

# Research and Practice of Rail Wear Information System Based on Big Data Technology

Chunyan Yang

Detection and Rescue Center of Shuozhou-Huanghua Railway, Cangzhou, Hebei, 062350, China

## Abstract

With the development of intelligent railway construction, a large number of railway professional data information is generated every day. In view of the continuous growth of massive data, how to meet the ability of extracting data in the massive railway professional data, mining valuable information through big data analysis, so as to apply the results based on data and analysis to decision-making behavior, rather than based on experience and intuition, to the operation and maintenance of equipment in enterprise has very high value. Shuozhou-Huangshan railway comprehensive inspection vehicle rail wear inspection system has been put into operation for two years, accumulating a large number of rail wear data, and has established a rail wear management information system with excel as the basic template. How to use the inspection data to systematically analyze the curve wear of Shuozhou-Huangshan Railway, and guide the on-site safe, economic and scientific rail 'state' through the analysis results is the research direction of rail wear information system based on big data technology. It makes possible for big data technology to support the practice of rail wear information system to analyze the influencing factors of rail wear, study the development rate of curve wear, predict the development trend of wear, and assist the realization of rail life management.

## Keywords

big data; railway wear; data analysis

## 基于大数据技术的钢轨磨耗信息系统研究与实践

杨春艳

朔黄铁路线路检测和救援中心, 中国·河北 沧州 062350

## 摘要

随着智能化铁路建设发展, 每天产生大量的铁路各专业数据信息, 针对海量数据持续增长, 如何在海量的铁路专业数据中满足提取数据的能力, 通过大数据分析挖掘出有价值的信息, 从而将基于数据和分析的结果应用于决策行为, 而非基于经验和直觉, 对设备的运维、企业的管理都有非常高的价值。中国朔黄铁路综合检测车钢轨磨耗检测系统上线应用两年, 积累了大量钢轨磨耗数据, 已建立起以 Excel 表格为基本模板的钢轨磨耗管理信息体系。如何利用检测数据对朔黄铁路曲线磨耗情况进行系统性分析, 并通过分析结果指导现场安全、经济、科学的钢轨“状态修”, 实现钢轨的寿命管理, 是基于大数据技术的钢轨磨耗信息系统研究的方向。它使得大数据技术支持钢轨磨耗信息系统实践分析钢轨磨耗影响因素, 研究曲线磨耗发展速率, 预测磨耗发展趋势, 辅助实现钢轨寿命管理成为可能。

## 关键词

大数据; 钢轨磨耗; 数据分析

## 1 引言

随着铁路信息化步伐的加快, 铁路运营系统的数据已具有海量性、多源性、数据类型多、处理要求快等特点。传统的铁路数据分析仍然采用人力以 Excel 表格为基本模板的定时定量分析钢轨磨耗指标, 并按月记录原平分公司管内钢轨侧磨发展情况。但随着数据量的逐月增加, Excel 表格功能已经远远无法满足对海量检测数据的梳理、统计、分析等需求, 也无法满足对磨耗数据更为深入的研究需要。分析钢轨磨耗

影响因素, 研究曲线磨耗发展速率, 预测磨耗发展趋势, 辅助实现钢轨寿命管理等研究过程均受到了技术瓶颈制约。同时经考察全球尚不存在针对磨耗检测数据开发的数据分析处理系统。因此, 如何便利大数据分析, 深度挖掘磨耗数据价值成为钢轨磨耗研究的关键。

## 2 钢轨磨耗数据分析的现状与面临的问题

中国朔黄铁路原平分公司管内线路属于典型山区铁路, 桥隧毗连、填挖相接、坡度大、曲线多、半径小是其主要特

点。经统计,原平分公司上行重载线路有曲线179条,共计94.182km,占原平分公司上行线路总长的36.8%; $R \leq 600m$ 以下的曲线75条,共计43.112km,占分公司上行曲线总长的45.8%; $600 < R \leq 800$ 的曲线35条,共计21.309km,占分公司上行曲线总长的22.6%。由于线路条件复杂,加之面临逐年增加的运量和更大轴重车辆的通行,小半径曲线钢轨磨耗发展呈现出加速趋势。这不仅缩短了钢轨使用寿命,增加了养护维修的工作量和费用,也给行车安全带来极大隐患。因此,对朔黄重载铁路曲线磨耗数据进行系统性管理和分析面临以下六类问题<sup>[1]</sup>。

#### (1) 检测数据量大

朔黄铁路综合检测车钢轨磨耗检测系统上线应用两年,积累了大量钢轨磨耗数据,综合检测车磨耗检测数据的采集规则为一米四个采集点,每月的原平上行179个曲线区段钢轨磨耗数据可达近百万条。

#### (2) 曲线磨耗计算量大

为了考察每条曲线是否达到规定的磨耗伤损级别,曲线钢轨磨耗值成为重要指标。其计算方式是对每条曲线钢轨磨耗最明显区段里程段内的磨耗值进行算数平均得到。但由于涉及的文件较多,计算过程会面临与数据分割同样的误差问题。

#### (3) 通过总重与换轨周期比较量大

每条曲线钢轨在上线后其磨耗值会随着累计通过总重的变化而变化,当磨耗伤损程度达到规定值后须下线。为了考察磨耗发展情况与自上线后钢轨上所通过的重量总和之间的关系,累计通过总重成为一项重要考察指标。其计算方式是在同一次换轨周期内,检测当日的调度日报统计通过总重数据与上线当日的调度日报统计通过总重数据差值。但由于每一条曲线的换轨时间起始点不同,需要根据不同钢轨换轨日期的时间进行通过总重数据抓取,且曲线钢轨的数目较多,人工计算统计效率低且可能出现误差的可能性较大。

#### (4) 磨耗数据与关联数据量大

如果仅把数据结构化的进行储存管理还无法充分体现出数据的价值,必须将数据库中的数据进行关联查询才能对数据间隐含的信息加以利用。要研究曲线磨耗发展与累计通过总重之间的关系就必须对曲线磨耗、通过总重和台账信息进行关联查询。

#### (5) 磨耗数据存储量大

随着综合检测车每月检测后的磨耗数据不断增加,对数据的安全性管理也随之增强。本项目计划利用数据库技术解决对海量数据的管理。数据库是依靠结构性查询语言SQL对各种数据进行导入,由于涉及的数据相当大,同时涉及的表也比较多,人工逐条导入数据很繁琐也很耗时。

#### (6) 磨耗数据展示量大

仅按照里程情况对每条曲线的磨耗数据进行罗列无法直观对磨耗伤损级别进行详细考察,进而降低了对伤损曲线级别的迅速判断。同时,如果利用Excel表格作图功能进行逐月磨耗考察,并不能灵活跨月份对磨耗发展情况进行观察。磨耗检测数据的可视化将便利以上需求。

如何利用检测数据对朔黄铁路曲线磨耗情况进行系统性分析,借助计算机来高效解决大数据量分析及入库工作,并通过分析结果指导现场安全、经济、科学的钢轨“状态修”,实现钢轨的寿命管理,大数据分析技术的应用成为一种选择<sup>[2]</sup>。

## 3 基于大数据技术分析的优势和实现

### 3.1 大数据分析: 数据管理, 数据分析, 数据共享

数据是记录信息的载体,是知识的来源。数据的激增意味着记录范围、测量范围和分析范围不断扩大,知识的边界也在不断延伸。如何收集、保存、维护、管理、分析、共享正在指数级增长的数据是我们必须面对的一个挑战。

针对特定的业务过程,处理离散的事务,数据成为一种过程记录。数据的不断积累的结果,不仅仅用于查询,而且应用于分析。怎样从各个独立的信息系统提取、整合有价值的信息,从而实现从数据到信息、从信息到知识、从知识到利润的转化。通过数据库应用,建立一种格式多源数据存储中心,即使通过不同平台采集的数据,可以按照统一的定义格式被提取出来,再通过清洗、转换、集成最后归流入海,加载进数据仓库<sup>[3]</sup>。

同时通过统一接口,为每个系统提供信息共享的机制,使得各种运输数据充分共享,数据之间通过各种方式来匹配和结合,充分发挥和挖掘现有系统和数据的潜力,把业务数据进行采集、整理和统计分析,形成信息管理的基础数据源,并在此之上开发出用户需要的特殊功能,进一步提高朔黄铁路的信息化管理水平。

将专业设备监测数据,建立分析模型,实现设备状态趋势变化可感知、劣化自动预警,自动生成预警维护工单。数据之所以要进行分析处理,关键在于这些数据量巨大,这些数据精确地记载了过去和现在发生的事,对面向未来的决策有着极其重大的意义和价值。在数据中找出关联,发现规律和验证假设,可极大提升决策的准确性和可靠性,使决策真正智慧。智慧决策不仅有助于发现设备运行规律,同时为“状态修”提供数据分析支撑。通过数据分析进行数据的可视化展现,进而将复杂的数据转变为简单的交互,可实现对数据的“全方位”“立体式”呈现。如区别于传统柱状图的玫瑰花图,用于展现二维数据;表现宏观数据趋势的日历图;用于层级数据结构展现的太阳辐射图及树图;表现节点关联关系的弦图等等,这些简捷、生动且具有艺术气息的绚丽图表,不仅让庞杂的数据变成具有特定业务指标并实现不同数据之间的联动分析,并最终为企业提供科学决策<sup>[4]</sup>。

### 3.2 专业应用:专业应用,共享平台,智慧服务

利用钢轨磨耗信息系统软件(如图1所示)管理检测数据,分析影响磨耗的关键因素和磨耗发展规律,建立钢轨磨耗发展预测模型,对钢轨的“状态修”和修理周期提出合理化建议,并最终实现延长钢轨使用寿命的目标。

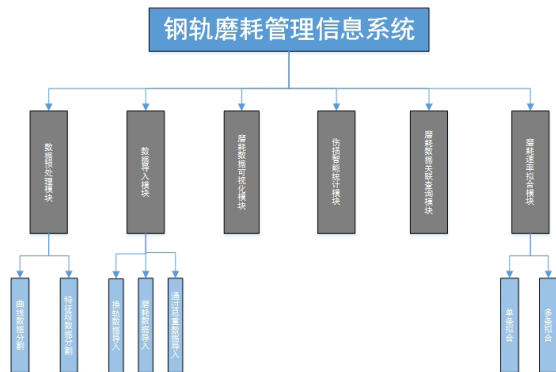


图1 钢轨磨耗管理信息系统功能结构图

#### (1) 数据分割自动化

综合检测车磨耗检测数据的采集规则为一米四个采集点,这代表每月的原平上行磨耗数据可达近百万条,而每次的检测数据都必须按照曲线所在的起始里程值将原始数据进行分割归档<sup>[5]</sup>。依靠人力去提取分析每条曲线的数据值,不仅耗时费力,还有可能导致提取数据时发生人工误差,所以计算方式是对每条曲线钢轨磨耗最明显区段里程段内的磨耗值进行算数平均得到(如图2所示)。

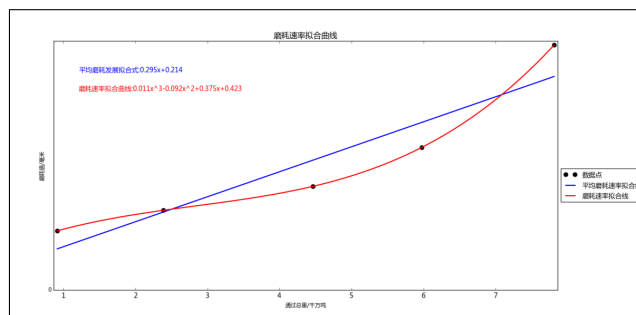


图2 钢轨磨耗速率拟合曲线

#### (2) 曲线钢轨磨耗值计算自动化

将数据提取分割管理是对数据进行分析的第一步。为了考察每条曲线是否达到规定的磨耗伤损级别,曲线钢轨磨耗值成为重要指标。其计算方式是对每条曲线钢轨磨耗最明显区段里程段内的磨耗值进行算数平均得到。但由于涉及的文件较多,计算过程会面临与数据分割同样的误差问题,所以也必须借助计算机来高效解决这部分工作,计算过程如图3所示。

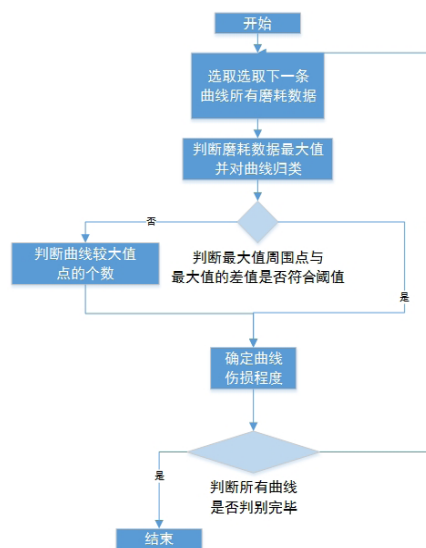


图3 曲线伤损判断流程图

#### (3) 累计通过总重计算自动化

每条曲线钢轨在线后其磨耗值会随着累计通过总重的变化而变化,当磨耗伤损程度达到规定值后须下线。为了考察磨耗发展情况与自上线后钢轨上所通过的重量总和之间的关系,累计通过总重成为一项重要考察指标。其计算方式是在同一次换轨周期内,检测当日的调度日报统计通过总重数据与上线当日的调度日报统计通过总重数据差值。但由于每一条曲线的换轨时间起始点不同,需要根据不同钢轨换轨日期的时间进行通过总重数据抓取,且曲线钢轨的数目较多,人工计算统计效率低且出现误差的可能性较大。

(4) 磨耗检测数据可视化

仅按照里程情况对每条曲线的磨耗数据进行罗列无法直观对磨耗伤损级别进行详细考察,从而降低了对伤损曲线级别的迅速判断。同时,如果利用 Excel 作图功能进行逐月磨耗考察,并不能灵活跨月份对磨耗发展情况进行观察。磨耗检测数据的可视化将便利以上需求 [6]。

(5) 数据导入数据库管理自动化

随着综合检测车每月检测后的磨耗数据不断增加,对数据的安全可靠性管理也随之增强。本项目计划利用数据库技术解决对海量数据的管理。数据库是依靠结构性查询语言 SQL 对各种数据进行导入,由于涉及的数据相当大,同时涉及的表也比较多。人工逐条导入数据很繁琐也很耗时,所以必须将数据导入数据库的过程实现自动化。钢轨磨耗管理信息系统数据库的 E-R 关系图,以及表中各字段的关联关系如图 4 所示。

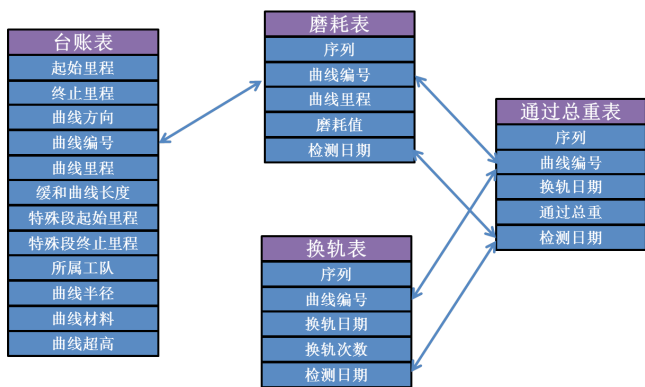


图 4 钢轨磨耗管理信息系统数据库的 E-R 关系图

(6) 曲线磨耗、通过总重和台账信息关联查询

如果仅把数据结构化进行储存管理那还无法充分体现出数据的价值,必须将数据库中的数据进行关联查询才能对数据间隐含的信息加以利用。

(7) 利用查询结果观察磨耗发展趋势

虽然利用数据库可以实现方便的关联查询,但要想考察钢轨磨耗整体的发展趋势就需要对查询后的数据进行二次处理。通过对查询后的数据进行建模分析以及对数据点进行曲线拟合将考察对象量化和直观化,从而进一步深入分析数据,挖掘数据中的价值。

以分割后每月每条曲线的磨耗数据为基础数据库,选择任意一条曲线进行观察,再选取其单月数据、连续或不连续的多月数据进行作图。将磨耗数据以折线图的形式呈现: X 轴为曲线里程(单位:公里 km), Y 轴为磨耗值(单位:毫

米 mm)。同时可以将朔黄曲线不同伤损程度磨耗标准值标注于图中。此功能方便了检测人员直观跟踪曲线磨耗的发展情况,且能第一时间对曲线的磨耗伤损程度进行判断。

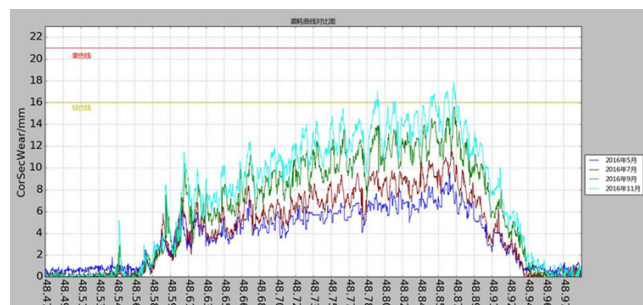


图 5 磨耗数据可视化设计示意图

4 结论与展望

使用本系统统计分析钢轨磨耗数据,可提高统计和分析工作效率,减少人为统计错误,增强信息传输的及时性和准确性,将技术人员从海量数据和繁琐统计等流程中解脱出来。同时软件系统分析的结果准确可靠,对指导安全生产、防止工作失误、指导设备状态修理有重要意义,其间接经济效益更为乐观。

钢轨磨耗管理信息系统软件的应用,通过断面磨耗分析系统,可对侧磨伤损程度进行问题地段筛选,对影响侧磨因素进行侧磨发展速率预判,对小半径曲线侧磨情况进行数据统计。更适用于重载铁路的运营需要,可大大降低各类行车事故的发生概率,提高铁路运输生产的效率。对公司的运输带来很大的经济效益和社会效益。

参考文献

- [1] 崔艳萍,马欣然.欧洲铁路行业安全管理的启示[J].铁道运输与经济.2015(4).59-63.
- [2] 郭春梅,孟庆森,毕学尧.服务器虚拟化技术及安全研究[J].信息网络安全.2011(9).23.
- [3] 宋振华.虚拟化技术中的存储管理问题研究[D].合肥:中国科学技术大学.2010.42-44.
- [4] 杜一磊.PaaS平台后端管理系统的设计与实现[D].北京:北京交通大学.2018.22-24.
- [5] 谢辉.嵌入式数据库同步系统的研究与实现[D].北京:北京交通大学.2011.24-25.
- [6] 吴勇军.一种 Oracle 数据库容灾系统的应用研究[D].青岛:中国石油大学.2010.26.