

文章编号: 1005-8451 (2016) 12-0031-04

大数据在铁路客运工作中的运用

宋 超

(上海铁路局 客运处, 上海 200071)

摘要: 大数据技术已经被各行各业广泛成功运用,并取得显著成绩,而铁路旅客运输(简称:客运)的大数据应用才刚刚起步,通过对铁路客运管理效率、客运营销和服务质量的深入分析,得出大数据在铁路客运营销、作业管理和服务质量等方面的应用场景,并给出铁路客运大数据建设的思路。

关键词: 铁路; 旅客运输; 大数据

中图分类号: U293 : TP39 **文献标识码:** A

Big data technology applied to railway passenger transportation

SONG Chao

(Department of Passenger Transport, Shanghai Railway Administration, Shanghai 200071, China)

Abstract: Big data technology has been widely applied in most of industries, and achieved remarkable achievements. However, the application of big data technology in railway passenger transportation is incipient. This article focused on the application of big data technology in railway passenger transportation management, marketing management and service quality. A series of application scenarios of big data technology in railway passenger transportation marketing, railway operation, service quality were presented. Some countermeasures about the big data construction in railway passenger transport were put forward.

Key words: railway; passenger transportation; big data

随着计算机、移动互联网等信息技术的快速发展,当今社会已经进入一个信息爆炸的时期,我们每时每刻都在产生、应用、分享大量的数据。全球知名咨询公司麦肯锡称“数据已经渗透当今每个行业和业务职能领域,并成为重要的生产要素”^[1],“大数据”应用已经成为政府、企业、学术界的热门话题,并在社会、经济和文化生活中被广泛应用。铁路作为这个时代的基础交通服务行业,尤其是高速铁路的大量运营以及人民生活水平的不断提高,使得社会对交通服务的要求越来越高、行业竞争也越来越激烈,因此,改进铁路旅客运输(简称:客运)服务质量、提高运输管理效率、增加铁路运营收入始终是铁路经营者面临的课题,而大数据时代的开启将为铁路解决这一课题提供全新的理念和手段。

1 大数据概念

1.1 大数据的产生

1989年, Gartner Group 的 Howard Dresner 提

收稿日期: 2016-06-02

作者简介: 宋 超, 高级工程师。

出“商业智能”^[2],通过商业智能将企业中现有的数据转化为知识、帮助企业做出明智的业务经营决策,其主要目标是将企业所掌握的信息转换成竞争优势,从而提高企业决策能力、决策效率、决策准确性。为了将数据转化为知识,需要利用数据仓库、联机分析处理(OLAP)工具和数据挖掘(Data Mining)等技术。而随着互联网的发展,企业收集到的数据越来越多、数据结构越来越复杂,一般的数据挖掘技术已经不能满足大型企业的需要,这就使得企业在收集数据之余,也开始有意识地寻求新的方法来解决无法存储和处理分析大量数据的问题。由此诞生了一个新的名词—“大数据”。“大数据”不仅仅指数据“容量大”,更大的意义在于通过对海量数据的交换、整合和分析,发现新的知识,创造新的价值,带来“大知识”、“大科技”、“大利润”和“大发展”。

1.2 大数据的特点及应用

大数据有着4“V”特征^[3],即Volume(容量大)、Variety(种类多)、Velocity(速度快)和Value(价值密度低)。Volume是指数据量巨大,数据技术的发展,使得我们日常生活、学习、工作产生大量数

据；Variety 是指数据种类繁多，如视频、音频、图片、地理轨迹、作业过程以及许多结构化数据信息；Velocity 是指数据产生快且处理速度也要快，满足实时性需求；Value 是指大数据的价值密度低，也是大数据的最关键特征，在大量数据里真正有价值的东西很少，单个数据价值低，关键是要发现有价值的数据或数据关联性、规律等，就像沙子淘金。

2 铁路客运管理工作中存在的问题

2015年底，我国高速铁路（简称：高铁）通车里程已达1.9万km，高铁客流也随之大幅增长，我国铁路运行已经跨入高铁时代。与高铁列车相配套的铁路客运技术装备，如：自助售票机、自助检票机、旅客广播和导向信息系统、互联网售票系统等日益现代化，有效地提高了客运管理效率和运营效益，但是，面对广大人民群众日益增长的服务质量需求和激烈的市场竞争，铁路客运仍有不少地方有待改善和提高，主要表现如下。

2.1 管理效率

为配合高铁运营以及新装备、新技术的运用，铁路部门加强了人才队伍建设，优化了生产力布局调整，取得了一定的成效，但是，这些措施都是基于建议、经验等模糊需求制定的，没有进行系统、精确的分析和论证，铁路客运一些基本运营管理方式、方法并没有改变，铁路客运的内部管理效率还有进一步提高的空间。

2.2 客运营销

我国高铁的投资主要来自于铁路债券、地方融资、银行贷款等，因此，高铁自开工建设起，铁路市场经营压力就随之产生，如何增加高铁的运营收入成为铁路运营部门持续追求的目标。目前，铁路企业收入以运输收入为主、多元经营为辅，铁路客运及其延伸服务是铁路企业收入的重要来源之一。在既有运能情况下，铁路客运收益遇到瓶颈限制，原有的营销手段已经不能满足日益激烈的市场竞争，因此，铁路企业必须勇于创新，研发和利用新营销手段，例如，收益管理等，深挖铁路经营潜力，才能产生更多的收益。

2.3 服务质量

虽然高铁运营凭借新装备、新技术的运用显著提高了铁路客运服务质量，但是仍然不能满足多样化的客运市场需求，从旅客购票、候车、乘车等客运基本服务到餐饮、购物等外延服务还有很大的提升空间，与航空公司、携程等知名服务企业的服务方式、服务质量相比，还有明显的差距。

3 铁路客运大数据的建设目标

大数据技术已经在商业营销、运营决策、旅游服务、安全管理等领域广泛应用，并取得显著成效，这给铁路客运营销、运营、管理等方面的创新提供了思路和方法。

3.1 信息共享

以客运大数据项目建设作为创新契机，打通信息壁垒，推进各系统整合共享。采用业务覆盖、专业整合思路，构建专业一体化系统，推进铁路客运领域信息系统纵向整合；规范基础信息标准，协调资源开放，推进与客运领域相关信息系统间横向互联互通。

3.2 客运营销

收益管理已经在航空领域广泛应用，并成为国内外各航空公司主要、关键的营销手段，而高铁客运运营已经基本进入自由市场竞争阶段，如何有效分析市场需求和自身供给、如何平衡市场需求和供给、如何细分客运市场并精准营销等问题恰恰是收益管理所能解决的。因此，利用铁路客票系统、客服系统（即咨询和投诉信息系统）、社会经济信息等数据开展客运收益管理将是提高铁路客运营销的有效方法和手段。另外，利用客票系统、餐饮服务系统、客服系统等开展客户关系管理、个性餐饮营销等大数据分析，也必将有效拓展铁路多元化经营发展，全面提高运营收益。

3.3 客运生产管理

在铁路客运运营管理中每天产生大量的工作信息，如人员、岗位作业内容、作业时间、作业地点、工作结果以及客流等信息，这些数据日积月累，隐含着大量有价值的规律和知识。铁路客运部门可以利用大数据技术对其进行挖掘和分析，找出日常客运作管理中疏忽、错误、多余等缺点或不足，并以此

进行管理流程优化和再造，提高客运作业工作效率。例如：对各种售票数据进行分析，以此对售票窗口、售票设备、售票人员的设置、配备进行优化和调整；对客运乘降作业数据和客流情况及需求进行分析，以此对车站客运服务流程进行梳理和优化。

3.4 客运服务

对铁路常旅客、各客运场所运营状况等数据进行分析和利用，参照社会其他服务行业服务的方式对旅客进行多样化和个性化服务，进一步提高铁路客运服务质量。例如：针对常旅客的咨询、投诉和建议等提供个性化的购票、乘车、餐饮等服务；针对季节客流情况以及商品和餐饮销售等情况，挖掘出潜在的需求规律，预先做好客运服务准备工作；对客运事故数据进行分析，找出同类事故发生的规律，提出防止类似事件发生的建议和措施。

4 铁路客运大数据系统建设思路

当前，铁路客运营销水平、客运生产管理效率、客运服务质量等方面的提高都离不开大数据技术的支持，而大数据技术在铁路客运中的应用尚处在起步阶段，因此，如何规划、建设铁路客运大数据系统成为当务之急。

4.1 加快客运信息化，丰富客运大数据

目前，铁路客运信息系统主要有客票发售和预订、旅客服务、客运服务等，而客运站务作业、乘务管理、设备管理、整备作业、餐饮保洁等方面还没有完善的信息系统。既有的信息系统也存在独立运行、信息数据分散、过程数据缺失等问题，因此，建立完备的铁路客运大数据应用系统需要开展的工作如下：(1) 加快客运站务作业、乘务管理、设备管理等方面的信息化建设，填补基本客运生产作业、管理、营销等方面空白；(2) 打通信息壁垒，持续推进各系统横向互联互通，实现信息数据共享；(3) 完善各信息系统，注重过程数据的保存，各专业系统可能都有明确的功能目标，不注重作业过程数据，而这一类数据价值很高，例如：旅客购票过程中的信息查询、客运人员作业过程等，

也应纳入存储和分析。

4.2 构建大数据平台

由于各专业系统都是以功能实现为主，存储空间有限，许多过期数据和过程数据未能及时存储，另外，许多外部数据也是客运业务经营分析的宝贵财富，因此，铁路部门要加快大数据平台的建设步伐，及时采集、整理各系统及外部的数据信息，为大数据的开发应用奠定数据基础。当前，大数据平台技术日趋成熟，如 Hadoop、Spark 等，已经被广泛使用。大数据处理流程基本分为数据采集、数据处理和集成、数据分析及数据解释 4 个阶段，根据客运业务需求，以 Hadoop 为基础，结合 Spark 数据处理和分析技术，构建客运大数据平台，其基本逻辑框架如图 1 所示。

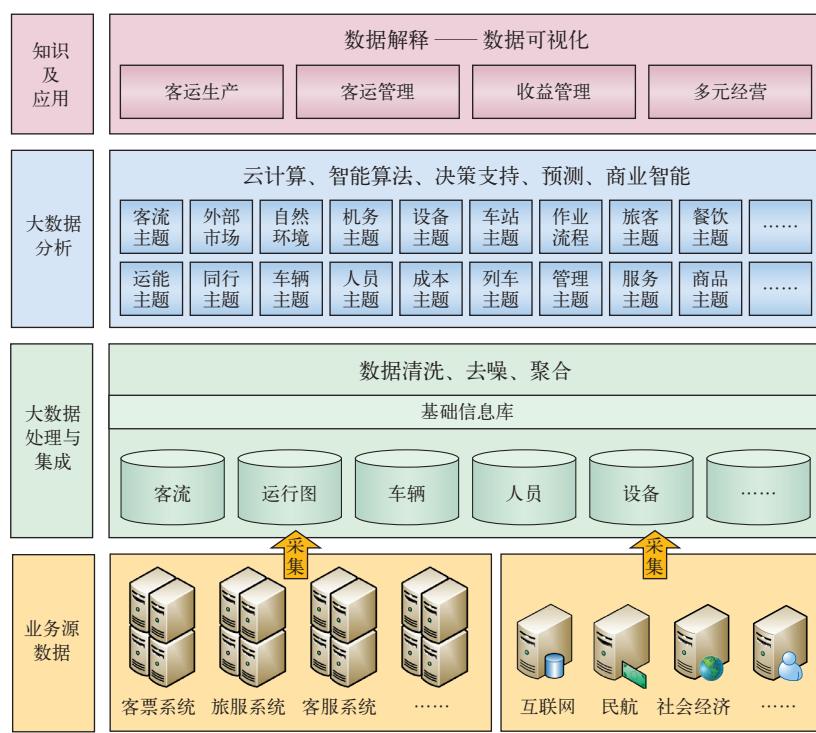


图1 客运大数据逻辑结构图

(1) 业务源数据。明确大数据平台所需的各种内外部数据，内部主要有：客票发售和预订系统、12306 客服系统、运行图系统、动车组运用系统等既有信息系统以及待建信息系统数据，外部主要有：国家和地方经济及人口信息、各地企矿业信息、各地学生信息、天气信息以及其他互联网信息等。

(2) 大数据处理与集成。由于客运大数据来源多，种类多，结构复杂，需要将这些复杂数据转换成单

一的或便于处理的结构，因此，大数据处理和集成主要对采集到的数据进行清洗、去噪和聚合，并分门别类地存储。

(3) 大数据分析。经过数据处理与集成的数据成为数据分析的原始数据，然后通过大数据平台云计算、智能算法、MapReduce 或弹性分布式数据集(RDD)等对客流、运行图、车辆、客服、外部等数据进行进一步处理和分析，得到有价值的主题数据，这也是大数据平台的核心部分。

(4) 知识及应用。通过数据分析和处理，利用数据可视化技术对数据进行解释和展示，帮助我们发现在客运生产作业、客运管理、收益管理、多元经营等方面潜在的价值和规律，全面提高铁路客运经营、管理和生产水平。

4.3 关键技术

(1) 网络安全。客运大数据平台有内网和外网两部分。内网主要是铁路客运信息系统、其他专业信息系统以及待建信息系统，主要包含铁路内部的数据资源；外网连接 Internet、运行平台门户系统等，包含各大官网、调查公司及互联网的数据资源。为保证数据的安全，在外部网中设外部数据存储，外部数据采集到外部数据存储，通过安全平台进行数据交换后才存储在大数据平台中，从而保证系统的安全性，其数据采集流程如图 2 所示。

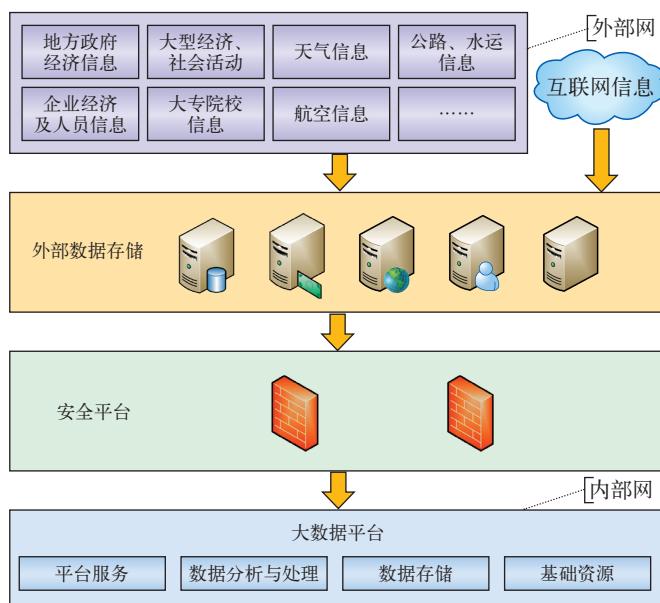


图2 数据采集结构图

(2) 旅客信息安全。由于一些客运信息系统中

包含旅客个人隐私信息，如旅客身份信息、联系电话、出行路径等，而大数据平台要求数据共享，所以一些使用者通过数据分析就可以轻易掌握旅客出行习惯和个人重要信息，这将带来个人隐私安全问题，为此，需要对部分关键数据进行“匿名处理”，例如，将集中信息流控制和差分隐私保护技术融入“云计算”的数据生成与计算阶段，从而防止计算过程中的数据隐私泄露。

5 结束语

大数据技术帮助人们从海量数据中发现知识，获取信息，寻找隐藏在大数据中的模式、趋势和相关性，揭示社会运行和发展规律。我国铁路处于快速发展中，铁路客运部门正在运用先进的大数据技术，积极创新，全面提高客运营销、管理和服务水平，从而开创铁路发展的新纪元。

参考文献：

- [1] MANYIKA J, CHUI M, BROWN B, et al. Big data : The next frontier for innovation, competition, and productivity [R/OL]. [2013-7-24]. Las Vegas : The McKinsey Global Institute.
- [2] 余长慧, 潘和平. 商业智能及其核心技术 [J]. 计算机应用研究, 2002 (9) : 14-16, 26.
- [3] 中国信息产业网. 大数据的四个典型特征 [EB/OL]. [2012-12-04]. <http://cyyw.cena.com.cn/a/2012-12-04/135458292978407.shtml>.
- [4] 高 勇. 啤酒与尿布 [M]. 北京 : 清华大学出版社, 2008 : 1-6.
- [5] 大数据观察. 金融行业 5 个大数据应用案例及案例解析 [EB/OL]. [2016-08-27]. <http://www.shuju.net/article/MDAwMDAwN6zg2.html>.

责任编辑 王 浩