
目录

摘要..... 1

二、问题分析..... 2

三、模型假设..... 2

四、符号说明..... 3

五、模型的建立..... 3

 （一）模型建立分析..... 3

 （二）建立过程..... 4

 ➤ 实验数据： 4

 ➤ 分析单一变量的影响： 5

 ➤ 分析多变量的影响： 6

 （三）数据求解..... 7

 问题二： 7

 ➤ 1、模型建立： 7

 ➤ 结果分析..... 8

六、 模型评价..... 8

七、参考文献..... 9

八、模型编程代码..... 10

摘要

随着无线宽带网络的升级，以及智能终端的普及，越来越多的用户选择在移动智能终端上用客户端 APP 观看网络视频，这是一种基于 TCP 的视频传输及播放，运营商可以提供多种多样的业务和服务，而用户更加关注自身的体验质量，用户体验质量成为用户满意度的关键。基于此，本文将深入探究用户体验评价变量（初始缓冲时延，卡顿时长占比）与网络侧变量（初始缓冲峰值速率，播放阶段平均下载速率，E2E RTT)之间的关系。

在建立模型之前，我们对数据做了如下准备的工作：一、对原始数据进行处理，画出初始缓冲时延与初始缓冲峰值速率，播放阶段平均下载速率，E2E RTT 的散点图，根据图像的走势采用函数进行拟合。二、分别对各个自变量两两之间进行相关系数的拟合，观察变量之间的交互性。三、基于采集到的用户体验质量数据，使用多元线性回归方法来建立了用户体验质量评价模型。

通过对以上模型的分析，我们对数据采取多次拟合来就，建立了体验评价变量与网络侧变量的函数模型，根据统计量 R^2 , F , p , s^2d 的结果，对其进行残差分析，检验模型的可用性，不断完善模型。

关键词： 统计回归分析 初始缓冲时延 残差分析 F 统计量

二、问题分析

本问题是关于建立用户体验评价变量（初始缓冲时延，卡顿 时长占比）与网络侧变量（初始缓冲峰值速率，播放阶段平均下载速率，E2E RTT)之间的函数关系的问题。

我们首先分析影响初始缓冲时延的因素，建立起这些相互影响因素的网络结构图。

其次我们利用题目所给数据，建立始缓冲时延与初始缓冲峰值速率，播放阶段平均下载速率，E2E RTT 之间的函数关系。我们采取统计回归来多次拟合的方法，分析初始缓冲峰值速率，播放阶段平均下载速率，E2E RTT 这些元素对用户体验评价变量的影响，分别建立相关的函数模型。最后利用建立好的模型来估测初始缓冲时延与卡顿占比之间的关系，进行残差分析，不断完善模型的正确性。

三、模型假设

经过小组讨论，我们认为影响用户视频观看体验的影响变量较多，所以我们首先保证变量的单一性影响原则，先研究单个变量对用户视频观看体验的影响，进而探究影响因素之间的关系，并以此为依据，作出如下假设：

① 初始缓冲峰值速率低，会导致下载速率受限，影响业务交互时延增大，初始缓冲时延大，为了简化模型，我们假设初始缓冲峰值对 E2E RTT 没有干扰作用，它们之间相互独立。

② 假设初始缓冲峰值、E2E RTT 以及播放阶段平均下载速率之间相互独立，对初始缓冲 时延的影响没有交互性，分别建立各个变量与初始缓冲时延的关系。

③ 根据各个变量与因变量的散点图走向，建立单变量与初始的函数关系。

卡顿占比跟卡顿时长成正相关，与视频播放时长成负相关。由于样本数据中播放阶段时长取固定值 30000ms，所以主要讨论在视频播放阶段的卡顿时长。

④ 我们假设播放阶段缓冲区没有数据时出现卡顿，卡顿时延在播放阶段出现。

⑤ E2E RTT 大，会导致视频业务解释交互时延增大，丢包以及用户感知，在第一问中，我们建立了初始缓冲时延与 E2E RTT 的关系，由此分析卡顿占比与初始缓冲时延的关系。

四、符号说明

表 1.符号说明

符号	含义	单位
y	初始缓冲时延	ms
x_1	初始缓冲峰值速率	Kbp/s
x_2	E2E RTT	ms
x_3	播放阶段平均下载速率	Kbp/s
H	播放阶段总时长	ms
k	视频码率	kbps
Z	卡顿占比	1

五、模型的建立

（一）模型建立分析

经过研究，我们得到以下三个结论：

- ① 在视频播放之前进行缓冲，能够降低播放中断概率。
- ② 视频播放中断的原因是视频分组不能及时有效的到达接收端进行解码，导致接收端播放队列为空。
- ③ 初始缓冲时延能够减少播放中断，但是过大的初始缓冲时延也会影响用户的观看体验。

因此,初始缓冲时延是衡量视频传输服务质量的重要指标，初始缓冲时延能够减少播放中断，但是过大的初始缓冲时延也会影响用户的观看体验。由此，我们得出结论初始缓冲峰值速率的大小，直接反映初始缓冲时延。

（二）建立过程

在问题分析中指出，我们首先需要获得各个影响变量与初始缓冲时延和卡顿时长占比的函数关系式，进而分析各因素之间的关系，最后建立函数模型，由此我们建立过程如下：

➤ 实验数据：

表 1 初始缓冲时延与初始缓冲峰值速率，E2E RTT 等的的数据

初始缓冲时延	初始缓冲峰值速率	播放阶段平均下载速率	E2E RTT	VMOS	初始缓冲下载数据量(byte)	视频质量得分 SQuality	初始缓冲得分 SLoading
7898	5503	68	1115	1	1564336	4.34	1.32
1724	15000	65	5708	3.77	1545277	4.33	3.63
1134	25000	47	5682	3.95	1543192	4.33	4.02
995	35000	38	5905	3.99	1571563	4.33	4.12
852	45000	38	6174	4.04	1602428	4.33	4.23
834	55000	38	5473	4.04	1550992	4.33	4.24
658	65000	36	5124	4.1	1600392	4.37	4.37
666	75000	35	6921	4.09	1633410	4.33	4.07

➤ 分析单一变量的影响：

1、分析 y 与 x1 的关系，首先利用 matlab 作出 y 对 x1 的散点图。

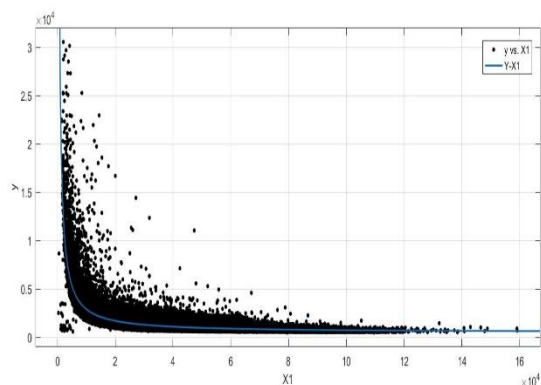


图1 y对x1的散点图

由此图，我们可以分析得出从图 1 可以发现，随着 x1 的增加，y 值有明显的指数递减趋势，拟合得出 y 与 x1 是反比例函数关系（其中 ε 是随机误差）。函数关系式如下：

$$y = \frac{a_1}{x_1} + b_1 + \xi \quad (1)$$

2、分析 y 与 x2 的关系，首先利用 matlab 作出 y 对 x2 的散点图。

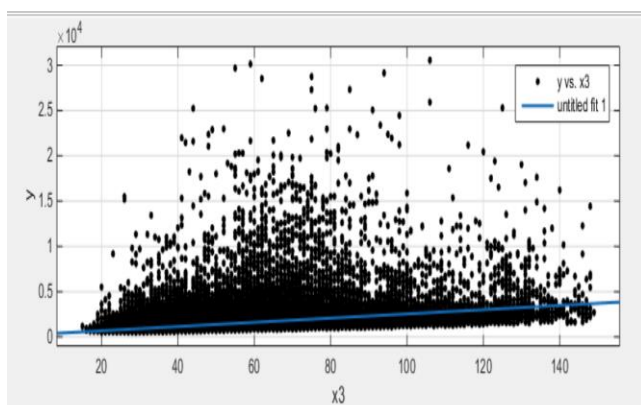


图2 y对x2的散点图

由实验数据拟合，y 与 x2 为正态分布函数。函数关系式如下：

$$y = a_2 \times e^{-\frac{(x_2 - b_2)^2}{c_1}} + \xi \quad (2)$$

3、分析 y 与 x_3 的关系，首先利用 matlab 作出 y 对 x_3 的散点图。

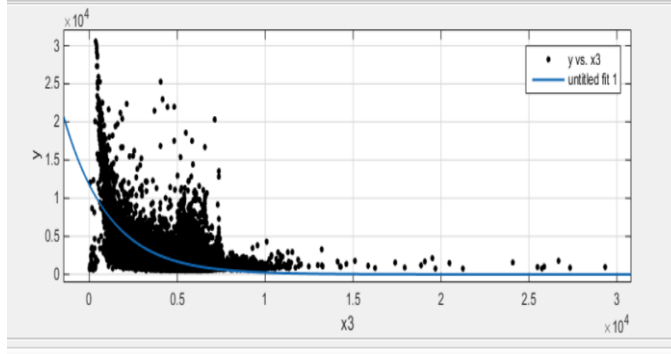


图3 y 对 x_3 的散点图

由实验数据拟合得出 y 与 x_3 为指数函数关系，函数关系式如下：

$$y = a_3 \times e^{b_3 \times x_3} + \xi \quad (3)$$

➤ 分析多变量的影响：

1、在模型（1）、（2）、（3）中，回归变量 x_1 、 x_2 、 x_3 对因变量 y 的影响都是互相独立的，即初始缓冲时延 y 的均值与初始缓冲峰值速率 x_1 的非线性关系由回归系数 a_1, b_1 确定，而不依赖于 E2E RTT x_2 ，播放阶段平均下载速率 x_3 。

$$y = \frac{a_1}{x_1} + b_1 + a_2 \times e^{-\frac{(x_2 - b_2)^2}{c_1}} + a_3 \times e^{b_3 \times x_3} + \xi \quad (4)$$

2、为了寻找因变量之间的关系，采用 excel 对它们进行相关系数的计算。为表明变量之间的关系，我们简单的用 x_1, x_2, x_3 的乘积代表它们之间的交互作用，于是将模型（4）增加三项。得到最终模型：

$$y = \frac{a_1}{x_1} + b_1 + a_2 \times e^{-\frac{(x_2 - b_2)^2}{c_1}} + a_3 \times e^{b_3 \times x_3} + a_1 \times x_1 \times x_2 + a_2 \times x_2 \times x_3 + a_3 \times x_1 \times x_3 + a_4 \times x_1 \times x_2 \times x_3 + \xi \quad (5)$$

（三）数据求解:

用 matlab 的统计工具对最终函数模型进行验证性分析, 得到的结果如下:

表2 统计数据

参数	参数估计	参数置信区间
a1	0.02628	-0.02662, -0.02594
b1	4.068	2456, 2487
a2	3043	2971, 3115
b2	4.068	3.887, 4.248
c1	4.589	4.449, 4.729
a3	1.174e+04	1.164e+04, 1.184e+04
b3	-0.0003881	-0.0003901, -0.0003862

表 2 显示, $R^2 = 0.924$ 指因变量 y (初始缓冲时延) 的 92.4% 可由模型确定, F 值远远超过 F 的检验临界值, p 远小于 α , 因而模型 (4) 从整体上来看是可用的。

问题二:

➤ 1、模型建立:

确定模型结构和相关性分析

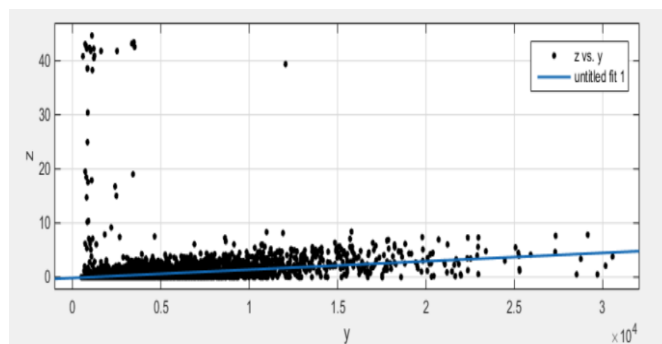


图4 z 对 y 的散点图

通过多次拟合, 发现 z 和 y 的相关性系数 $R^2=0.0736$, 说明卡顿占比与初始缓冲时延具有交互性, 其线性关系明显, 得到

$$F(z,y)=\alpha z+\beta y+\varepsilon \quad (6)$$

参数	参数估计值	参数置信区间
α	0.0001548	0.0001512, 0.0001583
β	-0.1699	-0.1766, -0.1632

➤ 结果分析

结果分析：表3显示， $R^2 = 0.951$ 指因变量 y （初始缓冲时延）的95.1%可由模型确定， F 值远远超过 F 的检验临界值， p 远小于 α ，因而模型（6）从整体上看是可行的。

六、 模型评价

该模型的优点是：
经过灵敏度分析，本文建立的数学模型在数据上的处理是快速的，计算得到的结果是可靠的，这说明本文建立的模型是正确的，是值得推广的。本文有效的网络视频质量评估体系可以解决主观评估苛刻的条件、人为因素影响、实施步骤复杂、代价昂贵、实时性不好等缺点，同时满足准确地反映视频的体质量的要求，而且评估质量更加精确可靠和客观，可以应用于实时视频通信中的质量评估。本文模型对评估体系进行分析优化，提高客户观看网络视频体验提供了完美的数据支持，成功解决网络视频业务的技术支持问题，具备快速便捷等多重优势，值得向社会推广。

七、参考文献

- [1] 赵伟丰. 基于RTT端到端的网络拥塞控制研究. 天津大学, 2013. 06
- [2] 陈楚雄, 柯江毅, 覃道满. 视频业务体验评估和优化提升探讨. 邮电设计技术. 2017. 02
- [3] 陈希宏, 金跃辉, 杨谈. 3 G网络中移动视频质量评估模型的研究, 计算机科学. 2015. 09
- [4] 李飞. 面向移动终端的视频用户体验质量评估模型研究, 西南交大. 2016. 5. 4
- [5] 杨付正, 万帅. 网络视频质量评估技术研究现状及发展动向, 通信学报. 2012. 04
- [6] Zhengyou Wang, Liying Li, Wan Wan, Zheng Wan, Cong Cai. A Study on QoS/QoE Correlation Model in Wireless-network, 978-616-361-823-8 © 2014 APSIPA.
- [7] 李晓龙. 无线环境下多媒体传输服务质量研究, 中国科学技术大学. 2014. 05
- [8] 付中南, 尚群, 公绪晓, 陈庆接, 周昌令. 基于用户体验质量的校园无线网络运行评价, 华中科技大学学报. 2016. 10
- [9] 华猛. 基于终端测量的用户体验质量评价的研究与设计, 北京邮电大学. 2014. 03
- [10] 刘继春. 面向用户体验的视频流业务质量评测方法, 北京邮电大学. 2013. 12
- [11] 焦阳. 无线网络中端到端视频流业务的用户体验质量预测及优化技术分析. 2015. 09
- [12] 林闯, 胡杰, 孔祥震. 用户体验质量 (QoE) 的模型与评价方法综述, 计算机学报. 2012. 01

八、模型编程代码

matlab 实例代码:

Y-X1:

```
function [fitresult, gof] = createFit(X1, y)
[xData, yData] = prepareCurveData( X1, y );
ft = fitype( 'power2' );
opts = fitoptions( 'Method', 'NonlinearLeastSquares' );
opts.Display = 'Off';
opts.Lower = [-Inf -1 -Inf];
opts.StartPoint = [6214.27813636366 -0.147293002006473
59.0073454326789];
opts.Upper = [Inf -1 Inf];
[fitresult, gof] = fit( xData, yData, ft, opts );
figure( 'Name', 'Y-X1' );
h = plot( fitresult, xData, yData );
legend( h, 'y vs. X1', 'Y-X1', 'Location', 'NorthEast' );
xlabel X1
ylabel y
grid on
```

Y-X2:

```
function [fitresult, gof] = createFit(X2, y)
[xData, yData] = prepareCurveData( X2, y );
ft = fitype( 'gauss1' );
opts = fitoptions( 'Method', 'NonlinearLeastSquares' );
opts.Display = 'Off';
opts.Lower = [-Inf -Inf 0];
opts.StartPoint = [30552 106 15.4163423583732];
[fitresult, gof] = fit( xData, yData, ft, opts );
figure( 'Name', 'Y-X1' );
h = plot( fitresult, xData, yData );
legend( h, 'y vs. X2', 'Y-X1', 'Location', 'NorthEast' );
xlabel X2
ylabel y
grid on
```

Y-X3:

```
function [fitresult, gof] = createFit(X3, y)
[xData, yData] = prepareCurveData( X3, y );
ft = fitype( 'exp1' );
opts = fitoptions( 'Method', 'NonlinearLeastSquares' );
opts.Display = 'Off';
opts.StartPoint = [3414.89573662046 -4.98287982878925e-05];
[fitresult, gof] = fit( xData, yData, ft, opts );
```

```
figure( 'Name', 'Y-X1' );  
h = plot( fitresult, xData, yData );  
legend( h, 'y vs. X3', 'Y-X1', 'Location', 'NorthEast' );  
xlabel X3  
ylabel y  
grid on
```