

# 基于弹幕分析的在线直播平台用户理解<sup>①</sup>



黄发良<sup>1</sup>, 谢国庆<sup>2,3</sup>, 陈子炜<sup>3</sup>

<sup>1</sup>(广西师范大学 广西多源信息挖掘与安全重点实验室, 桂林 541004)

<sup>2</sup>(福建省公共服务大数据挖掘与应用工程技术研究中心, 福州 350117)

<sup>3</sup>(福建师范大学 数学与信息学院, 福州 350117)

通讯作者: 黄发良, E-mail: [huangfliang@163.com](mailto:huangfliang@163.com)

**摘要:** 弹幕评论是网络直播平台与用户交互的主要方式之一, 借助弹幕行为的分析可以更有效地实现对网络直播平台的用户理解. 通过采集和利用 3 大热门直播平台 (斗鱼、熊猫与战旗) 的弹幕相关数据, 本文以假设验证的方式从用户属性与用户行为两个角度对在线直播平台用户进行分析与理解, 并建立基于用户行为特征时间序列的用户活跃模型对用户互动活跃度进行量化评估. 研究表明, 平台在线人数具有周期性变化的时间规律, 观众地域具有沿海发达城市集中分布的空间取向, 所提出的用户活跃模型能够对网络直播平台用户的行为活跃趋势做出合理的预测分析.

**关键词:** 用户理解; 弹幕分析; 在线直播; 社交媒体

引用格式: 黄发良, 谢国庆, 陈子炜. 基于弹幕分析的在线直播平台用户理解. 计算机系统应用, 2020, 29(2): 250–256. <http://www.c-s-a.org.cn/1003-3254/7241.html>

## Time-Sync Comments Analyzation for Understanding Subscribers to Live Streaming Services

HUANG Fa-Liang<sup>1</sup>, XIE Guo-Qing<sup>2,3</sup>, CHEN Zi-Wei<sup>3</sup>

<sup>1</sup>(Guangxi Key Lab of Multi-Source Information Mining & Security, Guangxi Normal University, Guilin 541004, China)

<sup>2</sup>(Fujian Engineering Research Center of Public Service Big Data Mining and Application, Fuzhou 350117, China)

<sup>3</sup>(College of Mathematics and Informatics, Fujian Normal University, Fuzhou 350117, China)

**Abstract:** It is an important communication way for webcast video watchers to produce and consume time-sync comments, which can be beneficial to understand the webcast video users. Based on data related to time-sync comment collected from 3 hot live streaming platforms (Douyu, Panda and Zhanqi), a hypothesis testing based method is proposed to analyze webcast video watchers from user attribute and user behavior, a user activity model is constructed based on user behavior feature time series analysis. Research results show that, the number of live streaming platform online users has obvious characteristics of periodic changes, source of live streaming platform online users tends to be distributed in inshore developed cities, and the proposed user activity model can effectively predict activity of users in live streaming platforms.

**Key words:** user understanding; time-sync comments analyzation; live webcast; social media

① 基金项目: 国家自然科学基金 (61962038); 广西多源信息挖掘与安全重点实验室开放基金 (MIMS17-01); 福建省自然科学基金 (2017J01497); 福建省教育厅 K 类科技项目 (JK2016007)

Foundation item: National Natural Science Foundation of China (61962038); Open Fund of Guangxi Key Lab of Multi-Source Information Mining & Security (MIMS17-01); Natural Science Foundation of Fujian Province (2017J01497); JK Science and Technology Program of Education Bureau, Fujian Province (JK2016007)

收稿时间: 2019-06-17; 修改时间: 2019-07-12; 采用时间: 2019-07-19; csa 在线出版时间: 2020-01-16

近年来,网络直播迅速发展成为一项新兴互联网文化产业.据中国直播榜网站在线数据显示:截止2017年2月,直播平台主播超39万,粉丝数超61亿人次,财富总量超3.8亿.网络直播一方面为主播达人们提供了展示与推销自我的平台,另一方面加强了网民的在线体验与及时感受.为了增加直播趣味性和互动性,各大网络直播平台纷纷采用弹窗弹幕作为用户实时交流的方式<sup>[1,2]</sup>.内容丰富且形式多样的弹幕数据中隐含着复杂的用户属性与用户行为,研究并理解在线直播平台用户具有弹幕内容审核与监控<sup>[3,4]</sup>、舆论热点预测<sup>[5]</sup>、个性化摘要标注<sup>[6]</sup>等多方面的应用价值.

在线直播用户理解已经引起研究者的广泛关注.文献[7]通过分析在线视频评论格式来研究用户评估与判断能力,举例了分体式和流式的评论布局特征对用户行为动机的影响.文献[8]则从弹幕文本内容入手,研究弹幕对网民视频认知的影响,并利用文本分类技术对视频内容进行快速的语义分析.在自我控制和社交能力上,用户可以提升其参与视频社区互动的活跃度<sup>[9]</sup>,拥有较高的互动活跃度可以促进用户自我控制和社交能力的发展.文献[10]提出一种基于弹幕文本的视频片段情感识别算法,建立用户兴趣度量模型,并以此进行视频片段的推荐.文献[11]利用句子级别的情感分析方法,建立基于情感词典的弹幕情感分析模型,对评论文本进行情感词抽取及情感值计算,并结合时间序列进行分析.文献[12]利用视频语义的时间依赖性、用户弹幕评论的交互性与用户偏好建立个性化的时序主题模型,有效提升视频标签算法的性能.文献[13]提出一个基于数据重建原则的时序概要模型,选择与视频最相关的代表性弹幕评论来生成视频的动态描述.然而绝大多数现有研究工作都是对弹幕的文化背景和社会影响力进行定性分析或利用文本挖掘技术建立弹幕-视频相关度量分析模型,而在用户行为的具体量化测量和用户活跃度的研究方面还有所欠缺.

为了对用户行为和弹幕语言进行更深入的探究,本文针对主流直播平台的弹幕进行分析和建模.在用户属性测量方面,本文根据用户的时空分布和接入设备的类型分析网络在线直播对社会的影响,借助弹幕文本长度分布、用户发送弹幕的频率分布以及用户个体的等级经验分布等估计用户群体在观看直播环境下即兴发言的属性特征.另一方面利用所抓取的网络在

线直播中弹幕及用户行为的相关数据,分析直播平台用户的交互特点和活跃度特征,提出基于网络在线直播的用户活跃度计算框架.本文提出的用户活跃度模型可以较好地适应当今新兴娱乐活动的弹幕直播,并对近期的用户行为理解和网络在线直播平台的用户活跃趋势做出合理预测和分析.

## 1 数据采集

本文利用网络爬虫程序对国内热门的3个主要直播平台(斗鱼、熊猫、战旗)进行弹幕数据的抓取.由于弹幕数据规模十分庞大且处于动态变化中,加之分布不均的特点,获取整个实时弹幕数据非常困难,因此在数据抓取的过程中,不同程度地运用了采样策略.为了保证弹幕数据的可靠性、丰富性和有效性,本研究样本时间跨度为2017年1月5日至2017年2月5日,对3个平台的6种主要直播类别:网络游戏、手机游戏、主机游戏、颜值/歌唱、户外直播、影视综艺的房间中随机抽取若干热门房间进行跟踪抓取.数据采集流程包含协议分析、编制抓包程序、自动抓取并存储于MongoDB数据库中,如图1所示.

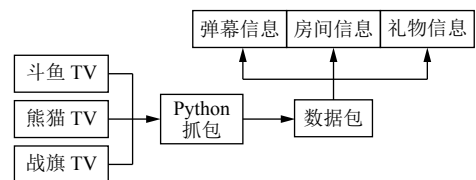


图1 数据采集流程

在各直播平台中,用户与主播互动主要包括“发送弹幕与赠送礼物”两种方式.其中弹幕内容为文本与原设定emoji表情;而礼物则包含了各种价值不等的礼物种类.对于这两种互动形式,我们对爬虫程序抓取的数据内容进行分类预处理,如表1所示,得出我们所需的各项数据:每场直播的房间信息、弹幕数据以及礼物信息.利用弹幕数据和打赏信息,可以方便地获取用户的各项特征以便描绘用户行为属性并对弹幕内容做出合适的统计分析.

表1 数据采集信息表

数据类别	采集内容
Room	Category, Room_id, Start_time, End_time
Barrage	Cur_time, Content, Person_num Fans_num, Weights, Plat, IP_addr, User_id, User_level
Gift	User_id, Cur_time, Combo, Price

通过对3个直播平台的观众人数与弹幕数的统计分析,绘制了直播平台人数与弹幕数的关系.如图2所示,其中深色为1月5日起至2月22日各平台的弹幕总数,浅色为同期时间各平台观看直播人数总和.可以看出,如今各平台的观众处于十万级别,但是弹幕量却达到百万级别,弹幕数与人数比例接近10:1,即每个观众平均发送10条弹幕.

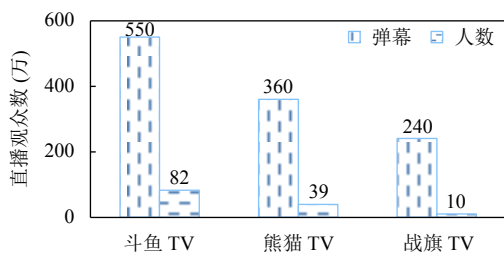


图2 3大直播平台数据分布图

## 2 分析与讨论

随着网络技术发展,网络使用用户激增,相关用户行为也变得复杂.根据网络中大量用户行为相关数据,建立相关用户行为模型,对用户行为理解进行探究,能够深入理解用户行为的特性,为企业新营销策略提供参考;同时能对用户行为言论监管和相关部门监管网络安全<sup>[14]</sup>提供重要作用.

不同于微博的测量<sup>[15]</sup>,而弹幕文本的动态时序特征与微博或者短评类文本相比较更加明显,具有更加短小的文体特征和多变的语义行为.在用户行为理解方面,Theilwall<sup>[16]</sup>探究了基于YouTube的用户和评论的相关特征,仅局限于相对静态的用户网络和YouTube视频,而并未涉及更加复杂多变弹幕直播中的用户行为网络.

为此,本文根据数据特点,将已抓取的数据分成用户属性与弹幕行为两个类别.用户属性包括用户观看直播时间、地点、使用的平台以及用户账号等级.用户行为包含活跃天数、打赏总额、弹幕数量以及弹幕长度与弹幕频率5个方面.在属性与行为两个方面里进行分析,在指定数据范围中总结出用户特点,根据这些特点,建立相关的用户行为模型.我们采用如图3所示的基于假设验证思路的研究.

### 2.1 用户属性驱动的用户理解

在社交网络中,用户属性对于研究用户群体分布和潜在影响力用户挖掘提供有效的特征依据.对于网

络在线直播而言,用户属性是由用户时空活跃分布、用户接入平台的类型和当前用户的经验等级分布等组成.针对当下网络在线直播的宏观特点以及新形式下大众娱乐特点,我们做出如下假设:

(1) H1a, 用户观看网络在线直播的时间分布出现周期性的变化规律,活跃时间大多集中于晚上.

依据:弹幕直播作为一种新型的休闲娱乐方式,用户可以自由安排自己的互动时间而主播也可以根据用户粉丝的日常活跃情况设置工作和直播时长.随着网络直播行业的发展,全民直播,人人参与的热情也在不断提高,只要一部智能手机就可以满足直播或是接入平台参与互动的要求,直播和互动门槛不断降低.在其他非工作和茶余饭后的休闲时间里,观众可以自由选择参与互动的的时间,而晚间时段成了这些人的最佳选择之一.主播也根据粉丝的时间周期性地安排直播时长,主动迎合观众的需求以吸引更多的粉丝从而增加获得更多收益的机会.

分析:根据已采集的数据样本,构建以每5分钟为时间差的时间序列样本模型,并根据在线用户数目构建出了用户人数的时间分布情况.如图4所示,在线用户数目随着时间呈现周期性变化,每个周期都会先后出现两个相对峰值,且每个周期的峰值分别出现在午间以及凌晨两个时间段,其中以凌晨时分的峰值为较高.大部分直播平台用户选择在下班休息这段时间进行直播或观看直播,在线直播作为新型的休闲娱乐方式已经和这部分受众的日常生活作息相互关系,这也就是平台用户在时间上具有明显的周期性的原因.在线直播的时间分布与用户普遍的休闲娱乐规律相重合,与H1a情况相符.

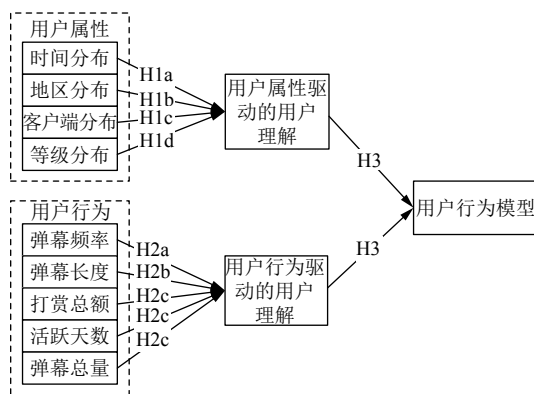


图3 分析流程

(2) H1b, 用户区域性分布不均匀,有可能集中分布于沿海发达城市.

依据: 用户可以通过 PC 网页或是移动客户端建立与直播平台的互动模式, 而且对网络依赖度十分强大. 其次, 根据《中国互联网络发展状况统计报告》<sup>[17]</sup>显示, 现如今国内的网络覆盖程度已经超过 50%, 沿海省份更是接近网络全覆盖. 再加上沿海省份发达的经济现状, 拥有巨大的人口数量. 综合以上几点, 可推测沿海省份的观众分布会大于内陆观众数量.

分析: 将用户的地域分布划分为境外、西北、东北、华中、西南、华北、华南以及华东 8 大地区, 得到如图 5 所示的用户地域分布情况. 在线直播平台用户主要集中华东和华南地区, 用户数量占全体观众数量的 56%, 体现了用户在区域上分布不均的显著特点. 针对用户地域分布存在的显著差异, 我们进一步找出了用户数量排名靠前的 10 个省份. 如图 6 所示, 可以明显看出, 用户数量在广东浙江这样的沿海发达地区最多, 这也例证了 H1b 的假设.

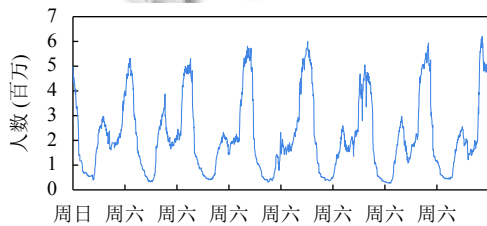


图 4 在线用户数目的时间分布

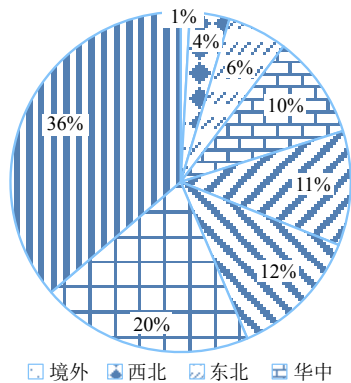


图 5 用户地域分布情况

(3) H1c, 用户更可能选择以网页端接入方式为主, 移动设备登陆为辅的交互观看方式.

依据: 如今大部分网络直播平台的前身是某些电脑社交、视频、语音网站, 在以往的运营中拥有一定程度的用户基础. 其次, 电脑网页端在使用起来方便快捷, 并且网页设计精美, 相对于其他端运行更稳定. 移

动端可以随时关注主播信息, 观看主播的节目, 并且如今智能手机的广泛应用, 使得移动端拥有巨大市场. 综合上述情况, 可推测用户观看直播的方式是以电脑网页端为主, 移动端为辅的交互方式.

分析: 为了准确定量了解用户在线观看直播平台的分布情况, 利用已采集的用户发送弹幕和打赏礼物数据, 生成在线用户接入平台分布图. 如图 7, PC-Web 平台的用户占比为 51%, 拥有较大的观众支持度. 同时, 47% 的用户使用 Android 平台与 IOS 平台观看直播. 在两个相对均衡的选择方式中, 以电脑网页为主移动端为辅的观看模式已经是主流模式, H1c 得证.

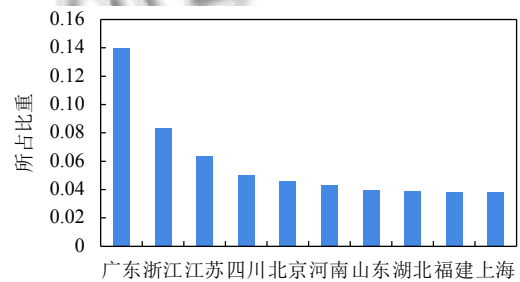


图 6 用户省份分布情况

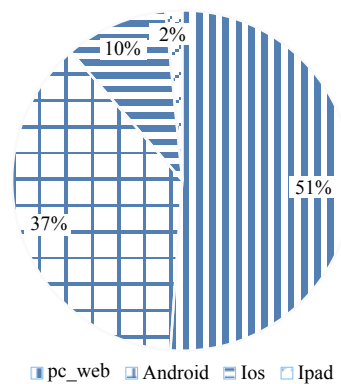


图 7 用户平台分布情况

(4) H1d, 用户平台经验等级分布呈低等级人数多, 高等级人数少的分布特点.

依据: 赠送虚拟礼物成了用户提升等级的最主要的途径之一. 普通用户通过账户充值购买一定数额的虚拟礼物打赏主播或者参与平台不定期的会员活动获取一定的经验来提升用户账号等级. 从社会学和心理学角度出发, 大多数用户倾向于不参与或者是不常参与和主播的礼物互动而是发送弹幕来参与直播的互动, 而有些富豪或者是主播的忠实粉丝就会利用打赏的方式取悦主播以达到表现自我的目的. 因此用户等级分

布可能呈现从低等级到高等级逐步减少的分布状况。

分析: 利用所采集的3个主流直播平台用户交互数据, 建立用户经验等级分布模型. 如图8所示, 用户等级主要集中在低等级部分, 且用户数量随着等级的增加而逐渐减少. 大部分用户主要集中于15级以下, 这也符合用户使用在线直播平台的娱乐目的: 趋向于直接观看直播和参与实时评论, 而较少赠送高额礼物的特点. 所赠送的礼物总价值越多等级上升越快, 经验等级为15级以上的用户虽然所占比例不多但却是平台直播间的忠实粉丝用户群体, 直接影响主播的收益高低.

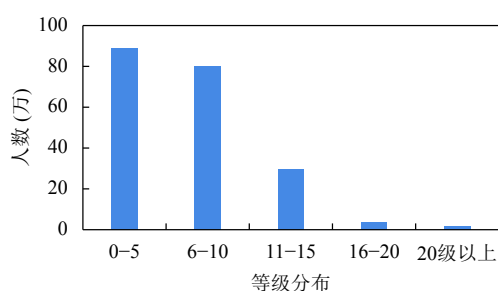


图8 用户等级分布情况

## 2.2 用户行为驱动的用户理解

网络直播作为网络时代的一种新风尚, 对人们的生活侵入程度越来越强. 用户群体往往对直播关注度高, 收看时间长, 弹幕发送频繁, 并通过赠送大量虚拟礼物满足互动消费需求, 自主性强. 因此, 用户行为相比较用户属性而言具有较高的灵活性与不确定性. 对于用户行为, 本文综合大量数据, 寻求其普遍规律, 做出如下假设:

(1) H2a, 用户发送弹幕的频率随时间的增长而逐步降低.

(2) H2b, 用户发送弹幕的长度更加短小并呈现集群分布的特点.

依据: 不同于微博文本, 用户所发表的弹幕评论通常小于32个字, 结构更加自由、形式更加多样、内容更加丰富, 具有更强的实时性. “在网上没人知道你是一条狗”, 这句话生动地反映了网络的虚拟性和匿名性. 匿名机制的存在, 使得用户可以肆无忌惮地发表自己的看法, 参与直播的实时评论, 因此弹幕更新相当频繁. 另一方面, 弹幕在表达个人的观点的同时利用直播屏幕上的滑动效果分享给其他用户, 这也限制弹幕的有效时长. 弹幕这一讲究快捷的特点, 无疑加速了短文本

的产生. 根据中国互联网络信息中心公布的第38次全国互联网发展统计报告显示, 我国网民以10~39岁的年轻群体为主. 因此, 在内容语义上, 弹幕更贴近网络化用语以及口语, 多使用特殊表达形式, 如数字或者表情的组合, 这就满足了用户群体实时弹幕互动的需求.

分析: 为了探究用户发送弹幕频率的分布情况, 我们对间隔时间在10分钟以内(以每分钟为间隔单位)的弹幕数量进行了统计. 如图9所示, 说明用户发送弹幕频率高, 在3分钟内弹幕频率均高达百万人次级别, 并且随间隔时间增长而逐渐减少, H2a得证.

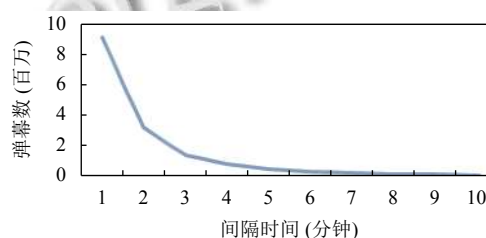


图9 用户发送弹幕频率情况

同理, 为探究弹幕长度特征, 我们统计了弹幕长度分布情况. 如图10所示, 弹幕长度以短文本为主, 1~5字左右的弹幕最为常见, 且随着文本长度的增长弹幕数量出现了明显的减少, 总体呈现出弹幕集群分布于短小处的特征, 这与H2b假设相吻合.

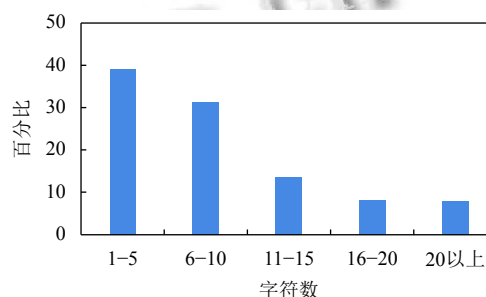


图10 弹幕长度分布情况

弹幕总体呈现出发送频率高, 简洁多样, 即时性强的特征, 一定程度上也切合了当前弹幕文化碎片化和互动性强的特点.

(3) H2c, 用户活跃度与用户历史发送弹幕量、打赏金额以及互动天数正相关.

依据: 首先对活跃用户进行定义: 在直播时间段中, 参与在线网络直播互动的用户定义为活跃用户.

作为直播平台的生命线, 用户活跃度是用户黏性

的反应,直接决定着一个直播平台的商业价值.用户活跃度的具体量化指标往往因应用场景而异,可以根据用户来访、互动情况、核心功能使用频率等综合确定.例如:吴慧等<sup>[18]</sup>从用户背景、社交关系、发表内容质量及社交行为4个方面来设计微博用户活跃度指标.张效尉等<sup>[19]</sup>从某社交网站用户群数据中,选取在一段时间内社交网络用户群中用户平均发布消息的数量作为衡量社交网络用户群用户活跃度指标.王锦坤等<sup>[20]</sup>在设计基于协同过滤的推荐算法时,提出用户活跃度与用户浏览项目数量成正比的模型.本文选择用户发送弹幕量、用户打赏金额以及用户互动时长(天数)作为直播平台用户的活跃度指标,主要出于这样的考虑:弹幕是直播用户对主播人与主播内容的评论,发送弹幕是直播用户观看的体验与参与度的表现;打赏金额是直播观众对主播服务质量的认可与肯定,是深度参与的重要体现;互动时间越长说明该直播用户对此直播平台提供的服务越认可,也是用户黏性的体现.

为此,本文假设用户历史发送弹幕量、打赏金额以及互动天数影响用户活跃度.

分析:为了提取更多的用户行为特征,我们统计了连续两周时间内用户在线天数、发送弹幕数量、赠送礼物总额、最大发送弹幕数量、最大赠送礼物价值以及这段时间内最大连续在线天数.由于用户活跃度存在个体差异性,故本文假设这一组数据能够用来表征不同的用户活跃度.为了验证用户活跃度和这些变量之间的相关关系,我们对这一组数据进行了主成分分析.主成分分析法能将较多的数据变量通过删除变量间的重叠部分得到较少的综合变量,从而将用户活跃度更加准确直观地表示出来.从表2可以看出,共有5个主成分.由于第一主成分相比其它主成分和这些变量之间具有更突出的相关关系,故抽取第一主成分F1作为用户活跃度的表征.根据主成分分析结果,可以得到在线天数和用户活跃度、弹幕数量、礼物总量、最大弹幕量、最大礼物量、最大连续在线天数的相关系数分别为0.724、0.762、0.218、0.838、0.212和0.792.因此,对于用户活跃度满足 $F1=0.724 \times T1+0.762 \times T2+0.218 \times T3+0.838 \times T4+0.212 \times T5+0.792 \times T6$ , $T1-T6$ 分别代表在线天数、弹幕数量、礼物总量、最大弹幕量、最大礼物量、最大连续在线天数.在线天数越长,发送弹幕量越大,赠送礼物越多,F1的值越大,用户活跃度越高,这个结果与H2c假设吻合.

### 2.3 用户属性与用户行为相融合的用户理解

从H1a以及H1b可以看出用户在线观看直播的时间以及地域方面有明显的时空分布特征,从H2a以及H2b可以看出弹幕长度以及频率特征,从H2c可以得到用户平台活跃度相关因素.在线直播主要给用户提供了即时性互动(即弹幕礼物)的全新观看直播体验.由于在线直播集中在深夜时段以及具有消费性的特点,用户人群具有明显的时空分布特征.弹幕礼物作为在线直播互动体验的媒介,承载着用户行为的产出.而用户平台活跃度作为用户互动的产物和在线直播平台衡量一个客户价值的重要标准,自然也是刻画用户人群的重要因素.因此,本文认为用户模型的构建可从用户属性和用户行为两个方面着手.

表2 成分相关性分析表

成分	1	2	3	4	5
在线天数	0.724	-0.053	-0.631	0.273	0.019
弹幕数量	0.762	-0.209	0.589	0.008	0.168
礼物总量	0.218	0.972	0.075	0.002	-0.002
最大弹幕量	0.838	-0.202	0.475	0.028	-0.176
最大礼物量	0.212	0.974	0.079	0.001	0.003
最大连续在线天数	0.792	-0.066	-0.535	-0.288	0.006

### 3 结论与展望

用户属性和用户行为共同驱动用户理解,构成网络在线直播背景下的用户模型.

用户属性方面,观众观看网络直播的时间选择上具有周期性变化规律,并且得出这与主播的开播时间与日常作息密切相关的结论.在沿海发达城市的用户相对集中,特别是人口多,经济发达同时拥有众多外来人口的城市,在观看直播的观众中占有较大比重.根据观看直播平台分布情况来观察,如今观众选择的直播平台具有“网页端为主移动端为辅”的主要特点.

而在用户行为方面,用户在直播平台的互动方式居多,其中以发送弹幕为主.用户所发送的弹幕内容具有“短文本,高频率”的特点.弹幕内容网络化,接近口语,贴切生活同时也容易产生误解与冲突.根据用户等级可以判断出赠送礼物的比重较少,并且用户等级从低到高呈减少的分布状态.而这两种互动方式与在线天数共同支持用户活跃度的理论推断.

为此,本文针对具体的活跃度给出定量的活跃度计算指标,提取影响用户活跃度的主要特征利用PCA技术根据影响力最大化方向提取主成分作为用户

活跃值的参考。

本次的实验依旧有存在部分不足。首先,在抓取数据方面,由于所使用的工具的局限性,只抓取到监测期间有进行互动的用户资料信息,而这一点就忽略了在监测期只进行观看没有参与互动的用户;其次,用户信息只停留在进行互动时的状态,互动后获得收益状态没有跟踪到;再者,选取的数据源较为局限,可能会出现数据偏差的问题。

### 参考文献

- 1 Geerts D, Vaishnavi I, Mekuria R, *et al.* Are we in sync? Synchronization requirements for watching online video together. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York, NY, USA. 2011. 311.
- 2 Liu LL, Suh A, Wagner C. Investigating communal interactive video viewing experiences online. In: Kurosu M, ed. Human-Computer Interaction. Cham: Springer, 2016.
- 3 李君贤. 网络直播中弹幕语言暴力机制的形成与消解. 西部学刊, 2016, (10): 59–60.
- 4 He M, Ge Y, Chen EH, *et al.* Exploring the emerging type of comment for online videos: Danmu. ACM Transactions on the Web, 2018, 12(1): 1–33. [doi: 10.1145/3180440]
- 5 Chen XP, Chen JY, Ma L, *et al.* Fine-grained video attractiveness prediction using multimodal deep learning on a large real-world dataset. International World Wide Web Conferences Steering Committee. Geneva, Switzerland. 2018. 671–678.
- 6 Chung CT, Hsiung HK, Wei CK, *et al.* Towards personalized video summarization using synchronized comments and Probabilistic Latent Semantic Analysis. 2014 IEEE 3rd Global Conference on Consumer Electronics. Tokyo, Japan. 2014. 414–415.
- 7 Shen Y, Chan H, Hung I. Let the comments fly: The effects of flying commentary presentation on consumer judgment. International Journal of Systems Science, 2014, 12(12): 1469–1475.
- 8 Schultes P, Dorner V, Lehner F. Leave a comment! An in-depth analysis of user comments on YouTube. 11th International Conference on Wirtschaftsinformatik. Leipzig, Germany. 2013. 659–673.
- 9 Hu M, Zhang ML, Luo NA. Understanding participation on video sharing communities: The role of self-construal and community interactivity. Computers in Human Behavior, 2016, 62: 105–115. [doi: 10.1016/j.chb.2016.03.077]
- 10 邓扬, 张晨曦, 李江峰. 基于弹幕情感分析的视频片段推荐模型. 计算机应用, 2017, 37(4): 1065–1070, 1134. [doi: 10.11772/j.issn.1001-9081.2017.04.1065]
- 11 郑颀颀, 徐健, 肖卓. 情感分析及可视化方法在网络视频弹幕数据分析中的应用. 现代图书情报技术, 2015, 31(11): 82–90.
- 12 Wu B, Zhong EH, Tan B, *et al.* Crowdsourced time-sync video tagging using temporal and personalized topic modeling. Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, NY, USA. 2014. 721–730.
- 13 Xu LL, Zhang C. Bridging video content and comments: Synchronized video description with temporal summarization of crowdsourced time-sync comments. AAAI Publications, Thirty-First AAAI Conference on Artificial Intelligence. 2017. 1611–1617.
- 14 肖云鹏. 在线社会网络用户行为模型与应用算法研究[博士学位论文]. 北京: 北京邮电大学, 2013.
- 15 樊鹏翼, 王晖, 姜志宏, 等. 微博网络测量研究. 计算机研究与发展, 2012, 49(4): 691–699.
- 16 Thelwall M, Sud P, Vis F. Commenting on youtube videos: From guatemalan rock to el big bang. Journal of the American Society for Information Science and Technology, 2012, 63(3): 616–629. [doi: 10.1002/asi.21679]
- 17 中国互联网络信息中心. 中国互联网络发展状况统计报告. 2018.
- 18 吴慧, 张绍武, 林鸿飞. 微博社交网络的用户影响力评价方法. 中文信息学报, 2017, 31(4): 184–190. [doi: 10.3969/j.issn.1003-0077.2017.04.026]
- 19 张效尉, 余云霞, 王伟. 社交网络群中用户活跃度分析与预测. 西南师范大学学报(自然科学版), 2018, 43(12): 115–121.
- 20 王锦坤, 姜元春, 孙见山, 等. 考虑用户活跃度和项目流行度的基于项目最近邻的协同过滤算法. 计算机科学, 2016, 43(12): 158–162. [doi: 10.11896/j.issn.1002-137X.2016.12.028]