

基于众包感知的 OTT 业务体验影响因素分析

徐小龙¹ 李克¹ 王海¹ 宋晓勤² 王征³ 吴雄³

(1. 北京联合大学智慧城市学院, 北京, 100101; 2. 南京航空航天大学电子与信息工程学院, 南京, 210016; 3. 中国电信股份有限公司上海分公司, 上海, 200122)

摘要: 随着移动互联网和 OTT(Over-the-top)业务的高速发展,传统以网络为中心的运维方式难以以为继,因此如何提高用户业务体验、变革传统网络质量的评价和优化方法迫在眉睫。本文围绕基于终端侧业务感知大数据进行网络和业务质量评价这一新型网络评价与运维作业模式,首先对影响端到端用户业务感知的因素进行了较全面的剖析;并利用从普通用户终端上采集的现网真实的海量业务感知数据,重点针对网页浏览这一代表性的 OTT 业务,从多个不同的维度深入研究影响业务感知的关键因素,以及关键业务感知指标间的关联关系,揭示了 OTT 业务感知的关键影响因素间的内在联系。研究结果对于进一步分析业务感知质差成因、合理构建用户体验质量(Quality of experience, QoE)映射模型等具有很好的参考价值。

关键词: 众包感知; 业务体验; 网络测量; OTT 业务

中图分类号: TP302.7 **文献标志码:** A

Crowdsourcing-Based Analysis on Impact Factors of OTT Service Perception

Xu Xiaolong¹, Li Ke¹, Wang Hai¹, Song Xiaqin², Wang Zheng³, Wu Xiong³

(1. College of Smart City, Beijing Union University, Beijing, 100101, China; 2. College of Electronics and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, 210016, China; 3. China Telecom Corp. Ltd. Shanghai Branch, Shanghai, 200122, China)

Abstract: Service perception analysis is the crucial solution to understand user experience and network quality, and the maintenance and optimization of the mobile networks. With the fast development of mobile Internet and OTT(over-the-top) services, the traditional network-centric mode of network operation and maintenance is no longer an effective way. In this situation, how to effectively evaluate and optimize user's service perception is getting more and more important. The key factors that impact on user's end-to-end OTT service perception is analyzed based on the so-called crowd sourcing based user perception monitoring methodology. Furthermore, the inner relationship among the key factors and the interaction between key quality indicators are evaluated from several aspects, by utilizing big data set of user perception collected from large amount of smart phone users in the real network. The results of analysis are highly valuable references for the improvement of user perception and network quality, network deployment and optimizations, etc.

Key words: crowdsourcing; service perception; network measurement; OTT service

引言

基于互联网的信息传输通常提供的是实时性和可靠性较差但灵活性较强的服务,相比而言,基于电信网络的业务则具有高度的稳定性和可靠性。为保持其竞争优势,运营商通常会投入大量人力、物力到网络质量的评价和优化,并逐步形成了以网络为中心的网络运维模式。这样带来的问题就是运营商通常会站在网络的视角看问题,重网络质量而轻业务体验。随着移动网络传输能力的快速提升以及智能终端的普及,近年来互联网阵营的 OTT(Over-the-top)业务得到了高速发展并逐步成为移动网络承载业务的主体,对电信运营商的传统自有业务形成了替代效应,最典型的例子就是微信对短信和语音业务的冲击。在这种新的形式下,移动网络有逐步沦为管道的危险,而仅仅通过保障网络质量已经难以保障良好的业务体验。因此,为了保障用户业务体验,降低客户流失风险,变革移动网络质量的测量、评价和优化方法迫在眉睫。

传统的移动网络运维主要依靠路测(Drive test, DT)和定点拨测(Call quality test, CQT)等手段、以网络关键性能指标(Key performance indicator, KPI)为目标来评价网络质量,更注重从网络层面而非业务层面进行评价和优化,网络质量和用户感知存在越来越大的剪刀差。为此,运营商逐步开始重视用户体验质量(Quality of experience, QoE)。研究人员也围绕业务关键质量指标(Key quality indicator, KQI)的数据采集和 QoE 的评价方法展开了相关研究^[1-8]。虽然用户感知越来越重要,但目前数据采集和评价的手段仍有一定局限性,主要是通过在网络侧部署探针的方式如深度包解析方法(Deep packet inspection, DPI)^[9]。一般来说,离用户越近则感知越准确,因此如何从用户侧获取感知信息尤为关键。运营商开始尝试采用基于终端侧用户感知大数据进行更加准确的网络和业务质量评价^[10],希望以此牵引各相关业务部门共同协作促进业务感知的提升,进一步用于指导精准化的市场营销和客户服务,称之为众包用户感知方法(Crowdsourcing-based user perception, CUP)。

业务感知评价的最终目的是指导网络的优化调整以提升业务感知水平,这首先需要针对影响业务感知的因素进行分析和定位。本文针对上述问题,以网页浏览类应用为研究对象,基于在现网中采集的海量真实终端侧业务感知数据,通过多种数据分析和统计学手段,系统地分析了影响用户业务感知的各类因素,并从多个维度分析了这些因素对业务感知的影响机制,以及各主要影响因素之间、感知指标之间的关联关系。

1 相关工作

下面对常用的网络测量方法、QoE 评价与分析进行简要分析,并对本文所采用的终端侧业务感知大数据的采集方法作重点阐述。

1.1 网络测量方法

网络测量与评价主要包括 3 个阶段:即数据采集、质量评价和数据分析。从数据采集的方式上来说,网络和业务测量方法主要分为两大类:主动测试和被动监测。从数据采集的位置来分,则包括终端侧采集和网络侧采集两大类。主动测试是最为传统也是最成熟的一种测量方式,其代表性方法是 DT/CQT,例如文献[11]采用的就是从终端侧发起的主动测试,而网络侧的数据采集方法中最具有代表性也是目前应用最广的就是 DPI 技术。除此之外,还可以通过用户问卷调查的方式获得手机用户的主观业务感知质量信息,其代表性的方法为净推荐值(Net promoter score, NPS)^[12]。

主动测试的优点在于测试环境完全可控和透明,而其缺点是采集的样本点少,且只是测试人员在有限时间、地点对用户业务行为的有限模拟,难以反映全体用户的真实业务行为。此外,由于测试人员行为上的差异性,导致测试结果的公立性和准确性有较大的不确定性。由于电信业务过程中往往产生海量数据,因此被动监测方式获取到的电信数据天然具有大数据的特征,如何借助于大数据的手段进行数

据分析得到了越来越多的重视。文献[13]中对于网络侧移动数据的采集和分析研究现状作了较为全面的论述,其中最具代表性的技术就是 DPI。其最大优势是可以实现对网内全量用户的数据采集和分析,但部署成本高昂且局限于本网内,难以进行跨运营商网络的对标分析,对业务的识别和业务特征采集的准确性也受到解析技术的限制会有一定偏差。此外,随着 https 技术的广泛应用,其业务识别能力也面临很大挑战。

随着智能终端以及搭载了各种传感器的可穿戴设备的兴起,基于众包的终端侧测量得到了学术和产业界的重视。文献[14]中将这类测量方法命名为移动众包感知(Mobile crowdsensing, MCS),文中将其分为两种类型,即参与式感知和机会感知。前者需要个人的主动参与以贡献感知数据(如个人上传照片或报告实施路况等),而机会感知则以被动的、自治的方式工作,通常不需要用户的主动参与。

本文所采用的 CUP 方法属于机会感知的范畴,重点关注如何从用户智能终端上采集用户的实际在网业务信息,并进行网络和业务质量的评价分析。该方法具有很强的客观性,数据准确性很高,能够较全面地反映网络和用户的真实情况,但其最大难点在于 APP 的部署。可以采用独立的 APP 形式,但更多的则是作为 SDK 插件与其他 APP 捆绑。另一难点是如何避免对用户隐私的侵犯。此外,由于一般是基于操作系统层面之上的数据采集,其获取的无线参数较为有限,对于更深层次的网络数据比如空口信令更是无法获取。

1.2 基于大数据的 QoE 评价与分析

海量网络和业务数据的获取为更准确的 QoE 评价和分析提供了更丰富的信息,也吸引了学术界的广泛关注。文献[3]中探索了将大数据分析用于网络优化以提升用户 QoE 的问题,描述了一个大数据驱动的移动网络优化的框架,讨论了相应的机器学习大数据分析方案和大数据特征。文献[15]介绍了基于随机矩阵理论和机器学习的统一数据模型,提出了一个将大数据分析应用于移动网络的框架,并描述了几个典型应用场景,包括信令大数据、业务大数据和无线信号大数据等。

文献[16]基于网络侧话单数据(Call detail record, CDR)构建了典型移动互联网业务的用户感知评估体系,并提出了一种基于统计的用户感知问题的多维度定界方法;文献[17]则在终端众包感知数据和 DPI 数据相结合的基础上提出了移动互联网业务 KQI 的评测方法,以及基于统计的终端/无线/核心网/SP 四段感知质差问题定界、定段、定位分析模型。

在 QoE 的影响因素分析方面,文献[5]中将 QoE 影响因素分为 3 个层面:服务层面、环境层面及用户层面。服务层面的影响因素则可以进一步包括网络层、应用层及服务层的参数。网络层的参数反映网络传输的状况,应用层的参数反映没有经过传输的服务的性能,服务层的参数确定了通信的语义、内容、优先级、重要性以及定价。环境层面的影响因素则包括自然环境、人文与社会环境以及服务运行环境等。用户层面的因素包括用户的期望、体验经历、用户体验时所处的身心状态和自身背景。但该文中对于具体业务场景下的影响因素以及这些因素与业务体验指标之间的相关性缺乏深入的分析,这正是本文的主要研究内容和贡献。

1.3 众包用户感知的数据采集

本文所采用的基于众包用户感知的数据采集与分析系统的系统架构如图 1 所示。

其中用户终端上安装了数据采集前端,通过在后台运行的方式监测用户的业务行为,并在设定的触发条件(比如用户用特定的浏览器 APP 访问了预定义的某个目标网站页面)下采集相关的业务感知和辅助信息并上传至云端的数据采集分析平台。分析平台对回

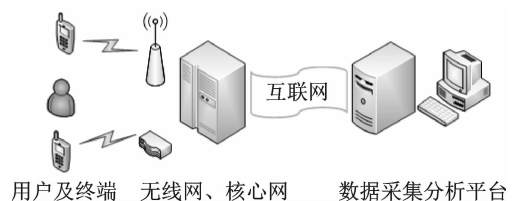


图 1 CUP 数据采集与分析系统的通用架构
Fig. 1 Common architecture of CUP data acquisition and analysis system

传的原始数据进行相应的预处理,包括无效数据的剔除和用户信息的脱敏。通常不采集用户的高度敏感信息,包括手机号码、短信文本等,仅采集终端标识(International mobile equipment identity, IMEI)和用户标识(International mobile subscriber identity, IMSI)作为后台统计分析的用户标识,并在预处理中进行脱敏处理,包括 IMEI/IMSI 信息的加扰和加密等。采集的信息通常包括以下 3 类:

(1) 业务感知信息

这是最核心的采集信息。一般针对最具代表性的 OTT 业务类型,如网页浏览、视频播放和游戏等。对每类业务根据业务特征以及与用户体验的相关性定义若干 KQI 指标。具体而言,对于网页浏览类业务的感知指标主要体现在 HTTP 请求各阶段的时延,可根据业务需要进行相应的定义。例如文献 [10] 中所定义的指标包括首包时延、页面打开时延,还可以结合速率等指标。本文所采用的指标具体定义如下:

(a) 首包时延(D_k)。用户发起网页浏览请求到收到目标服务器响应的第 1 个 HTTP 200 OK 报文包之间的所经历的时长,即

$$D_k = T_{200} - T_{req} \quad (1)$$

其中 T_{req} 为用户发起网页浏览请求的时间点, T_{200} 为收到服务器第一个 HTTP 200 OK 报文的时间点。首包时延体现了服务器对 HTTP 请求的响应是否有反应及响应的快慢。

一次网页浏览过程包括若干环节,即 DNS 解析、TCP 连接建立和 HTTP 交互,具体如图 2 所示。因此这里的首包时延可进一步分解为 3 个分段时延指标之和,即

$$D_k = D_{dns} + D_{tcp} + D_{get} \quad (2)$$

其中

$$\begin{cases} D_{dns} = T_{dns} - T_{req} \\ D_{tcp} = T_{tcp} - T_{dns} \\ D_{get} = T_{200} - T_{tcp} \end{cases} \quad (3)$$

式中: D_{dns} 为 DNS 解析时延,即从发起页面浏览请求时刻到终端收到完成 DNS 解析的时刻 T_{dns} 的时间差。需要注意的是,如果在用户发起网页浏览请求的时刻终端尚未建立起空口连接,则该时延还包含空口建立时延; D_{tcp} 为 TCP 连接建立时延,即从 DNS 解析完成到终端发出 TCP 连接确认的时刻的时间差; D_{get} 为 GET 请求时延,即从终端发出 TCP 连接确认到收到第一个 TCP 数据包(包含 HTTP 200 OK)的时刻 T_{200} 的时间差。

(b) 页面打开时延(D_p)。用户发起浏览请求到收到目标服务器的第一个 HTTP 响应(仅页面文本内容,不包含资源的二次加载)、整个 HTTP 页面下载完毕并渲染完成的时长。页面打开时延反映了用户从访问开始到页面打开所需要等待的全部时间,是影响用户体验的最直接的感知指标。页面时延是在首包时延的基础上增加接收响应时延,即

$$D_p = D_k + D_{res} \quad (4)$$

其中 D_{res} 为接收响应时延,是指从收到第一个响应数据包到终端发出 [FIN, ACK] (即接收完成) 的时刻的时间差,即

$$D_{res} = T_{res} - T_{200} \quad (5)$$

(2) 网络环境信息

每次监测到目标样本时,均同步采集当时的网络制式和基本网络参数(场强、信号质量和小区标识

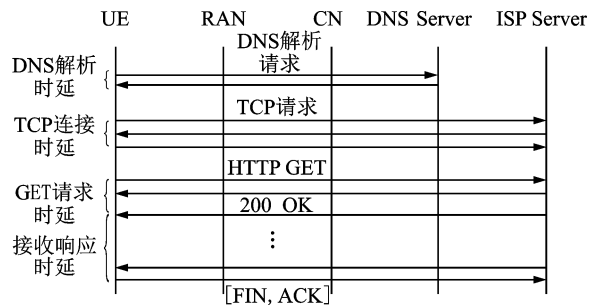


图 2 网页浏览的交互过程

Fig. 2 Signaling process of web browsing service

等),对不同的网络制式,具体的参数不同。例如对 LTE 网络,场强为 RSRP,信号质量为 RSRQ,小区标识为四元组{TAC,eNodeB ID, CI, PCI},这些信息表征了当时的网络质量的基本情况,这对于分析网络环境对业务感知质量的影响是重要的参考。

(3) 基础信息

主要包括终端信息、位置信息等。终端信息包括用户标识 IMSI、设备标识 IMEI、终端型号以及操作系统版本等,位置信息则包括采集业务感知样本时的终端经纬度、定位精度、定位方式(GPS 定位或网络辅助定位)。

2 OTT 业务感知质差的影响因素

对于一个 OTT 业务的业务过程而言,影响业务感知的因素涉及到端到端的各个环节,主要包括云、管和端 3 个部分,即业务服务器、移动网络(包括无线和核心两大部分)以及终端 3 大环节。具体可以将上述因素细分为 6 个维度进行分析,即:无线网、核心网、时间、终端、用户和 ISP 网站。

(1)无线网:无线网相比核心网和 ISP 网站而言,由于地理环境和区域网络负载的多变性,其对用户业务感知的影响最为明显和不确定,是影响业务感知的最主要环节,也是优化提升的主要目标。具体而言,良好的无线网络首先体现在网络覆盖,即覆盖区域内各处的信号强度足够,更进一步则是信号质量良好(干扰较低)。

(2)核心网:包括核心网设备和传输链路。核心网由于影响的面很广,且其传输环境可控,一般运营商会核心网的运行维护上提供非常高级别的保障,因此其性能稳定,很少会影响到用户的业务感知。

(3)时间(负载)因素:用户的业务发起具有随机性,同时在分布上也存在统计特征,这种用户业务需求的时间差异化特征对无线网、核心网和 ISP 网站都会形成冲击从而影响到业务体验。这种由于用户业务需求在时间上的不均衡性具体表现为网络设备负载的差异性。高负载时段一般来说会导致用户业务感知的明显下降。

(4)终端因素:智能终端尤其是安卓智能终端,其碎片化特征明显,品牌和机型众多,不同的硬件能力对用户的业务体验有明显的影响,相对而言,高端机型由于其良好的硬件配置和软件优化能力,其业务的体验会优于低端机型。

(5)用户因素:具体的用户由于其终端的个性化性能差异和使用习惯,以及不同个体对业务感知良好的心理认知和接受程度上的差异,反映在业务感知的评估结果上也会存在差异。

(6)ISP 因素:对于大型 OTT 业务提供商,为保障用户体验,一般会大量采用内容分发网络(Content distribution network, CDN)^[18]的方式将内容下沉以就近提供服务。因此,即使同一网站由于其网站 CDN 服务器布放的位置、处理能力和带宽的差异,也会对不同地区的用户业务感知带来差别。

对于基于 HTTP 协议的业务,域名的解析需要 DNS 服务器,一般布放在运营商网内,有的也会利用第三方服务器提供域名解析。DNS 服务器的处理能力和带宽也会影响网页浏览业务的用户体验。

3 业务感知影响因素的相关性分析

本节重点从无线网、时间(负载)、终端、ISP 等几个主要维度,以网页浏览类业务为例,利用从现网中实际采集的海量业务感知大数据进行分析,从而验证上述影响因素对用户体验的实质性影响。数据来源于普通用户智能终端并已做了脱敏处理。采集所涉及的目标网站包括新浪、百度和搜狐等 9 个国内主流网站,所有样本均在 LTE 网络下获得,总有效样本数为 7 434 030 个。

3.1 无线覆盖对业务感知影响的相关性分析

对于无线网络而言,良好的信号覆盖是业务感知的重要基础,其中最基本的指标是业务发生时所在位置的信号强度,其次是信号质量,即干扰水平,对 LTE 网络分别是参数 RSRP 和 RSRQ。信号强度和信号质量越好,其空口的数据交互成功率和时延也越好,进而带来更好的业务体验。

首先分析覆盖能力对于业务感知指标的影响,即信号强度和信号质量与业务感知指标之间的相关性。不同 RSRP 水平下的平均首包时延和页面打开时延的分段分布情况如图 3 所示。

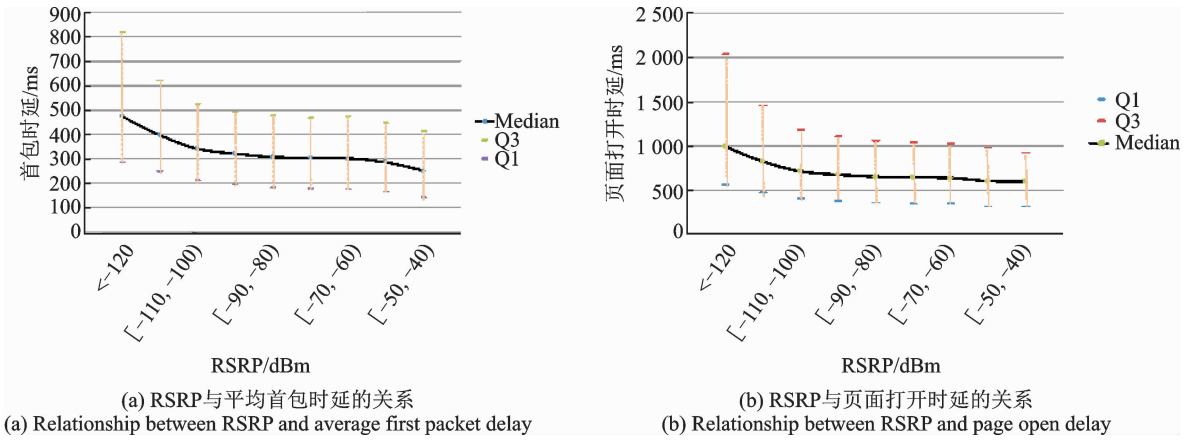


图 3 RSRP 与平均首包时延/页面打开时延的关系

Fig. 3 Relationship between RSRP and average first packet delay/page open delay

图 3 中 Median 曲线为中位数, Q1 和 Q3 分别为该 RSRP 分段上的 KQI 指标 1/4 分位数和 3/4 分位数。由此可以看出当场强小于 -110 dBm 时(对 LTE 网络而言,一般将场强低于 -110 dBm 的情况定义为弱覆盖),两个指标均出现了明显的恶化,这表明信号强度对浏览类业务的感知指标影响较大。对 RSRQ 和首包/页面打开时延的箱体图分析的结果与此类似。由此可见,无线环境质量对于端到端业务感知具有直接的影响,是业务感知保障优化工作中需要重点关注的因素。

3.2 网络负载对业务感知的影响

由于从终端侧无法获得网络负载信息,作为一种替代方案,分析一天不同时间段内的业务感知指标的差异,观察不同时间用户不同的业务使用强度对感知指标的影响。一般而言,一天之中不同时间段的业务使用强度是不同的。图 4 是根据业务行为数据分析得到的网页浏览类业务在不同时段的业务活跃度(即业务发起时间占比)。可以看出,在 23 点~次日 7 点之间的夜间时段,业务使用强度很低,也即网络负载很低。

将业务感知采样数据按照不同时段进行统计,可以得到图 5 的结果。从图中可以看出,感知指标的波动与业务活跃度分布的波动具有显著的相似性。这表明业务活跃度对于感知指标也具有显著的影响。

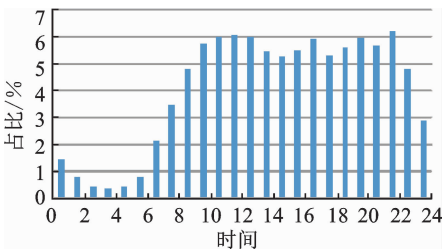


图 4 不同时段的网页浏览业务活跃度分布

Fig. 4 Distribution of web browsing service intensity at different time periods

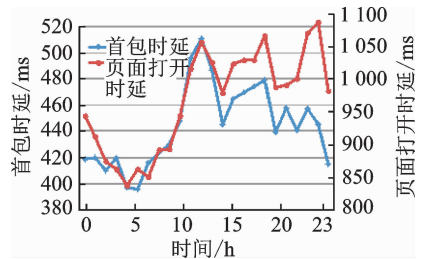


图 5 首包时延/页面打开时延在不同时段的水平

Fig. 5 Average first package delay and page open delay at different time periods

还可以进一步对上述因素间的相关性进行定量评价。相关性的量化评价的传统方法是计算 Pearson 相关系数(Pearson correlation coefficient, PCC)。PCC 是基于协方差和标准偏差的估计,两个样本集合的 Pearson 相关系数为

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (6)$$

其中 $\{X_i, i=1 \sim n\}$ $\{Y_i, i=1 \sim n\}$ 是所观察到的两组样本, \bar{X} 和 \bar{Y} 分别是其均值。

为了进一步验证该观察,同时采用另一种方法,即最大信息系数(Maximal information coefficient, MIC),以检测为两个变量的独立性关系^[19-20]。经证明, MIC 方法可以广泛地捕捉各类函数或非函数的关联性,其中一些甚至是通过其他方法如 PCC 方法、Kruskal 互信息估计器^[21]所无法发现的。MIC 的定义如下。

对于一个有限集合 $D \subset R^2$ 和正整数 x 与 y , 定义

$$I^*(D, x, y) = \max I(D |_G) \quad (7)$$

其中,求最大值是指对于 x 乘 y 分辨率的网格分割,通过调整 X 轴和 Y 轴的分割点所能计算得到的最大互信息值。有了在不同分辨率下得到的 I^* , 可以构造特征矩阵并求得有限集合 D 的特征矩阵。

定义二维集合 D 的特征矩阵 $\mathbf{M}(D)$ 是一个无穷矩阵, 即有

$$\mathbf{M}(D)_{x,y} = \frac{I^*(D, x, y)}{\log \min\{x, y\}} \quad (8)$$

经过标准化后的特征矩阵中元素的值均落在 $[0, 1]$ 上。

集合 D 的最大信息系数则可以表示为

$$\text{MIC}(D) = \max_{xy < B \ll n} \{M(D)_{x,y}\} \quad (9)$$

其中对于 $0 < \epsilon < 1$, 存在 $\omega(1) < B(n) \leq O(n^{1-\epsilon})$ 。 $B(n)$ 的建议值为 $n^{0.6}$ 。

根据 MIC 的定义可见, MIC 具有普适性、公平性和对称性等性质。所谓普适性,是指在样本量足够大(包含了样本的大部分信息)时,能够捕获各种各样的有趣关联,而不限定于特定的函数类型(如线性函数、指数函数或周期函数),或者说能均衡覆盖所有的函数关系。一般变量之间的复杂关系不仅仅通过单独一个函数就能够建模,而是需要叠加函数来表现。所谓公平性,是指在样本量足够大时能为不同类型单噪声程度相似的相关关系给出相近的系数。例如,对于一个充满相同噪声的线性关系和一个正弦关系,一个好的评价算法应该给出相同或相近的相关系数。可以看到,随着噪声程度的增加,这些方法对某些函数关系表现出的健壮性不一致。利用 MIC 可以大大减少数据清洗的麻烦,只要有足够的数据量可以代表总体信息,就可以直接用 MIC 计算分析。用 PCC 和 MIC 方法分别计算首包时延和页面打开时延与不同时段业务活跃度的相关性,结果如图 6 所示。显然,从图 8 可以看出,无论是首包时延还是页面打开时延,都表现出与业务活跃度显著的相关性。

3.3 网站与业务感知指标的关系

对于网页浏览业务而言,对于不同的网站由于其网页大小不同,需要下载的页面大小也不一样,因此,页面打开时延指标有可能受到页面大小的影响。为此将业务感知指标数据按照不同的网站进行统计,结果如图 7 所示。

从图 7 可以看出,不同网站的首包时延基本处于同一水平,这是因为首包时延主要取决于网站对终端侧网页请求的响应速度,与页面大小无关;而页面打开时延则与页面大小有明显的相关性,反映出页面打开时延与首页数据的传输时延相关性更高,与空口速率密切相关。因此,在评价页面打开时延的优劣时,如果对不同网站均采用相同的优劣判决门限则有失公平,因为不同区域各类网站的样本数分布

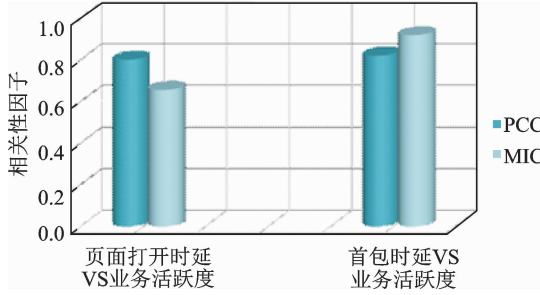


图 6 KQI 指标与业务活跃度的定量相关分析

Fig. 6 Quantitative correlation analysis between KQI index and service intensity

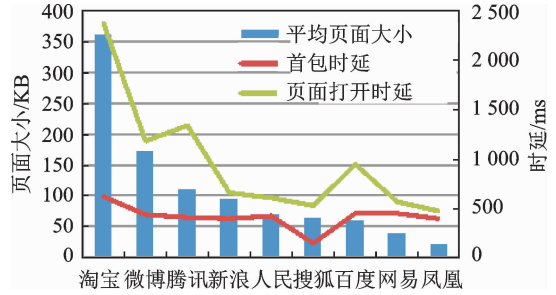


图 7 不同网站的 KQI 指标对比

Fig. 7 Comparison of KQI indexes of different web-sites

具有一定的差异性。因此,参考各网站的页面大小、本地用户对各网站的访问频度的偏差对页面打开时延指标的优劣判决门限进行加权归一化,以避免由于网站自身和用户访问偏好的差异性带来的评价结果的偏差。

假设本地区 n 个目标网站的页面大小分别为 $\{P_i, i=1 \sim n\}$, 平均页面大小为 $\bar{P} = \text{mean}(P_i)$, 各网站的用户访问量占比为 $\{\alpha_i, i=1 \sim n\}$ 且 $\sum_{i=1}^n \alpha_i = 1$, 设页面打开时延的基本劣质门限为 T_0 , 即

$$T_0 = 1.25 \sum_{i=1}^n \bar{D}_{pi} \cdot \alpha_i \quad (10)$$

其中 \bar{D}_{pi} 为第 i 个网站的平均页面打开时延, 则第 i 个网站的页面打开时延劣质门限可以按下式计算确定, 即

$$T_i = T_0 \cdot \beta_i \cdot \frac{P_i}{P} \quad (11)$$

其中 β_i 为控制参数。

3.4 终端资源负荷与业务感知指标的关系

终端的计算资源主要指 CPU 和内存。过高的 CPU 和内存占用会导致终端的计算能力和响应能力的下降, 影响到 APP 的运行。因此, 为了确定通过浏览器 APP 进行的网页浏览业务的用户体验是否受终端资源负荷的影响, 对业务感知指标与不同终端资源负荷的相关性指标即 MIC 进行计算, 如果如表 1 所示。可以看出, 首包时延和页面打开时延与 CPU 和内存之间无明显相关性, 即终端资源负荷对业务感知指标无明显影响。

表 1 终端资源负荷与首包/页面打开时延的 MIC 分析

Tab.1 MIC analysis of terminal resource load and first packet delay and page open delay

资源占用	首包时延	页面打开时延
CPU 占用	0.014	0.024
内存占用	0.012	0.012

3.5 业务感知影响因素的降维分析

对于业务感知指标而言, 在采集 KQI 指标样本的基础上, 进一步需建立 KQI-QoE 映射模型以便更直观地评价业务体验。为了减少 KQI 指标间的冗余对模型的影响, 需要对指标集合进行降维分析, 以确定最终用于构建模型的 KQI 指标集合。具体对网页浏览类业务而言, 一种方式是将首包时延和页面打开时延作为输入 KQI 指标集, 另一种方法则是将分段时延指标 $\{D_{dns}, D_{tcp}, D_{req}, D_{res}\}$ 作为输入, 这时需要对该指标集合中的 4 个分段指标进行降维分析。

最常用的降维分析是主成分分析法 (Principal component analysis, PCA)。其目的是用 1 组较少的

独立变量代替大量相关变量,同时尽可能保留初始变量信息,推导所得的变量称为主成分。探索性因子分析法(Exploratory factor analysis,EFA)也是一项用来找出多元观测变量的本质结构、并进行降维的技术,在教育和心理学评价中应用较多^[22]。它主要是为了找出影响观测变量的因子个数,以及各个因子和观测变量间的相关程度,其基本模型为

$$\mathbf{X} = \mathbf{A}\mathbf{\Xi} + \mathbf{\Delta} \tag{12}$$

其中 $\mathbf{X} = (x_1, x_2, \dots, x_k)^T$ 为 k 维观测变量向量, $\mathbf{\Xi} = (\xi_1, \xi_2, \dots, \xi_m)^T$ 为 m 维公共因子向量 ($m < k$), $\mathbf{\Delta} = (\delta_1, \delta_2, \dots, \delta_k)^T$ 为对应各观测变量的唯一性因子(表示该变量不能被公共因子所解释的部分)和观测变

量随机误差之和。 $\mathbf{A} = \begin{bmatrix} \lambda_{11} & \cdots & \lambda_{1m} \\ \vdots & \ddots & \vdots \\ \lambda_{k1} & \cdots & \lambda_{km} \end{bmatrix}_{k \times m}$ 为负载矩阵,各元素为对应的各公共因子的权重。分别应用

PCA 和 EFA 方法对可能影响页面打开时延的 4 个分段指标即 DNS 解析时延、TCP 连接建立时延、GET 请求时延和接收响应时延进行分析,结果如图 8 所示。由图可知,两条红线分别表示 PCA 和 EFA 的阈值,高于红线部分为 PCA 和 EFA 算法建议选取的因子数。由上图可知,EFA 建议提取 3 个因子,而 PCA 建议提取一个因子。考虑到高估因子通常比低估因子数的结果好,同时为了尽可能避免曲解“真实”情况,可以选择 EFA 法得到的因子结构进行建模分析。又因为 4 个分段指标对页面打开时延的影响可看作是独立的,因此选用正交旋转的方法提取因子,以便对结果做出更直观合理的解释,结果如表 2 所示。

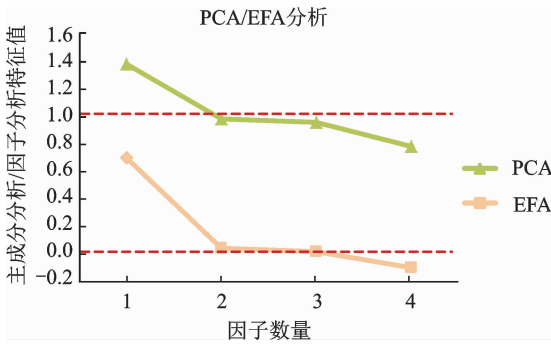


图 8 分段时延的 PCA/EFA 因素分析

Fig. 8 Analysis of PCA and EFA of segment delay

表 2 EFA 各因子得分 $\{a_{ij}\}$

Tab. 2 EFA factor scores $\{a_{ij}\}$

	ξ_1	ξ_2	ξ_3
D_{dns}	0.02	<u>0.97</u>	0.08
D_{tcp}	<u>0.99</u>	0.02	0.12
D_{req}	0.09	0.10	0.36
D_{res}	0.01	-0.04	<u>0.72</u>

由此可得

$$\xi_j = \sum_{i=1}^4 x_i a_{ij} \quad j = 1, 2, 3 \tag{13}$$

式中: ξ_j 代表第 j 个公共因子; x_i 分别代表 4 个分段时延观测变量; a_{ij} 代表每一个分段时延对应的权重。

由 EFA 计算结果进一步可知,第 1 主成分 ξ_1 的贡献率是 25%,其中 TCP 连接建立对页面打开时延的影响最大。其次是第 2 主成分 ξ_2 ,其贡献率为 24%,其中 DNS 解析对页面打开时延的影响最大。而 ξ_3 中则是以接收响应时延对页面打开时延的影响为最大。基于此,可采用 $\{\xi_j, j = 1, 2, 3\}$ 作为输入构建 KQI-QoE 映射模型。

4 结束语

本文针对传统移动网络运维和优化作业模式在 OTT 业务越来越占据主导地位的当前情况下所面

临的诸多问题,对影响端到端用户业务感知的因素进行了较全面的剖析。重点围绕网页浏览类业务,利用从普通用户终端上采集的现网真实的海量业务感知数据,从多个不同的维度深入研究影响业务感知的关键因素,以及关键业务感知指标间的关联关系,揭示了 OTT 业务感知的关键影响因素间的内在联系。分析发现,无线覆盖能力对首包时延和页面打开时延的影响较为突出,尤其是在弱覆盖场景下,KQI 指标会出现迅速恶化,揭示了无线覆盖对于业务感知的直接而重要的影响。其次对于时间维度上首包和页面打开时延所受的影响进行了定性和定量的相关性分析,结果均表明不同时段的业务活跃度水平与 KQI 指标间具有强相关性。网站维度的分析可以看出,首包时延与网站的关联性较低,而页面打开时延则受页面大小的影响很大。基于此,本文提出了一种页面打开时延劣质门限的计算方法,可以较好地解决不同网站页面大小以及用户访问偏好所带来的公平性问题。最后引入 PCA 和 EFA 分析方法研究分段时延用于 QoE 建模时的降维问题,并给出了相应的 KQI 指标选择建议。本文通过分析影响业务感知的关键因素,可以很好地用于指导用户业务感知优化工作,对于进一步分析业务感知质差成因、合理构建 QoE 映射模型等具有较高的参考价值。

参考文献:

- [1] Qiao Zizhi. Smarter phone based live QOE measurement [C]// Proc 15th Int'l Conf on Intelligence in Next Generation Networks (ICIN2011). Berlin, Germany:[s. n.], 2011:64-68.
- [2] Huang Fenghui, Zhou Wenan, Du Yu. QoE issues of OTT services over 5G network [C]//Proc 9th Int'l Conf Broadband and Wireless Computing, Communication and Applications (BWCCA2014). Guangzhou, China:[s. n.], 2014: 267-273.
- [3] Zheng Kan, Yang Zhe, Zhang Kuan, et al. Big data driven optimization for mobile networks towards 5G [J]. IEEE Network Mag, 2016, 30(1):44-51.
- [4] Kamel A, Al-Fuqaha A, Guizani M. Exploiting client-side collected measurements to perform QoS assessment of IaaS [J]. IEEE Trans Mobile Computing, 2015, 14(9):1876-1887.
- [5] 林闯,胡杰,孔祥震.用户体验质量(QoE)的模型与评价方法综述[J],计算机学报,2012,35(1):1-15.
Lin Chuang, Hu Jie, Kong Xiangzhen. A review of the models and evaluation methods of user experience quality (QoE) [J]. Journal of Computer, 2012, 35(1):1-15.
- [6] Aggarwal V, Halepovic E. Prometheus: Toward quality-of- experience estimation for mobile apps from passive network measurements [C]//Proc ACM HotMobile'2014. Santa Barbara, USA:[s. n.], 2014:26-27.
- [7] Oyman O, Singh S. Quality of experience for HTTP adaptive streaming services[J]. IEEE Communications Magazine, 2012, 50(4):20-27.
- [8] Singh K, Hadjadj-Aoul Y, Rubino G. Quality of experience estimation for adaptive HTTP/TCP video streaming using H. 264/AVC[C]//Proc 9th Annual IEEE Consumer Communications and Networking Conference. Las Vegas, Nevada, USA: IEEE, 2012:127-131.
- [9] Kumar S, Turner J, Williams J. Advanced algorithms for fast and scalable deep packet inspection[C]//Proc of ACM/IEEE Symposium on Architectures for Networking and Communications Systems. San Jose, USA:[s. n.], 2006: 81-92.
- [10] 中国电信集团公司技术标准[S]. 中国电信移动互联网业务感知测试 APP 功能规范,2015.
Technical Standard of China Telecom[S]. Functional Specification of China Telecom Mobile Internet Service Perception Test APP, 2015.
- [11] 张成伟,程文青,陈晓军.基于 Android 平台的 3G 移动网络测量研究及性能分析[J]. 计算机科学,2015,42(2):24-28.
Zhang Chengwei, Cheng Wenqing, Hei Xiaojun. Research and performance analysis of 3G mobile network based on android platform [J]. Computer Science, 2015, 42(2):24-28
- [12] 爱立信.用大数据提升运营商 NPS [J]. 电信工程技术与标准化,2015,28(3):92.
Ericsson. Enhancing carrier NPS with big data [J]. Telecommunication Engineering Technology and Standardization, 2015, 28(3):92.
- [13] Naboulsi D, Fiore M, Ribot S, et al. , Large-scale mobile traffic analysis: A survey [J]. IEEE Communications Surveys & Tutorials, 2016, 18(1): 124-161.
- [14] Ganti R, Ye F, Lei H. Mobile crowdsensing: Current state and future challenges[J]. IEEE Communications Magazine, 2011, 49(11):32-39.
- [15] He Ying, Yu Fei, Zhao Nan, et al. Big data analytics in mobile cellular networks [J]. IEEE Access, 2016, 4(1):

1985-1996.

- [16] 陈森, 陈超, 张小勇, 等, 基于大数据分析的移动互联网用户感知评估系统 [J]. 电信科学, 2015, 4: 1-8.
Chen Sen, Chen Chao, Zhang Xiaoyong, et al. Mobile Internet user perception evaluation system based on big data [J]. Telecommunication Science, 2015, 4: 1-8.
- [17] 马啸威, 曹维华, 李文云, 等, 移动互联网业务感知质量优化方法及系统 [J]. 广东通信技术, 2017, 2: 40-45.
Ma Xiaowei, Cao Weihua, Li Wenyun, et al. Optimization methodology and system for mobile Internet service perception [J]. Guangdong Telecommunication Technology, 2017, 2017(2): 40-45.
- [18] Gerber A, Hajiaghayi M, et al. To cache or not to cache: The 3G case [J]. IEEE Internet Comput, 2011, 15(2): 27-34.
- [19] Reshef D, Reshef Y, Finucane H, et al. Detecting novel associations in large data sets [J]. Science, 2011, 334: 1518-1524.
- [20] Reshef D, Reshef Y, Finucane H, et al. Supporting online material for detecting novel associations in large data sets [EB/OL]. <http://science.sciencemag.org/content/suppl/2011/12/14/334.6062.1518.DC1>. 2011.
- [21] Kraskov A, Stogbauer H, Grassberger P. Estimating mutual information [J]. Phys Rev E, 2004, 69: 1-16.
- [22] 孙晓军, 周宗奎. 探索性因子分析及其在应用中存在的主要问题 [J]. 心理科学, 2005, 28(6): 1440-1442.
Sun Xiaojun, Zhou Zongkui. Exploratory factor analysis and its main problems in application [J]. Psychological Science, 2005, 28(6): 1440-1442.

作者简介:



徐小龙 (1991-), 男, 硕士研究生, 研究方向: 数据挖掘, E-mail: logxiaoxu058@163.com。



李克 (1972-), 男, 通讯作者, 博士, 教授, 研究方向: 智能运维、移动大数据和感知无线电, E-mail: like@buu.edu.cn。



王海 (1991-), 男, 硕士研究生, 研究方向: 数据挖掘。



宋晓勤 (1973-), 女, 博士, 副教授, 研究方向: 认知网络、感知无线电和网络QoS。



王征 (1973-), 男, 硕士, 研究方向: 移动网络感知。



吴雄 (1975-), 男, 本科, 研究方向: 移动网络感知。

(编辑: 刘彦东)