



流媒体业务感知解决方案的研究及应用

殷诚, 许清, 奉致富, 陈彬, 周命祥, 曾德强

(中国移动通信集团设计院有限公司湖南分公司, 湖南 长沙 410003)

摘要: 随着短视频与直播业务在移动互联网流量中的占比持续增长, 用户对视频业务“连得上、启动快、播放优”的体验需求日益严苛。系统分析了短视频与直播业务的核心参数体系及其动态平衡模型, 并通过5G网络及Wi-Fi环境下的实测数据, 明确了不同业务场景(包括短视频播放、观众视角直播、主播视角直播、礼物特效及直播连线PK)的感知拐点门限。针对5G网络视频业务中用户感知劣化问题, 提出基于“双维八项”的流媒体感知评估体系, 通过斯皮尔曼秩相关系数建立KQI-KPI关联模型, 量化分析视频初缓时延、卡顿率等关键感知指标与无线网络参数的映射关系。最后, 从参数优化与特性功能应用两个方面提出感知提升方案, 包括弱场边缘能力提升、MAC下行性能优化、波束成形、业务识别保障、QoS调度保障等, 并通过现网应用验证了方案的有效性, 显著提升了视频下载速率、降低了播放缓存时延及RTT时延。

关键词: 短视频; 直播; 流媒体感知; 5G网络优化; KQI-KPI关联模型; 参数优化; QoS调度

中图分类号: TP393

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2025313

0 引言

近年来, 短视频与直播业务爆发式增长, 短视频及直播类应用已成为移动互联网流量的核心驱动力。工业和信息化部2024年统计数据显示, 视频类业务在移动网络总流量中的占比已接近60%, 其中短视频在视频类流量中的占比高达74.7%。以抖音、微信为代表的头部平台, 凭借日活跃用户数超7亿的庞大规模, 进一步巩固了其在内容消费领域的主导地位。短视频的即时性、高互动性特征, 以及直播业务的实时音视频传输需求, 对移动网络性能提出了前所未有的挑战。

1 流媒体相关基本概念与参数体系

1.1 视频流媒体核心参数定义

视频流媒体业务的质量与传输效率由以下核

心参数决定。

(1) 分辨率 (Resolution)

定义: 视频画面中像素点的数量, 以水平像素数 $\times$ 垂直像素数表示 (如720p为1 280 $\times$ 720)。分辨率越高, 图像越清晰, 但数据量越大。

典型场景: 抖音短视频以540 $\times$ 960 (比例9:16) 为主, 直播支持蓝光 (1 920 $\times$ 1 080)、超清 (1 280 $\times$ 720) 等多档调节。

(2) 码率 (Bitrate)

定义: 单位时间内传输的视频数据量 (单位: Mbit/s)。码率与画质正相关, 但受网络带宽限制。

编码标准影响: 相同分辨率下, H.265编码的码率需求为H.264的50%。例如, 1080p视频在H.264下需3~8 Mbit/s, 而H.265仅需1.5~4 Mbit/s。



(3) 帧率 (Frame Rate)

定义：每秒传输的帧数 (FPS)。帧率 ≥ 24 FPS 可保障人眼感知流畅性，抖音默认采用 30 FPS 以提升动态画面表现力。

卡顿阈值：帧率低于 20 FPS 时，用户感知卡顿概率增加 50%。

(4) 压缩编码技术

压缩编码技术指通过去除视频信号中的冗余信息和不可见细节，以及利用人眼对视频画面的特性进行压缩，从而减小视频文件的体积。

视频压缩比一般指压缩后的数据量与压缩前的数据量之比。所以实际传输速率低于计算出来的速率，需要用实际计算出来的速率除以压缩比。

以下为常见压缩编码方式及应用场景。

- MPEG-2：最早期的视频压缩编码标准之一，主要用于 DVD、数字电视等领域，压缩比约 25:1。
- MPEG-4：在 MPEG-2 的基础上增加了更多的压缩技术，适用于网络视频传输和流媒体应用，压缩比约 65:1。
- H.264/AVC：广泛用于在线视频与实时通信，压缩比约 100:1。
- H.265/HEVC：压缩效率提升 50%，适用于 4K/8K 超高清视频，但计算复杂度更高。

1.2 参数关联性与动态平衡模型

视频参数并非独立存在，其关联性直接影响

用户体验与网络资源消耗。

(1) 分辨率-码率-帧率三角关系

固定码率下：分辨率升高将导致单帧数据量减少，画质下降；反之，降低分辨率可提升单帧画质。

(2) 动态调整策略：抖音通过智能算法实时平衡三者，例如，在弱网环境下自动降为 480p 并减少帧率至 24 FPS，以维持播放流畅性。

(3) 压缩效率与带宽需求

H.265 优势：在相同画质下，H.265 较 H.264 节省 45%~50% 带宽。例如，4K 视频采用 H.265 时码率仅需 10~15 Mbit/s，而 H.264 需 20~30 Mbit/s。不同编码标准下的典型码率需求见表 1。

表 1 不同编码标准下的典型码率需求

分辨率	H.264 码率范围/ (Mbit·s ⁻¹)	H.265 码率范围/ (Mbit·s ⁻¹)
480p	0.8~2.0	0.4~1.0
720p	1.5~4.0	0.75~2.0
1080p	3.0~8.0	1.5~4.0
4K	20~30	10~15

2 各类业务实测感知分析

2.1 抖音短视频实测分析

抖音视频组采用预加载和缓冲机制，结合 Wi-Fi 限速和 5G 网络下 CQT 感知测试，确保高清模式视频流畅，下行速率不低于 2 Mbit/s，省流模式下行速率不低于 1 Mbit/s。视频播放缓冲理论模型如图 1 所示。

抖音 App 现场实测采样情况如图 2 所示。

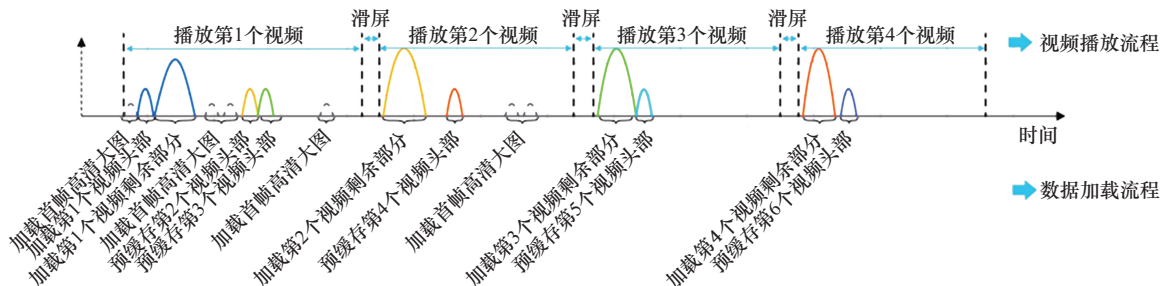


图 1 视频播放缓冲理论模型

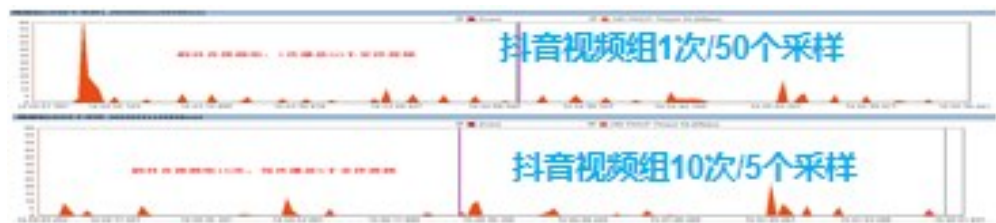


图2 抖音App现场实测

5G网络实测情况：省流模式下行速率0.53 Mbit/s时，经常卡顿，缓冲时间长，但依然能看视频。超清下行速率低于1 Mbit/s，出现卡顿，缓冲时间长，但依然能看视频。实测情况见表2及表3。

表2 5G网络实测情况

分辨率	RSRP/dBm	NR PDCP Thrput DL/ (Mbit·s ⁻¹)	感知情况
省流模式	-100	2.49	正常播放
省流模式	-105	2.07	正常播放
省流模式	-110	1.89	正常播放
省流模式	-115	0.53	经常卡顿，缓冲时间 长，但是能看
高清模式	-100	2.68	正常播放
高清模式	-105	2.46	正常播放
高清模式	-110	2.41	正常播放
高清模式	-115	0.96	经常卡顿，缓冲时 间长，但是能看

表3 Wi-Fi限速感知测试情况

分辨率	下行速率/ (Mbit·s ⁻¹)	现象（测试每个视频 观看约5 s时长）
高清	5	正常播放，无卡顿
高清	4	正常播放，无卡顿
高清	3	正常播放，无卡顿
高清	2	正常播放，无卡顿
高清	1.5	持续滑动视频，偶现卡顿
高清	1	持续滑动视频，大部分视频卡顿
省流	3	正常播放，无卡顿
省流	2	正常播放，无卡顿
省流	1	正常播放，无卡顿
省流	0.8	频繁卡顿，影响正常播放
省流	0.5	基本上每个视频都卡顿

结合 Wi-Fi 限速和 CQT 定点感知测试结果，抖音视频组感知拐点门限建议见表4。

表4 感知拐点门限建议

分辨率	PDCP_DL 速率门限/ (Mbit·s ⁻¹)
省流模式	≥1
高清模式	≥2

2.2 观众视角抖音直播实测分析

不同分辨率下观看抖音直播进行 5G 网络定点、DT 实测及 Wi-Fi 限速感知测试，结合 3 种感知测试结果，分别对不同分辨率的下载速率感知拐点门限进行判定，建议门限见表5。

表5 感知拐点建议门限

分辨率	PDCP_DL 速率门限/ (Mbit·s ⁻¹)
蓝光	≥5
超清	≥3
高清	≥2.5
标清	≥1.5

2.3 主播视角抖音直播实测分析

主播视角对直播业务进行 Wi-Fi 限速和 5G 网络实测，自己直播无法感受到是否存在卡顿；主播卡顿及延迟需要从观众视角关注，主播的画面与观众视角延迟 3 s 左右，在很好的无线环境下都会出现这种现象，与抖音服务器的内部处理有关联。Wi-Fi 限速感知测试情况见表6。

5G网络实测情况：蓝光上行速率1.61 Mbit/s，声音和画像延迟较高，30 s左右；超清上行速率1 Mbit/s左右，声音和画像延迟5 s左右；高清上行速率0.9 Mbit/s左右，声音和画像延迟5 s左右；标清上行速率0.7 Mbit/s左右，声音和画像延迟5 s左右，5G网络实测情况见表7。



结合 Wi-Fi 限速及现网定点测试, 不同分辨率直播场景拐点门限建议见表 8。

表6 Wi-Fi限速感知测试情况

分辨率	上行速率/(Mbit·s ⁻¹)	感知情况
蓝光	3	画面无卡顿, 声音延迟 3 s 左右
蓝光	2.5	画面无卡顿, 声音延迟 3 s 左右
蓝光	2	画面无卡顿, 声音延迟 3 s 左右
蓝光	1	画面偶现卡顿, 声音延迟 10 s 左右
蓝光	0.8	画面基本处于一直卡顿, 声音延迟 12 s 左右
超清	3	画面无卡顿, 声音延迟 3 s 左右
超清	2	画面无卡顿, 声音延迟 3 s 左右
超清	1.5	画面无卡顿, 声音延迟 3 s 左右
超清	1	画面无卡顿, 声音延迟 5 s 左右
超清	0.8	画面频繁卡顿, 声音延迟 10 s 左右
超清	0.5	画面一直卡顿, 声音延迟 11 s 左右
高清	2	画面无卡顿, 声音延迟 3 s 左右
高清	1	画面无卡顿, 声音延迟 3 s 左右
高清	0.8	画面无卡顿, 声音延迟 5 s 左右
高清	0.5	画面频繁卡顿, 声音延迟 10 s 左右
标清	2	画面无卡顿, 声音延迟 3 s 左右
标清	1	画面无卡顿, 声音延迟 3 s 左右
标清	0.8	画面无卡顿, 声音延迟 3 s 左右
标清	0.6	画面偶现卡顿, 声音延迟 8 s 左右
标清	0.5	画面频繁卡顿, 声音延迟 13 s 左右

表8 拐点门限

分辨率	PDCP_UL 速率门限/(Mbit·s ⁻¹)
蓝光	≥2.5
超清	≥1.5
高清	≥1.0
标清	≥0.8

2.4 抖音礼物特效速率实测分析

选取蓝光“苍海一舟”直播间进行礼物特效进行下载速率关联测试, 其中特效绚丽程度越高, 屏占比越大, 产生的下载速率越高, 其中“化蝶飞”“嘉年华”的特效在直播过程中所需要的下行速率在 9 Mbit/s 左右。直播间礼物特效如图 3 所示。



图3 直播间礼物特效

表7 5G网络实测情况

分辨率	RSRP/dBm	NR PDCP Thrput UL(实测)/(Mbit·s ⁻¹)	感知情况
蓝光	-100	2.01	约有延迟 5~10 s
	-105	1.61	画面与声音, 延迟较大, 时长约 30 s, 勉强能播
	-110	0.88	画面与声音延迟 1 min 左右, 且明显卡住不动, 基本无法使用
	-115	0.19	卡顿, 无法直播使用
超清	-100	1	画面与声音, 延迟 5 s 左右, 可以正常直播
	-105	0.94	画面与声音, 延迟 5 s 左右, 可以正常直播
	-110	0.99	画面与声音, 延迟 5 s 左右, 可以正常直播
	-115	0.23	卡顿, 无法直播使用
高清	-100	0.84	画面与声音, 延迟 5 s 左右, 基本可以正常使用
	-105	0.82	画面与声音, 延迟 5 s 左右, 基本可以正常使用
	-110	0.87	画面与声音, 延迟 5 s 左右, 基本可以正常使用
	-115	0.22	卡顿, 无法直播使用
标清	-100	0.69	画面与声音, 延迟 5 s 左右, 可以正常使用
	-105	0.67	画面与声音, 延迟 5 s 左右, 可以正常使用
	-110	0.68	画面与声音, 延迟 5 s 左右, 可以正常使用
	-115	0.29	频繁卡顿, 延迟约 2 min

不同抖音直播礼物特效速率评估见表9。

表9 抖音直播礼物特效速率评估

直播礼物名称 (特效画面)	PDCP层礼物 前均值/(Mbit·s ⁻¹)	PDCP层礼物时 均值/(Mbit·s ⁻¹)
化蝶飞	5.6	9.8
大飞机	5.2	8.9
玫瑰之约	5.4	8.8
抖音1号	5.5	8.8
人气票1000连发	5.5	8.8
嘉年华	5.4	8.5
梦回紫禁城	5.4	8.1
永生花	5.1	7.6
龙抬头	5.2	7.2
真的爱你	5.4	7.3
比心兔兔	5.5	6.9
冰蓝玫瑰	5.4	6.7

2.5 抖音直播连线与PK速率实测分析

选取2人直播PK、4人直播PK验证,分析发现从观看视角而言,直播间下行速率与PK人数及在线人数无明显关系,而是与观看直播间清晰度有关。

不同PK人数的速率分析:进入2人直播PK的一方直播间观看PK,清晰度为蓝光,下行速率6 Mbit/s以上,从直播画面进入PK的另一方直播间观看,清晰度为超清,下行速率4 Mbit/s左右。不同PK人数的速率分析如图4所示。

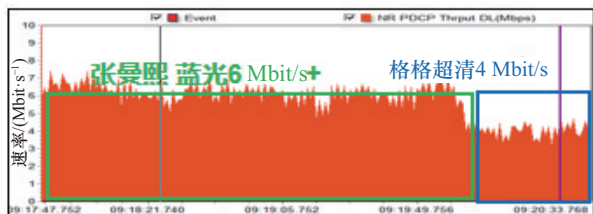


图4 不同PK人数的速率分析

对比2人和4人PK直播间,选取蓝光直播间进行对比,速率差异不大,不同蓝光PK直播间速率对比如图5所示。

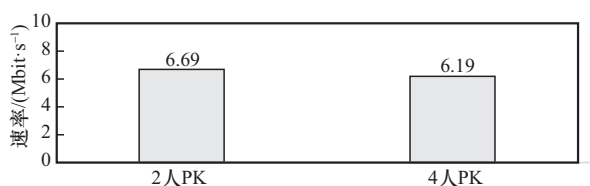


图5 不同蓝光PK直播间速率对比

3 流媒体感知评估体系

针对5G网络视频业务中用户感知劣化问题,提出基于“双维八项”的流媒体感知评估体系。通过斯皮尔曼秩相关系数建立KQI-KPI关联模型,量化分析视频初缓时延、卡顿率等关键感知指标与无线网络参数的映射关系,最终形成闭环优化机制。实际应用表明,该体系可有效定位感知瓶颈问题。

3.1 重点指标定义

随着移动视频流量占比超过70%,用户对视频业务“连得上、启动快、播放优”的体验需求日益严苛,用户体验质量评价指标见表10。

从源头出发确定5项重点KQI指标定义以及对应算法。

(1) 视频流媒体初始播放成功率=(传输足够初缓时间播放数据的次数/合成视频请求总次数)/N×100%

初缓时间:首次播放视频所需的缓存时间。不同网站模型略有不同,需要分网站做不同的配置。

足够数据:流媒体码速×初缓时间。

(2) 视频流媒体初始缓存时延= Σ (传输足够初缓时间播放数据的时刻-合成视频请求时刻)/N

通过计算公式可知初始缓冲时延指的是客户端在开始播放视频之前(一般为1~3 s)的缓冲区域中准备足够的数据流,初始缓冲平均时长如图6所示,而初始缓冲时延则主要受到视频流媒体服务器的带宽和数据处理能力等因素影响。



表 10 用户体验质量评价指标

用户体验质量	评价指标	计算方法	定义
连得上 视频能否打开	初始播放成功率/%	$\text{TCP 连接建立成功率} \times \text{首 GET 请求成功率} \times \text{初始缓存成功率}$	用户发起视频流媒体业务并完成初始缓存, 成功开始播放的概率
启动快 视频可否加速加载	视频播放等待时长/初始缓冲时延/ms	$\Sigma (\text{传输足够初缓时间播放数据的时刻} - \text{合成视频请求时刻}) / N$	用户从发起视频流媒体业务, 到缓冲完成并开始播放之间的等待时长
	RTT 时延/ms	下行 RTT = TCP ACK 时间点 - SYN ACK 时间点 上行 RTT = SYN ACK 时间点 - TCP SYN 时间点	能够反映出数据在网络中传输的速度和稳定性统计数据发生到接收端确认的总时长
播放优 播放是否有卡顿现象, 卡顿是否频繁	卡顿占比/%	视频卡顿时长/视频播放时长	统计视频流媒体播放期间, 卡顿时长的占比
	速率码率比	视频下载速率/下载视频码率比值	统计视频流媒体播放期间, 速率码率比值关系
	视频流下载速率/(kbit·s ⁻¹)	视频下载数据量/视频下载时间	整个视频流媒体过程中, 视频流媒体数据的下载速率

(3) 往返时延 (RTT)

下行 RTT = TCP ACK 时间点 - SYN ACK 时间点

上行 RTT = SYN ACK 时间点 - TCP SYN 时间点

表示从发送端发送数据开始, 到发送端收到来自接收端的确认 (接收端收到数据后便立即发送确认), 总共经历的时延。

RTT 是衡量网络传输性能的重要指标之一, 上下行 RTT 示例如图 7 所示, 能够反映出数据在网络中传输的速度和稳定性。通常情况下 RTT 越短, 网络传输的速度就越快, 反之则越慢。因此, 通过监测 TCP 中的 RTT 时延, 可以初步判断网络的性能如何。

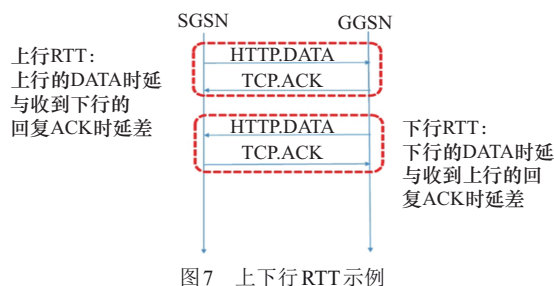


图 7 上下行 RTT 示例

(4) 视频流媒体下载速率 = $\Sigma [(\text{视频下载数据量} \times 8) / (\text{视频传输时间} \times 1000)] / N$

通过算式可知视频业务中客户端与 Web 服务器成功建立 TCP 连接的前提下, Web 服务器成功响应客户端发起的 HTTP Get 请求, 且响应消息全部收取的下行流量及时长累计的和, HTTP 下载过程如图 8 所示。

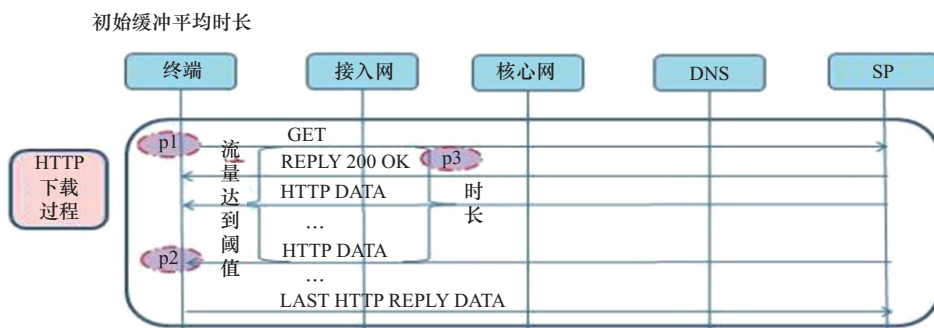


图 6 初始缓冲平均时长

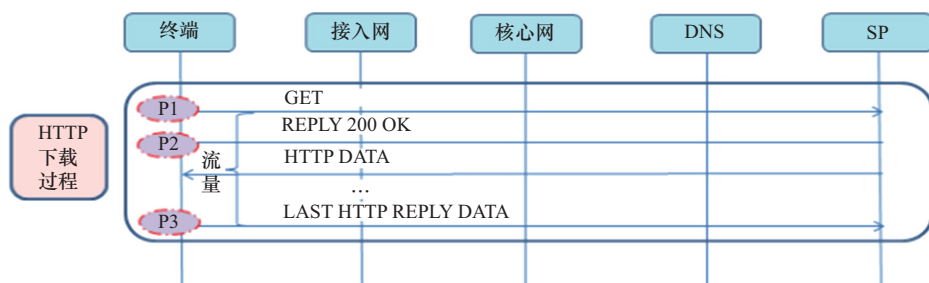


图8 HTTP下载过程

(5) 视频流媒体速率码率比=用户视频播放过程中下载速率/视频码率

视频下载的速率与视频码率的比值，无单位。

视频速率码率比可以反映视频的卡顿情况，速率码率比过低小于1，可能会导致卡顿情况。

(6) 视频流媒体卡顿占比=视频卡顿时长/视频播放时长

卡顿时长：已下载内容时长（当前下载时戳-初始时戳）≤用户已观看内容时长（当前时间-初始播放时间），则判断出现卡顿，记录出现卡顿时

戳和缓冲完成下次开始正常播放的时戳，两者之差为卡顿时长（广告、用户暂停、拖拽导致卡顿不算在内），HTTP下载过程卡顿示例如图9所示。

播放时长：流媒体播放总时长是指视频业务中客户端与Web服务器成功建立TCP连接的条件下，Web服务器成功响应客户端发起的HTTP Get请求，且响应消息全部收全的时长累计和，如图10所示。

然而传统网络优化方法存在以下痛点。

- (1) KPI指标与用户感知脱节。
- (2) 多维度参数关联性缺乏量化分析。

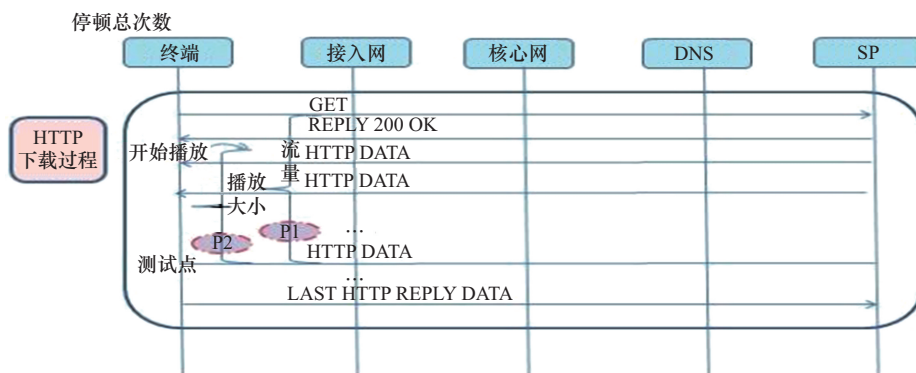


图9 HTTP下载过程卡顿示例

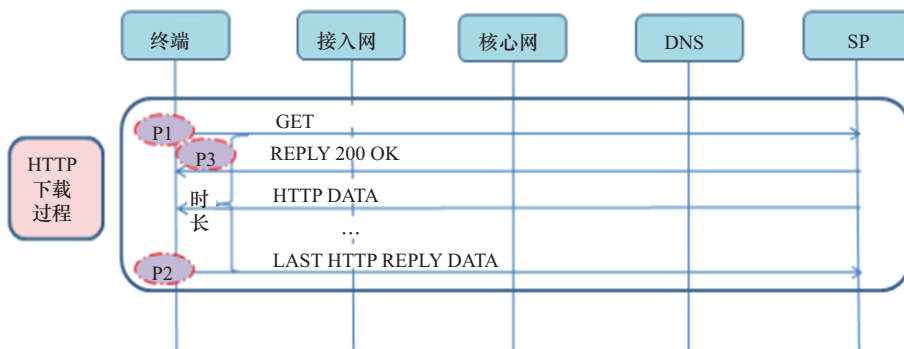


图10 HTTP下载过程播放示例



(3) 优化措施缺乏数据驱动依据。

通过构建流媒体感知评估体系, 解决“端-管-云”全链路感知问题的精准定位。

3.2 评估体系设计

流媒体评估体系实现流程如图11所示。体系架构分为五层。

(1) 数据采集层: 抓取源数据, 包括无线侧指标(如MCS、CQI)、核心网参数(如RTT、TCP重传率)及QoE日志、MR数据、XDR话单多终端播放日志, 形成原始数据池, 采用Tukey's Fences方法剔除大于3IQR的异常样本。

(2) 特征处理引擎: 对原始数据进行时间对齐(精度±50 ms)、异常值清洗及特征衍生(如计算速率波动率、时延变化梯度), 生成标准化特征矩阵。

(3) 关联分析模块: 基于斯皮尔曼秩相关分析, 量化KQI(如卡顿率)与KPI(如256QAM占比)的非线性关系, 识别关键影响因子($|\rho| > 0.7$ 的强关联项), 斯皮尔曼秩相关分析算法如下:

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \quad (1)$$

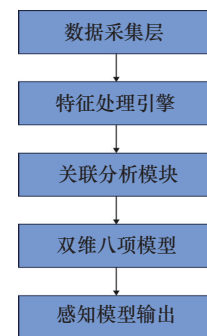


图11 评估体系实现流程

其中, d_i 为KQI-KPI排序差异, n 为样本数。通过1 000次Bootstrap重采样计算 p 值进行显著性验证, 建议存在关联强度分级见表11。

表11 关联强度分级

ρ 绝对值范围	关联等级	优化优先级
≥ 0.7	强相关	紧急
$0.5 \sim 0.7$	中等相关	高
< 0.5	弱相关	观察

根据斯皮尔曼相关系数评估, 结合网络实际情况以及重点评估指标, 得出结论其中端到端5项指标与网管KPI容量、无线质量及速率类指标关联性较强, 关联性强的指标可以作为后续重点关注指标项, 评估体系实现流程如图12所示。

评估大类	评估小类	视频播放成功率/%	视频下载速率/Kbit·s ⁻¹	视频播放平均缓存时长/ms	视频播放卡顿时长占比/%	视频TCP业务异常结束率/%	视频TCP上行RTT平均时延/ms	视频TCP下行RTT平均时延/ms	视频平均码率/Kbit·s ⁻¹	视频速率码率比
端到端性能指标	视频播放成功率/%	1.000	0.278	-0.089	-0.077	-0.156	0.002	0.084	-0.175	0.321
	视频下载速率/Kbit·s ⁻¹	0.278	1.000	-0.31	-0.124	-0.218	0.031	-0.013	-0.168	0.898
	视频播放平均缓存时长/ms	-0.089	-0.31	1.000	0.013	0.122	0.000	0.288	0.156	-0.241
	视频播放卡顿时长占比/%	-0.077	-0.124	0.013	1.000	0.002	0.019	-0.081	-0.042	-0.09
	视频TCP业务异常结束率/%	-0.156	-0.218	0.122	0.002	1.000	0.154	0.07	0.213	-0.279
	视频TCP上行RTT平均时延/ms	0.002	0.031	0.000	0.019	0.154	1.000	-0.062	-0.135	0.096
	视频TCP下行RTT平均时延/ms	0.084	-0.013	0.288	-0.081	0.07	-0.062	1.000	-0.063	0.014
	视频平均码率/Kbit·s ⁻¹	-0.175	-0.168	0.156	-0.042	0.213	-0.135	-0.063	1.000	-0.385
接通类	视频速率码率比	0.321	0.898	-0.241	-0.09	-0.279	0.096	0.014	-0.385	1.000
	无线接通率_EC/%	0.043	-0.027	0.042	-0.026	-0.063	-0.171	-0.04	-0.008	-0.012
保持类	无线掉线率_EC/%	-0.09	-0.117	0.020	0.048	0.087	0.1	0.029	0.017	-0.122
	5G切换成功率_EC/%	0.064	0.068	-0.03	-0.011	-0.041	-0.052	-0.016	-0.041	0.073
无线质量类	低CQI占比_EC/%	0.054	0.025	0.056	-0.053	-0.035	-0.034	0.425	-0.054	0.032
	RRC连接重建比例_EC/%	0.058	0.085	-0.022	-0.031	0.056	0.133	0.292	-0.031	0.087
	上行MAC层传输块错误率_EC/%	0.133	0.161	0.066	-0.075	-0.043	-0.052	0.549	-0.078	0.177
	下行MAC层传输块错误率_EC/%	-0.059	-0.075	-0.015	0.073	0.027	0.095	-0.08	-0.012	-0.075
	上行HARQ重传比率_EC/%	0.059	0.065	0.063	-0.018	0.002	0.018	0.408	-0.055	0.068
	下行HARQ重传比率_EC/%	-0.041	-0.053	0.003	0.045	0.042	0.1	0.082	-0.027	-0.049
	5G小区下行MCS优良率_EC/%	-0.213	-0.347	0.033	0.088	0.136	-0.041	-0.306	0.169	-0.391
	平均RRC连接用户数_EC/个	-0.377	-0.581	0.054	0.192	0.304	0.076	-0.25	0.26	-0.657
容量类	PDCH信道占用率_EC/%	-0.318	-0.553	0.118	0.113	0.327	0.037	0.146	0.252	-0.629
	上行PRB平均利用率_EC/%	-0.366	-0.583	0.084	0.17	0.313	0.063	-0.146	0.255	-0.657
	下行PRB平均利用率_EC/%	-0.323	-0.55	0.101	0.107	0.335	0.032	0.121	0.27	-0.635
	5G上行用户平均速率_EC/Mbit·s ⁻¹	0.027	0.104	-0.104	0.009	-0.005	0.033	-0.255	0.013	0.106
速率类	5G下行用户平均速率_EC/Mbit·s ⁻¹	-0.159	-0.151	-0.101	0.122	0.019	0.051	-0.689	0.084	-0.188
	下行256QAM编码比例_EC/%	-0.115	-0.167	-0.032	0.055	0.018	-0.004	-0.508	0.065	-0.193
	上行256QAM编码比例_EC/%	-0.089	-0.037	-0.109	0.101	0.045	0.156	-0.499	0.038	-0.041
	上行平均MCS_EC/个	0.007	0.028	-0.083	0.035	0.013	0.066	-0.346	0.006	0.036
	下行平均MCS_EC/个	-0.132	-0.172	-0.061	0.073	0.017	-0.018	-0.602	0.082	-0.205
	5G小区下行平均RANK_EC	-0.19	-0.21	-0.051	0.136	0.089	0.093	-0.587	0.106	-0.236
干扰类	小区上行平均干扰电平	-0.108	-0.085	0.083	0.151	0.074	0.014	-0.055	0.046	-0.008

图12 评估体系实现流程

(4) 双维八项模型：双维八项模型如图 13 所示，该模型结合了斯皮尔曼评估体系与手动关联的 KPI 指标趋势图，从端到端评估指标维度和网管评估指标维度两大方面入手。其中，端到端评估指标包括视频下载有效速率、视频播放平均缓存时长、下行平均速率、视频 TCP 下行 RTT 平均时延、视频播放卡顿时长占比等 5 项关键指标。网管评估指标则涵盖上行平均速率、上行平均 MCS、下行平均 MCS、下行 256QAM 编码比例、低 CQI 占比、下行平均 RANK 等 8 项关键指标中的三项。这一体系为后续的用户感知分析与指标提升提供了坚实的框架基础，实现体验量化评分。

(5) 感知模型输出：作为整个体系的核心输出模块，该接口通过三级处理机制将分析结果转化为可落地的网络优化方案。目前，已完成视频播放平均缓存时延模型、下行 RTT 时延模型等关键算法框架的搭建，为网络质量评估提供了量化支撑。

端到端平台数据指标与 KPI 指标进行关联分析：视频播放平均缓存时延与上/下行速率质量相关性指标成反比关系，视频有效速率与 5G 下行速率质量相关指标成正比关系，从而确定 TDD 网

络中视频播放平均缓存时延模型，如图 14 所示。

端到端平台数据指标与 KPI 指标进行关联分析：下行 RTT 时延与上/下行小区用户速率、上/下行 MCS、上/下行 256QAM 占比、下行平均 Rank 指标成反比，与低 CQI 占比成正比关系，从而确定 TDD 网络中下行 RTT 时延模型，如图 15 所示。

端到端平台数据指标与 KPI 指标进行关联分析：视频播放平均缓存时延与上/下行速率质量相关性指标成反比关系，视频有效速率与 5G 下行速率质量相关指标无相关性，从而确定 FDD 网络中视频播放平均缓存时延模型，如图 16 所示。

端到端平台数据指标与 KPI 指标进行关联分析：下行 RTT 时延与上/下行小区用户速率、上/下行 MCS、上/下行 256QAM 占比、下行平均 Rank 指标成反比，与低 CQI 占比成正比关系，从而确定 FDD 网络下行 RTT 时延模型，如图 17 所示。

4 感知提升方案及应用效果

通过感知评估体系筛选出差小区后，可以首先从“基础优化”“参数优化”“特性应用”三大方面进行提升，其中基础优化涉及常规的故障排

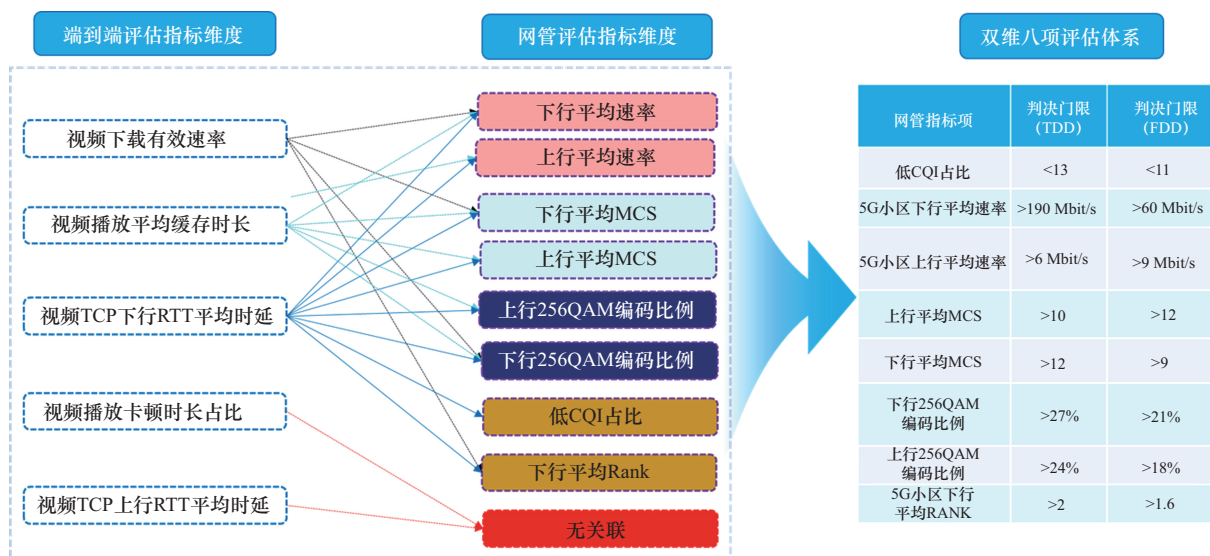


图 13 双维八项模型

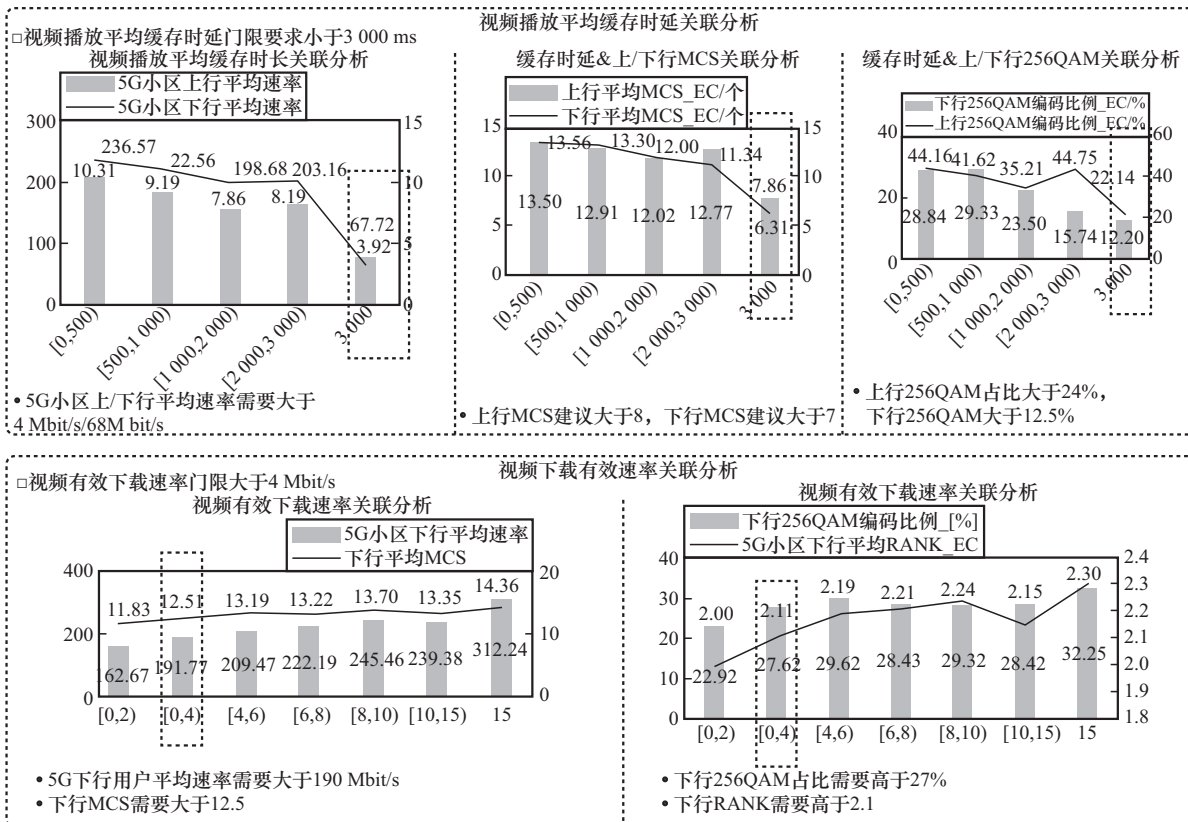


图14 TDD网络中视频播放平均缓存时延模型

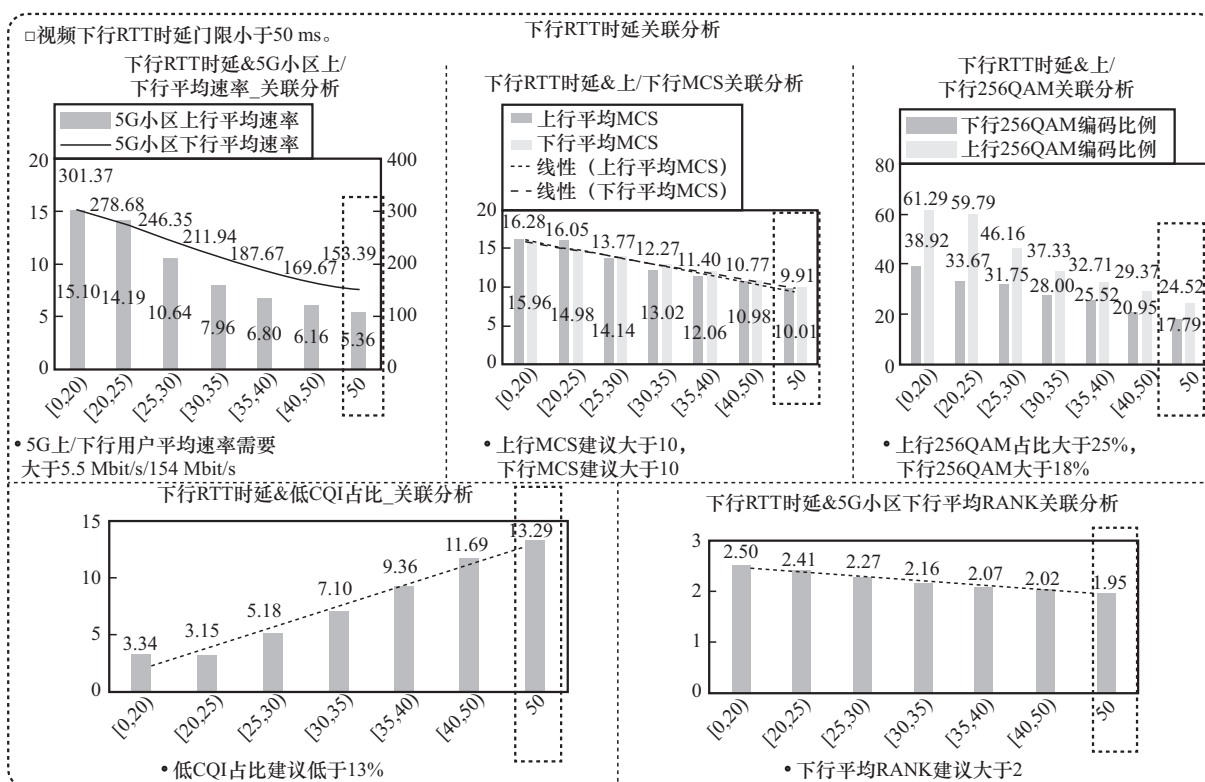


图15 TDD网络中下行RTT时延模型

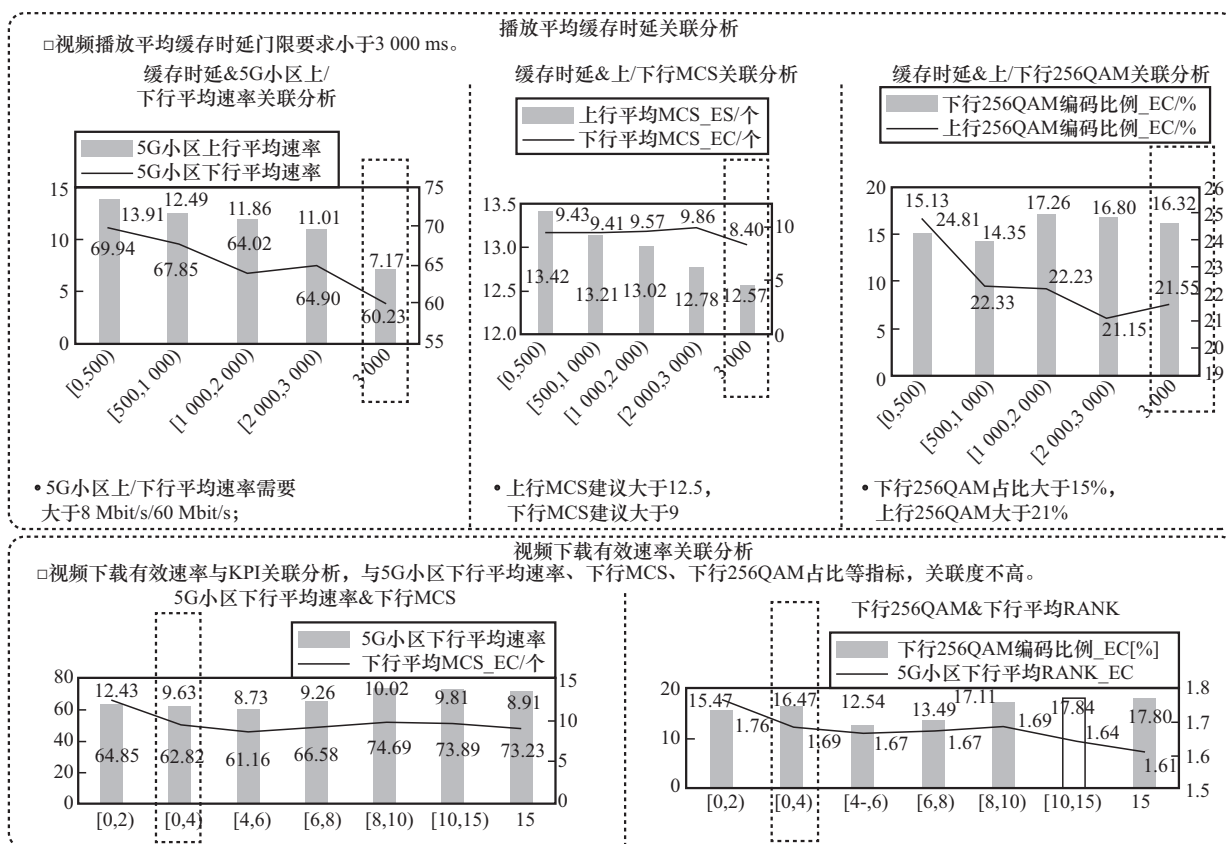


图16 FDD网络中视频播放平均缓存时延模型

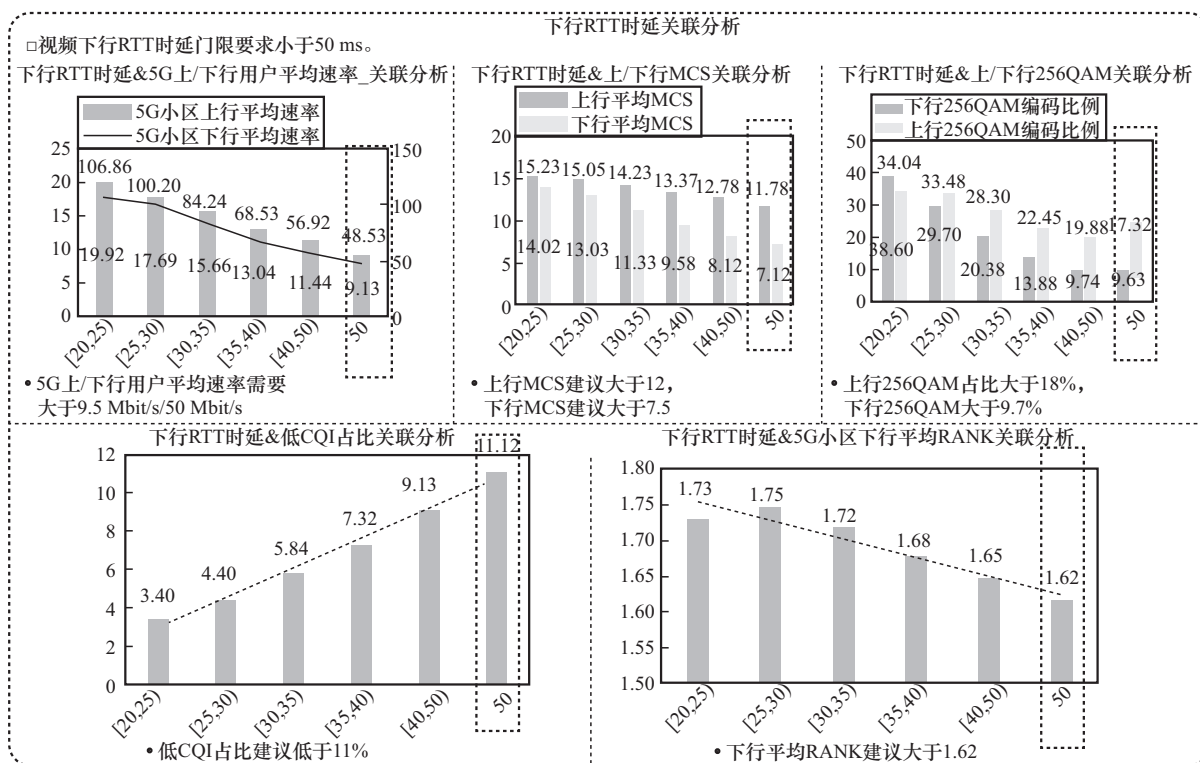


图17 FDD网络下行RTT时延模型



查、覆盖优化、容量优化、干扰优化、切换优化等,在此不做重点阐述,本文主要从参数和特性功能两方面展开感知优化提升的方案研究。

4.1 参数优化

参数优化是通过对网络关键参数进行精细调整,以提升视频流媒体业务的感知质量。本文针对不同场景和业务需求,制定了一系列科学合理的参数优化策略,主要从弱场边缘能力提升参数、MAC下行性能参数优化和波束成形参数调整进行了研究,并在现网进行了部署应用。

4.1.1 边缘弱场性能提升

(1) 方案原理

下行采用PDCCH码本赋型增强控制信道覆盖,结合干扰随机化降低邻区干扰,提升PDSCH接收质量;上行通过CP/DFT波形自适应、频选调度与干扰规避策略,在弱场环境下提升发射功率与频谱效率,并通过远端PRB自适应、多列导频Ruu重构及预调度技术,实现资源分配与解码性能的精准优化。

下行PDCCH码本赋型:通过波束赋型提高终端接收PDCCH的信号强度,从而提高接收性能,减少差点DCI漏检,降低CCE聚合等级。

下行干扰随机化:降低邻区PDSCH的干扰,提升PDSCH信道质量,保证下行PDSCH更好地接收性能。

CP/DFT波形自适应:上行弱场环境下,CP波形切换到DFT波形可以获得更高的发射功率,提升边缘弱场接收性能。

上行频选加上行干扰随机化:好点按照模3走干扰随机化,差点时对IOT高的PRB避开调度,达到调度时避开上行邻区,以及其他干扰的目的,提高接收信号质量,获得较好的接收性能。

远端PRB自适应:自动调整上行PRB调度带宽来平衡上行PRB数,上行MCS的关系,提升上行频谱效率。

上行多列导频Ruu重构:通过对上行均衡算

法IRC的干扰矩阵Ruu采用重构法计算,提高干扰矩阵的计算精度,从而提高上行解码性能。

上行预调度:上行的稀疏调度会导致MCS偏高不能及时修正,从而导致误码增加,通过预调度能够增加上行的调度次数,降低误码,改善吞吐量。

涉及的弱场优化参数见表12。

表12 弱场优化参数

参数名称	参数值
上行远点PRB自适应	开
上行预调度	下行触发上行预调度
上行干扰随机化	开
上行频选	开
CP/DFT波形自适应	srs端口与波形切换同时进行
上行Ruu重构	多列导频重构
下行干扰随机化	开

(2) 应用效果

某地市部署弱场性能提升参数,采用网速管家APP进行感知测试,参数部署后视频类感知时延<50 ms占比从33%提升到93%;其他游戏类、网页类感知时延指标均有提升,总体感知时延低于50 ms占比从36.84%提升到65.79%。感知参数修改后,上行平均用户速率从1.79 Mbit/s提升到2.66 Mbit/s,下行平均用户速率从16.11 Mbit/s提升到18.97 Mbit/s。用户在该区域的视频下载速率和播放缓存时延等关键指标得到了显著提升,上行FTP速率修改前后对比如图18所示。

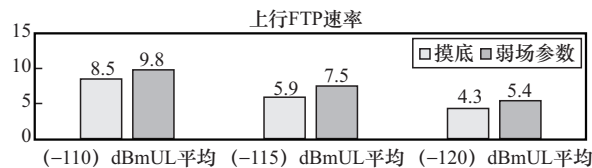


图18 上行FTP速率修改前后对比

下行FTP速率修改前后对比如图19所示。

弱场感知测试时延低于50 ms占比(-115 dBm)修改前后对比如图20所示。

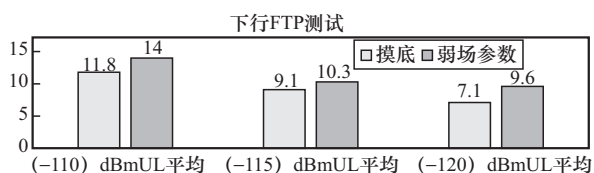


图19 下行FTP速率修改前后对比

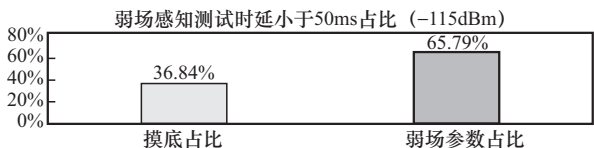


图20 弱场感知测试时延小于50ms占比(-115dBm)修改前后对比

上下行平均用户速率修改前后对比如图21所示。

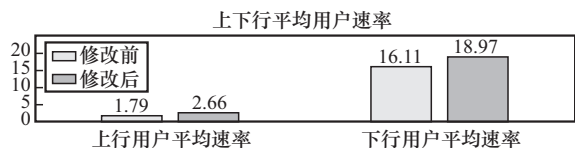


图21 上下行平均用户速率修改前后对比

网络游戏时延修改前后对比如图22所示。

4.1.2 MAC下行性能及波束成形参数优化

(1) 方案原理

媒体接入控制(MAC)层作为数据链路层的重要组成部分,其性能参数的优化对视频流媒体业务的传输效率有着直接影响。通过调整MAC层的调度算法、重传机制等参数,可以提高数据传输的可靠性和效率。例如,优化调度算法可以根据视频流媒体业务的特点,为其分配更多的传输资源,减少数据传输的等待时间;调整重传机制可以在保证数据可靠传输的前提下,减少重传

次数,降低时延。

波束成形技术通过调整天线的辐射方向图,将信号能量集中到特定方向,提高信号强度和质量。根据不同场景下的无线传播环境,对波束成形参数进行优化。在弱场环境下,通过优化波束成形参数,可以增强信号覆盖,提高用户的接收信号强度,从而改善视频流媒体业务的播放质量。例如,调整波束的宽度和方向,使其更好地覆盖用户所在区域,减少信号衰减和干扰。

① HARQ确认与修正结合了前向纠错(FEC)和自动重传请求(ARQ),用于提高数据传输的可靠性,确认与修正开关则直接影响了HARQ的性能和效率。

提高资源利用率:通过及时修正过激活状态,避免长时间占用频谱资源或其他硬件资源,从而提高资源利用率。

降低传输时延:修正开关可以加快错误检测和恢复的速度,减少由于重传或等待确认造成的传输时延,信令传输的稳定性和可靠性非常有益。

② SRB采用码本调度灵活性:通过启用码本调度开关,系统可以根据不同的业务类型和需求,灵活选择合适的资源分配方案,满足多样化的业务需求和提高网络性能至关重要。

效率:使用预定义的码本进行资源调度可以简化调度过程,减少计算量和信令开销,有助于降低网络时延并提高数据传输效率。

可预测性:码本调度使得资源分配更具可预测性,因为系统可以基于码本中的方案进行快速

RSRP=-115dBm	游戏网	摸底	弱场参数	直播网	摸底	弱场参数	电商网	摸底	弱场参数	视频网	摸底	弱场参数	资讯网	摸底	弱场参数
	王者荣耀	优秀	良好	抖音	良好	优秀	拼多多	良好	优秀	腾讯视频	优秀	优秀	腾讯新闻	良好	良好
	英雄联盟	良好	良好	快手	良好	优秀	京东	优秀	优秀	优酷	良好	优秀	UC头条	良好	良好
	和平精英	良好	良好	小红书	良好	优秀	美团	良好	良好	爱奇艺	良好	优秀	今日希望	优秀	优秀
	绝地求生	良好	优秀	微视	优秀	优秀	淘宝	优秀	优秀	bilibili	良好	优秀	趣头条	良好	优秀
	开心消消乐	良好	良好	虎牙直播	良好	优秀	唯品会	良好	良好	好看视频	良好	优秀	新浪新闻	良好	优秀
	穿越火线	良好	良好	斗鱼	良好	优秀	小红书	良好	良好	芒果TV	优秀	优秀	凤凰新闻	优秀	优秀
	欢乐斗地主	优秀	优秀	YY直播	优秀	优秀	天猫	优秀	优秀	搜狐视频	优秀	良好	汽车之家	优秀	优秀
	Dota2	良好	良好	美拍	良好	优秀	饿了么	优秀	良好						

图22 网络游戏时延修改前后对比



决策,对于保证信令传输的稳定性和可靠性非常有益。

MAC下行性能及波束成形优化参数表见表13。

表13 MAC下行性能及波束成形优化参数表

参数名称	原值	修改值
仅考虑过激活检测 HARQ 确认参与修正开关	关闭	打开
下行连错次数门限	16	8
业务码本下开关	关闭	打开
CQI 修正的频谱效率平滑因子	50	10
功率调整时频谱效率高门限	70 000	160 000
波束成形开关	关闭	打开

(2) 应用效果

省A对19个700M小区进行MAC性能及波束成形参数部署,部署后V-Max端到端平台即时通信下行RTT平均时延从85.24 ms下降至65.25 ms,下行平均RTT时延修改前后对比如图23所示;省B部署2 800多个700 M小区,SEQ端到端平台视频下行RTT高时延占比从3.1%下降到2.13%,下行RTT高时延占比如图24所示。对现网进行MAC性能及波束成形参数部署,有效降低了视频下行RTT平均时延和高时延占比。通过优化MAC层的调度算法和波束成形参数,使得基站能够更高效地利用无线资源,提升数据传输效率,从而降低视频业务的下行时延。

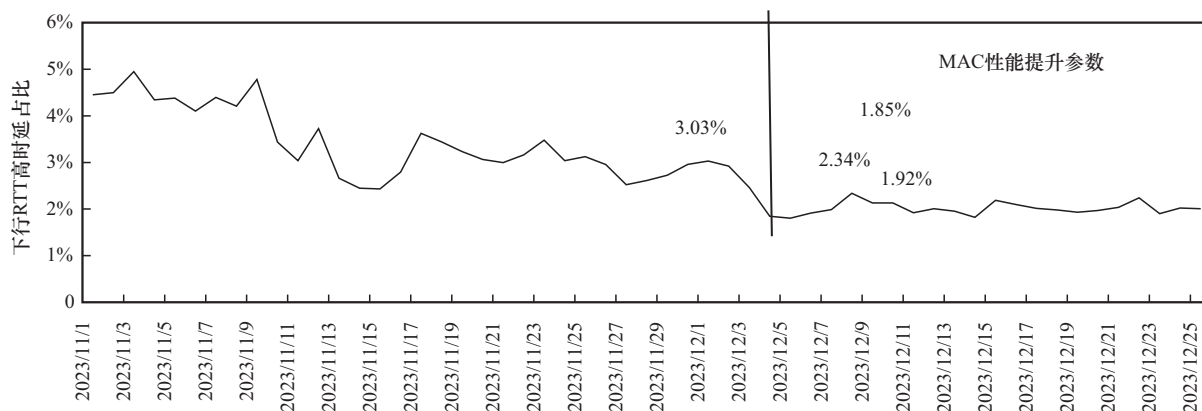


图24 下行RTT高时延占比

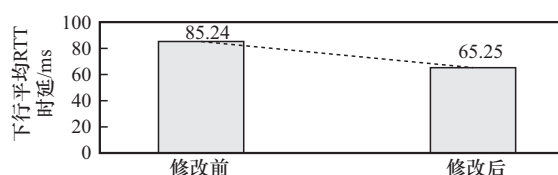


图23 下行平均RTT时延修改前后对比

4.2 特性功能

引入特性功能技术,为视频流媒体业务提供了更加优质的服务保障。本文研究的主要的特性功能主要包含业务识别保障特性、Qos调度保障特性和TCP代理。

4.2.1 业务识别保障特性

(1) 方案原理

为了实现对不同业务类型的差异化优化,首先需要准确识别业务类型。预先配置不同类型的特征IP或特征包,可以实现对视频、游戏、即时通信等业务的识别。在业务识别的基础上,针对识别出的业务类型应用相应的保障策略。例如,对于视频业务,可以调整调度优先级,保障其带宽需求;对于游戏业务,可以降低时延,提高实时性。通过业务识别与保障策略的实施,能够显著提升用户感知体验,业务保障原理如图25所示。

(2) 应用效果

通过端到端平台获取考核质差业务的IP地址,基站识别成功后,用户立即应用对应的保障

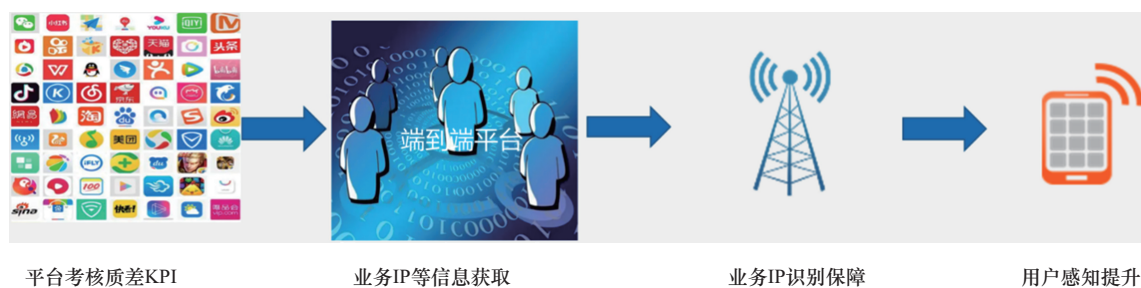


图25 业务保障原理

策略，好点用户调度优先级不调整。负荷不高区域，提升调度优先级作用不大，需要配置合理的保障速率及调度资源进行保障。某省部署了106个2.6G小区，视频类IP段30个，流量占比80%，V-Max端到端平台分析功能特性部署前后的视频感知指标，下载速率、播放缓存时延、速率码率比等关键指标均有小幅度的增益，端到端平台统计指标见表14。

4.2.2 Qos调度保障特性

(1) 方案原理

随着5G用户持续增加，5G用户的各种业务需要进行区分和感知提升，针对不同5QI等级，无线网差异化的设置相对调度比例，包括下行相对调度比例和上行相对调度比例，保障用户感知。

针对不同的业务类型，目前可以通过在核心网侧、基站侧扩展5QI配置来保障用户需求，需合理配置NGBR速率，降低对普通用户的影响。其基于业务类型的Qos差异化保障原理如图26所示。

(2) 应用效果

省公司市场部拟针对直播场景的需求，设计上行流量保障产品，提供更好的5G网络和服务质量来满足主播们的高清直播需求。为此制定了基于5G PCC的上行直播业务保障策略方案。

策略方案如下。

- 核心网和基站增加直播5QI flow 70。
- 上行设置不同的保障GRB速率进行验证。
- 采用8部终端：QCI6卡2主用1备用、QCI9

表13 端到端平台统计指标

状态	视频播放成功率	视频下载速率/ (Mbit·s ⁻¹)	视频下载有效速率/ (Mbit·s ⁻¹)	视频播放平均卡顿次数/次	视频播放平均缓存时长/ms	视频播放卡顿时长占比	视频TCP上行RTT平均时延/ms	视频TCP下行RTT平均时延/ms	视频速率码率比
部署前	98.96%	5.66	15.53	0.09	992.53	3.46%	23.95	47.73	15.73
部署后	99.09%	5.84	16.66	0.09	877.15	3.53%	24.03	50.96	16.97
增益	0.14%	0.18	1.13	0.00	-115.38	0.07%	0.08	3.24	1.24

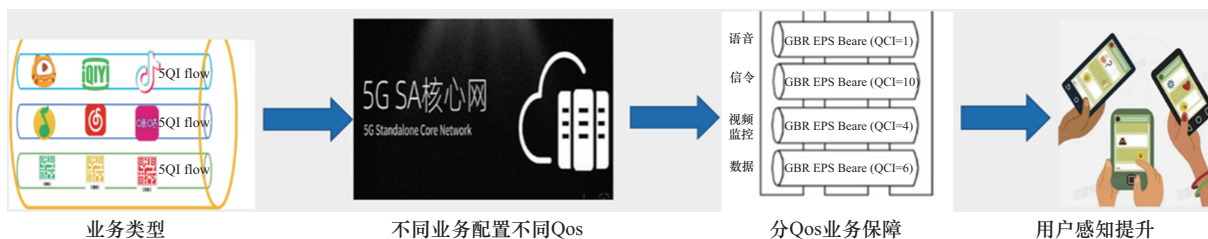


图26 基于业务类型的Qos差异化保障原理图



卡2主用1备用、QCI70卡5主用2备用。

- 选取好点和中点进行测试验证。

Qos 调度保障参数见表 14。

表 15 Qos 调度保障参数

参数名称	推荐值	备注
5QI 值为 70 的 5QI 优先级设置	90	根据各地运营商的要求配置 (新增)
QCI 值为 6 的 QCI 对应优先级的 10 倍值	90	根据各地运营商的要求配置 (新增)
上行 NGBR 最小保障速率	6	根据现场需求配置 (新增)
上行优先级排队算法	newPF	根据各地运营商的要求配置 (新增)
下行调度算法类型	newPF	根据各地运营商的要求配置 (新增)

将 5QI70 终端保障速率设置为 8 Mbit/s, 上行速率高于最低保障速率时, 不会启动保障, 5QI70 与 5QI6 优先级相同, 5QI70 与 5QI6 用户资源均分, 5QI70 与 5QI6 平均速率接近, 对比情况如图 27 及图 28 所示。

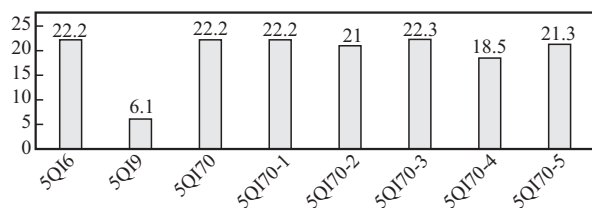


图 27 5QI70 终端保障速率设置 8 Mbit/s 时好点的 5QI 对比

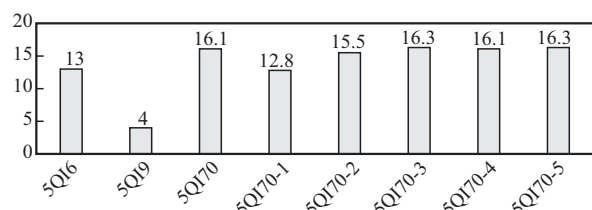


图 28 5QI70 终端保障速率设置 8 Mbit/s 时中点的 5QI 对比

将 5QI70 保障速率设置为 64 Mbit/s, 上行速率低于最低保障速率时, 启动保障, 5QI70 会资源抢占较多, 导致 5QI6 和 5QI9 的速率较低, 对比情况如图 29 及图 30 所示。

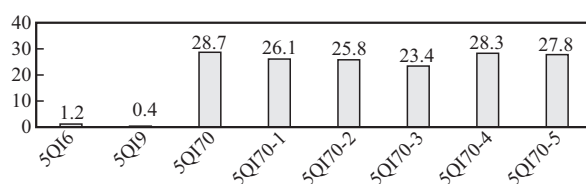


图 29 5QI70 终端保障速率设置 64 Mbit/s 时好点的 5QI 对比

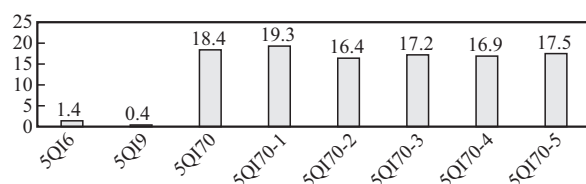


图 30 5QI70 终端保障速率设置 64 Mbit/s 时中点的 5QI 对比

4.2.3 TCP 代理特性

(1) 方案原理

TCP 代理是提升端到端时延的策略, 在 gNB 网网上, 通过增加 TCP 代理功能, 向服务器屏蔽无线环境传输特性差异, 通过向服务器端尽快发送确认, 将数据缓存在 gNB, 并由 gNB 根据空口质量向客户端以尽可能快的速度发送数据, 最大限度地利用空口传输带宽。TCP 代理通过慢启动加速、屏蔽 UE 侧发送的 ACK 等功能, 提升信令平台的网页时延等性能指标。例如, 在弱场环境下, TCP 代理可以通过优化数据传输路径, 减少数据传输过程中的丢包和重传, 从而提升数据传输效率, 降低端到端时延。TCP 代理原理流程如图 31 所示。

(2) 应用效果

在地市 A 开启了基于 TCP 代理的网页支付业务微信钱包、支付宝, 视频业务快手、抖音, 进行重点业务保障。采用“网页浏览下行平均 RTT 时延 (ms)”代表网页时延、“视频下载有效速率 (Mbit/s)”代表视频速率、“视频 TCP 上行 RTT 平均时延 (ms)”代表视频时延。支付业务开启 TCP 代理前后时延对比如图 32 可见, 网页时延缩短 27.7%; 视频业务开启 TCP 代理前后指标对比如图 33 所示, 视频下载速率提升 1.5%、视频时延缩短 12.43%, 时延感知提升较明显。

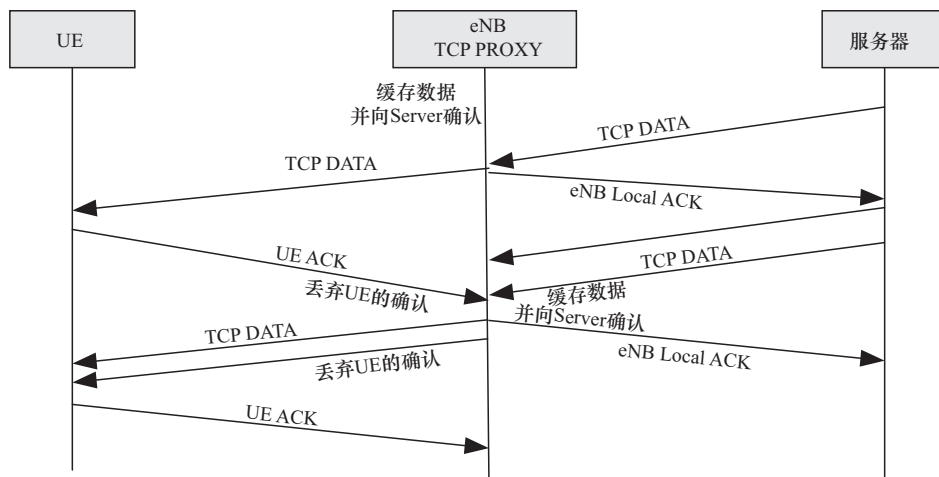


图31 TCP代理原理流程

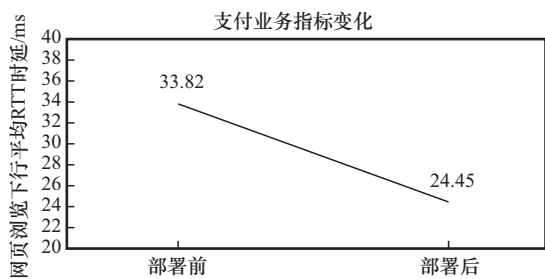


图32 支付业务开启TCP代理前后时延对比

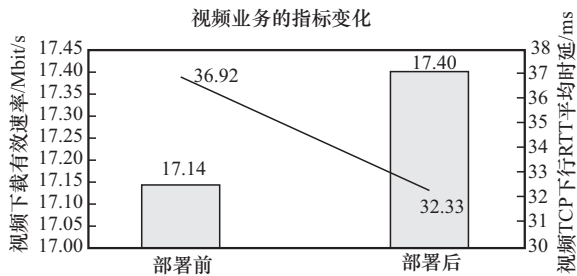


图33 视频业务开启TCP代理前后指标对比

5 结束语

5.1 总结

本文围绕短视频与直播业务爆发式增长背景下移动网络性能面临的挑战，系统性地开展了以下研究工作。

(1) 构建了流媒体核心参数体系与动态平衡模型：明确了分辨率、码率、帧率及压缩编码技术等核心参数的定义及其关联性，揭示了H.265等高效编码技术在节省带宽方面的优势。

(2) 完成了多业务场景的实测感知分析：通过5G网络及Wi-Fi限速下的CQT、DT测试，定量分析了抖音短视频、观众/主播视角直播、礼物特效及直播连线PK等业务在不同网络条件下的感知表现，并提出了各业务场景下的感知拐点门限建议值。

(3) 提出了“双维八项”流媒体感知评估体系：通过斯皮尔曼秩相关系数建立了KQI-KPI关联模型，实现了视频初缓时延、卡顿率等关键感知指标与无线网络参数的量化映射，为精准定位感知瓶颈提供了方法论。

(4) 制定了感知提升方案并验证了应用效果：从参数优化（弱场性能提升、MAC下行性能及波束成形优化）和特性功能（业务识别保障、QoS调度保障、TCP代理）两个维度提出了具体优化方案。现网应用结果表明，所提方案能有效提升视频下载速率、降低播放缓存时延及RTT时延，显著改善了用户感知。

5.2 展望

尽管本文在流媒体业务感知评估与提升方面取得了一定成果，但仍存在进一步深入研究的方向。

(1) AI驱动的自适应码率算法研究：未来可探索基于终端实时感知与网络状态预测的AI自适应



应码率算法, 实现更精准的动态参数调整, 进一步提升用户体验并节约网络资源。

(2) 端到端全链路感知监控: 当前研究侧重于无线接入网, 后续可将研究范围扩展至核心网、传输网及互联网内容源, 构建覆盖“端-管-云”的全链路感知监控与优化体系。

(3) 6G网络与流媒体业务融合探索: 面向6G时代, 研究新型网络架构(如空天地一体化网络)及关键技术(如语义通信)对超高清、沉浸式视频业务的支持能力与优化策略。

(4) 多业务差异化保障策略深化: 随着业务形态日益丰富, 需要进一步细化不同业务(如

VR/AR直播、云端游戏)的QoS需求, 并研究在网络切片等技术下的资源动态分配与保障机制。

(5) 用户体验主观评价与客观指标的深度融合: 引入更多用户体验主观评价数据, 与客观KQI/KPI指标进行深度融合分析, 构建更精准、全面的用户体验评估模型。

参考文献:

- [1] 李凡, 刘伟. 一种TCP代理方法、基站及装置: CN119484423 A[P]. 2025-02-18.
- [2] 苏金树. 高速网络TCP加速关键技术研究[D]. 长沙: 国防科学技术大学, 2007.
- [3] 中华人民共和国工业和信息化部. 端口控制协议代理(PCP proxy)技术要求: 2024[S].