

大数据技术在 CTC 系统运行中的应用研究

王文凭

中国铁路乌鲁木齐局集团有限公司乌鲁木齐电务段, 新疆 乌鲁木齐 830011

[摘要]由于高速铁路网络不断扩展, 调度集中系统运行数据急剧增加。利用大数据技术高效地采集、处理、应用这些数据, 就成了提高 CTC 系统运维和调度指挥效率的关键。文章对 CTC 系统数据特点进行了系统的分析, 根据目前的应用情况, 分析了数据标准不统一、融合难、质量差、存储压力大、安全风险高等问题, 从建立统一标准和融合平台、完善数据质量管理、提高处理能力和推进决策转化、加强安全防护四个方面提出优化措施, 以期给大数据技术在 CTC 系统中深度应用提供借鉴。

[关键词]CTC 系统; 大数据技术; 数据融合; 数据质量; 数据安全

DOI: 10.33142/sca.v9i7.20031

中图分类号: U284.5

文献标识码: A

Research on the Application of Big Data Technology in CTC System Operation

WANG Wenping

Urumqi Electric Power Depot of China Railway Urumqi Group Co., Ltd., Urumqi, Xinjiang, 830011, China

Abstract: Due to the continuous expansion of the high-speed railway network, the operation data of the centralized dispatch system has increased sharply. Efficiently collecting, processing, and applying this data using big data technology has become the key to improving the efficiency of CTC system operation and scheduling command. The article systematically analyzes the characteristics of CTC system data, and based on the current application situation, analyzes the problems of inconsistent data standards, difficult fusion, poor quality, high storage pressure, and high security risks. Optimization measures are proposed from four aspects: establishing a unified standard and fusion platform, improving data quality management, enhancing processing capabilities and promoting decision transformation, and strengthening security protection, in order to provide reference for the deep application of big data technology in CTC systems.

Keywords: CTC system; big data technology; data fusion; data quality; data safety

引言

调度集中系统属于铁路调度指挥的关键技术装备, 担负着列车运行追踪、站场显现、运行安排管控以及进路控制的任务, 同行车安全和运行速度有着直接联系。近几年来 CTC 系统在路网覆盖率、自动化程度、数据集成方面都有了很大的提高。但是随着系统规模变大、设备种类繁多, 运维数据呈指数级增长, 传统的以人工经验为主的运维数据处理方式已经不能满足智能化发展的需要。CTC 系统要处理上千个轨道电路、信号机和列车实时状态数据, 日均数据量可以达到几十亿条。怎样依靠大数据技术来达成海量数据的高效采集、融合处理以及深入挖掘, 成了当下 CTC 系统智能化创建亟待解决的重要难题。因此对大数据技术在 CTC 系统应用中存在的问题进行系统分析, 并提出相应的优化对策, 有重大的理论研究价值和工程实践意义。

1 CTC 系统及其数据特征分析

CTC 系统是以铁路总公司、铁路局、车站为三级结构, 有条不紊、信息畅通的指挥体系。铁路总公司层面主要对全国铁路运输进行宏观调控和计划安排, 其主要的的数据是长期战略规划以及跨局协调的指示。铁路局调度中心属于中层指挥节点, 有管内列车运行计划编制、调整、实时监控任务。车站子系统属于基层执行单元, 用自律机来完成进路选排、冲突检测、控制输出等主要任务。CTC 系统数据流向有计划数据下行通道、控制命令下行通道、状态数据上行通道这三种主要的通道。状态数据上行通道所采集的列车动态、信号设备状态、进路执行结果等数据用冗余网络的形式, 在 4s 之内完成上行传输。从数据类型上讲, CTC 系统运行时所产生数据包含设备状态数据、运行图数据、报警日志数据、维护工单数据、轨道电路状态数据、信号机状态数据等种类, 来源不同、格式各异^[1]。

铁路信号设备每天产生的监测数据量达到数百万条,数据量大,数据来源是 CTC、道岔状态、轨道电路状态等各个异构数据源。就数据采集而言,各个系统、各个厂家的设备所采用的采集方式及频率存在差异,一些重要的指标每隔三秒钟就会更新一次,这就意味着它的实时性需求很高。数据存储要保证 7*24h 不间断运行,对数据存储的可靠性和可用性有很高的要求。运维过程中积累的历史数据量年年增大,给大数据技术应用提供了丰富的数据支撑,但是也给数据管理、分析能力提出了严峻的考验。

2 大数据技术在 CTC 系统应用中存在的问题

2.1 数据采集标准不统一与融合难度大

目前 CTC 系统中各类设备、各种厂家的产品数据采集方式、格式定义、通信协议都不相同,没有统一的数据采集标准体系。铁路线路建设过程中各个标段由不同的企业施工,完工后所使用的监测设备、数据格式各不相同,每个标段只能建立自己的监测平台,不能共享到统一的中控系统中。由于信号设备多样、接口种类多、协议兼容性差,造成信号监测系统不能完全、及时地获得现场数据。由于各个子系统、各个专业之间数据标准的不统一,造成各个子系统、各个专业之间数据不能在同一个平台上汇聚、共享,形成严重的数据孤岛。另外,不同类型的设备各种指标的采集方式也各不相同,前端采用普通请求或者 Ajax 请求同后端交互,兼容性以及协同性都比较差。数据量大、数据源异构、数据冗余高、冲突证据处理慢等问题都严重影响到大数据技术在 CTC 系统中发挥它的作用。

2.2 数据质量与准确性不足

数据质量是决定大数据分析结果是否可靠的重要因素。但是由于现场电磁干扰、设备自身稳定性、采集链路质量等因素的影响,CTC 系统采集到的数据中存在数据缺项、数据错误、无效数据较多等问题。由于现场干扰、系统自身不稳定等原因造成监测数据缺失或者无效数据较多,数据质量差。在铁路运行现场,网络状况复杂,电磁干扰严重,关键数据采集系统链路质量不稳定,容易造成数据采集、传输不准确,数据丢包时有发生。除此之外,信号集中监测系统可以实时显示设备运行状态以及报警

信息,但是数据校验、清洗和治理缺少系统的手段^[2]。目前大部分故障诊断还使用单一设备、单一参数阈值的检测方式,准确度不高,并且没有考虑到故障的影响范围以及信息之间的联系,容易造成故障节点的误判。数据质量不好不但会降低故障诊断的准确率,还会妨碍大数据技术在预测预警、智能决策等高级应用领域的发展。

2.3 数据存储处理压力与成果转化效率低

由于监测内容不断增多、传感器的功能越来越强、数据存储的时间越来越长,CTC 系统的监测数据量呈指数级增长。由于监测内容不断增多、传感器功能越来越完善、监测数据的存储周期越来越长,监测数据量呈指数级增长,但是目前还没有相应的高效压缩和并行化处理技术。传统的以关系型数据库、单机处理为主的模式已经不能满足海量时序数据高速写入、实时查询的要求。该铁路局 CTC 系统日均亿级数据处理量和 7×24h 连续运行保障,但是目前缺少高效的并行化处理方式,数据压缩效率低、查询响应慢等问题越来越突出。与此同时,数据分析成果转化为实际运维决策的转化速度也迫在眉睫。虽然各个运维单位已经开始了数据驱动的管理模式,但是仍然以事后统计、简单报表为主,缺少了面向预测预警和智能诊断的深入分析能力。大量的检测监测数据、维修数据、试验数据没有得到充分利用,数据分析结果不能有效地转化为指导设备检修、优化资源配置、辅助调度决策的实际应用,造成了数据丰富、信息贫乏的局面。

2.4 数据安全与隐私保护风险突出

大数据技术被广泛应用以后,在提高数据处理能力的同时,也产生了巨大的数据安全问题。轨道交通领域的数据要保护的有乘客个人信息(身份信息、支付记录、出行轨迹),运营核心数据(列车调度指令、信号系统参数),设备运行状态(供电系统、轨道监测数据)^[3]。一旦泄露或者被恶意篡改,就会给铁路运输安全和公众隐私造成极大的危害。但是 CTC 系统及其相关联的监测系统在网络隔离、访问控制、数据加密等各方面的防护能力还存在不足。部分系统建成之后不再经常对操作系统版本进行升级以及安全补丁的更新,杀毒软件的病毒库、入侵检测的特

表 1 CTC 系统主要数据类型及特征

数据类别	数据来源	数据格式	数据体量	采集周期
设备状态数据	服务器、工作站、自律机、网络设备	结构化日志	极大	实时 (≤4s)
运行图数据	调度指挥系统、数据库	结构化/半结构化	较大	实时更新/分钟级
报警日志	综合维护系统、监测系统	结构化文本	大	事件触发
维护工单	运维管理系统	结构化/半结构化	中	作业触发/人工录入
轨道电路数据	轨道电路监测模块	结构化时序数据	极大	实时
信号机状态	信号机监测系统	结构化数据	大	实时

征库不能实现即时更新,从而使得系统存在漏洞以及病毒等安全隐患。随着 CTC 系统同外界的交互越来越多,攻击面也越来越大,由公共信道渗透到铁路信号系统的情况也越来越严重。目前的防护手段对于大规模、有组织的网络攻击来说仍然比较薄弱,不能满足日益增长的数据安全保护的要求。

3 大数据技术在 CTC 系统应用中的优化对策

3.1 建立统一数据标准与多源融合平台

针对数据采集标准不统一、多源异构数据融合困难等问题,要加快 CTC 系统数据标准体系的建设,制定出包含数据采集、格式定义、接口协议、传输流程等各个方面的行业统一标准。就标准的制定而言,应该把各个厂家、各个专业的数据规范加以整合,参照铁路信号集中监测系统等行业标准,创建统一的数据格式和传输途径。平台建设上要创建起统一的多源数据融合平台,整合分布式消息队列,RESTful 接口这些大数据汇集技术,塑造通信,信号专业数据的标准化采集共享流程,达成多源异构数据在统一平台上的高效汇聚和协同处理。设计出数据采集、共享一体化总线架构形式,实现通信、信号专业数据的标准化采集共享流程,根据多源数据应用层次上不同的实时性要求灵活接入、高效共享数据。以云边协同技术为依托,按照铁路总公司、铁路局、车站三级结构来设计和部署,把云中心部分功能下沉到靠近数据源的边缘侧,实现本地数据预处理和故障诊断。

3.2 完善数据质量管理机制

为了保证数据质量以及准确性,应该创建起包含数据全生命周期的管理机制。从数据采集角度来讲,加强前端感知设备的抗干扰性、采集稳定性,用冗余采集、校验的方式减少数据丢失、差错的情况,从而保证原始数据的真实性、可靠性、完整性。在数据处理环节,对于轨道交通信号系统监测数据量大、数据类型繁杂、压缩效率低等状况,可以参照基于分布式并行压缩策略以及动态网格局部离群概率的离群数据清洗和混合修复去噪技术方案,创建包含压缩、解压、清洗、修复等各个环节的信号监测数据快速精化处理方法体系。使用流式数据处理的方式来完成各个数据源的数据清洗、数据验证工作,排除异常值、重复记录等不正常的处理数据,并保证数据的质量。另外根据不同的设备数据特点来设计不同的清洗规则和质量控制措施,以达到更好的治理效果。另外还要建立数据质量评价指标体系以及定期核查制度,从完整性、准确性、一致性、时效性四个方面对数据实施持续的监控,定期出具数据质量报告,找出存在的问题并加以追溯,及时改正。

经过以上数据质量管理机制的系统执行,可以给之后的大数据分析挖掘、智能决策提供高质量、高可信的数据支撑,进而充分发挥出大数据技术在 CTC 系统运维方面的实际作用。

3.3 提升数据处理能力与推进决策转化

就数据存储处理的压力大、转化的效率低而言,应该加强大数据处理能力以及智能化分析水平的提高。从数据存储的角度来说,可以使用分布式数据库以及时序数据库相融合的混合存储结构,从而达到对大量时序数据进行高效写入、压缩保存并快速查询的目的。建立包含压缩、解压、清洗、修复等全过程的信号监测数据快速精化处理方法体系,利用所建立的车地通信测试平台完成了轨道交通信息传输系统安全性描述、建模、证明和性能验证,为快速、准确的故障诊断和预警打下了基础。从数据分析角度来讲,采用机器学习、深度学习等智能算法创建依靠多源数据融合的设备状态评价模型以及故障预估模型,达成由被动应对转向主动预防的目的。采集并融合 CTC 系统多源异构数据,从中提取出体现系统故障的特征以及隐藏的信息,用控制图算法和隐马尔可夫算法对系统的状态进行实时监测与评价,从而达到对系统微小异常的检测和潜在故障的预警的目的。利用智能分析平台以及数据可视化工具把数据分析结果转变为可以实施的运维建议和决策支持信息,促使运维工作由原来的依靠经验驱动向数据驱动转变,充分发挥大数据的价值。

3.4 强化数据安全防护体系

就数据安全和隐私保护风险而言,要创建包含采集、传输、存储、处理、共享等全生命周期的数据安全防护体系。网络隔离方面严格依照 CTC 系统独立组网、封闭运行的安全规定,在生产网络和查询系统、外部系统之间设置网络隔离设备,达到有效安全隔离的目的^[4]。从访问控制角度出发,用角色化的细粒度权限管理来对数据访问的行为进行全程跟踪与审计,保证数据流经过过程里全部链条都是可以追溯的。数据加密主要是采用加密存储、加密传输的方式将敏感数据保存起来或者传递出去,从而防止数据在传输和存储过程中被截取、篡改。从安全管理角度出发,创建数据分类分级管理机制,确定各个安全级别数据的保护标准及管理职责。同时积极采用主动防御技术,改变以前只依靠防火墙、入侵检测等传统手段的被动防护方式,创建起安全管理中心支撑的安全计算环境、安全区域边界以及安全通信网络三重防护体系,从而达成对全网安全策略的集中管控并加以动态回应。

4 结语

大数据技术给铁路 CTC 系统智能化升级以及运维效

率提高赋予了关键的技术支撑。本文对 CTC 系统数据特征进行系统的分析之后,从采集标准不统一、多源异构数据融合难、数据质量与准确性差、存储处理压力大、数据安全风险高等五个方面入手,提出建立统一数据标准和多源融合平台、完善数据质量管理机制、提高数据处理能力和加强数据安全防护体系等有针对性的优化对策。未来由于大数据、人工智能、云边协同这些新的技术不断融合并深入到系统各个方面之中,CTC 系统逐步进入了数据收集、智能化分析的大数据化阶段,从而给高速铁路安全高效运行和智慧铁路建设提供更加有力的技术支持。

[参考文献]

[1]靳俊,费振豪,冯振国.调度集中(CTC)系统安全防护体

系与技术[J].中国铁路,2023,1(10):68-75.

[2]程忆佳,周李昭轩.站界口 CTC 车次跟踪异常问题研究[J].铁道通信信号,2026,62(1):101-105.

[3]韩玥祺.基于国铁 CTC 的市域铁路调度指挥系统公交化功能提升研究[J].铁路通信信号工程技术,2025,22(11):104-108.

[4]施莉娟,司继盛,费振豪,等.基于 CTC 数据的高铁区间可达最大通过能力计算模型[J].铁道通信信号,2025,61(10):31-38.

作者简介:王文凭(1987—),毕业于新疆大学科学技术学院计算机科学与技术专业,研究方向为 CTC 系统运用维护。