

文章编号: 1005-8451 (2017) 10-0005-04

铁路分线成本大数据分析平台方案研究

吴艳华¹, 杨连报¹, 孙 晶¹, 白 璐²

(1. 中国铁道科学研究院 电子计算技术研究所, 北京 100081;

2. 北京交通大学 软件学院, 北京 100044)

摘 要: 从分线成本核算开展数据分析的目标入手, 搭建数据分析的总体架构, 描述数据分析与分线成本核算系统间的关系, 引入聚类、分类、回归分析、关联分析等数据挖掘算法, 从铁路总公司、铁路局/公司和特定线路等不同用户角度, 分析铁路线路公益性/经营性业务特征、不同成本类型间分布特征、各类统计工作量的分布规律、收支盈亏情况等, 为合理定义铁路线路业务性质、运输产品优化、经营决策等提供数据和辅助决策支持。

关键词: 铁路; 经营管理; 成本; 大数据分析

中图分类号: U29 : TP39 **文献标识码:** A

Big data analysis platform for railway sub-line cost

WU Yanhua¹, YANG Lianbao¹, SUN Jing¹, BAI Lu²

(1. Institute of Computing Technologies, China Academy of Railway Sciences, Beijing 100081, China;

2. School of Software, Beijing Jiaotong University, Beijing 100044, China)

Abstract: From the goal of sub-line cost accounting to carry out data analysis, this article built the overall structure of data analysis, described the relationship between data analysis and sub-line cost accounting system, introduced the data mining algorithms such as clustering, classification, regression analysis and correlation analysis, analyzed the characteristics of public welfare / business characteristics of railway lines, the distribution characteristics of different cost types, the distribution rules of various kinds of statistical workload, the profit and loss situation of the railway and so on, the railway company / the railway station / the company and the specific line. It was provided data and support decision support for the rational definition of the nature of the railway line business, transport product optimization, business decision-making.

Keywords: railway; management; sub-line cost; big data analysis

在铁路政企分开的新形势下, 铁路企业财务管理及核算工作的要求越来越高、越来越细。2013年7月, 根据《财政部对中国铁路总公司国有资产与财务监督管理暂行办法》中“总公司应建立公益性与经营性运输、线路单独核算机制”的规定, 铁路总公司财务部组织开展了分线成本核算问题研究。随后, 中国铁路总公司制定了《铁路运输企业分线成本核算管理暂行办法》(简称: 办法), 办法要求: 在基层单位将日常发生的成本费用信息核算到不同线路, 能够直接区分发生在具体线路的成本信息直接列支(单位直列(例如: 客专公司), 直接将该单位的成本费用核算到该单位的线路上; 科目直列, 从核算

系统中获取科目直接列支在线路上的核算数据, 生成分线科目直列数据), 无法直接区分的成本根据发生在不同线路上的相关工作量按照分线成本核算规则进行核算。

根据办法的要求, 铁路总公司组织开发分线成本核算系统: (1) 在铁路总公司层面, 统一相关的分线成本核算基础字典、分线成本核算模型同时发布供全路使用, 并与相关业务系统对接, 实现业务生产统计信息及时获取, 形成全路的分线成本核算基础信息、分线成本核算相关的工作量信息、分线成本核算结果信息, 为铁路运输企业分析、决策提供依据; (2) 在铁路基层单位层面, 根据分线成本核算模型以及相关工作量信息, 通过系统快速、准确地将成本费用信息核算至线路, 同时将分线成本核算基础数据与分线结果上传铁路局、铁路总公司,

收稿日期: 2016-05-04

基金项目: 中国铁道科学研究院电子计算技术研究所基金(DZYF16-16)。

作者简介: 吴艳华, 副研究员; 杨连报, 在读博士研究生。

为铁路局、铁路总公司收审、汇总提供数据支持。

分线成本核算系统自上线使用以来,为全路的线路成本核算提供了基本支持。系统已经积累了近3年的历史数据,随着分线成本管理工作的进一步开展,如何利用这些数据为管理决策提供科学的数据支撑,成为分线成本管理工作的需要开展研究的重点方向之一。

目前,国内外各领域针对成本管理、成本分析等纷纷开展数据分析相关研究。

1 成本核算相关研究

1.1 铁路成本核算相关研究

自铁路政企改革伊始,就有学者就会计成本核算、公益性服务、作业成本法等开始研究。邹美凤等^[1]总结铁路企业成本核算现状,指出铁路企业成本核算尤其是公益性业务成本核算存在的问题,提出将作业成本法应用于铁路运输企业公益性业务成本核算,并构建适合铁路运输企业公益性业务成本核算的框架体系;管红舞^[2]就加强铁路运输企业成本管理、建立分线成本核算模型,准确核算铁路各条线路盈亏情况等进行研究;黄雪^[3]梳理现有的成本核算方法、国内外铁路公益性服务会计核算的先进经验,以及各行业的公益性服务会计核算方法,分析有利于铁路公益性服务会计核算的先进经验,构建我国铁路公益性服务的会计核算体系,说明铁路公益性服务会计核算办法。

1.2 其它领域成本核算相关研究

国内在公路交通、医疗等方面,也都有学者关注于成本核算分析相关问题。丁岳维^[4]等通过对我国公路运输企业进行质量成本核算必要性的分析,说明质量成本核算关系到企业运营效益,并对公路运输企业质量成本构成及核算方法进行探讨,提出该类企业质量成本核算应遵循的原则;郑大喜^[5]通过梳理成本核算在医院管理中的应用,总结医院开展成本核算的意义,分析开展成本核算过程中遇到的难点,寻找解决难点的方法,从而更好地发挥成本分析评价和成本控制职能,降低医疗服务成本和医院营运成本。

本文从分线成本核算开展数据分析的目标入手,

搭建数据分析的总体架构,描述数据分析与分线成本核算系统间的关系,引入聚类、分类、回归分析、关联分析等数据挖掘算法,从铁路总公司、铁路局/公司和特定线路等不同用户角度,分析铁路线路公益性/经营性业务特征、不同成本类型间分布特征、各类统计工作量的分布规律、收支盈亏情况等,为合理定义铁路线路业务性质、运输产品优化、经营决策等提供数据和辅助决策支持。

2 总体架构

基于铁路分线成本核算的数据分析工作,通过采集系统中的历史数据和实时增量数据,以及区域经济、运输线路长度等外部数据,经过数据脱敏、治理等工作,面向业务分析应用构建数据模型,形成不同的数据主题、数据域,并通过各类数据分析算法形成数据分析结果,展现或推送给分线成本核算系统。总体架构,如图1所示。

3 分线成本分析应用场景

3.1 分线路主要构成成本分析

铁路运输成本主要包括铁路运输过程中发生的各种耗费,按财务制度的要求可以划分为主营业务成本、管理费用、财务费用和营业外支出;按要素分为工资、材料、燃料、电力、折旧、外购劳务和其他。

不同线路的运输成本由于其承担的运输任务或主要作业的不同,其构成也不同。通过数据分析可以按客、货、行、包找出各线路的主要构成成本;通过线路、部门、资产或时间等多维度对主要成本进行横向、纵向对比分析,挖掘数据背后的规律。

3.2 公益性/经营性线路区分

铁路运输分为经营性运输和公益性运输。经营性运输主要是面向市场、提供服务、取得相应效益而组织的运输行为,是铁路运输企业最基本的收益来源。公益性运输则是以服务公众为目的,高成本低收益或基本没有收益的运输行为。

针对高成本低收益或基本没有收益的线路,通过计算单位成本、单位收入等指标,找出其成本较高的原因。同时也结合所管辖区域线路经过区域的GDP、环境、人口等数据,区分该线路是经营性还

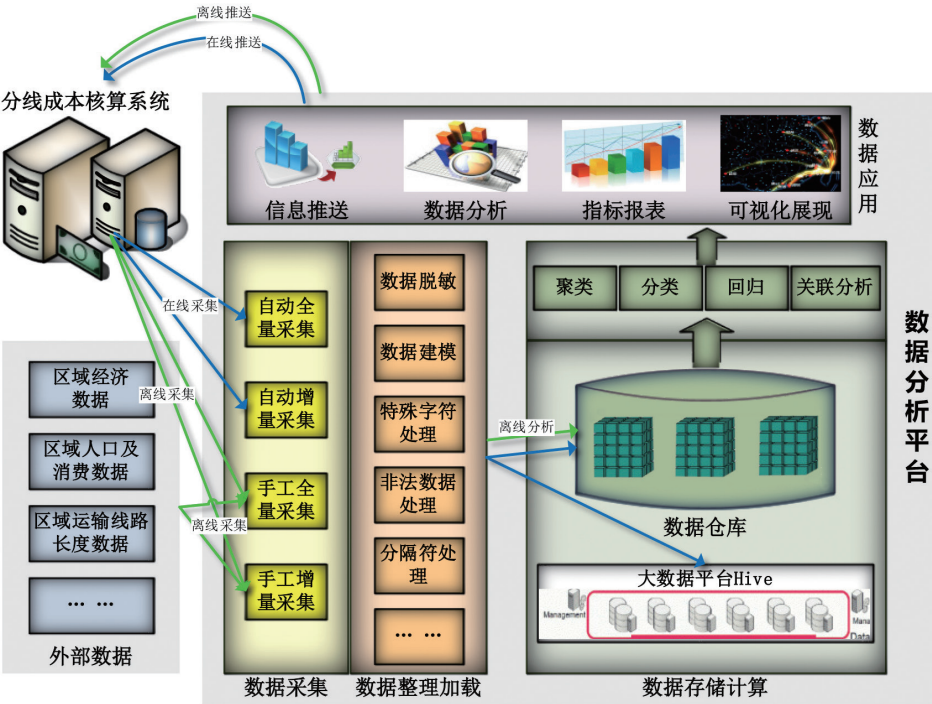


图1 数据分析总体架构图

是公益性线路。

3.3 公益性/经营性线路收支预测

根据公益 / 经营性线路的历史数据，综合成本支出、运营收入、走行区域的环境、人口、经济等因素，构建预测模型，预测未来的成本和收入，为铁路预算提供数据支撑，为传统公益性 / 经营性线路的区分。

3.4 公益性线路成本控制

在满足公益性线路服务大众或完成国家交给的特定任务的前提下，挖掘公益性线路主要构成成本的影响因素，减少资源消耗，实现成本控制。以经营线路的收入最大化为目标，通过估算某条线路的成本和收入，对开行列车编组等进行调整，对所设硬座、硬卧车厢数进行估算，实现经营性线路最大化收益。

4 分线成本分析模型方法

分线成本分析模型主要是应用聚类、分类、回归、基于时间序列的预测、关联分析和序列发现等算法。其中，聚类算法主要是根据线路的收入、支出、工作量等属性信息，进行自动聚类分析，再根据聚类得到的结果，进行标签化。分类算法是属于监督学习，需要先指定分类，然后用来区分归类，主要用在已知分类的基础之上，实现对新的数据分类。回

归分析和基于时间序列的预测主要是基于成本、收入、工作量等历史数据，进行趋势预测。关联和序列分析主要是分析影响成本、收入的因素，分析其关联关系，找出主要因素。算法功能和应用场景，如表 1 所示。

5 外部数据接入

线路分线数据分析不仅需要目前分线系统所有的分线成本、工作量数据。还需要线路客运收入、货运收入等数据，此外，还需要结合分析线路所走行区域的 GDP 数据、人口数据、交通数据等。分线数据分析的数据来源与构成，如表 2 所示。

表 1 分线成本分析模型及应用场景

算法分类	功能	应用场景
分类	根据一些变量的数据进行计算，依照结果将数据分类	公益性/经营性线路区分
聚类	将数据分群，找出群体间的差异和群内成员的相似性	公益性/经营性线路区分
回归	用一些列数据预测一个数值的可能值	公益性/经营性线路收支预测、公益性线路成本控制、经营性线路利益最大化
基于时间序列的预测	与回归类似，不同的是数据带有时间特性	公益性/经营性线路收支预测、公益性线路成本控制、经营性线路利益最大化
关联	找出某一事件或数据中会同时发生的事情或出现的数据	分线线路主要构成成本分析、关联关系分析
序列发现	与关联类似，不同的是序列中事件的相关性以时间因素来做分割	分线线路主要构成成本分析、关联关系分析

6 结束语

基于分线成本核算系统数据、外部相关数据，开展数据分析工作，将为铁路线路性质定义和管理提供有利的决策支持。但是，由于外部数据分散于不同的业务系统，且部署位置也各不相同，因此，外部数据接入是数据分析开展的难点。

基于铁路分线成本核算的数据分析，可以分两步实施：

(下转 P16)

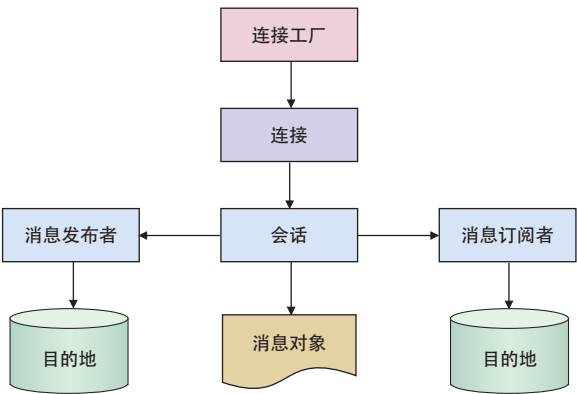


图7 路网播控平台消息总线工厂设计模式

息中间件技术，提出了基于消息总线的路网播控平台与线路 PIS 通信架构，实现了具有高扩展性和松耦合性的路网播控平台系统。基于消息中间件集群技术的路网播控平台与线路 PIS 接口的消息总线已应用于 PCC 系统以及线路 PIS，达到了路网播控平台对消息处理可靠性、安全性、可扩展性和松耦合性的要求，满足了路网播控平台灵活快速的适应不同线路 PIS 接口的数据通信需求。

参考文献：

责任编辑 徐侃春

[1] 阚庭明, 王富章, 魏 奇, 等. 城市轨道交通乘客信息系统集成播控平台的设计研究 [J]. 科技通报, 2013, 29 (7): 208-213.

[2] 于孝安, 孙同庆, 汪晓臣, 等. 城市轨道交通乘客信息系统路网控制中心设计的研究 [J]. 铁路计算机应用, 2013, 22 (5): 55-60.

[3] 杨正洪, 郑齐心, 吴 寒. 企业云计算架构与实施指南 [M]. 北京: 清华大学出版社, 2010.

[4] 曾宪杰. 大型网站系统与 Java 中间件实践 [M]. 北京: 电子工业出版社, 2014.

[5] 中国城市轨道交通协会. 关于规范和促进城市轨道交通传媒经营健康发展的指导意见: 中城轨 [2014]015 号 [Z]. 中国城市轨道交通协会, 2014, 9.

[6] 北京市交通委员会. 北京市轨道交通乘客信息系统应用规范 [S]. 北京: 北京市交通委员会, 2014.

[7] (美) 斯穆特 (Smoot, S. R.). 私有云计算: 整合、虚拟化和面向服务的基础设施 [M]. 潘 怡. 北京: 机械工业出版社, 2013.

[8] 汪晓臣, 于 鑫, 阚庭明, 等. 轨道交通路网编播中心系统建设方案研究 [J]. 铁路计算机应用, 2013, 22 (9): 43-46.

(上接 P7)

表2 分线数据分析数据来源与构成

序号	数据类型	主要信息系统	主管部门
1	公用基础数据	铁路主数据管理平台	信息化处
2	业务数据	机车管理信息系统	机务处
3	业务数据	工务管理信息系统	工务处
4	业务数据	电务管理信息系统	电务处
5	业务数据	通信管理信息系统	电务处
6	业务数据	车辆管理信息系统	车辆处
7	业务数据	供电管理信息系统	供电处
8	业务数据	房屋管理信息系统	土房处
9	业务数据	货票管理信息系统	货运处
10	业务数据	客票管理信息系统	客运处
11	业务数据	铁路物资管理信息系统	物资处
12	业务数据	非运输企业房地产管理系统	多经处
13	业务数据	非运输企业酒店管理系统	多经处
14	统计分析数据	运输收入管理系统	收入处
15	统计分析数据	统计信息系统	计划统计处
16	统计分析数据	资金清算系统	财务处
17	统计分析数据	预算管理系统	财务处
18	外部数据	国家统计局网站	国家统计局

(1) 以铁路总公司的数据分析需求为主，通过离线采集内、外部数据的方式，基于单台硬件设备

支撑数据分析应用；

(2) 基于铁路数据服务平台，逐步搭建起独立于分线成本核算系统之外的数据分析服务功能，通过自动采集和离线手工采集等多种数据采集方式，搭建数据仓库和大数据平台混合的数据分析平台，分析结果自动推送至分线成本核算系统中。

参考文献：

[1] 邹美凤, 于富生. 公益性业务成本核算—以铁路运输企业为例 [J]. 财会通讯, 2015 (8): 66-68.

[2] 管红舞. 关于铁路运输企业分线成本核算有关问题的探讨 [J]. 财务会计, 2015 (9): 156-157.

[3] 黄 雪. 铁路公益性服务会计核算研究 [D]. 北京: 北京交通大学, 2013.

[4] 丁岳维, 李卫东. 公路运输企业质量成本核算的探讨 [J]. 铁道运输与经济, 2007, 6 (29): 29-31.

[5] 郑大喜. 新医改形势下公立医院加强成本核算与控制的思路探讨 [J]. 医学与社会, 2010, 5 (23): 37-39.

责任编辑 徐侃春