

文章编号: 1005-8451 (2016) 09-0028-03

# 基于大数据平台的铁路旅客群体分析应用研究

吕晓艳, 刘彦麟, 单杏花, 李仕旺

(中国铁道科学研究院 电子计算技术研究所, 北京 100081)

**摘要:** 任何企业都无法吸引市场中所有的客户, 至少不能以相同的方式吸引所有客户<sup>[1]</sup>。大数据技术关于要全体不要抽样, 要效率不要绝对精确, 要相关不要因果的处理数据理念<sup>[2]</sup>, 在很大程度上为企业进行用户群体分析提供了有力的技术基础。本文基于铁路客运大数据平台, 对铁路旅客进行群体分析, 研究群体分析的应用, 以期实现对客运产品与服务的优化调整, 从而提升铁路客运市场优势。铁路换乘旅客群体分析案例的结果表明, 群体分析对于指导铁路客运管理与组织优化调整效果良好。

**关键词:** 旅客运输; 大数据; 群体分析

**中图分类号:** U293 : TP39 **文献标识码:** A

## Application research on railway passenger group analysis based on big data platform

LV Xiaoyan, LIU Yanlin, SHAN Xinghua, LI Shiwang

( Institute of Computing Technologies, China Academy of Railway Sciences, Beijing 100081, China )

**Abstract:** Any enterprise cannot attract all the customers in the market, at least not in the same way to attract all customers. The data processing concept of big data technologies, that is about focusing on the whole, efficiency and relevant, not on the sampling, absolute accuracy, cause and effect, etc., provides a strong technical foundation in passenger group analysis. Based on the big data platform of railway passenger transport, this article analyzed the group of railway passengers, studied on the application of group analysis, in order to implement the optimization and adjustment of passenger transportation products and services, enhance railway passenger market advantage. The result of transfer passenger group analysis indicated that it was good for guiding the management, organization, optimization and adjustment of railway passenger transport.

**Key words:** passenger transportation; big data; passenger group analysis

任何市场中, 消费者在需求、资源、时间、地点、购买态度和购买行为等方面都存在很大的差别<sup>[1]</sup>。通过将庞大的用户群体依据目标划分为不同的用户子群体, 用于调整与优化产品和服务, 使其有效满足用户子群体独特的需求, 从而实现企业价值, 这是现代企业通过市场定位达到稳定与提升市场份额的重要营销策略。

大数据技术支持下的群体分析, 是通过对消费者在消费过程中数据的积累与分析, 实现消费群体在数据上的可分性, 从而实现差异化消费群体, 进一步使企业为不同消费群体创造独特的产品与服务, 达到市场准确定位、营销策略精准实施、效益最大获取的目的。群体分析不同于一般用户分析, 其特

点在于群体特征锐化、个体特征钝化, 即强调群体的共性特征, 模糊群体中的个性特征。本文以铁路客运旅客群体为研究对象, 以大数据技术为支撑, 从客运产、销业务角度出发, 采用数据挖掘相关处理方法, 实现铁路旅客的群体分析。通过旅客群体分析, 对既有客运产品与服务进行评估, 提出调整与优化建议, 以达到提升铁路客运整体运行效益的目的。

### 1 铁路旅客群体分析

铁路客票发售和预订系统积累多年的生产与交易数据, 为铁路旅客群体分析奠定了高质量的数据基础。通过旅客的访问轨迹、交易轨迹、运输轨迹数据和形成的铁路用户画像数据, 支撑铁路旅客群体分析。铁路旅客群体分析从维度构建上主要包括用户维、产品与服务维和时空维:(1) 用户维包括用户属性分群分析、用户行为分群分析(主要为交

收稿日期: 2016-06-15

基金项目: 中国铁道科学研究院科研专项课题(研发中心)(J2016X005)。

作者简介: 吕晓艳, 副研究员; 刘彦麟, 工程师。

易行为和出行行为) ; (2) 产品与服务维包括用户出行需求分析、保持力分析、损失率分析等 ; (3) 时空维则是根据旅客与铁路客运各类业务办理的时间、地点和渠道等时空特征进行群体分析。通过这3个维度的旅客数据交叉复用, 实现铁路客运大数据的新价值, 实现大数据战略在铁路客运运输行业的应用布署。

铁路群体分析层次结构图如图1所示, 基于客运大数据平台和铁路旅客画像, 根据铁路旅客群体分析业务主题需求, 分析并提取相关数据, 进行数据分析, 从而实现对分析业务的支撑。目前的铁路旅客群体分析广泛应用在客运产品优化调整分析、售票组织策略分析、售票管理分析和异常用户分析等客运生产前端的分析。

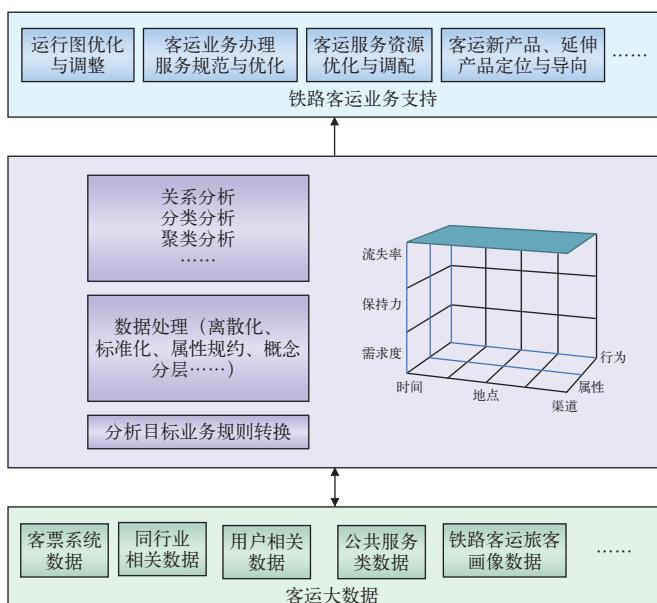


图1 铁路群体分析层次结构图

## 2 铁路旅客群体分析应用案例

为提升铁路客运服务水平而进行的旅客群体性分析主要服务于运行图优化与调整、客运业务办理服务规范与优化、客运服务资源优化与调配、客运新产品以及延伸产品的定位与导向等。文章以提高票额组织管理水平为目标, 以对铁路旅客群体进行联程分析为群体分析案例, 进行铁路旅客群体分析应用分析。

### 2.1 客运数据分析

客运数据采集、分析及结果应用按以下步骤进

行 : (1) 以旅客出行客运大数据为基础, 面向换乘业务, 抽取、组建换乘群体用户和出行数据, 定义客运换乘业务数据规则 : 同一证件号、同城换乘3日内的所有京广线车站乘车的动车组用户。(2) 分析换乘群体的换乘购票规律、换乘购票渠道、换乘间隔、换乘产品、换乘用户群体集等5个方面, 用于支撑综合分析。(3) 将分析结果提交客运管理生产部门、客运营销部门和车站客运主管部门, 实现产品调优、组织调优和服务调优。

以某一时段京广线本线车站间动车组乘客换乘为例, 对此间在京广线车站间经过一次换乘的铁路旅客进行数据提取, 通过对换乘数据的预售、换乘和换乘目的地的分析, 得到如图2所示的京广换乘群体特征图。

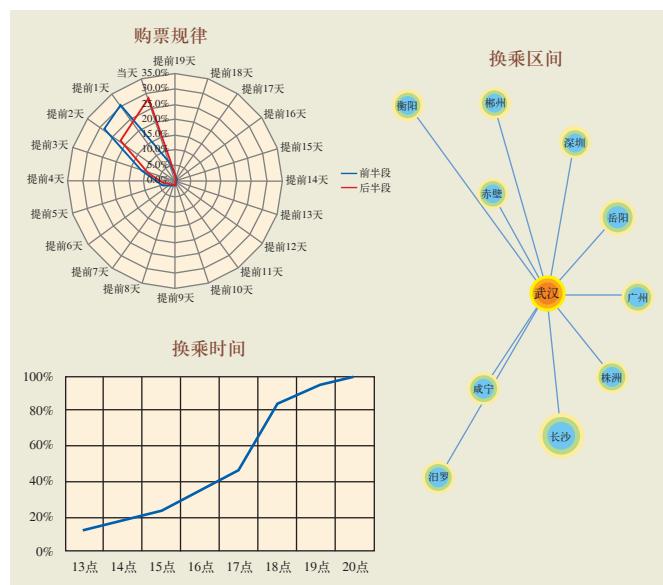


图2 换乘用户群体特征分析

- (1) 北京经武汉站的换乘用户较高；
- (2) 换乘旅客一般提前2天购买北京出发的车票以及要换乘车票；
- (3) 在当日购买后半段车票时, 主要集中在开车1 h 内在车站窗口购票；
- (4) 换乘时间一般集中在当日的18:00 ~ 19:00之间；
- (5) 后半段乘车区间主要以武汉为中心, 辐射至武广间的动车组列车, 重点流向是长沙。

上述结论可以支持铁路客运管理部门在产品结构、售票组织和旅客服务等方面进行调整。

## 2.2 分析结果应用

### 2.2.1 产品结构调整

北京出发至京广线沿线车站的动车组换乘旅客，在武广深间一次乘车抵达目的地的需求没有完全满足，特别是至岳阳、长沙，综合京广线运行图铺画特征，可考虑调整北京至广州深圳方向列车在武汉至广州深圳间的经停车站。

### 2.2.2 售票组织调整

就绝对换乘人数分析，北京经过武汉，换乘至长沙这一高峰换乘区间，换乘数量不足以支撑新增开行列车，但通过大数据分析和换乘人数预测，票额组织调整可以实现让北京至长沙间的旅客通过北京至长沙的列车完成运输需求，让换乘流通过票额组织成直通流。

### 2.2.3 旅客服务调整

在换乘成为武汉车站的一个必要业务时，可以通过大数据分析，获得换乘群体集中换乘时间和换乘车次，对应武汉站可以通过构建车站的武广换乘专用通道和旅客引导专员，实现换乘客流的顺畅运输，提升旅客的铁路出行体验。

综上所述，基于客运大数据的用户群体特征分析，在当前建立以旅客为中心的客运营销体系下，其提取的数据价值显而易见，分析结果应用于指导生

产，产生的经济效益与社会效益明显。但大数据分析结果在庞大的铁路客运体系中的及时流转与控制流转是应用的一个局限和难点，也是我们下一步研究的重点。

## 3 结束语

铁路客运大数据分析是铁路信息化和铁路营运管理部门当前的研究热点和重点。重拾客运数据价值，检查铁路客运运营管理有效性和合理性，指明客运管理与组织的优化调整方向，实现大数据技术指导下的铁路客运高效生产是铁路客运大数据平台的重要意义所在。本文从铁路用户群体分析应用角度出发，对目前基于客运大数据平台的群体分析应用流程、架构进行阐述，分析结果用于指导铁路客运生产，证明了将大数据蕴含价值转化为服务于旅客的信息是优化铁路客运管理与组织的一个重要方法。

### 参考文献：

- [1] Philip Kotler, Gary Armstrong. 市场营销原理 [M]. 赵平,译. 北京: 清华大学出版社, 1999, 10: 156-179.
- [2] 缔元信. 用户分群画像: 抽样“猜想”让位于大数据“观察” [EB/OL]. <http://www.thebigdata.cn/YingYongAnLi/12697.html>, 2014-12-05.

责任编辑 杨利明

(上接 P16)

## 3.5 风险发现

基于用户画像的风险标签，可以设计铁路售票或客运组织过程的风险控制模型，识别风险用户，防范12306互联网售票过程中抢票、倒票和囤票行为，防范客运组织过程中的闯闸、动车组列车上吸烟、异常退签获利等行为。

## 3.6 应急指挥

综合运用客运大数据平台数据模型，可以设计客运应急指挥模型，提高应急事件识别能力，建立应急指挥流程和步骤，提高应急指挥的处置能力，提升突发事件发生时铁路旅客的出行体验。

## 4 结束语

客运是铁路行业的核心支柱产业，客运大数据

技术的运用将成为未来客运增运增收的信息化支撑手段。铁路客运大数据平台的建设不可能一蹴而就，从基础计算资源和存储资源的投入，到画像系统的设计和调整；从路内外数据标签的获取，到各应用场景的技术实现，都将是一个不断持续迭代的过程。在此过程中，可以根据应用的紧迫性，边设计边建设边应用，逐步取得相应的效果。

### 参考文献：

- [1] 邵明豪. 数据预处理技术的具体实现形式研究 [J]. 网络安全技术与应用, 2009 (6): 52-53.
- [2] 陆嘉恒. 大数据挑战与NoSQL数据库技术 [M]. 北京: 电子工业出版社, 2013.

责任编辑 王浩