

基于大数据的无线通信用户行为分析与精准传播策略研究

张浩¹ 张诗淇²

1. 中国电信股份有限公司唐山分公司, 河北 唐山 063000

2. 澳门城市大学, 澳门 999078

[摘要]为解决无线通信行业用户规模扩张与传播效率不匹配、资源配置同质化等问题, 本篇文章提出一套基于大数据的全流程用户行为分析与精准传播技术体系。文章整合无线通信网络日志、终端交互数据、业务消费记录等多源异构数据, 构建“数据采集-预处理-特征工程-模型训练-策略输出”的闭环架构。在技术实现上, 采用改进的 K-Means++算法优化用户分群精度, 结合 CNN-LSTM 融合模型挖掘用户行为时序特征与场景关联规律, 创新性提出“动态画像+实时适配”的精准传播机制。实证结果表明, 该体系可将用户行为识别准确率提升至 89.7%, 传播转化率较传统模式提高 34.2%, 资源投入成本降低 27.5%, 能为无线通信企业的精细化运营提供技术支撑与实践路径。

[关键词]大数据; 无线通信; 用户行为分析; 精准传播

DOI: 10.33142/sca.v8i12.18759

中图分类号: TP39

文献标识码: A

Research on Wireless Communication User Behavior Analysis and Precise Propagation Strategy Based on Big Data

ZHANG Hao¹, ZHANG Shiqi²

1. Tangshan Branch of China Telecom Corporation, Tangshan, Hebei, 063000, China

2. City University of Macau, Macau, 999078, China

Abstract: In order to solve the problems of mismatched user scale expansion and propagation efficiency, as well as homogeneous resource allocation in the wireless communication industry, this article proposes a full process user behavior analysis and precise propagation technology system based on big data. The article integrates multi-source heterogeneous data such as wireless communication network logs, terminal interaction data, and business consumption records to construct a closed-loop architecture of "data collection - preprocessing - feature engineering - model training - strategy output". In terms of technical implementation, an improved K-Means++ algorithm is adopted to optimize the accuracy of user grouping, combined with the CNN-LSTM fusion model to mine the temporal characteristics of user behavior and scene correlation rules. An innovative precise propagation mechanism of "dynamic profiling + real-time adaptation" is proposed. The empirical results show that this system can improve the accuracy of user behavior recognition to 89.7%, increase the transmission conversion rate by 34.2% compared to traditional models, reduce resource investment costs by 27.5%, and provide technical support and practical paths for the refined operation of wireless communication enterprises.

Keywords: big data; wireless communication; user behavior analysis; accurate dissemination

引言

随着 5G 技术的规模化部署与数字经济的深度发展, 无线通信行业进入“存量竞争”与“增量创新”并存的新阶段, 用户需求呈现出个性化、场景化、动态化的显著特征。传统基于群体统计的粗放式传播模式, 因缺乏对用户个体行为的深度洞察, 导致传播内容与用户需求错位、资源配置效率低下、用户体验提升乏力等问题日益突出, 难以适应行业高质量发展的需求。基于此, 立足大数据技

术与机器学习算法, 构建多维度、深层次的用户行为分析体系, 创新提出适配无线通信场景的精准传播策略, 可为无线通信企业优化资源配置、提升传播效率、增强用户黏性提供技术参考, 同时为大数据技术在通信行业的深度应用拓展实践路径。

1 大数据驱动的无线通信用户行为分析技术体系

1.1 多源数据采集与预处理技术

无线通信用户行为数据具有来源分散、类型多样、体

量庞大、噪声密集等特征,其采集与预处理的质量直接影响后续分析结果的可靠性^[1]。本文构建的多源数据采集框架涵盖四类核心数据:网络侧数据(包括基站接入日志、网络流量数据、信号强度变化记录、切换成功率等),主要利用5G核心网的UPF(用户面功能)、AMF(接入和移动管理功能)模块实时采集;终端侧数据(包括终端型号、操作系统、硬件配置、电池状态、应用安装与使用记录等),凭借终端SDK与设备管理平台协同采集;业务侧数据(包括通话时长、短信收发量、流量套餐类型、增值业务订阅记录、消费金额等),从业务支撑系统(BSS)与运营支撑系统(OSS)中提取;场景侧数据(包括用户地理位置信息、时间维度数据、场景标签如通勤、办公、居家等),根据GPS定位、基站三角定位与行为特征关联推导获取。

为保障数据质量,设计四级预处理流程:第一级数据清洗,采用基于统计规则的异常值检测算法(如 3σ 准则)剔除网络波动导致的异常数据,根据缺失值插补模型(KNN插值法)补充缺失数据,利用重复数据检测算法(基于MD5加密的记录去重)删除冗余信息;第二级数据集成,采用联邦学习框架实现多源数据的分布式融合^[2],在不泄露原始数据的前提下完成数据关联,利用Schema映射技术解决异构数据结构不一致问题;第三级数据转换,将非结构化数据(如应用使用日志)通过TF-IDF算法转化为结构化特征向量,对数值型数据采用Min-Max归一化处理,对分类数据采用独热编码转换,确保数据格式满足建模要求;第四级数据规约,采用主成分分析(PCA)与局部线性嵌入(LLE)结合的混合降维算法,在保留核心特征的前提下,将高维数据维度降低60%以上,提升后续模型训练效率。

1.2 用户行为特征工程与建模

用户行为特征的提取与建模是行为分析的核心环节,因此一定要兼顾特征的全面性与针对性。基于预处理后的多源数据,构建“三维度”特征体系。基础属性特征包括用户性别、年龄、终端类型、套餐等级等静态特征,反映用户的固有属性;通信行为特征包括通话频率、时长分布、流量消耗规律、增值业务使用偏好、网络接入时段等动态特征,体现用户的核心通信需求;场景关联特征包括地理位置分布、场景切换频率、不同场景下的行为差异等情境特征,揭示用户行为的场景依赖性。进行特征交叉与衍生处理后,生成如“通勤时段4G/5G网络切换频率”“居家场景视频业务使用时长占比”等深度特征,丰富特征维度。

在行为建模阶段,采用“聚类分群+时序预测”的双

模型架构:首先,针对传统K-Means算法对初始聚类中心敏感、易陷入局部最优的缺陷,引入改进的K-Means++算法,通过基于密度的初始中心选择策略提升聚类稳定性,结合轮廓系数与Calinski-Harabasz指数自适应确定最优聚类数,将用户划分为高频通信型、流量消费型、增值业务偏好型、基础需求型、低活跃型五类群体,聚类准确率较传统算法提升15.3%;其次,针对用户行为的时序关联性,构建CNN-LSTM融合预测模型,利用CNN网络提取用户行为的空间特征(如不同业务间的关联关系),依托LSTM网络捕捉行为的长短期时序依赖(如节假日流量消费规律),结合注意力机制强化关键时间节点的特征权重,实现对用户未来15d内通信需求、流量消耗、业务办理意向的精准预测,将预测误差控制在12%以内。

1.3 动态用户画像构建

传统静态用户画像在适应无线通信用户行为的动态变化方面显得越来越力不从心,故本文构建了“基础画像+实时画像”的双层动态画像体系。基础画像基于用户长期行为数据构建,涵盖用户基本属性、核心需求偏好、长期消费能力等稳定特征,采用增量学习算法每季度更新一次;实时画像聚焦用户短期动态行为,包括实时网络状态、当前所在场景、近期行为变化趋势、临时业务需求等动态特征,基于流计算框架(Flink)实现秒级更新。

画像融合采用加权融合策略,采取层次分析法(AHP)确定基础画像与实时画像的权重系数,其中基础画像权重占比60%,实时画像权重占比40%,确保画像既反映用户长期稳定特征,又能快速响应短期动态变化。为提升画像的实用性,建立画像标签体系,包括一级标签5类、二级标签23类、三级标签89类,标签粒度细化至“工作日早高峰通勤场景5G流量高消耗用户”“夜间居家场景视频会员订阅意向用户”等精准维度,为后续精准传播提供明确的目标导向。

2 基于用户行为分析的精准传播策略设计

2.1 差异化内容定制策略

基于用户分群结果与动态画像,构建“需求-内容”精准匹配机制,做到“因人施策”。针对高频通信型用户,重点推送语音套餐升级、亲情网业务、高清通话增值服务等内容,突出通信质量提升与资费优惠;考虑到流量消费型用户的需求,推荐大流量套餐、定向流量包、流量叠加包等产品,并结合用户流量消耗高峰时段推送限时优惠活动;针对增值业务偏好型用户,根据其具体偏好(如视频、音乐、游戏、阅读)推送对应会员服务、专属内容权益、业务组合套餐,同时基于协同过滤算法推荐关联度高的其

他增值业务；就基础需求型用户而言，以简洁实用的基础套餐优化、账单清晰化服务、网络故障快速报修通道等内容为主^[3]，避免过度营销；低活跃型用户方面，推送回归激励礼包、个性化套餐降级方案、专属客服一对一咨询等内容，激活用户活跃度。

内容呈现形式根据用户终端类型与使用习惯差异化设计。对年轻用户群体（流量消费型、增值业务偏好型），采用短视频、H5 互动页面、直播带货等生动直观的形式；对中老年用户群体（基础需求型、高频通信型），以文字短信、语音播报、图文结合的简洁形式为主，确保信息传递的有效性。

2.2 动态渠道选择与时机优化策略

渠道选择方面，基于用户场景特征与渠道偏好数据，构建多渠道协同传播体系。对用户行为进行详尽分析之后，识别各用户群体的偏好渠道——高频通信型用户偏好短信、客服电话渠道；流量消费型用户集中在 APP 推送、小程序弹窗、短视频平台广告；增值业务偏好型用户对社交媒体推广、直播带货、KOL 推荐接受度更高；基础需求型用户倾向于营业厅线下推广、社区服务点宣传；低活跃型用户需通过多渠道联动（短信+APP 推送+线下海报）提升触达率。

传播时机优化采用“时序预测+场景匹配”的双重驱动机制。基于 CNN-LSTM 预测模型输出的用户行为高峰时段，结合场景特征确定最佳传播时机：对通勤场景用户，选择早 7:00~9:00、晚 17:00~19:00 的通勤时段，利用车载终端、手机 APP 推送相关内容；对办公场景用户，避开工作高峰时段，选择午间 12:00~13:30、下班后 18:00~20:00 推送；对居家场景用户，集中在晚 20:00~22:00 的休闲时段推送视频、娱乐类增值业务；对节假日场景用户，提前 3~5d 推送流量套餐、出行相关通信服务等内容。同时，引入强化学习算法，根据用户对不同时机传播的反馈（如点击、办理、忽略）实时调整传播时机参数^[4]，持续优化触达效果。

2.3 传播效果评估与闭环优化机制

为真正做好传播效果的评估，构建多维度、量化的传播效果评估体系，涵盖传播触达率、内容点击率、业务转化率、用户留存率、资源投入回报率五大核心指标。传播触达率衡量信息成功传递的用户比例，反映渠道覆盖的有效性；内容点击率体现用户对传播内容的兴趣度；业务转化率是核心评估指标，反映传播策略对用户行为的引导效果；用户留存率关注传播后用户活跃度的持续情况；资源投入回报率以计算传播带来的收益与投入成本比值的

方式评估策略的经济性。

最后，建立“评估-分析-优化”的闭环优化机制。利用大数据分析平台实时采集各传播环节的指标数据，采用层次分析法（AHP）确定各指标的权重系数，计算综合效果得分；针对效果不佳的环节，深入分析原因，如触达率低可能源于渠道选择不当，转化率低可能是内容与用户需求不匹配；基于分析结果，对传播策略进行针对性优化，包括调整内容形式、更换传播渠道、优化推送时机、更新用户画像标签等，以增量迭代的方式持续提升传播效果。以及，将优化后的策略反馈至模型训练环节，更新模型参数，形成“数据-模型-策略-效果-数据”的良性循环^[5]，有始有终、始终如一。

3 实证分析

3.1 实验数据与环境

实验数据来源于某区域无线网络的用户数据，选取 10 万用户连续 3 个月的多源数据作为样本，包括网络日志数据（约 500GB）、终端交互数据（约 120GB）、业务消费数据（约 80GB）、场景关联数据（约 50GB）。样本用户涵盖不同年龄、性别、职业、消费能力群体，确保样本的代表性。

实验环境配置如下：硬件采用 Intel Xeon Gold 6330 处理器，64GB 内存，10TB SSD 存储；软件环境基于 Ubuntu 20.04 操作系统，采用 Hadoop 3.3.4 构建分布式数据处理集群，使用 Python 3.8 作为开发语言，借助 TensorFlow 2.8 框架实现机器学习模型训练，采用 Spark 3.3.0 进行大数据分析。

3.2 实验结果与分析

3.2.1 用户行为分析模型性能

对比本文提出的改进 K-Means++ 聚类算法与传统 K-Means、DBSCAN 算法的性能，结果显示，改进 K-Means++ 算法的聚类准确率为 89.7%，较传统 K-Means 算法提升 15.3%，较 DBSCAN 算法提升 11.2%；聚类迭代次数减少 28%，运算效率显著提升。在用户行为预测方面，CNN-LSTM 融合模型的平均预测误差为 11.8%，较单一 LSTM 模型降低 4.5%，较 CNN 模型降低 6.3%，表明融合模型能更全面地捕捉用户行为的时空特征。

3.2.2 精准传播策略效果

将本文提出的精准传播策略与传统粗放式传播策略进行对比实验，结果如下：精准传播策略的传播触达率达 92.3%，较传统策略提升 18.5%；内容点击率为 27.6%，提升 31.4%；业务转化率达 15.8%，较传统策略的 11.7% 提升 34.2%；用户留存率在传播后 30d 内达 82.1%，提升

23.7%；资源投入回报率为1:3.8，较传统策略的1:2.4提升58.3%，资源投入成本降低27.5%。实验结果表明，本文提出的基于大数据的用户行为分析与精准传播策略，在提升传播效果、优化资源配置方面具有显著优势。

3.2.3 模型适应性与稳定性

通过改变样本规模（5万、10万、15万用户）与数据维度（基础数据、基础+业务数据、全维度数据），测试模型的适应性与稳定性。结果显示，当样本规模从5万增加至15万时，改进K-Means++算法的聚类准确率波动不超过2.1%，CNN-LSTM模型的预测误差波动控制在1.5%以内；当数据维度从基础数据增加至全维度数据时，模型性能持续提升，且在全维度数据下趋于稳定，表明模型具有良好的适应性与稳定性，能够应对不同规模与维度的用户数据处理需求。

4 结束语

综上所述，本文围绕无线通信行业精准传播的核心需求，构建了一套基于大数据的用户行为分析与精准传播技术体系，为用户行为分析提供了高质量的数据支撑，也为大数据与机器学习技术在通信行业的深度应用提供了实践参考，具有重要的理论价值与工程意义。未来可结合5G-A与6G技术的发展趋势，融入空天地一体化网络数据、物联网终端交互数据等新型数据维度，优化用户行为

特征体系与建模方法，进一步适应未来通信网络的发展需求。随着技术的持续迭代，大数据驱动的精准传播将成为无线通信行业高质量发展的核心驱动力，为行业创造更大的商业价值与社会价值。

【参考文献】

- [1]杨小均.基于大数据的学生行为分析与策略[J].中国管理信息化,2025,28(17):191-194.
- [2]荆军昌,张志勇,宋斌,等.融合用户传播风险和节点影响力分析的虚假信息传播控制方法[J].智能系统学报,2024,19(2):360-369.
- [3]刘阳.大数据时代新闻可视化的传播策略[J].东西南北,2024(6):59-61.
- [4]张宇,吴雨晨,朱乾光,等.大数据分析技术下无线通信网络信号异常诊断探究[J].中国宽带,2025,21(9):82-84.
- [5]梁超.基于大数据的新媒体传播效果评估与优化策略探讨[J].传播力研究,2025,9(1):13-15.

作者简介：张浩（1979.1—），男，本科毕业于吉林大学，无线通信专业，研究生毕业于电子科技大学，软件工程专业，现就职于中国电信股份有限公司唐山分公司，云中台解决方案技术经理及产品经理，通信工程师；张诗淇（2006.4—），女，澳门城市大学本科在读，计算机科学专业。