

3G网络中移动视频质量评估模型的研究

陈希宏¹ 金跃辉¹ 杨 谈^{1,2}

(北京邮电大学网络与交换技术国家重点实验室 北京 100876)¹

(北京邮电大学软件学院 北京 100876)²

摘要 随着3G网络技术的不断发展和广泛应用,移动视频业务比以往更受用户的关注。与传统的有线网络视频业务相比,移动视频的传输条件不太稳定,更容易产生误码;移动终端的视频播放性能更容易受到设备硬件的限制,这就要求有更适合移动终端的视频编码方式。此外,不同类型的视频内容、用户的兴趣爱好等因素也会对用户观看视频的体验产生不同的影响。以上因素给移动视频服务提供商在业务质量的评估以及用户体验的提升方面提出了巨大的挑战。目前在移动视频质量评估的研究中,主要采用基于服务质量(Quality of Service, QoS)的评价方法,但是这些方法没有考虑用户主观体验参与在内的诸多因素,因此并不是一种非常有效的评价方法。针对影响移动视频用户体验质量的主客观因素,研究了无线参数、终端设备参数和视频编码参数对移动视频质量的影响,提出了基于用户体验质量(Quality of Experience, QoE)的视频质量评价方法。

关键词 移动视频质量, 用户体验质量, 主观评价, 层次分析法

中图分类号 TP393 文献标识码 A DOI 10.11896/j.issn.1002-137X.2015.9.018

Study on Quality Assessment Model for Mobile Videos over 3G Network

CHEN Xi-hong¹ JIN Yue-hui¹ YANG Tan^{1,2}

(State Key Laboratory of Network and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China)¹

(School of Software Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China)²

Abstract With the rapid development of 3rd generation network technology, the mobile video service has attracted more attention of the mobile users than ever. Compared to the traditional cable network video service, mobile video service is more sensitive to the transport condition and there can be more probabilities of bit error. And the performance of mobile video is limited to the configuration of mobile devices' hardware, which means it's necessary to choose an appropriate way to encode video for mobile videos. Besides, some other factors can affect the quality of user experience for mobile videos as well, such as video contents, user's interests and so on. All these factors make it difficult for the mobile video service providers to evaluate the quality for mobile videos, as well as leading to an improvement of the mobile video experience for users. At present, most of the researches on assessment of video quality depend on quality of service. But it is not an effective method, because this method doesn't take the factors of user experience into consideration. Therefore we proposed a new subjective assessment of quality for mobile video, and in the research we adopted a user intervention approach when collecting data and focusing on user experience in order to achieve scientific and effective strategies to evaluate the quality for mobile video service.

Keywords Mobile video quality, QoE, Subjective assessment, Analytic hierarchy process

1 引言

随着3G无线网络技术的兴起和广泛应用,无线通信网络已经具备承载多媒体业务的能力,而智能手机和平板电脑的普及以及移动视频应用的迅速发展,使得移动视频业务更具有互动性,也进一步提高了移动视频业务的使用量。

与有线网络相比较而言,3G无线网络状况不太稳定,除去网络层的传输速率的波动之外,手持设备的移动性将导致

手机的信号强度发生变化,进而影响到网络传输速率;此外,无线网络环境比有线网络更为恶劣,因此数据传输的误码率也会更高,而高压缩的码流对传输错误非常敏感,甚至还会造成错误向后扩散;手机终端的性能也决定了移动视频业务的编解码处理方式不能过于复杂。因此,面向无线网络的移动视频业务对视频编码技术和传输技术提出了更大的挑战。

移动视频业务不断增长,已经成为包括移动运营商在内的服务提供商的主流业务之一。随着用户需求的不断变化,

到稿日期:2014-05-17 返修日期:2014-07-17 本文受国家863项目(2011AA01A102),国家973项目(2009CB320505),中央高校基本科研业务费专项资金(2014RC0501)资助。

陈希宏(1990-),男,硕士,主要研究方向为移动互联网软件, E-mail: xh-chen@163.com;金跃辉(1965-),女,硕士,教授,主要研究方向为网络测量与评估、移动互联网应用;杨谈(1982-),女,博士,讲师,主要研究方向为网络测量与性能分析、社会网络分析、移动互联网。

服务的多样性也越来越明显,在直播、点播、下载等基本业务类型的基础上,出现了微视、快拍等基于用户原创形式的视频类型。就目前的移动视频业务而言,用户主要使用视频点播和直播业务,包括电影、音乐、体育、动漫、直播、原创、电视剧、综艺娱乐新闻等类型。此外,喜好不同的用户在观看同一个视频时,其体验也会明显不同。这些因素都将影响到移动视频业务的用户体验。

为了能在日益激烈的竞争中获得更多的用户认可,各大服务提供商都致力于找到一种以用户满意度为标准的视频质量评价方法。目前主流的方法是采用基于服务质量即 QoS 的评价方法^[1]。QoS 主要包括网络的时延、丢包率、抖动等指标,但是这些指标只考虑了网络层的性能,是一种客观评价方法,依赖于监测网络层性能参数^[2-4]、应用层性能参数^[5]或者用户的行为方式来进行评价。现有的大量技术和标准基于客观方法评估应用的质量,例如音频质量评估^[5]、视频质量评估^[6],因而客观方法具有一定的局限性。基于用户体验质量即 QoE 的主观评价方法是一种以用户主观感受为标准的视频的评价方法^[7,8]。MOS(Mean Opinion Score)^[1]方法是目前使用最广泛的主观用户体验评价方法,但它具有一定的局限性,如文献^[9]中所述。除了以上两种方法,文献^[10]中综合了主观和客观的方法,通过结合测量得到的网络层性能参数和用户在体验质量下降时提交给测量系统的反馈来对 QoE 进行评判。以上方法均未从视频质量整体评估体系出发进行研究,没有充分考虑用户体验,也没有针对移动场景进行特定分析。除了网络层性能指标以外,QoE 还包括了无线参数、移动设备的性能、视频编码方式、用户的观看习惯等因素^[7]。此外,不同的用户在观看不同类型的视频内容时也对视频质量有不同的要求。这些指标分布在不同的层次上,在以 QoE 为基础的视频质量评价方法研究时^[11],需要考虑这些因素。

本文针对移动视频的特点,围绕影响 QoE 的主客观因素,主要研究无线参数、终端设备参数和视频编码方式对移动视频质量的综合影响,提出了一种基于无线参数、视频编码参数和终端设备参数的移动视频质量评价方法。本文第 2 节主要介绍影响移动视频质量的关键因素;第 3 节主要介绍研究所需的用户体验数据集的收集和处理;第 4 节主要介绍移动视频质量评估模型的建立;第 5 节是模型的求解与验证过程;最后是本文的研究总结和下一步工作计划。

2 影响移动视频质量的关键因素

2.1 无线参数

移动视频业务从源到终端,需要经过视频源的压缩编码、有线网络和无线网络的传输、移动终端的接收和解码 3 个基本过程。其中需要经过无线网络的传输是移动视频业务与传统多媒体视频业务最大的区别。无线网络自身的复杂性和特殊性,使得在无线网络传输过程中,移动视频业务的质量对无线网络参数非常敏感。

3G 无线网络的空中接口测试很大程度上是对物理层相关指标的测试,而物理层的性能也是影响移动网络的质量好坏的最大因素。3G 无线网络参数主要包括了接收终端信号强度(Received Signal Strength Indication, RSSI)、接收信号功率(Received Signal Code Power, RSCP)、发射功率(Tx Power, Tx)、接收功率(Rx Power, Rx)、信干比即手机可用导

频的信号强度与干扰信号的强度之比(Signal to Interference Radio, SIR)、信噪比即每一个码片接收到的可用导频的能量强度与手机接收到的其他小区的干扰能量之比(Energy Chip/Interfere Other Cell, Ec/Io)等等。表征 3G 无线网络质量的指标有很多,在无线网络优化测试和网络性能测试中比较常见的就是以上提出的 RSSI、RSCP、Ec/Io、Rx_Power、Tx_Power、SIR 6 项指标,这些指标对 3G 网络的影响更为直接有效,其物理意义也更能体现出手机等移动终端所处的 3G 网络的质量。

2.2 编码参数

移动视频的质量与视频的失真程度密切相关,失真程度越低,视频的质量就越好。当视频在经过网络传输之前,视频的失真由有损压缩编码引起,因此视频的压缩失真与采用的压缩标准、压缩速率等因素直接相关,可以说视频编码参数决定了视频源经过压缩编码、通过网络传输前的失真程度,直接决定了视频服务器上的视频源的原始质量。

所谓视频编码方式就是指通过特定的压缩技术,将某种视频格式的文件转换成另一种视频格式文件的方式。常见的视频编码方式有 H. 261、H. 263、MPEG-1/2、MPEG-4、JVT、H. 264/AVC 等。H. 264 集中了以往标准的优点,并吸收了以往标准制定中积累的经验,采用简洁设计,使它比 MPEG-4 更容易推广。H. 264 创造性地使用了多参考帧、多块类型、整数变换、帧内预测等新的压缩技术,使用了更精细的分像素运动矢量和新一代的环路滤波器,使得压缩性能大大提高,系统更加完善,因此更加适合移动视频的压缩编码^[12]。在编码参数中,影响视频质量的参数主要有帧率、分辨率、颜色数、压缩比、关键帧。其中颜色数的性能对于现在的手机视频来说设置相差不大,而关键帧这个因素一般是在传输过程中丢失而造成视频质量下降^[13]。本研究选取帧率、分辨率、压缩比作为测试的参数集。其中由于压缩比反映在视频的编码速率也就是比特率上,而比特率还能包含文件大小等更多的信息,因此采用视频比特率取代压缩比作为一个指标^[14]。在一些在线视频播放的研究中,也提出可以利用终端的编码器,在播放时将参数反馈到视频服务器上^[15],进而对视频编码参数进行即时调整来优化视频质量。

2.3 终端参数

移动终端设备的性能直接决定了视频经过网络传输、解码之后的视频失真程度。在一款硬件配置高、屏幕显示效果好的手机上往往能够获得非常好的视频观看体验。影响移动终端的视频性能的参数有很多,包括屏幕尺寸、屏幕材质、屏幕分辨率、CPU 主频、RAM 大小等等。而对于视频质量影响较为明显的参数为 CPU 主频、屏幕尺寸、屏幕材料、屏幕分辨率和 RAM 大小^[16-18],下文将对这 5 项参数进行分析,研究它们对视频质量的影响。

3 数据收集与处理

3.1 测试环境

首先需要准备视频源。根据第 2 节所述,将收集无线参数、编码参数和终端参数环境下用户观看视频的体验。其中无线参数将在 3G 无线网络环境下进行测试和收集,编码参数和终端参数将在本地环境下进行测试和收集。

在线视频源主要是用于无线参数和视频内容测试。视频源来自于主流视频网站的视频,例如优酷网、乐视网、搜狐视

频等。将测试视频的原始链接地址保存下来,按照一定的格式存储在 csv 格式的文件中。

根据在线测试时所采用的视频源,将它们下载到本地后转换成不同格式的视频,导入移动设备。

其次是终端设备。根据需要测量收集的参数的不同,测试时设备的使用情况也不一样,在无线参数、视频内容测试中使用一台相同型号的终端,而编码参数、终端参数测试时需要使用多台不同型号的终端。测试时所使用的设备如表 1 所列。

表 1 测试使用的设备

终端设备厂商	型号
HUAWEI	U8860
Motorola	DEFY+
HTC	G17
MEIZU	MX2
Xiaomi	2S

最后是测试软件,整个测试过程中需要使用两个测量软件,它们分别是 VideoAnalyzer 和 PHU GENEX。在测量 3G 无线参数时需要同时使用两款软件,测量编码参数和终端参数时只需要使用 VideoAnalyzer 软件。各项测试中具体使用软件的情况如表 2 所列。

表 2 测试使用的软件

参数类型	采集软件
无线参数	PHU GENEX 和 VideoAnalyzer
编码参数	VideoAnalyzer
终端参数	VideoAnalyzer

VideoAnalyzer 测试软件是我们自主开发的软件,是一款专门针对移动终端视频质量主观评价和数据收集而设计的评分工具;支持 2.3 以上的 Android 系统,支持市面上 95% 以上格式的视频的播放,并在播放过程中支持用户打分评价,能够将用户评价的分数存储到本地数据库中以供数据分析。在用户进行视频内容测试、编码参数和终端参数的测试过程中,VideoAnalyzer 可以自动化收集这些环境参数并将其保存在数据库中。

在测试设备上安装 VideoAnalyzer 软件,安装完成后打开软件,主界面如图 1(a)所示。在主界面下方的功能按钮中,选择中间的批量添加功能,导入上面提到的 csv 视频链接文件。在主界面点击 3 个菜单中的一项,如 3G Para Test,可以进入视频列表,界面如图 1(b)所示。

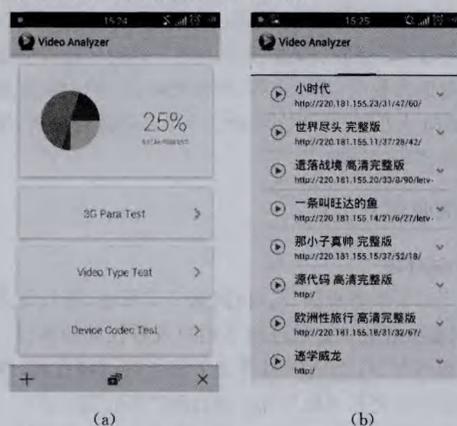


图 1 VideoAnalyzer 测试软件

PHU GENEX 无线参数测量软件是一款移动测量软件,这个工具能够测量 3G 无线参数的变化情况。如图 2 所示,3G 功能测试界面上可以看到 PHU 软件能够测量 3G 无线参数的变化,它能够测量结果保存在一定格式的文件中。在整个主观测试的过程中,让 PHU 在后台保持运行。

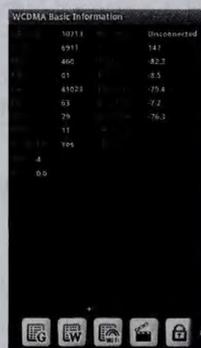


图 2 PHU 无线参数测试软件

3.2 测试方案

在测试方案中,主要包括用户体验主观评价标准和观看体验的测试流程。此处采用的是人工参与视频质量评分的主观评价方法,比起其他研究中的客观评价法^[6],此方法围绕用户体验的主观评分,将参数与用户主观评分进行关联分析,更体现了用户体验为主的目标。用户观看视频时的主观体验容易受到多种主观因素的影响,而这些主观因素与无线参数、编码参数和终端参数等客观因素密切相关,在测试数据的时候用户的主观体验评分将受到主观因素的影响。3G 网络移动视频质量评估测试需要用户在 3G 网络的环境下观看移动视频,其它的观看条件可参考 Rec. ITU-T P. 911 标准提出的视频质量主观测试标准。

在影响用户体验的主观因素中,视频内容是与用户体验密切相关的一个主观因素。视频的内容从题材上可以分为新闻、电影、电视剧、动漫、体育、娱乐等等;从形式上可以分为点播、直播、下载、上传等等^[18]。随着社交网络的兴起,诸如视频分享的一些视频分享业务也备受用户青睐^[19]。对于不同类型的视频,用户在观看时的关注点也不相同,通常视频质量的属性包括等待时间、画质、音质、流畅度、音画同步性。例如用户观看电影时,用户对于等待时间的容忍程度较高;而在观看新闻短片时,对于等待时间的容忍程度则会降低许多。此外,用户的兴趣爱好等背景因素也会对用户体验造成一定的影响。例如足球迷在观看球赛直播时对视频质量的要求与非足球迷不一样。因此在画质、音质、音画同步和流畅度方面,不同类型的视频内容和用户自身的兴趣爱好,都会对用户体验造成不同的影响^[20]。因此,在进行数据收集的过程中,需要将不同用户的主观体验的实际评分值作为参考。

在评价标准方面,无线参数、编码参数和终端参数测试需要用户在观看视频时对视频整体的观看体验进行一次主观评分。而视频内容方面,如第 2 节所述,采用的是属性评分法,所以需要用户对等待时间、画质、音质、流畅度、音画同步性分别进行主观评分之后,再进行整体的视频观看体验评分^[17]。在测试流程方面,用户首先都要学习主观评价标准并对视频主观评价有一个初步的认识,再使用 VideoAnalyzer 观看视

频,同时进行主观评分^[21]。在主观评价的过程中主要采用表3所列的评价标准。

表3 主观评价标准

评分值	主观体验
5	非常满意
4	比较满意
3	一般
2	比较差
1	非常差

整个测试过程将严格按照测试流程进行。如上文所述,测试的一般流程是,首先了解评分标准,接着进行观看视频的体验,最后进行正式的评分。其中在视频内容测试的过程中,需要对视频进行属性评分之后再整体评分。

因此,测试的具体操作过程如下所示:

第1步 评分标准介绍。向测试人员介绍评分标准,并说明测试的流程。

第2步 根据评价标准,形成移动视频观看体验。让测试人员观看视频,根据前一步的评分标准,形成初步的评价体验。

第3步 正式测试,观看视频,根据主观体验评分。在完成前两步之后,测试人员可以在移动设备的 VideoAnalyzer 软件中自由选择视频进行观看,并根据主观体验进行评分。

3.3 结果处理

按照上面的测试方案,主观评分的数据采集过程中,除了需要人工参与主观评分的工作以外,还需要终端设备进行参数的采集。然而,终端设备上采集的参数需要进行量化后才能用来进行评估和分析。此外,还需要对量化之后的数据进行归一化处理。归一化是一种简化计算的方式,即将有量纲的表达式变换为无量纲的表达式,使其成为纯量,目的是避免具有不同物理意义和量纲的输入变量不能平等使用的问题。在移动视频主观评测的结果集中,数据量化和归一化可以同时进行。首先将数据按照各自的取值范围划分成5-10个区间,根据不同的区间赋予不同的量化值,各量化值分别为1-5之间的整数或者是以0.5为步长的小数。然后将各参数量化数值的最大值作为分母,其他的量化值除以该最大值就可以做到归一化。为了保证各参数的测试结果具有实际意义,当各测量结果超出表4-表8中规定的参数取值范围时,可直接将其剔除。

在影响移动视频质量的3类因素中,各类参数量化方法不一样。针对这3类因素及其物理含义,按照各自的测试结果集中的取值范围制定了下面的归一量化标准,表4列出无线参数归一量化标准,表5列出终端参数的标准,表6-表8列出编码参数的标准。下面的终端参数标准还需要根据市面上的移动终端实际情况进行修改和更新,增加更高端的配置参数,剔除过时的低端参数。

表4 无线参数归一量化标准

RSSI	Ec/Io	归一量化值
(-113, -95)	(-22, -17)	0.2
[-95, -85)	[-17, -12)	0.4
[-85, -75)	[-12, -8.5)	0.6
[-75, -65)	[-8.5, -4.5)	0.8
[-65, -50)	[-4.5, -1)	1.0

表5 终端参数归一量化标准

CPU	RAM	Dpi	Size	Material	归一量化值
800	N/A	N/A	2.4	N/A	0.5
1024	512	200	3.2	TFT26	0.6
1228	768	200-225	3.5	TFT1600	0.7
1433	1024	225-250	4.0	SLCD	0.8
1024	512	250-300	4.3	OLED	0.9
(not Android)	(not Android)				
2*1220+	1500+	300+	4.7+	AMLOED+	1.0

表6 编码参数归一量化标准——比特率

Bitrate	归一量化值
256	0.25
512	0.5
1024	1

表7 编码参数归一量化标准——帧率

Frame	归一量化值
12	0.3
15	0.5
25	0.8
30	1

表8 编码参数归一量化标准——分辨率

Resolution	归一量化值
128×96	0.1
160×120	0.2
176×144	0.3
320×240	0.4
352×240	0.5
352×288	0.6
400×240	0.7
480×320	0.75
480×360	0.8
640×480	0.85
720×480	0.9
1280×800	0.95
1920×1080	1

例如测试出的一组结果中,无线参数为(RSSI, Ec/Io) = (-95dB, -13dB), 编码参数为(resolution, framerate, bitrate) = (640×480, 25Hz, 512kbps), 终端参数为(CPU, RAM, 屏幕尺寸, 屏幕材质, 屏幕分辨率) = (1GHz, 1.4GB, 4.3寸, AMOLED, 1024*768), 归一量化之后各项参数变为(RSSI, Ec/Io) = (0.4, 0.6), (resolution, framerate, bitrate) = (0.85, 0.5, 0.5), (CPU, RAM, 屏幕尺寸, 屏幕材质, 屏幕分辨率) = (3, 4.5, 4.5, 5, 4.5)。这样就完成了测试数据集的量化和归一化。

按照上述的处理方法,将不同环境中采集的各个参数环境下的用户主观体验值作量化归一处理,得到整体的数据结果集。然后把整体测试结果集分成训练集和验证集两部分,训练集用来构建模型,计算模型的系数;验证集用来验证模型的准确性。

4 评估模型

4.1 改进的层次分析法

层次分析法 (Analytic Hierarchy Process, AHP) 由美国运筹学家 Thomas 首次提出,是用来解决多指标综合问题的一种方法。层次分析法是指将一个复杂的多目标决策问题作为一个系统,将目标分解为多个目标或准则,进而分解为多指

标(或准则、约束)的若干层次,通过定性指标模糊量化方法算出层次单排序(权数)和总排序,以作为目标(多指标)、多方案优化决策的系统方法。它结合定性和定量的方法,同时减少了主观因素的影响,确保系统和模型的合理性,提供了一个科学的评价分析方法。

在层次分析法的处理过程中,将决策问题按总目标、各层子目标、评价准则直至具体的实施方案(或者指标)的顺序分解为不同的层次结构,然后用求解判断矩阵特征向量的办法,求得每一层次的各元素对上一层次某元素的优先权重,最后再用加权的方法逐层向上归并各实施方案对总目标的最终权重,最终权重最大者即为最优方案。

“优先权重”是一种相对的量度,它表明各实施方案在某一特点的评价准则或子目标,表示优越程度的相对量度,以及各子目标对上一层目标而言重要程度的相对量度。这个相对量度需要由专家参与评定。层次分析法比较适合于具有分层交错评价指标的目标系统,而且目标值又难于定量描述的决策问题。因此具体用法是采用专家打分法构造判断矩阵,求出其最大特征值及其所对应的特征向量,归一化后,即为某一层指标对于上一层次某相关指标的相对重要性权重。

层次分析法在描述同一层指标关系时,需要依赖专家的经验,这增加了层次模型建立的难度和随机性。所以本文模型将采用一种实际数据驱动的改进的层次分析法,即在对同一层的指标进行分析时,通过采集的实际测试数据进行多元线性回归分析,来判断各个指标之间的相对数学关系。这样定量地计算出各个指标的权重系数,解决了原来层次分析中需要依赖专家打分的问题,使得层次模型更加客观、准确。

4.2 构建层次模型

根据层次分析法将一个决策问题按照总目标、评价准则到各具体指标的分解顺序,把移动视频质量的评估值作为第一层目标层;无线参数、编码参数、终端参数作为第二层准则层;具体的各个参数作为第三层指标层,例如 RSSI、Ec/Io、分辨率、帧率、比特率、CPU、RAM、屏幕尺寸、屏幕材质、屏幕分辨率。整体层次模型如图 3 所示。

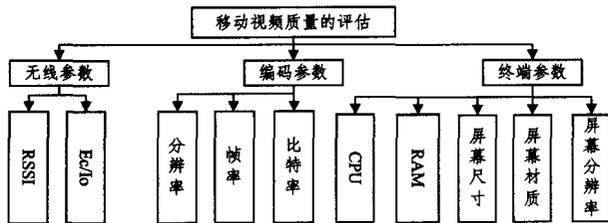


图 3 层次评估模型

根据上面构建的层次评估模型,将移动视频质量评估问题分解成为分析无线参数、编码参数和终端参数 3 个准则的子问题,并进一步将无线参数、编码参数和终端参数这 3 项准则分解到各准则下的参数指标上。

5 模型的求解和验证

5.1 层次模型求解

上述层次模型将各个指标按照不同类型划分成了目标层、准则层、指标层。在进行模型构建时,首先建立指标层对准则层的影响模型,再建立准则层对目标层的影响模型,最后推导出指标层对目标层的影响模型。

在模型求解时,主要采用的是多元线性回归的方法。多元线性回归方法是指一种现象是与多个因素相联系的,受到多个因素共同作用的影响。由多个自变量的最优组合共同预测或估计因变量,比只用一个自变量进行预测或估计更有效,更符合实际。

在实际的多元线性回归分析中,需要对数据源进行归一化处理,然后计算多元线性回归系数,检验回归系数的合理性,并不断调整系数和自变量,使用验证集检验模型的准确率,最终得出有效的模型。下面将首先分别建立指标层各项参数对准则层中 3 个准则的影响模型,接着再构建准则层对目标层的影响模型,最后推导出指标层中的各个参数对目标层移动视频质量的影响模型。

5.1.1 多元线性回归分析

多元线性回归分析包括了模型的构建和模型系数的检验两个步骤,其中模型的构建包括数据集的归一化处理 and 多元线性回归计算,模型系数的检验需要判断模型的拟合优度、回归方差、回归系数的显著性和共线性。

首先,对于各项参数,通过测试可以得到它的一系列值 \vec{m} ,前面第 3 节的数据收集中,对结果进行了归一化处理,具体方法是在这个向量中找出绝对值最大的那个元素 m_{\max} ,将 \vec{m} 的所有的元素除以 m_{\max} 后得到归一化的参数向量 \vec{m}' 。

其次,计算多元线性回归模型的系数。具体方法是将归一化的各项参数值向量 \vec{m}' 和用户的主观评价结果导入,进行多元线性回归分析,采用逐步剔除自变量的方法进行回归计算。通过多元线性回归法可以得出到模型汇总表、回归系数表和方差分析表。这些表中分别包括了模型的相关系数 R 、拟合优度 R^2 、各个参数的回归系数 B 、显著程度 sig 、方差膨胀系数 VIF ,以及模型的回归方差和残差。

最后,需要对计算出来的系数进行检验。在拟合模型的计算结果中,相关系数 R 和拟合优度 R^2 都是在 $(0,1)$ 之间,并且要求相关系数和拟合优度都是越接近 1 越好,至少要达到 0.5 以上才算基本合格,越大说明模型越能够符合实际情况;然后是显著程度 sig ,显著程度 sig 小于 0.1 才算符合要求,并且显著程度越小,说明模型能够解释实际情况就越明显;接下来是方差膨胀系数 VIF ,方差的膨胀系数表示的是多个参数的共线性,越大则共线性越明显,说明几个参数中需要剔除某些关联性较大的参数;最后是回归方差和残差,二者之和为总方差,回归方差占总方差的比例越大,则说明模型能够解释的实际数据越多,模型越可靠。

5.1.2 构建层次模型

在层次总模型中,将目标层、准则层、指标层的各个参数表示成数学符号,各自的说明如表 9 所列。

根据上文的模型,抽象出来的层次模型表示为:

$$QoE_{predict} = \alpha_0 \cdot M_1^{\alpha_1} \cdot M_2^{\alpha_2} \cdot M_3^{\alpha_3} \quad (1)$$

式(1)中的 3 个准则可表示为:

$$\begin{cases} M_1 = \sum_{i=1}^2 m_{1i} \cdot x_{1i} + x_{10} \\ M_2 = \sum_{i=1}^3 m_{2i} \cdot x_{2i} + x_{20} \\ M_3 = \sum_{i=1}^5 m_{3i} \cdot x_{3i} + x_{30} \end{cases} \quad (2)$$

在上面的两个公式中,式(1)表示准则层对目标层的影响

模型,因为在准则层对目标层的影响模型中,3个准则对于目标层而言都是必不可少的,因此它们之间的关系为乘积。上面的系数向量 $\vec{\alpha}=(\alpha_0, \alpha_1, \alpha_2, \alpha_3)$ 中 $\alpha_1, \alpha_2, \alpha_3$ 表示的是3个准则对于目标层的影响程度, α_0 为常数项。

式(2)表示的是指标层对准则层的影响模型。其中各个数学变量的系数分别是指标层中各项参数对相应准则的影响因子,也就是多元线性回归模型的回归系数和常数,分别表示为:

$$\begin{cases} \vec{B}_1=(x_{10}, x_{11}, x_{12}) \\ \vec{B}_2=(x_{20}, x_{21}, x_{22}, x_{23}) \\ \vec{B}_3=(x_{30}, x_{31}, x_{32}, x_{33}, x_{34}, x_{35}) \end{cases} \quad (3)$$

向量 \vec{B}_1 表示的是RSSI、Ec/Io对无线参数的回归系数,向量 \vec{B}_2 表示的是分辨率、帧率、比特率的回归系数,向量 \vec{B}_3 表示的是CPU主频、RAM大小、屏幕尺寸、屏幕材质、屏幕分辨率的回归系数。

表9 层次模型中的数学符号含义

符号	说明
QoE	用户体验值的实际值
QoE _{predict}	目标层中用户体验值的评估值
QoE _i	测试结果集中每一条记录的用户体验实际值
QoE _{i-predict}	测试结果集中每一条记录的用户体验评估值
M ₁	无线参数
M ₂	编码参数
M ₃	终端参数
m ₁₁	RSSI
m ₁₂	Ec/Io
m ₂₁	分辨率
m ₂₂	帧率
m ₂₃	比特率
m ₃₁	CPU主频
m ₃₂	RAM大小
m ₃₃	屏幕尺寸
m ₃₄	屏幕材质
m ₃₅	屏幕分辨率

构建模型的过程也就是求解式(1)和式(2)系数的过程。在构建模型时,使用数据集中的训练集,首先构建指标层对准则层的影响模型,得出各项参数对各自的准则的回归系数,即求出式(2)的系数;接着再构建准则层对目标层的影响模型,得出各准则对目标层的回归系数,即求出式(1)的系数;最后将求出系数之后的式(2)代入式(1),推导出指标层对目标层的影响模型。

首先构建的是指标层对准则层的影响模型。根据多元线性回归方法,计算回归模型的回归系数,检验通过之后,就可以得到各个参数的回归系数。指标层的各个参数对于准则层的各个准则影响的回归系数和常数项为式(3)。

在无线参数多元线性回归分析的过程中,发现RSSI和Ec/Io的多元线性回归模型的拟合优度很低,无法准确解释大多数的实际数据,因此无法采用多元线性回归模型构建无线参数的影响模型。但是从另外一个角度分析发现,当RSSI高于-93dB时,移动视频的用户体验评分值分布均匀,即用户体验与RSSI和Ec/Io没有明显的线性回归关系;而当RSSI低于-93dB时,移动视频业务几乎不可用,用户无法获得用户体验。因此,在影响模型中,无线参数准则的回归系数

$\vec{B}_1=(x_{10}, x_{11}, x_{12})$ 没有意义。可以将层次总模型中的无线参数准则抽取出来作为分段函数条件,即当RSSI高于-93dB时,根据模型中的编码参数和终端参数可以评估移动视频的用户体验;而当RSSI低于-93dB时,用户将不会获得移动视频业务的体验。

因此无线参数的回归系数向量 $\vec{B}_1=(x_{10}, x_{11}, x_{12})$ 没有意义。无线参数的影响模型为分段函数:

$$M_1 = \begin{cases} 1, & m_{11} \geq -93\text{dB} \\ 0, & \text{other} \end{cases} \quad (4)$$

除去无线参数以外,可以计算出编码参数和终端参数各自的多元线性回归系数,结果为:

$$\begin{cases} \vec{B}_2=(0.972, 3.004, 0.168, 1.355) \\ \vec{B}_3=(0, 0.40, 0.10, 0.15, 0.20, 0.15) \end{cases}$$

因此第三层指标层中各个参数对准则层的准则影响模型即式(2)可以表示为:

$$\begin{cases} M_1 = \begin{cases} 1, & m_{11} \geq -93\text{dB} \\ 0, & \text{other} \end{cases} \\ M_2 = \sum_{i=1}^3 m_{2i} \cdot x_{2i} + x_{20} \\ M_3 = \sum_{i=1}^5 m_{3i} \cdot x_{3i} + x_{30} \end{cases} \quad (5)$$

其中的回归系数分别为:

$$\begin{cases} (x_{20}, x_{21}, x_{22}, x_{23})=(0.972, 3.004, 0.168, 1.355) \\ (x_{30}, x_{31}, x_{32}, x_{33}, x_{34}, x_{35})=(0, 0.40, 0.10, 0.15, \\ 0.20, 0.15) \end{cases}$$

故计算出指标层对准则层的影响模型式(2)的系数之后,可以将其表示为式(5),它反映了指标层对准则层的影响关系。

此时,当无线参数为分段函数时,层次总模型式(1)也相应地分为两种情况,可以表示为式(6):

$$QoE_{predict} = \begin{cases} \alpha_0 \cdot M_2^{\alpha_2} \cdot M_3^{\alpha_3}, & m_{11} \geq -93\text{dB} \\ 0, & \text{other} \end{cases} \quad (6)$$

其中,无线参数的系数 α_1 已经从公式中剔除。接下来,只对当 $QoE_{predict}$ 不为0的情况进行分析,也就是只需要计算向量 $\vec{\alpha}=(\alpha_0, \alpha_2, \alpha_3)$ 的结果。

接下来构建准则层对目标层的影响模型。上文中已经将式(1)表示为式(6),式(6)中的第一部分即 $QoE_{predict}$ 不为0时,表示为式(7):

$$QoE_{predict} = \alpha_0 \cdot M_2^{\alpha_2} \cdot M_3^{\alpha_3} \quad (7)$$

上面的系数向量 $\vec{\alpha}=(\alpha_0, \alpha_2, \alpha_3)$ 中 α_2, α_3 表示的是编码参数和终端参数这两个准则对于目标层的影响程度, α_0 为常数项。在构建准则层对目标层的影响模型时,首先对上面的式(7)两边取对数,得到下面的式(8):

$$\lg(QoE_{predict}) = \alpha_2 \cdot \lg M_2 + \alpha_3 \cdot \lg M_3 + \lg \alpha_0 \quad (8)$$

这样就转化成为了求解多元线性回归方程的系数。采用上面相同的计算方法求出系数向量 $\vec{\alpha}=(\alpha_0, \alpha_2, \alpha_3)$ 后,便可构造准则层对目标层的影响模型。

经过多元线性回归计算和回归系数的验证,式(7)中的各项准则参数的回归系数计算结果为 $\vec{\alpha}=(10^{-0.628}, 0.278, 1.047)$,代入式(6)为:

$$QoE_{predict} = \begin{cases} 10^{-0.628} \cdot M_2^{0.278} \cdot M_3^{0.047}, & m_{11} \geq -93\text{dB} \\ 0, & \text{other} \end{cases} \quad (9)$$

式(9)就是计算出系数之后的式(1),反映了准则层对目标层的影响关系。

可以看到,经过回归分析和数学计算,求出了层次模型式(1)和式(2)的系数,即分别得到式(9)和式(5)。

最后,把式(5)代入式(9),可以推导出指标层对目标层的影响模型,通过这个评估模型可以预测各项指标参数环境下的用户体验值。

5.2 模型的验证

对模型进行验证时,主要分析预测结果的准确率(Accuracy)和平均绝对百分误差(Mean Absolute Percent Error, MAPE)。准确率是指,将验证集中的各项影响因素代入模型,计算出每条记录的评估值 $QoE_{i-predict}$,再与验证集中每条记录的实际用户体验值 QoE_i 比较,其中评估值 $QoE_{i-predict}$ 与实际值 QoE_i 相等的数目 M 占验证集总数量 N 的百分比。平均绝对百分误差的计算方式为,将验证集中各项影响因素代入模型,计算出每一条记录预评估值 $QoE_{i-predict}$,与每条记录的实际用户体验值 QoE_i 相减得到的绝对值,再除以实际值 QoE_i 得到绝对百分误差,最后将验证集中所有记录的绝对百分误差求和,再除以验证集的总数量 N 求出相对误差率的平均值 MAPE。

因此当验证集中有 N 条数据时,准确率 Accuracy 和平均绝对百分误差 MAPE 的计算公式分别为式(10)和式(11)。

$$Accuracy = \frac{M \times 100}{N} \% \quad (10)$$

$$MAPE = \frac{\sum_{i=1}^N (|QoE_{i-predict} - QoE_i| / QoE_i)}{N} \quad (11)$$

在上面构建数学模型的过程中,得到指标层对准则层影响的模型的系数,即式(3),以及准则层对目标层视频质量影响模型的系数 $\vec{a} = (a_0, a_1, a_2, a_3)$,进而得到了模型的两个公式,即式(5)和式(9)。

在第3节数据收集和和处理中,我们采集了不同终端下的3000条测试结果,包含了各项指标参数下的用户体验评分值。接着把数据结果进行了归一量化处理,并将处理之后80%的结果作为训练集,经过5.1节中的计算,求出了评估模型的系数。最后采用剩下的20%的数据,即验证集中的600条数据,来对评估模型的准确性进行验证。具体做法是把测量、处理所得到的各项指标(包括无线参数中的RSSI,编码参数中的分辨率、帧率、比特率,以及终端参数中的CPU主频、RAM大小、屏幕分辨率、屏幕尺寸、屏幕材质)代入上面得出的评估模型式(9)和式(5),预测出验证集中每一条记录的用户体验评估值 $QoE_{i-predict}$,再将这个预测值与测试中的用户体验实际值 QoE_i 进行比较,代入上面的式(10)和式(11)从而计算出 Accuracy 和 MAPE。

经过测试,有323条经过评估模型预测出来的 $QoE_{predict}$ 与用户体验的实际值 QoE 相同,因此评估模型得出的准确率达到了53.83%,平均绝对百分误差为10.62%,表明预测出的用户体验与真实结果比较接近,证明了所构建的评估模型具有一定的参考价值。

结束语 本文从用户体验的角度出发,提出了影响移动视频质量的关键因素,并创新性地提出按照参数指标的类型,构建分层次的模型来评估移动视频的质量。测试数据来源广泛,采用了主、客观评价方法相结合的思想,使得测试数据集更加符合用户实际使用移动视频业务的情况。在收集数据的过程中,采用了人工参与的主观评分法,收集了各项参数指标环境下的用户体验值。在评价模型系数的计算过程中,采用了多元线性回归方法来分析多因素对一个目标共同作用的影响,成功地将移动视频质量评估的问题分解到层次模型各层当中,通过逐层建立多元线性回归模型和数学推导的办法得出模型底层各个指标对移动视频业务质量的影响,最终构建出一个科学有效的移动视频质量评估模型。

本文研究了3G无线参数、编码参数、终端参数对移动视频质量的影响。而在移动视频业务中,网络传输层的参数,例如时延、丢包和带宽,对移动视频质量也有着一定的影响。网络传输层参数对移动视频的用户体验最直接的影响就是流畅度,这与观看移动视频的用户体验密切相关。此外视频内容对用户体验也有一定的影响,可以引入SI(Spatial Information)以及TI(Temporal Information)来反映视频内容。因此本研究的下一步工作将在用户观看移动视频的同时,引入视频内容对用户的影响,考虑更多类型的无线网络,例如Wi-Fi、LTE等,并测量网络传输层的一些关键参数,例如时延、丢包、带宽等,分析它们对移动视频质量的影响关系,集成到本文中已经提出的层次模型中,对现有的评估模型做出有效的扩展,使得移动视频质量层次评估模型更加全面、准确。

参考文献

- [1] 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 Xiang-zhen. Survey on Models and Evaluation of Quality of Experience[J]. Chinese Journal of Computers, 2012, 35(1): 1-15
- [2] Chen Kuan-ta, Huang Chun-ying, Huang P, et al. Quantifying Skype user satisfaction[J]. ACM SIGCOMM Computer Communication Review, 2006, 36(4): 399-410
- [3] Chen Kuan-ta, Huang P, Lei Chin-laung. Effect of Network Quality on Player Departure Behavior in Online Games[J]. IEEE Transactions on Parallel and Distributed Systems, 2009, 20(5): 593-606
- [4] Huang Te-yuan, Chen Kuan-ta, Huang P. Tuning Skype's redundancy control algorithm for user satisfaction[C]// INFOCOM 2009. IEEE, 2009: 1179-1187
- [5] Rix A W, Beerends J G, Hollier M P, et al. Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs[C]// 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2001 (ICASSP'01). Salt Lake City, UT, 2001: 749-752
- [6] Chikkerur S, Sundaram V, Reisslein M, et al. Objective Video Quality Assessment Methods: A Classification, Review, and Performance Comparison[J]. IEEE Transactions on Broadcasting, 2011, 57(2): 165-182
- [7] Balachandran A, Sekar V, Akella A, et al. Developing a predictive model of quality of experience for internet video[C]// Pro-

- ceedings of the ACM SIGCOMM 2013 conference on SIGCOMM. Hong Kong, China: ACM, 2013, 36(4): 339-350
- [8] Chan A J, Pande A, Baik E, et al. Temporal quality assessment for mobile videos[C]//Proceedings of the 18th Annual International Conference on Mobile Computing and Networking. ACM, 2012: 221-232
- [9] Knoche H, De Meer H G, Kirsh D. Utility curves: mean opinion scores considered biased[C]//1999 Seventh International Workshop on Quality of Service, 1999 (IWQoS'99). IEEE, 1999: 12-14
- [10] Chen Kuan-ta, Tu Chen-chu, Xiao Wei-chen. OneClick: A framework for measuring network quality of experience[C]//INFOCOM 2009. IEEE, 2009: 702-710
- [11] Choi S G, Jong Kim H. A study on a QoS / QoE correlation model for QoE evaluation on IPTV service[C]//The 12th International Conference on Advanced Communication Technology, 2010 (ICACT). IEEE, 2010: 2, 1377-1382
- [12] Zoran Bojković, Zoran Miličević. Subjective Video Quality Assessment in H. 264/AVC Video Coding Standard [J]. Telfor Journal, 2012, 4(2): 111-115
- [13] Tan H L, Li Z, Tan Y H, et al. A Perceptually Relevant MSE-Based Image Quality Metric [J]. IEEE Transactions on Image Processing, 2013, 22(11): 4447-4459
- [14] 臧伟. 基于无线网络环境的手机视频码率自适应技术研究[J]. 软件, 2012, 33(4): 81-83, 88
Zang Wei. Research of Mobile Video Rate Adaptive Method Based on Wireless Network Environment [J]. Computer Engineering & Software, 2012, 33(4): 81-83, 88
- [15] 邵阳, 马越, 郭捷. 移动互联网在线视频播放优化研究与实现 [J]. 微型电脑应用, 2009, 25(11): 10-12
Shao Yang, Ma Yue, Guo Jie. The Research and Implementation of the Optimization of Online Video Player Based on Mobile Network [J]. Microcomputer Applications, 2009, 25(11): 10-12
- [16] Jumisko-Pyykk S, H J, Kkinen. Evaluation of subjective video quality of mobile devices [C]//Proceedings of the 13th annual ACM international conference on Multimedia. Hilton, Singapore: ACM, 2005: 535-538
- [17] Garcia M, Canovas A, Edo M, et al. A QoE Management System for Ubiquitous IPTV Devices [C]//Third International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies, 2009 (UBICOMM'09). Sliema, 2009: 147-152
- [18] Jumisko S H, Ilvonen V P, Vaananen-Vainio-Mattila K A. Effect of TV content in subjective assessment of video quality on mobile devices [C]//Electronics Imaging 2005. International Society for Optics and Photonics, 2005: 243-254
- [19] 余忆晗. 手机视频用户满意度调查研究 [J]. 城市建设理论研究 (电子版), 2013(21)
Yu Yi-han. The Research of Satisfaction of Mobile Video Users [J]. ChengShi Jianshe LiLun Yan Jiu, 2013(21)
- [20] Dobrian F, Sekar V, Awan A, et al. Understanding the impact of video quality on user engagement [J]. ACM SIGCOMM Computer Communication Review, 2011, 41(4): 362-373
- [21] Watson A, Angela S M. Measuring perceived quality of speech and video in multimedia conferencing applications [C]//Proceedings of the Sixth ACM International Conference on Multimedia. ACM, 1998: 55-60

(上接第 77 页)

- [15] 彭震宇. 最大独立集和最小弱定点覆盖问题求解及其应用研究 [D]. 无锡: 江南大学, 2008
Peng Zhen-yu. Maximum independent set and minimum weak vertex cover problem solving and its application research [D]. Jiangnan University, 2008
- [16] 林宇, 程时端, 邬海涛, 等. IP 网端到端性能测量技术研究的进展 [J]. 电子学报, 2003, 31(8): 1227-1233
Lin Yu, Cheng Shi-duan, Wu Hai-tao, et al. The achievement of end-to-end performance measurement technologies in ip networks [J]. Wuxi: Acta Electronica Sinica, 2003, 31(8): 1227-1233
- [17] Dorigo M, Maniezzo V, Colomni A. Ant system: optimization by a colony of cooperating agents [J]. IEEE transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 1996, 26(1): 29-41
- [18] Dorigo M, Gambardella L M. Ant colony system: a cooperative learning approach to the traveling salesman problem [J]. IEEE Transactions on Evolutionary Computation, 1997, 1(1): 53-66
- [19] Dorigo M, Caro G D. Ant colony optimization: a new meta-heuristic [C]//Proceedings of the 1999 Congress on Evolutionary Computation, 1999 (CEC 99). 1999
- [20] 王芳, 李美安, 段卫军. 基于动态自适应蚁群算法的云计算任务调度 [J]. 计算机应用, 2013, 33(11): 3160-3162
Wang Fang, Li Mei-an, Duan Wei-jun. Cloud computing task scheduling based on dynamically adaptive ant colony algorithm [J]. Journal of Computer Applications, 2013, 33(11): 3160-3162
- [21] 杨星, 马自堂, 孙磊. 云环境下基于改进蚁群算法的虚拟机批量部署研究 [J]. 计算机科学, 2012, 39(9): 33-37
Yang Xing, Ma Zi-tang, Sun Lei. Research on Extended Ant Colony Optimization Based Virtual Machine Deployment in Infrastructure Clouds [J]. Computer Science, 2012, 39(9): 33-37
- [22] 李成兵, 郭瑞雪, 李敏. 改进蚁群算法在旅行商问题中的应用 [J]. 计算机应用, 2014, 34(S1): 131-132, 165
Li Cheng-bing, Guo Rui-xue, Li Min. Application of improved ant colony algorithm in travelling salesman problem [J]. Journal of Computer Applications, 2014, 34(S1): 131-132, 165
- [23] Waxman B M. Routing of multipoint connections [J]. IEEE Journal on Selected Areas in Communications, 1988, 6(9): 1617-1622
- [24] 吴庆洪, 张纪会, 徐心和. 具有变异特征的蚁群算法 [J]. 计算机研究与发展, 1999, 36(10): 1240-1245
Wu Qing-hong, Zhang Ji-hui, Xu Xin-he. An ant colony algorithm with mutation features [J]. Journal of Computer Research and Development, 1999, 36(10): 1240-1245
- [25] 叶志伟, 郑肇葆. 蚁群算法中参数 α, β, ρ 设置的研究 [J]. 武汉大学学报 (信息科学版), 2004, 29(7): 597-601
Ye Zhi-wei, Zheng Zhao-bao. Configuration of parameters α, β, ρ in ant algorithm [J]. Geomatics and Information Science of Wuhan University, 2004, 29(7): 597-601