



中国视频消费用户体验白皮书

移动短视频和移动直播（观看端）

中国通信标准化协会互动媒体专委会 (TC613)

中关村现代信息消费应用产业技术联盟

前言

本文件由世界超高清视频产业联盟视频体验工作组组织制订，并负责解释。

本文件发布日期：2023年3月29日。

本文件由中国通信标准化协会互动媒体专委会提出并归口。

本文件归属世界超高清视频产业联盟、中国通信标准化协会互动媒体专委会、中关村现代信息消费应用产业技术联盟。任何单位与个人未经联盟书面允许，不得以任何形式转售、复制、修改、抄袭、传播全部或部分内容。

本文件主要起草单位：

中国信息通信研究院、中国联合网络通信有限公司、中国电信集团有限公司、中国移动通信有限公司研究院、华为技术有限公司、优酷信息技术(北京)有限公司、北京快手科技有限公司、北京火山引擎科技有限公司、北京爱奇艺科技有限公司、北京密境和风科技有限公司、上海兆言网络科技有限公司、聚好看科技股份有限公司、二六三网络通信股份有限公司、智者四海(北京)技术有限公司、北京流金岁月传媒科技股份有限公司、上海艾策通讯科技股份有限公司、中移(杭州)信息技术有限公司、中兴通讯股份有限公司、上海交通大学、北京理工大学、北京邮电大学、德科仕通信(上海)有限公司、中恒达(北京)软件测评科技有限公司、北京市博汇科技股份有限公司、中关村现代信息消费应用产业技术联盟。

本文件主要起草人：

宋利、贾武、罗传飞、杨崑、张世俊、汤聪华、宋祖平、武亮平、王亚军、缪旭阳、朱家悦、张春蕾、吴敬芳、李静、梅大为、杜依菲、马英武、郑川川、余大利、徐静雯、王志航、高玮泽、唐赓、钱雷、段涛、任子健、史东平、刘帅、赵益、艾芳、吴静、郭锦程、周凯旋、刘文泉、王勇、杨琛、黄挺、尹芹、王金东、陶国军、翁冬冬、闫石、吴雪波、陈劫联、钟琳、张家斌、姜卫平、陈红、刘璇。

免责声明：

- 1, 本文件免费使用，仅供参考，不对使用本文件的产品负责。
- 2, 本文件刷新后上传联盟官网，不另行通知。

目录

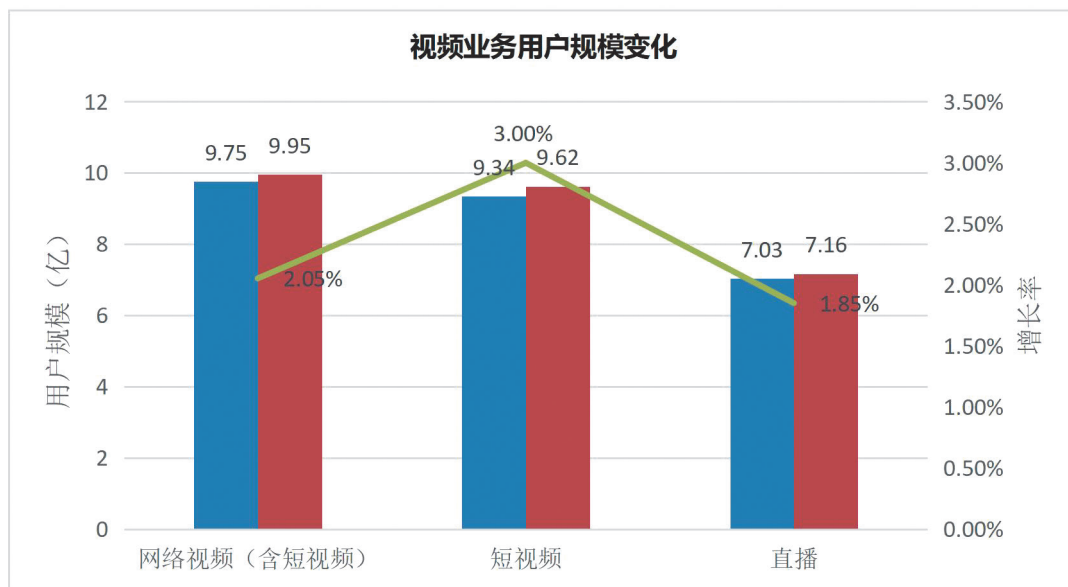
1. 移动视频发展现状及趋势	1
1.1 移动视频已成为移动互联网主要应用	1
1.2 移动视频业务形态发生很大变化	3
1.3 移动视频对推动产业升级发挥重要作用	4
1.4 当前业界缺乏移动视频体验评价体系和标准	5
2. 定义统一的移动视频体验评测方案对产业健康发展至关重要	5
2.1 移动视频服务下影响用户留存的因素呈现多样性	5
2.2 OTT视频服务创新，提升用户体验	7
2.3 统一的体验评测标准是产业共同且迫切的需要	11
3. 移动视频业务体验评测方案及现状	11
3.1 短视频业务	11
3.2 直播视频（观众端）	17
4. 展望	21
4.1 5G 时代的视频技术基本趋势	21
4.2 内容主导的移动短视频业务体验新特性	23
5. 附录	25
5.1 术语解释	25
5.2 引用	27

1. 移动视频发展现状及趋势

1.1 移动视频已成为移动互联网主要应用

手机上网已经成为我国网民的主要上网形式。截至 2022 年 6 月，我国网民规模为 10.51 亿，我国手机网民规模达 10.47 亿，比例高达 99.6%^[1]。在时长方面，网民人均每周上网时长为 29.5 个小时，较 2021 年 12 月提升 1.0 个小时。

视频业务使用频率是仅次于即时社交通讯的主要应用。如图 1 所示，截止 2022 年 6 月，视频业务中，视频用户总数（含短视频）为 9.95 亿，相比 2021 年 12 月增长 2.05%；其中短视频用户规模为 9.62 亿，较 2021 年 12 月增长 3%；网络直播用户规模为 7.16 亿，较 2021 年 12 月增长 1.85%^{[1][2]}。



来源：CNNIC第49次和第50次《中国互联网络发展状况统计报告》整理

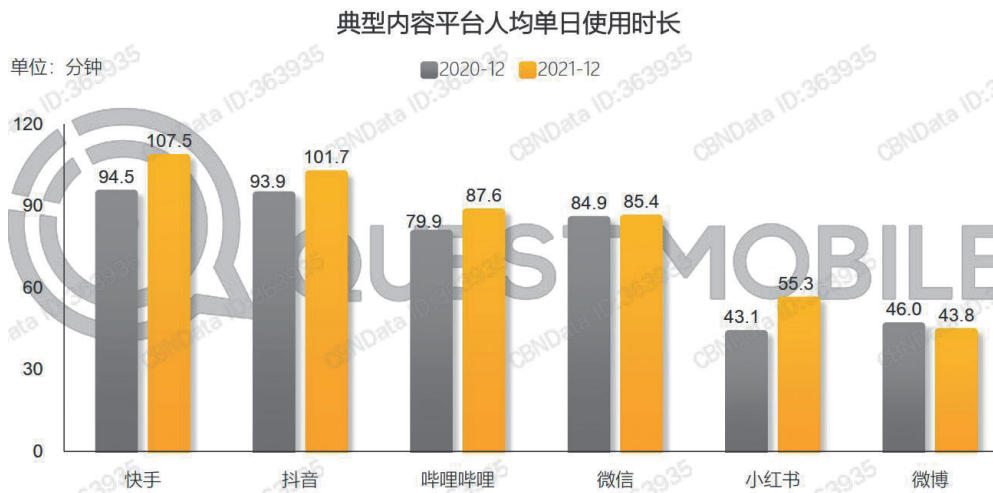
图1 视频业务用户规模数统计

视频是用户沉浸时长最多的应用。如图 2 所示，截至 2022 年 6 月，我国网民的人均每周上网时长为 29.5 个小时，较 2021 年 12 月提升 1.0 个小时。我国网民平均每天上网时长超过 4 小时)，而其中，有接近一半的时间是花在短视频类平台上，如抖音、快手。



来源：CNNIC第50次《中国互联网络发展状况统计报告》

图 2 视频业务月活用户数统计

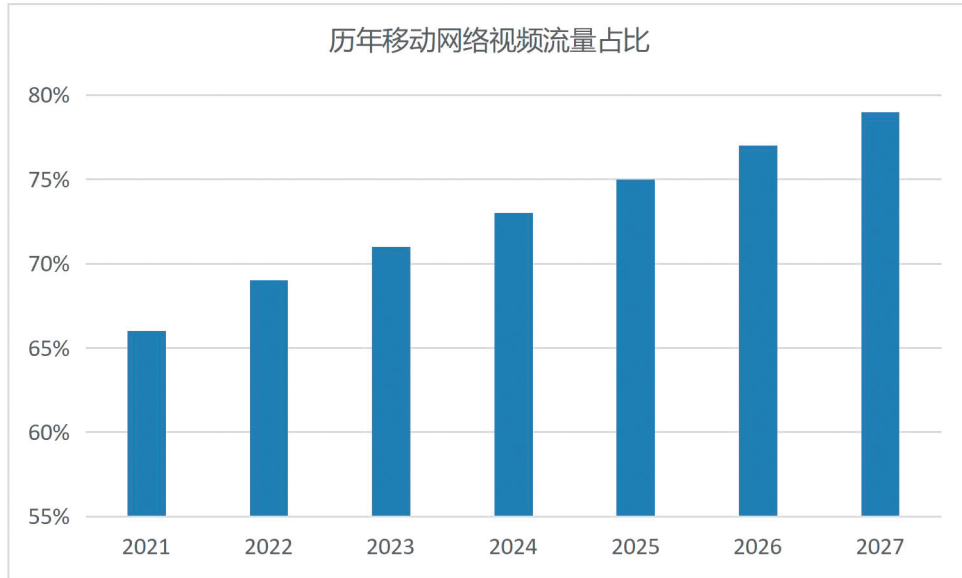


注：人均单日使用时长，指在统计周期内，所有用户启动该App所消耗的总实际有效使用时长。实际有效使用时长是指该App程序界面处于前台激活状态的时间，App在后台运行的时间，不计入有效使用时间。

来源：《QuestMobile2021中国移动互联网年度大报告》

图 3 用户日使用时长平台统计

此外，视频类业务已经成为现阶段移动网络流量的主要来源。根据2022年6月《爱立信移动市场报告》，视频类业务流量目前占全部移动流量的66%，预计到2027年将增长到79%^[3]。

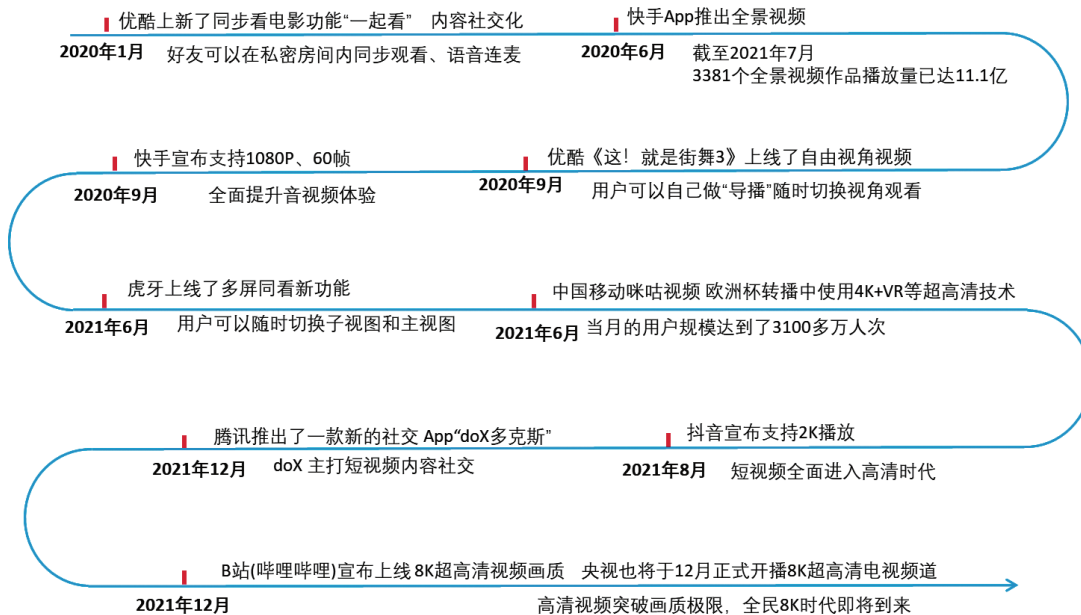


来源：2019年~2022年爱立信《移动市场报告》

图4 移动网络中视频流量占比统计

1.2 移动视频业务形态发生很大变化

如图5所示，从2020年开始各视频应用平台就纷纷上线差异化功能提升用户观看体验。我们看到视频内容的清晰度不断提升，交互性也越来越强，同时整个视频内容也逐步从平面转向3D化。



来源：互联网公开数据整理

图5 视频应用发展趋势

移动视频的业态发展可以归结为以下几个方面：

- A. 超高清化：超高清化的表现在于视频分辨率的提升。如手机移动端逐步从推出 1080p 内容源的观看逐步过渡到 2K, 4K 甚至更高。尤其是移动端视频应用平台支持 180°/360°全景视频以及自由视角视频的播放，用户在手机移动端观看 1080P 以上分辨率内容源也有了强烈诉求。
- B. 空间化：视频的空间化是相比于平面视频最显著的变化。移动端观看平面视频往往是基于内容源制作拍摄时提供的固定有限个视角。而空间化的视频，如 180°/360°全景视频、自由视角视频等可以让用户通过手机移动端触屏以及陀螺仪进行操作，选择视频内容源中想要的视角进行观看。当然，空间化也带来了视频的交互化，用户的交互程度逐步提升。
- C. 交互化：视频交互化分位两个方面。一是用户与视频内容的交互，即上述提到的由于视频的空间化发展带来的人与内容的交互。另一方面是用户与用户的交互，如优酷推出的“一起看”，腾讯的短视频内容社交。这类更多的是以视频内容为载体，打通不同用户之间的社交障碍，提升视频观看的趣味性，进一步提升用户的观看体验。

1.3 移动视频对推动产业升级发挥重要作用

2021 年，短视频用户规模持续增长，行业依然保持稳定增长态势。短视频平台一方面加速布局知识领域，推动知识传播；另一方面不断与传统产业融合，创造出更大的经济价值。短视频推动知识传播，成为信息传播的重要渠道。2020 年以来，各大短视频平台一方面大力扶持内容创作者，鼓励泛知识内容产出；另一方面积极开发出诸如视频合集的新功能和直播课等新形式，打造多层次、立体化的知识图谱。在广度上，平台知识内容已涵盖生活、教育、人文、财经、军事等众多领域，充分满足用户多元化需求；在深度上，平台通过推出视频合集等功能、打造名校名师直播公开课等形式，促进知识体系化传播，提升知识学习深度。2021 年，抖音上线了四期“萌知计划”，投入百亿流量扶持知识创作者，鼓励创作更多适合青少年人群学习的知识内容；快手推出了两季大型直播活动“快手新知播”，为用户提供全新的认知角度与获取知识渠道。短视频与农产品上行、文旅产业深度融合，激发经济活力。一是短视频应用助力农产品销售。源头农户、商家通过短视频、直播来宣传和推介优质农产品，为农产品进城打开销路。数据显示，2021 年 1 至 10 月，快手有超过 4.2 亿个农产品订单经由直播电商从农村发往全国各地，农产品的销售额和订单量与 2020 年同期相比，分别增长了 88%和 99%。此外，短视频平台还为农民和乡村创业者提供专业培训，保障农产品短视频、直播销售模式的

可持续发展。二是短视频应用激发文旅产业活力。在文化产业层面，短视频平台通过加强流量扶持、提高变现能力、打造开放平台及开展城市合作等方式，培养挖掘年轻一代对非物质文化遗产的了解和好奇心，帮助发掘“非遗”的文化和市场价值。在旅游业层面，短视频平台不断加强与西安、重庆、南京等城市的合作，吸引旅游景点入驻宣传，助力城市形象传播和推广，带动旅游业发展。

1.4 当前业界缺乏移动视频体验评价体系和标准

目前业界在 DVB、IPTV 和 OTT 互联网电视领域已经形成了较完善的视频用户体验评价体系和标准（如：YD/T 3776-2020、YD/T 3777-2020、YD/T 3778-2020、YD/T 3779-2020 等宽带视频服务用户体验评估系列行业标准），并且基于该标准的监测探针产品和系统已经在广电和电信运营商现网中广泛部署，有效保障电视大屏用户获得良好的体验质量。然而，在移动短视频及 UGC 直播业务领域，业界尚未形成较完善的用户体验评价体系、测试方案和相关标准。本白皮书旨在提出一套行之有效的移动短视频和 UGC 直播业务的测试指标、测试方案、测试数据，从而为推动相关标准的制定提供强有力的数据支撑和实践依据。

2. 定义统一的移动视频体验评测方案对产业健康发展至关重要

2.1 移动视频服务下影响用户留存的因素呈现多样性

随着5G网络建设，互联网基础设施的升级，近十年来，各类视频应用掀起了视频互联网化的浪潮，视频的制作和分享正在大规模地从传统的广播电视（广电）向互联网、OTT迁徙。移动视频服务的多样性和视频形态的多样性导致影响用户体验以及留存的因素呈现多维化趋势。

在ToC（即面向大众消费者）侧，用户对超高清视频的观看体验有了更高质量追求。以“帧享”为例，它是由优酷发起，视频平台、影视制片商和终端设备商联合推出的超高清解决方案，通过“帧享超高清云+端+智能”相结合的技术，将视频生产和终端播放联动起来，打造极致影音播放效果。数据统计表明，相比观看

1080P分辨率的用户，观看“畅享”超高清的用户，其会员转化率提升了67%。

在短视频领域，能否快速打开感兴趣的短视频是影响用户留存与播放数据的直接因素。优酷在短视频场景中，对视频首屏时延进行了优化提升。优化后，Android端的单用户VV（Video View 视频播放次数）增长20%+；iOS端增长29%+，双端均提升到60%+。

在ToB（即面向企业用户）直播领域，如图6所示，企业直播由最初的单向直播向互动直播演进，不论是电商直播时主播回答观众对产品的疑问，或是数字会展中参展商与潜在客户的一对一互动，突破多人观看的局限，互动性的不断增强让企业的直播活动衍生出丰富多样的玩法。企业直播构筑了企业与用户双向沟通的桥梁，成为一个新交互窗口，借助直播企业客户可以展示自身形象、传达能力与定位，向受众传递生动的“名片”。ToB直播中的许多元素都会影响企业客户留存拉新，比如画面流畅度、功能是否简洁明了、主播形象与素养等。



来源：互联网公开图片

图6 企业直播演进历程

在内容全网分发的时代，要在视频领域取得突破，非常重要的一个方面是平台是否拥有适合视频内容生长的氛围。就知乎而言，作为国内著名的知识类社区，以问答为主要形式的解释类中视频与其社区氛围高度契合。知乎内容视频化趋势的加强，促进了优秀内容的创作和高效流通，具有典型“获得感”特征的优质视频领域，日活跃用户渗透率达到40%。

移动互联网时代，直播行业也变得更加体系化。目前直播领域如火如荼，涉及娱乐、游戏、电商、社交等人们看到和听到的一场场生动直播，涉及到信号的采集，传输，最终在手机上呈现，每个环节都影响着用户的视听体验。比如在电商行业，相当数量的主播以及机构已经意识到直播体验对于成交的重要性，并在分辨率、码率、首屏时延、端到端延迟、卡顿率等维度的提升优化进行了重点投入。

2.2 OTT 视频服务创新，提升用户体验

根据 2021 年 6 月份发布的《2021 中国网络视听发展研究报告》[4] 显示，我国 9.44 亿网络视听用户里，28.2% 会选择倍速观看视频，尤其是 00 后用户群体有近四成选择倍速观看方式。同时，如图 7 所示，火山引擎发布的调研结果表明，当用户的首屏时延超过 200ms 时，用户退出当前播放视频的概率会越来越大。最理想是首屏时延控制在 100ms 以内。以上数据表明，视频观看形式出现越来越多样的形式变化，因此研究交互情况下的视频体验指标具有积极的现实意义。



来源：互联网公开数据

图 7 火山引擎视频首帧时延与用户行为调研

为提升用户体验、创新内容形态，不同平台会采用不同措施来推动用户的留存拉新。比如，优酷率先在端侧落地了“自由视角”能力，用户在观看过程中可自由选择观看角度。以《这！就是街舞 第四季》为例，自由视角总 VV (video view, 视频播放量) 近 150 万，总 PV (page view, 页面浏览量) 达 140 万+，显著提升了用户观看的趣味性与沉浸感。如图 8 所示，优酷参与的“科技冬奥”重点专项——“冰雪项目-交互式多维度观赛体验技术与系统”在北京冬奥成功投入使用。



来源：互联网公开图片

图 8 优酷自由视角现场拍摄方案

用户在观看视频时，越来越习惯“弹幕”这一互动形式，通过弹幕表达观感、与其他用户进行交流也逐渐成为一种观影习惯。然而当弹幕数量过多时会遮挡画面，导致以观影为主的用户在体验上不及弹幕之前。基于此，不少平台也推出了“弹幕穿人”的功能。以某视频为例，从用户反馈来看，超过半数的用户认为弹幕穿人符合自己的需求；从客观数据表现来看，弹幕穿人上线后用户普遍提升了弹幕开启行数，更乐于通过弹幕进行互动分享了。

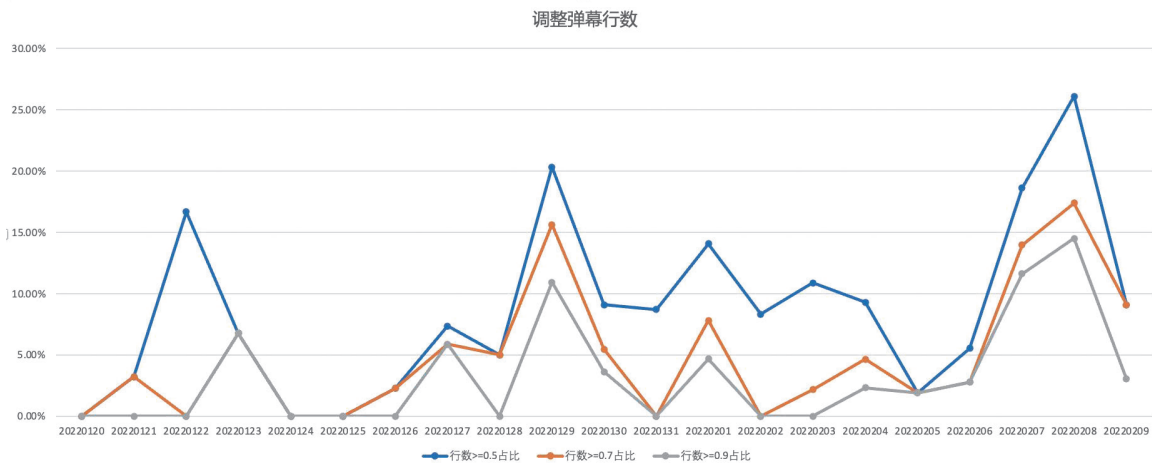


图 9 弹幕穿人投放时间点（2022 年 1 月 27 日）之后，用户在打开调整弹幕行数面板之后，更大可能性选择更多弹幕行数。

用户对于超高清的追求正在越来越迫切，越来越普及。“畅享”在超高清视频体验上，除了对视频源进行质量评估，还构建了一套完备的设备认证体系，对海量设备的芯片解码和运算能力、屏幕的刷新率、分辨率、亮度、色域以及显示效果进行认证。针对不同厂商对终端设备的后处理能力不同的现状，尤其是中低端设备的画质处理可能会导致最终的显示效果变差的情况，畅享利用云端海量的算力和复杂的算法进行分析，将分析的算法信息给到终端上，终端设备结合源上内容信息做画质处理。畅享通过云+端+智能的云端联动方式确保用户在终端上观看到最佳的效果，数据表明畅享超高清的会员转化率要高于 1080p。



图片来源：互联网公开图片

图 10 《长安十二时辰》1080p vs 帧享（需放大观看）

电商直播相比于如泛娱乐的其它视频应用，成像环境包括光照更加纷繁复杂，成像设备更加参差不齐，用户明显会更倾向于码率低但画质高的直播。如今淘宝作为头部电商平台，正致力于通过长期的技术投入和创新，为平台上的消费者提供比肩广电级别的视频画质体验。通过和直播间成像环境、特定场景、硬件、传输配合的全链路优化，淘宝直播技术用很低的码率支持高清 1080P 带货直播，甚至可将平均码率控制在 1Mbps 左右。



图片来源：互联网公开图片

图 11 淘直播某茶具行业案例：1080p@30fps,码率~0.8Mbps

直播场景中直播流的首屏性能也是影响用户留存与播放数据的直接因素。花椒直播专门对视频的秒开性能进行了专项优化，打开视频直播所需时间降低到优化前的 25% 以下，用户体验有了很大提升，上线后一个月的数据分析表明，iOS/Android 用户的人均直播观看数环比增加了 7.2%。

在 ToB 领域，从 2018 年企业直播元年，直播成为企业的必备能力；到 2021 直播营销元年大量企业认可了直播价值，开启直播营销，短短几年，直播破圈加速飞快发展。与此同时，企业客户对直播的互动体验要求也越来越高，特别是对于电商直播和在线教育等强互动直播场景，互动体验几乎可称为衡量直播成功与否的最关键环节。263 云通信通过自研 NRTC 实时音视频底层技术发布低延迟直播，深度优化的直播可将端到端时延最低控制在 200ms 以内。优化后，263 云通信毫秒级的无延迟直播在教育培训、带货直播、赛事直播等领域客户数量大幅增长。

2020 年在新冠防控政策影响下，线上生活由原先短时期的例外状态成为了常态，人类现实生活逐步向虚拟世界迁移，虚拟与现实的边界变得日益模糊，虚拟直播也正是此时进入大众视线。虚拟直播呈现的将不仅仅是现实世界的声音与画面，而是基于真实人物叠加虚拟环境的直播，甚至是由数字人物和虚拟环境组成的 XR 直播。高度沉浸式的直播体验除了具备极强的视觉冲击力，更可以让观众从旁观者变成参与者，企业与目标群体之间的交互将存在更多的可能性。在 263 云通信打造的 3D 虚拟直播场景中，它借助虚拟现实技术构建和监测虚拟场景，把人物行为与虚拟场景合二为一，通过绿幕抠像、实时渲染、超低延迟的互动跟踪等技术，实现虚拟成像与直播现场同台输出，打造出虚实相融的虚拟现实环境，实现了观众从“在线”到“在场”的重要转变，同时也为虚拟世界与现实世界搭建了互动入口，让这两个世界产生更多的交融与连接。如图 12 所示，目前该项虚拟直播已经在医药、金融、教育等领域得到了广泛应用。



图片来源：互联网公开图片

图 12 3D 虚拟直播在教育领域的应用

随着虚拟直播与 VR/AR、互联网、云计算、物联网以及人工智能等新兴技术的进一步融合，势必会迎来交互硬件+沉浸式内容+全场景应用的大融合，实现更为深度的虚实互动，开启虚拟现实产业的商业化新纪元。

2.3 统一的体验评测标准是产业共同且迫切的需要

综上所述，在当下移动视频形态如此多样化的形势下，用户在不同维度上对观看体验的需求展现出来了多元化。同时，平台方针对不同需求的各种创新型能力的研发和提供，使目前用户体验变成了更加广义的且多维化的指标，主观层面包括互动性、易用性、沉浸感、清晰度等，客观层面包括首屏时延、卡顿率等。基于此，统一的体验评价体系、统一的体验评测标准是产业共同且迫切的需要。

3.移动视频业务体验评测方案及现状

3.1 短视频业务

3.1.1 短视频业务流程

如下图13所示，短视频业务体验包括三大块：采集、转码以及消费。

在短视频的采集部分，采集端上传的分辨率码率决定了该视频的最高质量，当前终端侧上传片源以1080P为主，最高支持高动态范围（High-Dynamic Range, HDR）的4K分辨率。

采集部分：服务器端形成不同的分辨率版本，供不同的播放端条件选择。

转码部分：在云端服务器上，存储了大批量采集上来的视频内容。同时基于云的算力对同一视频内容进行不同分辨率、码率版本的转换。同时借助云端的人工智能（Artificial Intelligence, AI）算力对视频内容进行审核。

消费部分：消费侧用户从内容分发网络（Content Delivery Network, CDN）侧获取片源播放，由于是主动推送，所以待推送的优先会放CDN，CDN省内命中率会达90%以上。CDN未命中的场景会回源站取片源，不同地域CDN热点视频存在差异，省/城市级不会再下沉细分。

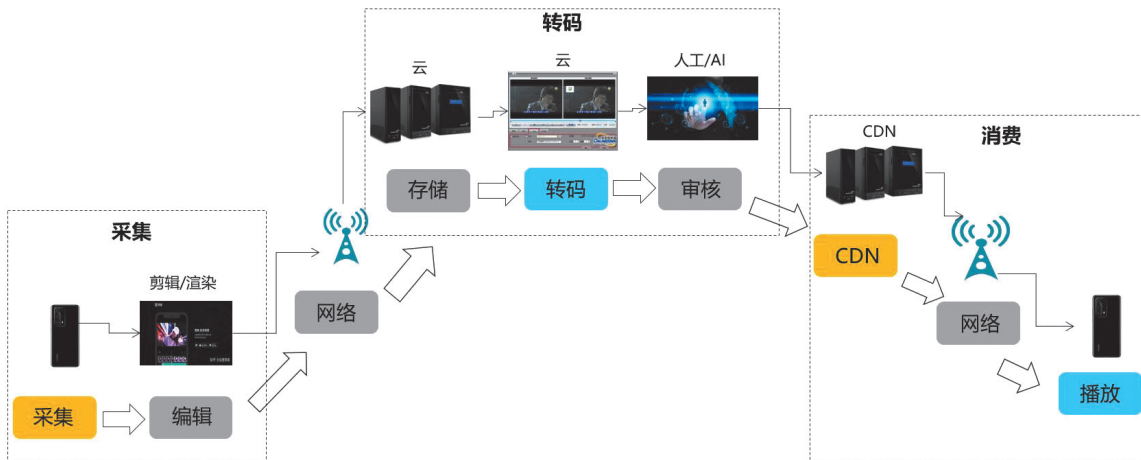


图 13 短视频生产消费流程

当用户发出观看请求时，流媒体服务系统就会根据用户喜好，将存放在片源库中的节目信息主动推送，以视频和音频流文件的方式，通过高速传输网络传送到用户终端。短视频一般采用单播网络来实现，单播是利用一种协议将IP数据包从一个信息源传送到一个目的地，此时信息的接收和传递只在两个节点之间进行。

本文中短视频业务的评测主要聚焦在消费部分的用户体验评测。OTT平台通过在终端侧应用中进行体验指标的打点采集，来获取用户的体验以及行为。当用户观看每一条短视频时，应用侧将会记录下当前用户的体验情况并上报至OTT的数据服务器。本文中的体验数据指标即来源于这类OTT应用的用户监测数据。

3.1.2 短视频体验评测关键指标

短视频通过互联网服务商的移动视频业务系统实现，本文件涉及的短视频涵盖了如下主流网络传输协议：HTTP/TCP，QUIC/UDP等，流媒体传输协议有DASH、HLS等等；采用以下视频编码格式：H.264、H.265、AV1等等；显示设备：智能手机、平板电脑等屏幕尺寸在10英寸以下的移动终端。

A. 基于单一KQI指标的体验评测

用户使用短视频的过程包括启动终端/播放器，通过滑屏选择观看内容，观看等操作。因此影响短视频用户体验的关键指标，从如下三个方面：视听体验质量、呈现体验质量以及交互体验质量中提取呈现。

- a) 视听体验质量，主要与内容清晰度、编码参数等有关，这里提取出码率作为短视频视听质量的关键评测指标。
- b) 呈现体验质量，即用户的收视/收听体验，取决于观看过程中出现的节目信号质量，影响因素包括卡顿、音画不同步等，可以通过视音频信息的传输性能和质量损伤等客观指标进行衡量，此处提取出卡顿作为短视频呈现体验质量的关键评测指标。

c) 交互体验质量，即用户的互动过程体验，取决于系统对用户交互操作的响应速度，涵盖了平台、网络、终端性能指标，此处以短视频观看过程中的首屏时延作为交互体验质量的关键评测指标。

从短视频业务的微观模型以及交互方式上，可以将呈现体验质量中的卡顿和交互体验质量中的首屏时延归一处理。

B. 基于综合体验得分的体验评测

当然，除了基于单一指标开展体验评测，短视频用户体验评测还可以综合考虑视听体验质量、呈现体验质量、交互体验质量等各方面因素。用户的视频体验质量得分，综合了视听体验质量、呈现体验质量、交互体验质量，是真实反映用户体验感知的平均主观评分（MOS）。详细遵循标准T/INFOCA6-2022 - 移动短视频用户体验质量（QoE）评测方法。

3.1.3 短视频体验评测结果现状

A. 5G短视频业务体验评测整体情况

5G短视频业务从首屏时延、码率以及分辨率三个单KQI维度的分布看当前用户的整体表现。短视频业务的用户体验数据来源于联盟多家短视频业务提供商，数据覆盖全国22个省份，5个自治区以及4个直辖市（两个特别行政区以及中国台湾省暂不涉及），数据量总体规模达到万级。在码率中位数1.1M的基础下，短视频的首屏时延仍有超过50%无法满足400ms的基础体验要求。分辨率上有70%的用户可以体验到720P及以上的超高清视频，但仍有提升的空间。

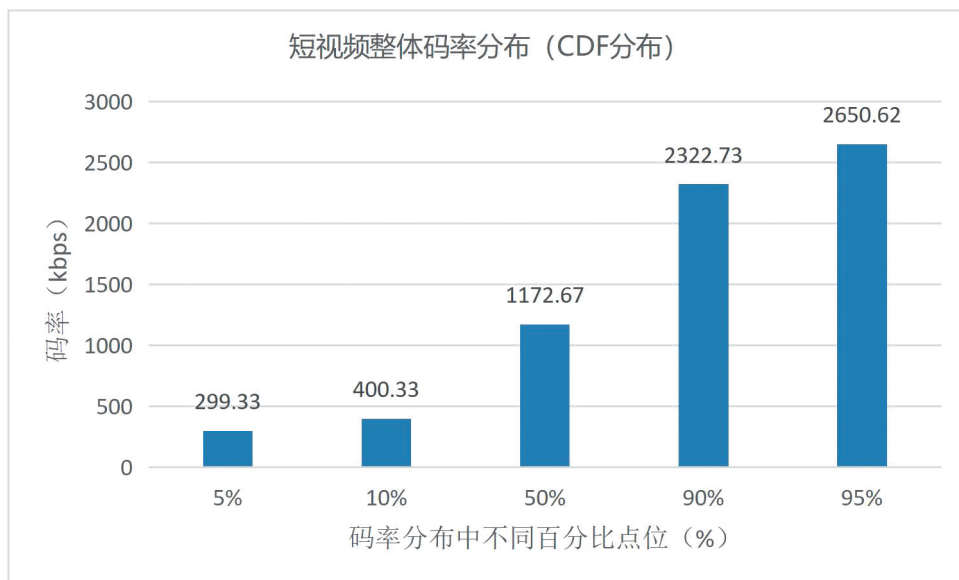


图 14 短视频码率整体分布图

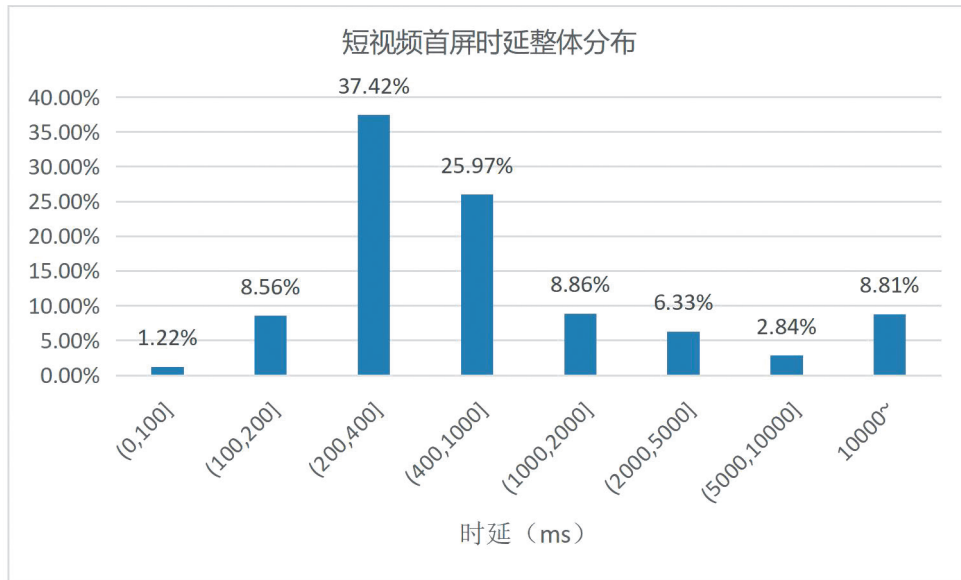


图 15 短视频首屏时延整体分布

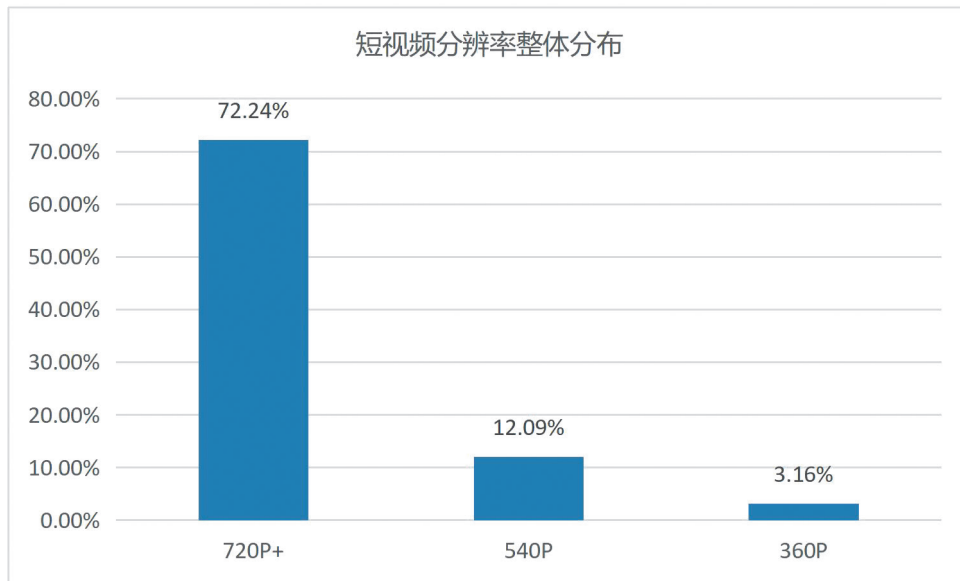


图 16 短视频分辨率整体分布

B. 4G与5G网络下的体验差异化对比

从码率上，4G与5G网络下码率并无明显差异，主要原因是各短视频平台基于网络的码率自适应选择策略各有差异。

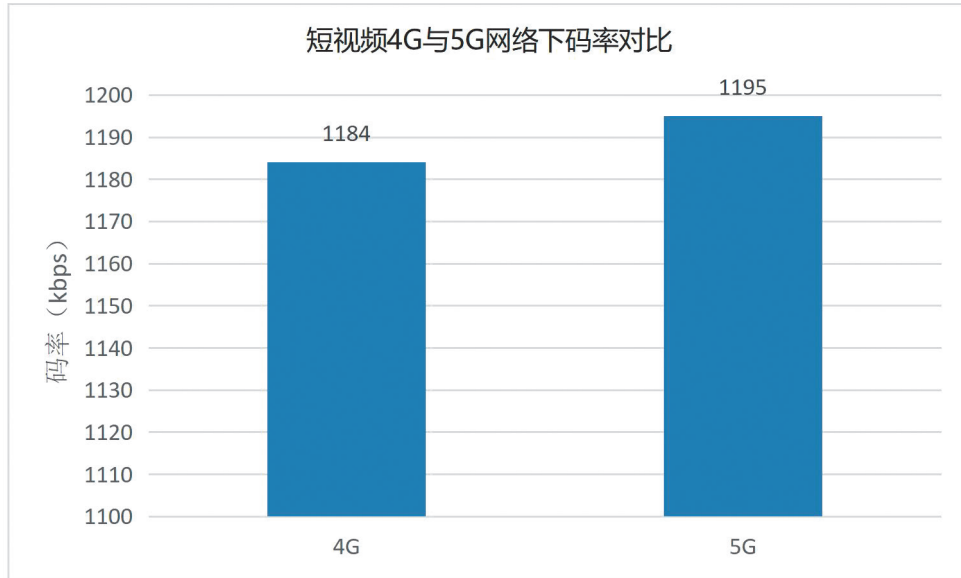


图 17 4G 与 5G 网络下码率对比

4G与5G网络首屏时延对比，5G网络下体验要明显好于4G网络，5G网络下时延降低25%。从这里可以明显看出5G网络带来的巨大优势。但5G网络下首屏时延的均值在300ms左右，距离前文提到的用户期待的100ms极致体验，仍存在较大差距。

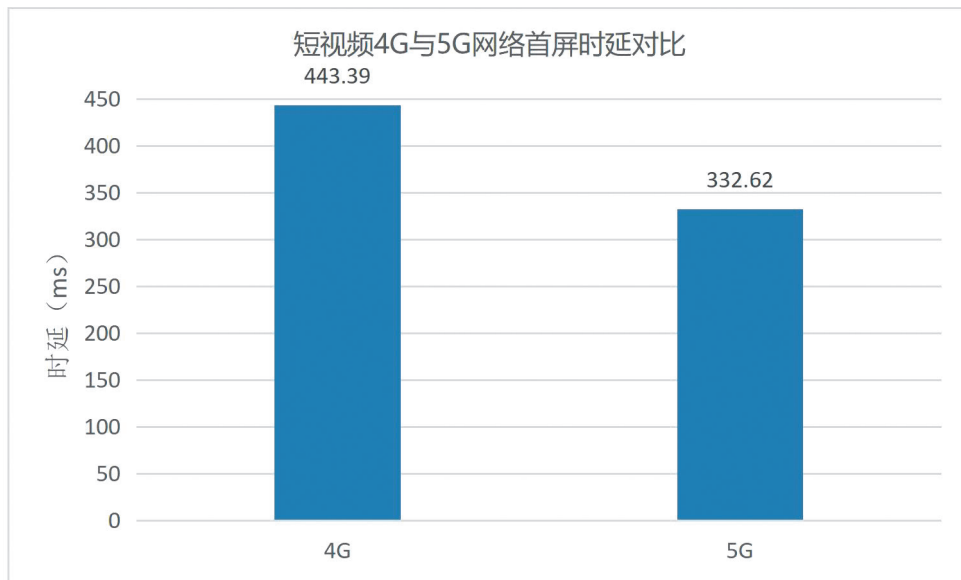


图 18 4G 与 5G 网络下首屏时延对比

从短视频记录中提取出首屏时延在1秒以内的记录，得出4G与5G网络下体验质差（首屏时延超过400ms）占比对比。5G网络相比4G网络质差的占比降低24%。

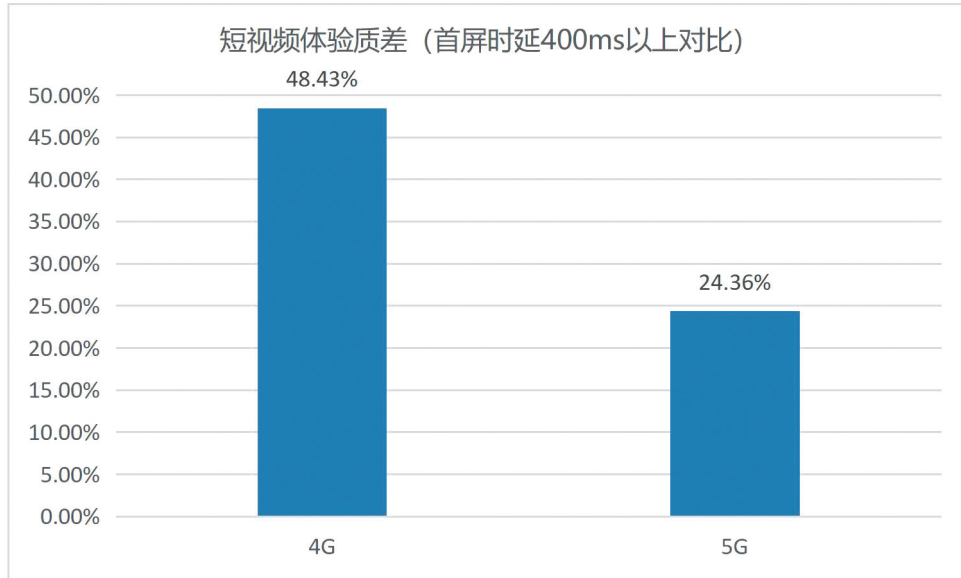


图 19 4G 与 5G 首屏时延对比

C. 各运营商之间5G体验测评对比

从运营商维度进行横向比较，就单个关键质量指标(Key Quality Indicator, KQI)维度，整体首屏时延对比，运营商B < 运营商A < 运营商C。同步看400ms以上的首屏时延占比，运营商B < 运营商A < 运营商C。在短视频的码率表现上，各家差异并不明显。

综合首屏时延以及码率的对比，从现有数据样本分析，运营商B网络下的短视频业务体验最优。

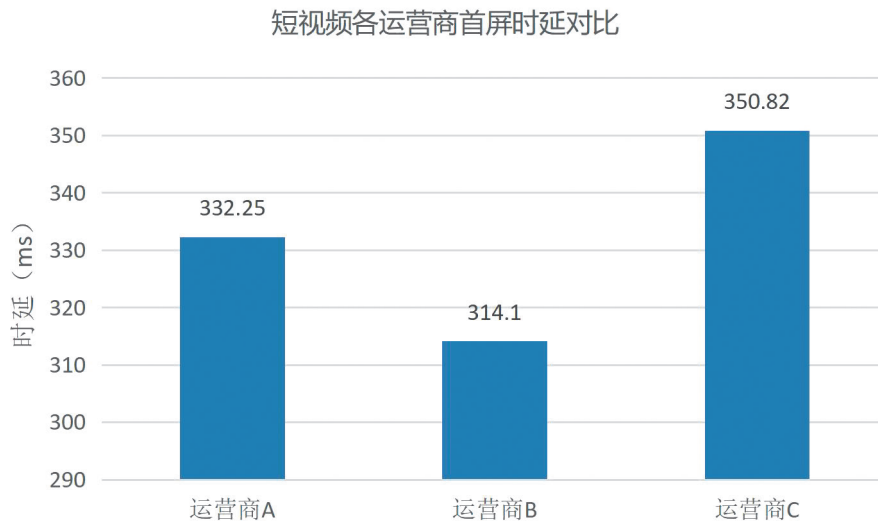


图 20 不同运营商 5G 网络下首屏时延对比 (首屏时延数值优化)

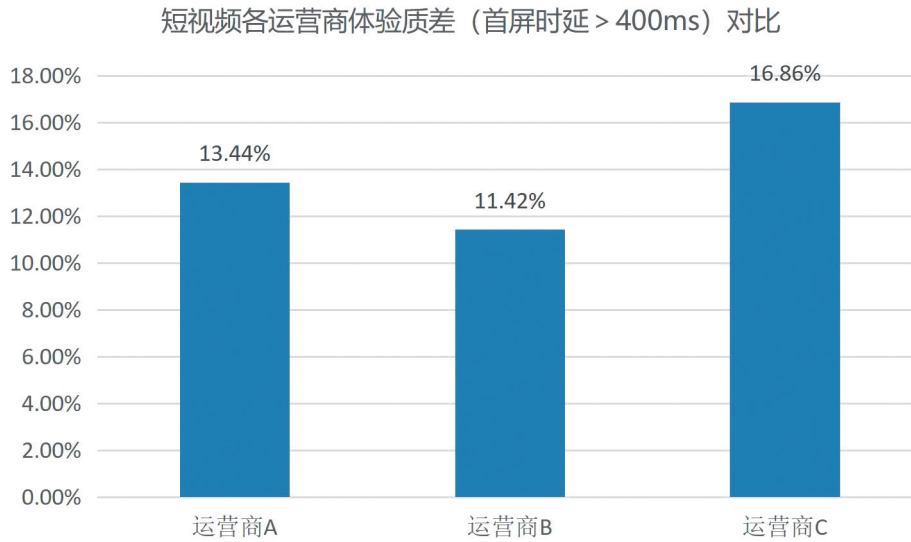


图 21 不同运营商 5G 网络下 400ms 以上首屏时延占比对比

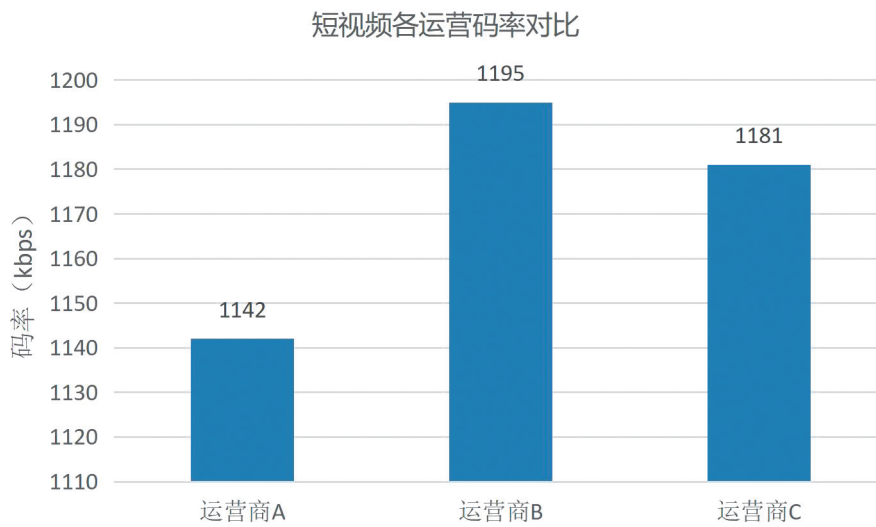


图 22 不同运营商 5G 网络下码率对比

3.2 直播视频（观看端）

3.2.1 直播视频业务流程

移动直播的过程为主播端将产生的音视频流实时地发送到音视频云平台，观众实时接收主播的音视频流进行观看。移动直播过程中主播和观众间会频繁地进行实时音视频互动，或者主播之间进行连麦PK，互动性比较

强，对端到端的时延要求比较高。直播中的用户角色分为主播端和观众端。

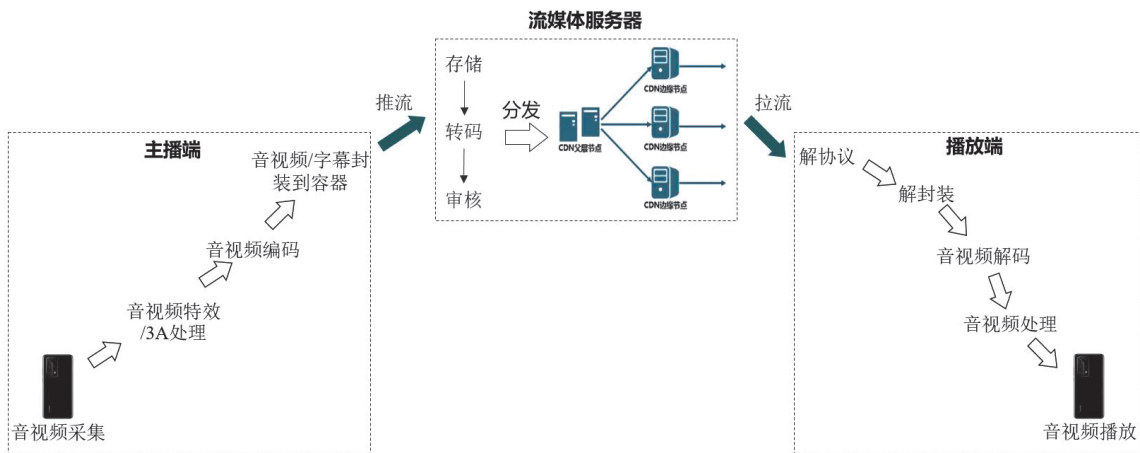


图 23 直播业务流程图

本文中描述的直播视频的业务体验评测主要面向直播观众端侧的用户体验评测。直播（观看端）的业务体验数据采集原理与短视频业务类似，也是在终端侧OTT应用中打点统计获得。

3.2.2 直播（观看端）体验评测关键指标

A. 基于单一KQI指标的体验评测

直播中观众端实时地接收主播端产生的音视频流进行观看。直播中衡量观众端体验的主要指标从视听体验质量、交互体验质量、呈现体验质量三个因素中进行提取。视听体验质量包含音频体验质量和视频体验质量。

- 视听体验质量，取决于音视频的清晰度、流畅度等因素，此处以直播（观看端）的分辨率、码率作为主要体验评测指标。
- 交互体验质量，直播（观众端）交互体验质量也包含观众端在直播间的跳转时延、观众进入直播间的首屏时延等。观众端在直播间的跳转时延是指在主播连麦PK场景下（如主播A与主播B连麦PK）主播A直播间的观众从主播A的直播间跳转到主播B的直播间所需的时间。观众进入直播间的首屏时延是指从观众端点击进入主播直播间到直播间首帧画面成功播放显示的时间。此处重点关注观众端用户观看的体验，提取首屏时延作为主要体验评测指标。
- 观众端呈现体验质量，可以音频卡顿率、视频卡顿率、花屏、音画同步时延等指标进行衡量，此处以观看端视频的卡顿率作为呈现体验质量的主要体验评测指标。

B. 基于综合体验得分的体验评测

与短视频一样，除了基于单一指标开展体验评测，直播（观看端）用户体验评测还可以综合考虑观众端视

听体验质量、交互体验质量、呈现体验质量，是真实反映移动直播观众端用户体验感知的平均主观评分

(MOS)。详细遵循标准T/INFOCA6-2022 - 移动直播视频用户体验质量 (QoE) 观众端评测方法。

3.2.3 直播（观看端）体验评测结果现状

A. 直播（观看端）业务体验评测整体情况

对于直播（观看端）的用户体验，重点分析了5G网络下其码率与首屏时延的分布表现。直播（观看端）业务的用户体验数据来源于联盟多家短视频业务提供商，数据覆盖全国22个省份，5个自治区以及4个直辖市（两个特别行政区以及中国台湾省暂不涉及），数据量总体规模达到万级。通过分析该数据发现，5G下整体码率比短视频要高，其中1M~2M码率区间占比最高，约46%。但首屏时延整体偏大，基本是200ms起步。

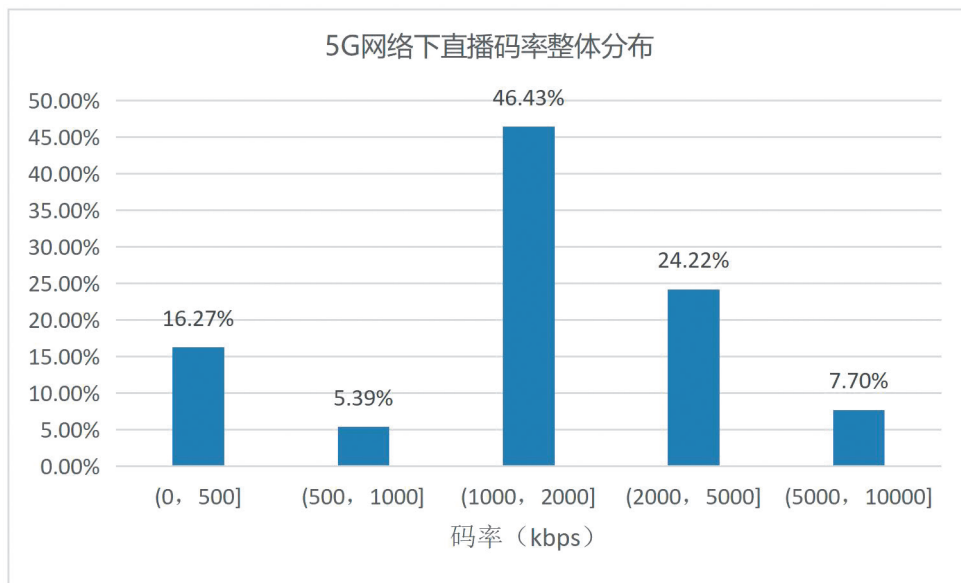


图 24 直播码率整体分布

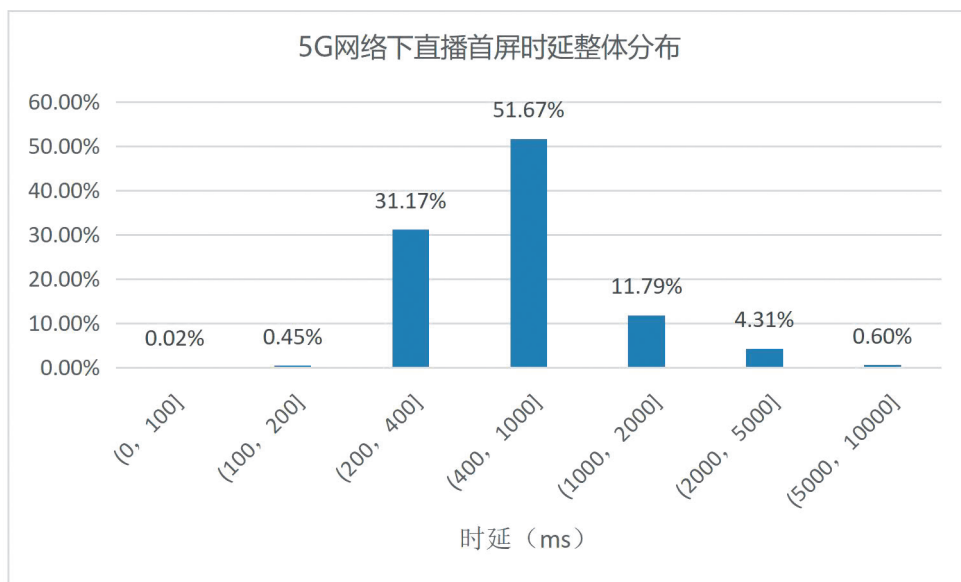


图 25 直播首屏时延整体分布

B. 各运营商5G体验评测对比

同样看直播（观看端）的业务体验表现，在首屏时延方面，各运营商体验质差（首屏时延大于400ms）对比，运营商A最差，运营商B与C相当。从1M以上高码率占比看，运营商B占比最高。显然是由于运营商B网络条件更优能够让用户体验更高码率的直播视频业务。

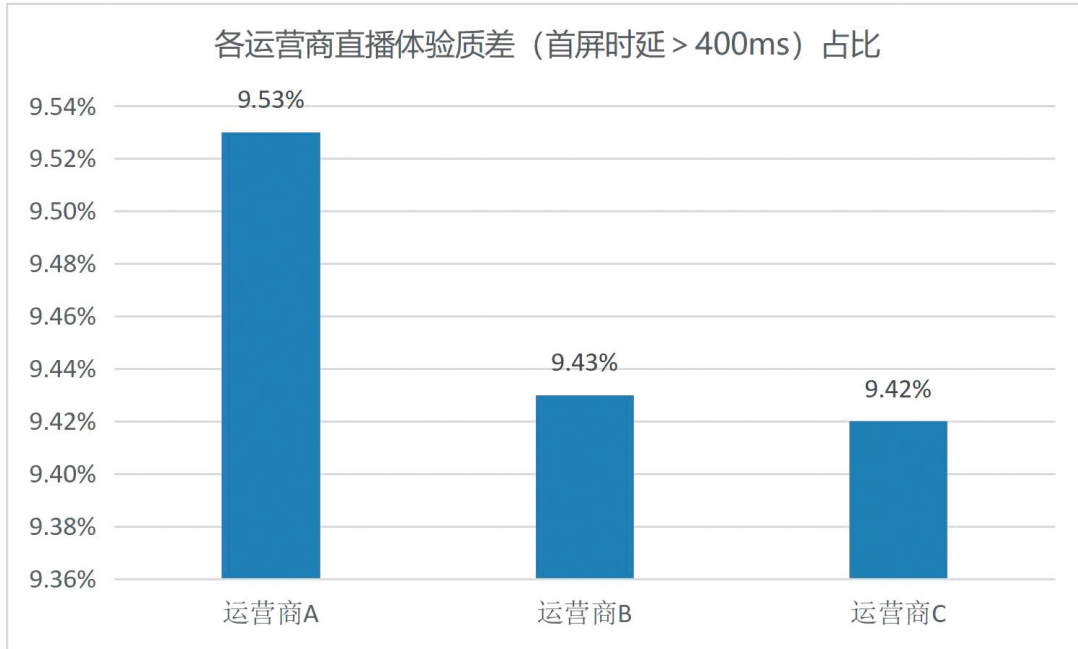


图 26 不同运营商 5G 网络下直播 400ms 以上首屏时延占比对比

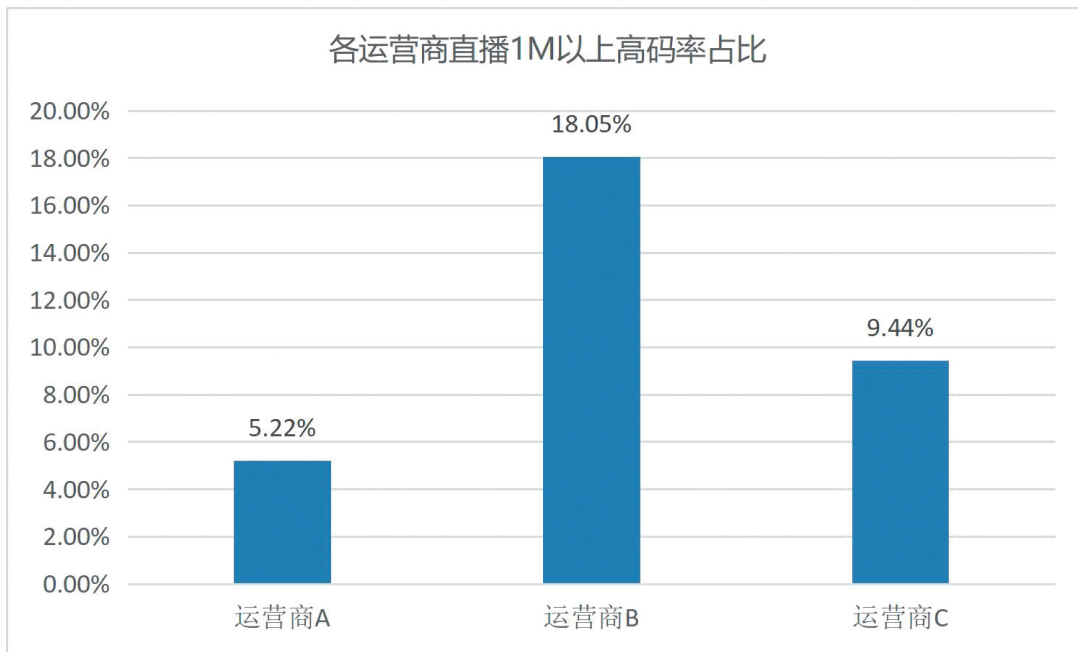


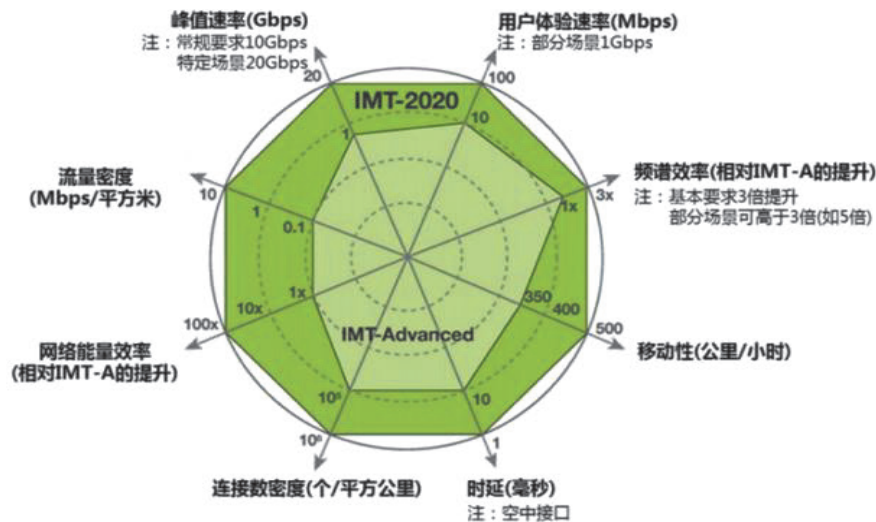
图 27 不同运营商 5G 网络下 1M 以上高码率占比

受限于直播（观看端）业务的数据采集样本量，直播（观看端）的体验评测仅包含了整体数据情况描述以

及运营商之间的体验对比两个维度，4G 与 5G 网络下的体验差异对比以及城市间的体验对比并未涉及。

4.展望

5G（第 5 代移动通信技术）被认为是信息领域自万维网问世以来最大的技术突破。业内人士认为，5G 将推动第四次工业革命，变革人们的生活与生产方式。5G 关键技术主要体现在超高性能无线传输技术和高密度无线网络技术中。如下图所示，与以 IMT-Advanced 为代表的 4G（第 4 代移动通信技术）相比较，5G 传输速率增加了 10 到 100 倍，峰值传输速率为 10 Gbps，端到端延迟可达毫秒量级，连接的设备密度增加 10 到 100 倍，频谱效率提高了 5 到 10 倍，在 500km/h 的运动速度中保证良好的用户通信体验。5G 网络性能的提升不仅突破了时空限制，更实现了人与物的互联，是跨时代的飞跃。



图片来源：《5G经济社会影响白皮书》

图 28 4/5G 关键性能对比

4.1 5G 时代的视频技术基本趋势

5G 以超越 4G 至少十倍的峰值速率、毫秒级的传输时延和千亿级的连接能力，深度融合人工智能、虚拟和增强现实等先进技术，成为基础性的生产力，颠覆并重构网络传输、内容生产及终端接收各个环节，为视频赋能，带来信息传播交互更极致的体验。5G 时代，视频将向更实时、高清、沉浸、交互的方向发展。

4.1.1 超高清视频流传输

5G 网络可有针对性地解决超高清视频大数据量传输的痛点。当前，4G 网络带宽有限，4K/8K 视频传输和在线播放无法取得预期效果，用户在观看视频时常常会受到播放卡顿、花屏或视频分辨率自动降低等影响。5G 更快的传输速率和更大的网络容量将彻底解决这一问题，用户在下载和观看时都能获得更好的体验，超高清视频或成为 5G 视频标配。除了视频画面清晰度大幅提升外，5G 能够为超高清视频提供稳定的实时传输能力，比传统线缆传输更加灵活，不受空间限制，能满足更灵活的超高清视频传输需求，也就是说，5G 高带宽、低延时的优势将能同时实现超高清视频的移动化和实时播放，超高清视频直播业务迎来广阔应用前景。2022 年 1 月 6 日，北京冬奥列车暨高铁 5G 超高清演播室在京张高铁清河站上线，利用 5G 技术，实现奥运赛事直播随时随地高清直播。

4.1.2 虚拟现实视频及互动体验

XR（扩展现实）是 VR（虚拟现实）、AR（增强现实）和 MR（混合现实）的总称，将在 5G 技术支持下释放全部潜能。高质量的 XR 体验对网络传输能力和网络时延要求极为苛刻。4G 环境下，渲染能力不足、互动体验不强、终端移动性差等，使其在 2015、2016 年短暂的喧嚣后归于沉寂，未实现大规模商用。5G 有望弥补 XR 的应用短板。一是高带宽、低延时的特性，意味着信道传输容量大幅提升，能够提高虚拟世界与真实世界的交互效率，保障 XR 的业务体验。二是 5G 支持下的边缘计算及云计算技术，能解决 XR 用户端依赖高性能计算硬件设备的问题，保证人们可以在易携带、高性能的终端设备上享受 XR 内容，实现无绳化、轻量化，以全新方式在生活与工作中与虚拟世界交互。5G 与增强现实或虚拟现实技术结合使用，为视频体验带来许多新的可能。5G+360°VR 实时全景视频可以开启“上帝视角”，让每一名用户都可以自主选择观看角度、画面和具体内容，如在体育比赛或活动直播中锁定自己喜欢的人物来观看等。此外，5G 可赋能全息技术，异地同步的 5G+VR 直播视频，让不同地点的用户置身同一时空环境，在远程互动教学、会议、医疗等诸多场景中可广泛使用。2020 年 4 月 15 日起，中央广播电视总台央视频 5G 新媒体平台联合中国电信，推出“珠峰十二时辰”系列慢直播，依托中国电信“5G+云网”技术，通过 4K 高清画面和 VR 视角，首次向全国观众 360 度全景呈现珠峰 24 小时实时景观变化，万千网民足不出户就能以“身临其境”方式欣赏到珠峰的壮美与险峻。另有研究表明，5G 将有利于触觉反馈装备的研发和使用，将为 AR/VR/MR 视频增加新的感知维度，进一步升级娱乐体验。爱立信公司的消费者调查报告指出，5G 时代消费者每周在家之外，或多花 3 个小时在移动设备上观看视频，其中三分之一的时间将使用 VR/AR 眼镜，且在全球范围内，有 50% 的消费者认为到 2025 年，所有人都将佩戴 AR 眼镜。

4.1.3 内容生态“超视频化”

5G 与物联网、人工智能等新技术将共同驱动一个泛媒介、泛视频的时代，5G 技术上的优势强势改变着视频的传播生态和内容格局，视频或成为主要社会表达，占据内容传播的主力位置。5G 与 4G 相比最重要的进化，在于推动世界进入万物互联时代。基于多用户共享接入（Multi-User Shared Access, MUSA）等 5G 核心技术，低成本、低功耗的 5G 海量连接得以实现。5G 连接数密度达到 100 万个/平方公里。高强度的万物互联，让人们可以利用各种设备作为互联网的接口与虚拟世界连接，做到无所不传，也有望让视频深入垂直行业，如智慧城市、智慧医疗、智慧农业等，成为信息记录、传输交互的新载体，令万物“视”联走进现实，丰富视频使用场景。另外，5G 与 4K/8K、AR/VR、AI 等技术融合，动作捕捉、实时互动、虚拟人物等新视频技术加速迭代渗透，虚拟世界与现实世界的界限进一步模糊，将催生更多样的视频内容形态。专家预测，5G 将会影响所有的移动互联网应用业务朝着“视频流”化趋势发展，包括虚拟现实等类型的“超视频化”方向发展，极大地丰富人们的工作生活、娱乐体验。

4.2 内容主导的移动短视频业务体验新特性

目前面向公众客户移动视频，主要是在各种新媒体平台（如抖音、快手等）上播放的、适合在移动状态和短时休闲状态下观看的、高频推送的视频内容（以竖屏方式观看内容为主），时长为几秒到几分钟不等。

用户点播视频的过程包括启动终端/播放器，选择观看内容，进行点播、观看等操作。影响点播视频用户体验的因素来自终端的软硬件、传输协议、编解码、网络、平台、内容源质量等多个方面；其中需要考虑的关键因素可以被归纳到视听体验质量、用户的交互体验质量和观看体验质量等三个方面。

实时音视频互动直播的过程为主播端将产生的音视频流实时地发送到实时音视频云平台，观众实时接收直播的音视频流进行观看。实时音视频互动直播中主播和观众需要频繁地进行音视频互动，或者主播之间进行连麦PK，互动性比较强，对端到端的时延要求比较高。直播中的用户角色分为主播和观众。

主播端进行直播的过程为主播端将产生的音视频流实时地发送。衡量主播端质量的主要指标有音视频质量、交互质量、主播端传输质量。

直播中观众端实时地接收主播端产生的音视频流进行观看。直播中衡量客户端体验的主要指标有音频体验质量、视频体验质量、交互体验质量、观众端呈现体验质量。

“内容为王”。用户观看移动视频时的体验，越来越体现出内容主导的趋势，即如何提升视频内容对于用户的吸引力。这将涉及以下方面：一是用户对于视频内容的价值评估，包括实用指导性和感性体验；二是用户

对于主播的认可度和忠诚度；三是用户的兴趣偏好。对于这些内容主导的视频体验新动向，需要挖掘新的采集参数如用户观看同一节目源、同一主题类视频、同一主播的次数、观看中是否有倍速操作等来进行全面评估用户的综合体验。

此外，随着视频业务向XR、元宇宙等新的方向演进，需要采集更多的设备交互、用户行为等方面的关键参数，进一步拓展用户体验评估模型。

5. 附录

5.1 术语解释

下列的术语和定义适用于本文件：

■ 移动互联网

移动互联网强调使用蜂窝移动通信网（如4/5G）接入互联网，因此常常特指手机终端采用移动通信网接入互联网并使用互联网业务；而无线互联网强调接入互联网的方式是无线接入，除了蜂窝网外还包括各种无线接入技术（如Wi-Fi）

■ 移动视频业务

服务商通过各类移动终端（手机、PAD）提供的各类视频播放业务，包括Wi-Fi和移动网络（4G/5G）接入等不同形式。

■ 短视频

短视频是指在各种新媒体平台上播放的、适合在移动状态和短时休闲状态下观看的、高频推送的视频内容，几秒到几分钟不等。内容融合了技能分享、幽默搞怪、时尚潮流、社会热点、街头采访、公益教育、广告创意、商业定制等主题。由于内容较短，可以单独成片，也可以成为系列栏目。

■ 直播视频

直播视频分为两类，一类是在网上提供电视信号的观看，例如各类体育比赛和文艺活动的直播，这类直播原理是将电视（模拟）信号通过采集，转换为数字信号输入电脑，实时上传网站供人观看，相当于“网络电视”；另一类是“网络直播”：在现场架设独立的信号采集设备（音频+视频），再通过有线或无线网络上传至服务器，发布至网址供人观看，观众也可以跟主播进行实时互动。除特殊情况下另作说明外，本白皮书特指后者。

■ 点播加载时长

用户使用点播服务观看无片头广告的视频内容时，从用户触发播放操作到终端播出首画面的时间间隔。

■ 频道切换时长

用户使用电视直播服务时，从现有直播频道切换到另一个直播频道，通过遥控器等操控设备触发切换操作到完成新的直播频道画面出现的时间间隔。

■ 播放时长

用户观看视频过程中，用户观看视频时长总和/播放总次数，采用全部测试样本的统计平均值。

■ 卡顿时长

用户观看视频过程中，发生画面卡顿时长总和/卡顿总次数，采用全部测试样本的统计平均值。

■ 卡顿次数

用户观看视频过程中，发生卡顿次数总和/播放总次数，采用全部测试样本的统计平均值。

■ 卡顿时长占比

用户观看视频过程中，发生卡顿时长总和/播放总时长，采用全部测试样本的统计平均值。

■ 卡顿率

用户观看视频过程中，发生卡顿的播放次数总和/播放总次数，采用全部测试样本的统计平均值。

■ 无卡顿率（零卡顿率）

没有发生卡顿的播放次数总和/播放总次数，采用全部测试样本的统计平均值。

■ uVES 评分

根据uVES标准规定的计算方法，对从测试终端获取的视频源质量（分辨率、帧率、编码方式、码率、画面复杂度）、交互体验（点播加载时长、频道切换时长）、播放体验（点播卡顿率、直播花屏率）的数据，综合计算得出评分；可以用来衡量用户观看视频的主观体验。

■ 平均下载速率

从视频播放终端统计的某个时间段内的“媒体下载数据量/下载时长”，通常以每秒比特数（bit/s或bps）来衡量，其中下载时长不包括没有实际下载数据的空余时间。

■ 质差率

在一个统计周期内，用户观看视频的体验质量低于质差阈值（v）的时间超过规定的时间，则将其定义为质差用户。质差率=质差用户数/总用户数

■ 业务范围

本白皮书所涵盖的视频业务目前包括：移动蜂窝网下通过移动终端（手机、Pad）访问的短视频，长视频及直播业务。

■ 数据来源

本白皮书采用的与国内行业发展相关的数据来源于工业和信息化部正式发布的数据、视频体验联盟成员提供的数据以及业界公开报告的数据（在引用的地方有标注）。

本白皮书中采用的移动视频业务相关数据来自联盟成员中视频服务商提供的现网抽样统计数据，广视频业务/IPTV业务/OTTV业务的相关数据来自联盟成员的现网抽样点测的获取数据。

■ 用户体验质量抽样统计方法

由联盟成员中视频服务提供商统计各自平台上连续一段时期内的移动终端在线服务的数据，然后对多个视频服务提供商来源的数据进行加权平均计算。

■ 抽样点测的组织实施

为了验证联盟单位提供的统计数据的有效性，联盟组织成员单位实施了用户体验质量的抽样点测。由上海交通大学、德科仕通信（上海）有限公司、上海艾策通讯科技股份有限公司为主负责测试流程定义和提供数据分析工具，由北京邮电大学网络和交换国家重点实验室协助数据整理分析，历时近2个月完成了点测数据的采集、数据分析和效验工作。

5.2 引用

[1] 第50次《中国互联网络发展状况统计报告》，

<http://www.cnnic.net.cn/NMediaFile/2022/1020/MAIN16662586615125EJOL1VKDF.pdf>

[2] 第49次《中国互联网络发展状况统计报告》，

<https://www.cauc.edu.cn/jsjxy/upfiles/202203/20220318171634656.pdf>

[3] 2022年6月《爱立信移动市场报告》，<https://www.ericsson.com/zh-cn/reports-and-papers/mobility-report/reports/june-2022>

[4] 《2021中国网络视听发展研究报告》，<http://www.cnsa.cn/attach/0/2112271351275360.pdf>

[5] QuestMobile2021中国移动互联网年度大报告，<https://www.questmobile.com.cn/research/report-new/222>

[6] 北京师范大学新闻传播学院、字节跳动（中国）公共政策研究院《5G时代短视频长视频的应用场景、发展趋势》

[7] 工信部《2021年通信业统计公报》,

https://www.miit.gov.cn/gxsj/tjfx/txy/art/2022/art_e8b64ba8f29d4ce18a1003c4f4d88234.html

[8] 工业和信息化部、教育部、文化和旅游部、国家广播电视总局、国家体育总局。《虚拟现实与行业应用融合发展行动计划（2022—2026年）》2022年10月28日,

https://www.miit.gov.cn/zwgk/zcwj/wjfb/tz/art/2022/art_775aaa3f77264817a5b41421a8b2ce22.html

[9] 工业和信息化部电子信息司《虚拟现实与行业应用融合发展行动计划（2022—2026年）》解读,

https://www.miit.gov.cn/zwgk/zcjd/art/2022/art_acb9482eac994486913b592cebf71b81.html

