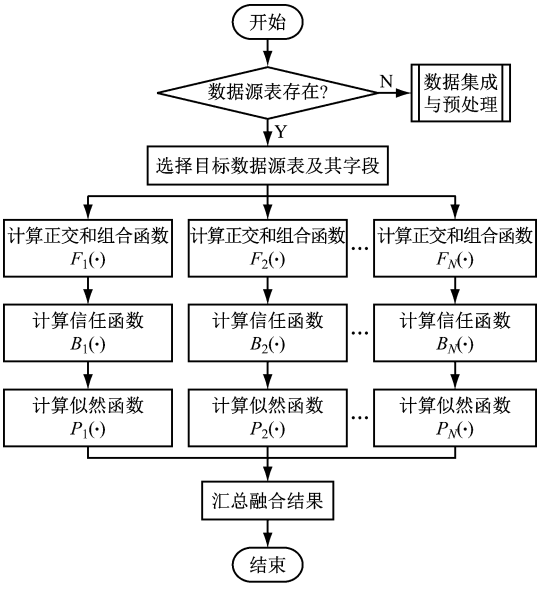

图 4　数据融合流程

Fig. 4　Data fusion process

(2) 主服务器(Master)根据数据表字段与节点个数,调用 splite()方法将数据集平均分割给各从服务器(Slave)进行融合计算。

(3) 计算基本概率分配函数。设总样本空间为 Ω,基本概率分配函数 $f(\cdot)$满足 2 个约束条件:

$$\begin{cases} f(\varnothing)=0 \\ \sum_{A\subseteq\Omega} f(A)=1 \end{cases} \tag{2}$$

式中:$f(\varnothing)$为空命题$\varnothing$为真的概率;A 为 2^{Ω} 中的 1 个或多个命题;$f(A)$为命题 A 为真的概率。

(4) 计算正交和组合函数 $F(\cdot)$。设 $f_1(\cdot)$,$f_2(\cdot)$为 2 个不同的基本概率分配函数,正交和组合函数 $F=f_1(\cdot)\oplus f_2(\cdot)$符合 $F(\varnothing)=0$,则

$$F(\Omega)=k^{-1}\sum_{A\cap C=\Omega} f_1(A)f_2(C) \tag{3}$$

$$k=1-\sum_{A\cap C=\varnothing} f_1(A)f_2(C)=\sum_{A\cap C\neq\Omega} f_1(A)f_2(C) \tag{4}$$

式中:C 为 2^{Ω} 中的命题;k 为证据冲突因子,当 $k=1$ 时,表示命题 A,C 完全冲突;$0<k<1$ 时,表示可对命题 A,C 进行数据融合。

(5) 计算信任函数 $B(\cdot)$。$B(A)$表示对命题 A 为真的信任度。

$$B(A)=\sum_{B\subseteq A} f(B) \tag{5}$$

(6) 计算似然函数 $P(\cdot)$。$P(A)$表示对命题 A 为非假的信任度。

$$P(A)=1-B(\neg A) \tag{6}$$

式中$\neg A=\Omega-A$。

(7) 主服务器重写 reduce()方法,并将从服务器的融合结果汇总并写入 HDFS 中。

2.4　多源异构数据可视化技术

Highcharts 是基于 Web 应用程序、用 JavaScript 语言编写的可视化交互图表库。首先通过初始化函数 Highcharts.chart()创建图表,然后通过构造函数绑定图表对应的全局样式、绘图区、图表事件等配置项,最后通过 JavaScript 读取纯文本数据文件或通过 Ajax 请求数据接口读取与处理数据。利用 Highcharts 数据可视化组件的 jquery.min.js,highcharts.js 等静态资源库,将厂区地图、视频监控、工艺流程、生产统计报表等数据通过丰富的可视化图表(包括仪表盘、速度仪、散点图、漏斗图等)展现给管理决策层,便于其观看、分析与决策。

3　选煤数据中台应用

该数据中台已在山西焦煤集团有限责任公司洗选加工部、二级子公司和分属选煤厂试运行,实现了数据定义与系统集成、数据处理、数据融合与数据可视化展示 4 大数据管理功能。选煤中台主要功能模块如图 5 所示。

图 5　选煤数据中台功能模块

Fig. 5　Function module of coal preparation data center platform

(1) 数据定义与系统集成管理。该管理模块构建了 MDM 和 ESB 2 大系统。MDM 系统通过汇总各业务部门需用的主数据及相应的属性字段,编制了包括单位类、产品类和通用基础类等主数据目录,便于提供集成的主数据字段映射关系。ESB 系统通过拓展 Rest 服务接口,成功将已上线的山西焦煤集团有限责任公司主数据管理系统、统一认证管理平台、协同 OA 系统、煤炭销售管理系统等集成,由 ESB 系统统一进行服务监控和消息转发,满足了集团异构系统互通互联的集成需求。

(2) 数据处理管理。提供了归一化、相关系数矩阵和噪声异常点检测程序,实现了计划(数/质量计划、生产计划、采购计划等)、生产(工艺流程、生产指标等)、能源(能源消耗分析、选煤节能对标等)、安全(双预控、瓦斯、消防等)、环保("三废"管理等)等业务数据的清洗与处理。

(3) 数据融合管理。通过多源异构数据融合子系统,选煤厂成功实现了多故障类型下压滤机等设备的故障原因分析,避免了只通过单一机理和传感器数据无法准确判断复杂工况下设备故障的问题。压滤机是最常用的选煤厂煤泥脱水关键设备,在压滤生产系统中具有重要作用。压滤机发生故障时,受复杂运行环境的影响,如仅通过单一传感器监测数据对常见故障进行原因分析,必然忽视了多种故障因素共同作用的影响。利用多源异构数据融合子系统定义样本空间 h_1(压滤机拉钩传动链轴承温度过高)、h_2(压滤机拉钩传动链轴承异响)2 种异常状态。将 2 种状态下轴承处温度、压力数据集通过噪声异常点检测程序过滤掉异常的"脏数据"后,进行归一化处理,并根据专家知识库给出样本空间的基本概率分配函数,见表 3。

表 3 样本空间基本概率分配函数

Table 3 Basic probability assignment functions of sample space

传感器	基本概率分配函数			
	$f(\varnothing)$	$f(h_1)$	$f(h_2)$	$f(h_1,h_2)$
温度传感器	0	0.64	0.02	0.34
压力传感器	0	0.20	0.09	0.71

将温度数据集与压力数据集对应的基本概率分配函数进行融合,由式(3)、式(4)得出 $k=0.946\ 4$,$F(h_1)=0.689$,$F(h_2)=0.049$。$F(h_1)$最大,表明压滤机拉钩传动链轴承温度过高的可能性大。由式(5)得出 $B(h_1)=0.689$,$B(h_2)=0.049$,由式(6)算出 $P(h_1)=0.951$,$P(h_2)=0.311$,$B(h_1)>B(h_2)$,$P(h_1)>P(h_2)$,根据当前样本集推断出压滤机处于拉钩传动链轴承温度过高状态的可能性最大。

(4) 数据可视化管理。通过 Highcharts 数据可视化组件,将选煤关键指标、运营情况等数据通过速度仪表盘、雷达球等形象化图表展示给管理决策层。如图 6 所示,可视化界面左侧实时滚动刷新各选煤厂月度技术指标、四耗指标和煤炭主营业务等关键数据;中间显示生产车间的视频监控画面和调度室的选煤工艺流程图;右侧显示月度生产指标、产量与完成率等数据。

4 结论

(1) 基于选煤企业数据管理的现状、存在的问题与需求,运用 Hadoop 生态圈大数据技术,提出了基于Hadoop生态圈的选煤数据中台总体架构,研究了数据定义与集成、数据处理、数据融合与数据可视化展示等关键技术。

图 6 Highcharts 数据可视化

Fig. 6 Highcharts data visualization

(2) 应用结果表明,基于 Hadoop 生态圈的选煤数据中台可实现主数据定义标准与系统集成接口规范化,提高了选煤数据处理能力,实现了多源异构选煤数据的融合共享、数据实时交互式的可视化展示。

参考文献(References):

[1] 刘银志,赵廷钊,毛善君,等. 大型煤矿集团管理智能化系统研究[J]. 工矿自动化,2020,46(12):25-30.
LIU Yinzhi, ZHAO Tingzhao, MAO Shanjun, et al. Research on intelligent management system of large coal group[J]. Industry and Mine Automation, 2020, 46(12):25-30.

[2] 董永胜,陈为高,侯佃平,等. 智能化选煤厂研究与建议[J]. 工矿自动化,2021,47(增刊 1):26-31.
DONG Yongsheng, CHEN Weigao, HOU Dianping, et al. Research and suggestions on intelligent coal preparation plant[J]. Industry and Mine Automation, 2021, 47(S1):26-31.

[3] 王国法,王虹,任怀伟,等. 智慧煤矿 2025 情景目标和发展路径[J]. 煤炭学报,2018,43(2):295-305.
WANG Guofa, WANG Hong, REN Huaiwei, et al. 2025 scenarios and development path of intelligent coal mine[J]. Journal of China Coal Society, 2018, 43(2):295-305.

[4] 疏礼春. 智能煤矿数据中台架构及关键技术研究[J]. 工矿自动化,2021,47(6):40-44.
SHU Lichun. Research on the architecture and key technologies of intelligent coal mine data middle platform[J]. Industry and Mine Automation, 2021, 47(6):40-44.

[5] 马小平,代伟. 大数据技术在煤炭工业中的研究现状与应用展望[J]. 工矿自动化,2018,44(1):50-54.
MA Xiaoping, DAI Wei. Research status and application prospect of big data technology in coal industry[J]. Industry and Mine Automation, 2018, 44(1):50-54.

[6] 匡亚莉.智能化选煤厂建设的内涵与框架[J].选煤技术,2018(1):85-91.
KUANG Yali. The intension and framework for the construction of intelligent coal preparation plant[J]. Coal Preparation Technology,2018(1):85-91.

[7] 张磊.信息自动化技术在选煤厂的应用探讨[J].煤炭工程,2018,50(增刊 1):125-127.
ZHANG Lei. Discussion on the application of information automation technology in coal preparation plant[J]. Coal Engineering,2018,50(S1):125-127.

[8] 高晓茜,杨永胜,白龙,等.石圪台选煤厂信息化调度管理[J].洁净煤技术,2019,25(增刊 1):161-163.
GAO Xiaoqian,YANG Yongsheng,BAI Long,et al. Information dispatching management of Shigetai coal preparation plant[J]. Clean Coal Technology,2019,25(S1):161-163.

[9] 孙小路,周春侠,张永志,等.选煤生产过程标准数据平台建设及其关键技术[J].煤炭工程,2021,53(1):38-42.
SUN Xiaolu, ZHOU Chunxia, ZHANG Yongzhi, et al. Construction of standard data platform for coal preparation process and its key technologies[J]. Coal Engineering,2021,53(1):38-42.

[10] 刘贤康.基于 Hadoop 生态圈的工业数据平台设计与研究[D].武汉:华中科技大学,2019.
LIU Xiankang. Design and research of industrial data platform based on Hadoop ecosphere[D]. Wuhan: Huazhong University of Science and Technology, 2019.

[11] 杜小勇,卢卫,张峰.大数据管理系统的历史、现状与未来[J].软件学报,2019,30(1):127-141.
DU Xiaoyong, LU Wei, ZHANG Feng. History, present, and future of big data management systems [J]. Journal of Software,2019,30(1):127-141.

[12] 曹现刚,罗璇,张鑫媛,等.煤矿机电设备运行状态大数据管理平台设计[J].煤炭工程,2020,52(2):22-26.
CAO Xiangang,LUO Xuan,ZHANG Xinyuan,et al. Design of big data management platform for operation status of coal mine electromechanical equipment[J]. Coal Engineering,2020,52(2):22-26.

[13] 王万丽,孙超,宿国瑞.基于云平台的煤矿安全智能管控信息平台设计[J].煤炭工程,2019,51(6):52-56.
WANG Wanli,SUN Chao,SU Guorui. Design of coal mine safety intelligent control information platform based on cloud platform[J]. Coal Engineering,2019,51(6):52-56.

[14] 刘纪芹,史开泉.大数据分解-融合及其智能获取[J].计算机科学,2020,47(6):66-73.
LIU Jiqin, SHI Kaiquan. Big data decomposition-fusion and its intelligent acquisition[J]. Computer Science,2020,47(6):66-73.

[15] 陈金环,王涛,李帅,等.基于改进 D-S 证据理论的智能安全传感器[J].信息技术与信息化,2021(2):232-235.
CHEN Jinhuan, WANG Tao, LI Shuai, et al. Intelligent security sensor based on improved D-S evidence theory[J]. Information Technology and Information Technology,2021(2):232-235.