

文章编号:1007-130X(2024)09-0001-07

基于 QoS-QoE 预测的传输瓶颈定位*

马心宇^{1,2}, 李彤^{1,2}, 曹景堃^{1,2}, 吴波³, 孙永谦⁴, 赵乙⁵

(1. 中国人民大学数据工程与知识工程教育部重点实验室, 北京 100872;

2. 中国人民大学信息学院, 北京 100872; 3. 腾讯科技有限公司, 北京 100080;

4. 南开大学软件学院, 天津 300350; 5. 北京理工大学网络空间安全学院, 北京 100081)

摘要:在实时音视频传输中, QoS 指标反映服务端可感知的网络情况, QoE 指标直接体现用户对视频业务的满意程度, 尽管 QoE 指标是服务提供商更为关注的指标, 但是由于接口适配和用户隐私保护等问题, 云服务提供商往往不能实时获得 QoE 数据, 因此无法及时对可能发生的 QoE 异常进行预测并采取优化。由于 QoS-QoE 存在一定映射关系, 提出一种基于服务端 QoS 指标实现对 QoE 指标进行瓶颈检测的模型, 可以减少运维人员定位的工作量, 提高网络优化效率。模型使用不平衡决策树进行 QoS-QoE 预测, 实现 QoE 异常检测。使用 LSTM 回归模型进行因果分析, 实现瓶颈定位。实验表明该模型对 QoE 异常检测有较高准确率, 并且可以发掘传输过程中对传输结果影响较大的 QoS 指标。

关键词: QoS-QoE; 异常检测; 因果分析; 实时通信

中图分类号: TP391.41

文献标志码: A

doi: 10.3969/j.issn.1007-130X.2024.09.001

Transmission bottleneck localization based on QoS-QoE prediction

MA Xin-yu^{1,2}, LI Tong^{1,2}, CAO Jing-kun^{1,2}, WU Bo³, SUN Yong-qian⁴, ZHAO Yi⁵

(1. Key Laboratory of Data Engineering and Knowledge Engineering, Renmin University of China, Beijing, 100872;

2. School of Information, Renmin University of China, Beijing 100872;

3. Tencent Technology Company Limited, Beijing 100080;

4. College of Software, Nankai University, Tianjin 300350;

5. School of Cyberspace Science and Technology, Beijing Institute of Technology, Beijing 100081, China)

Abstract: In real-time audio and video transmission, QoS (Quality of Service) metrics reflect the perceived network conditions at the server side, while QoE (Quality of Experience) metrics directly embody the satisfaction level of users with video services. Although QoE metrics are of greater concern to service providers, cloud service providers often cannot obtain QoE data in real-time due to issues such as interface adaptation and user privacy protection, making it difficult to predict and optimize potential QoE anomalies in a timely manner. Given the existing mapping relationship between QoS and QoE, this paper proposed a model that utilizes server-side QoS metrics to detect bottlenecks in QoE metrics, aiming to reduce the workload of operation and maintenance personnel and improve network optimization efficiency. The model employs an imbalanced decision tree for QoS-QoE prediction to achieve QoE anomaly detection. Furthermore, an LSTM regression model is utilized for causal analysis to locate bottlenecks. Experiments show that this model achieves high accuracy in QoE anomaly detection and can identify QoS metrics that significantly impact transmission outcomes.

* 收稿日期: 2023-12-29; 修回日期: 2024-02-20

基金项目: 国家自然科学基金(62202473, 62302244); 中国人民大学建设世界一流大学(学科)基金、腾讯基础平台技术犀牛鸟专项研究计划

通信作者: 李彤(tong.li@ruc.edu.cn)

通信地址: 100872 北京市海淀区中关村大街 59 号中国人民大学信息楼

Address: Information Building, Renmin University of China, 59 Zhongguancun Street, Haidian District, Beijing 100872, P. R. China

Key words: QoS-QoE; anomaly detection; causal analysis; real-time communication

1 引言

内容分发网络 CDN(Content Delivery Network)是一种全网内容加速服务,旨在解决因分布、带宽和服务器性能而导致的访问延迟问题。通过 CDN,用户可以就近获取所需内容,从而提高用户访问网站的响应速度和成功率。实际上,CDN 业务是由 CDN 供应商、业务提供商和终端用户(即网民)共同完成的。业务提供商从 CDN 供应商那里租用带宽和服务等资源,根据网站实际用户占用的 CDN 带宽量以及选用的增值服务类型向 CDN 供应商支付相关费用。终端用户通过视频点播、直播、下载、游戏等服务间接使用 CDN 服务,成为 CDN 服务的最终体验者。通过 CDN,用户可以享受到更快的视频加载速度、更稳定的直播体验以及更快速的下载和游戏体验。

在多 CDN 供应商的场景^[1]中,业务提供商租用多个 CDN 供应商的云服务,将企业网站的业务流量分配给这些供应商。每个周期,业务提供商会根据自身的标准对供应商的业务表现进行评分,并得出所租用的 CDN 供应商的排名,如图 1 所示。在下一个周期,当业务提供商再次将业务流量分配给 CDN 供应商时,会根据排名进行分配。排名高、业务能力强的 CDN 供应商将获得更多的业务流量。通过这种方式,业务提供商可以在不同的 CDN 供应商之间进行动态调整,以确保网站的业务流量能够得到高效的分发和加速。这种多 CDN 供应商的场景可以提高业务的可靠性和性能,并为用户提供更好的访问体验。

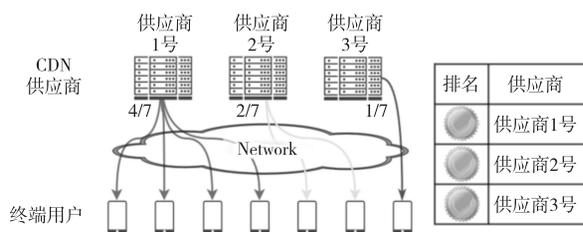


Figure 1 Multi-CDN provider scenario

图 1 多 CDN 供应商场景

对于视频业务提供商,周期性的评分标准就是 QoE(Quality of Experience)指标,即终端用户对移动网络提供的业务性能的主观感受。QoE 指标可以通过接近量化的方法来表示终端用户对业务与网络的体验和感受,并反映当前业务和网络的质量

与用户期望间的差距。例如,直播业务的 QoE 指标包含首帧时延、拉流成功率、百秒卡顿时长和视频渲染卡顿次数等。而对于 CDN 供应商,能够直接获取的是 QoS(Quality of Service)指标,即服务质量,包含网络本身的所有指标。QoS 数据由发送方产生,是服务器端能够第一时间获取的信息,通常包含往返时延 RTT(Round Trip Time)、丢包率、窗口大小和网络拥塞等指标。

本文将针对视频业务下的多 CDN 供应商场景,通过研究 QoS 与 QoE 的映射关系,实现对 QoE 指标的异常检测,并且分析造成该 QoE 异常的 QoS 瓶颈指标。本文可以为网络运维人员提供一个高效的瓶颈检测工具,减少运维人员定位的工作量,提高网络优化效率,让优化变得有的放矢。

2 相关工作

在当前的多 CDN 供应商场景下,云服务提供商需要关注的最重要的指标之一就是 QoE 指标。这个指标直接影响月度评分,进而影响了流量多少和最终收益。因此,如果能够获取实时的 QoE 指标,那么就可以准确了解传输策略会对视频业务形成怎样的效果。

然而,由于以下 2 点原因,云服务提供商无法实时获取 QoE 数据:

(1)接口适配问题。

由于视频业务方同时租用多个云服务供应商的云服务,不同云服务供应商的内部接口互不相同,导致业务方无法与全部云服务供应商配合,因此也无法设计一套全部适配的 QoE 接口提供给云服务供应商。

(2)用户隐私问题。

QoE 指标是反映客户端播放质量的指标,因此数据收集在客户端完成,而这些 QoE 指标中往往包含设备和当地网络信息,但是在一些国家和地区,用户隐私保护法律和法规非常严格,从而成为阻碍获取 QoE 数据的一个原因。

现有的传输优化方案结合 QoE 数据有较大的收益,比如 XLINK^[2]利用 QoE 反馈来控制数据包在多路径传输中的重新注入时机和路径,使下载时间、卡顿和首帧时延都有明显优化。然而这类结合 QoE 的传输优化方案都需要做双端的修改,这一点在多 CDN 供应商场景下是无法完成的。CDN

供应商只对传输策略(直接反映在 QoS 上)有控制权,对于更改传输策略后会对 QoE 有怎样的影响并不能确定,也无法解释更改某一传输参数后 QoE 是否会被优化。

随着基于语音和视频的网络应用服务的发展,人们开始不仅仅满足于知道网络 QoS 参数,更加关注网络服务的好坏,即 QoE 评价指标。在 QoS 与 QoE 的映射工作方面,倪萍等^[3]提出 IQoE2QoS(Improved QoE to QoS)算法,采用模糊理论的方法计算 QoE 到 QoS 的映射,并通过仿真表明,IQoE2QoS 算法对用户经验的分类准确度是线性回归算法的 2~3 倍;吴玉峰等^[4]提出一种基于 BP(Back-propagation)神经网络的模型,该模型使用抖动和丢包 2 个网络 QoS 参数作为输入层神经元,基于 VQM(Video Quality Metric)算法的 QoE 评价价值作为输出层神经元,完成 QoS 到 QoE 映射;姚会刚等^[5]提出一种分别使用变学习率梯度下降和弹性 BP 2 种算法的关系映射模型,能够利用网络层 QoS 指标数据在线实时评估网络视频 QoE;Orosz 等^[6]研究了主观质量评估与传输异常之间的相关性,测量了 QoS 参数和客观视频性能指标(视频质量指标)之间的相关性;Chen 等^[7]全面考察了视频质量评估方法的演变,并分析了其特点、优势和缺点。

本文拟解决的挑战是在服务端实现面向 QoE 的单端优化。通过在服务端利用 QoS 数据预测 QoE 指标,发现 QoE 异常时刻并及时调整传输策略,到达与双端修改的优化策略相同的效果。

3 模型设计

3.1 背景知识

实时音视频关注的重点之一是数据传输的及时性。当 2 个视频帧间的间隔大于 200 ms 时,用户就会肉眼可见地感受到卡顿。

实时音视频的 QoE 指标示例如表 1 所示,其中端到端时延是指 IP 数据包从离开源点时算起一直到抵达终点时的时间;首帧时延是用户从操作播放相关动作到首帧渲染出来的耗时,即用户从 App 上感知到的耗时;拉流成功率表示用户能够成功拉到流的比例;百秒卡顿时长和百秒卡顿次数统计的是最近 100 s 内 7 层播放器的缓冲区中没有数据的情况,直观地描述了用户观看过程是否流畅;视频渲染百秒卡顿则是统计相邻 2 个视频帧间播放间隔大于 200 ms 的情况。

CDN 云服务商可以直接获取的是与实时音视频数据发送有关的 QoS 数据,通常借助接收端传输层的确认信息^[8]得到,其中包含与网络丢包情况有关的 loss_events_times;与时延有关的 rtt、min_rtt、max_rtt;与发送数据量有关的 max_cwnd 和 in_pacing;与拥塞控制算法阶段相关的 last_slow_start_status、sent_bytes_during_slow_start。本文涉及到的部分 QoS 指标及其含义和示例如表 2 所示。

Table 2 Example of real-time audio/video QoS metrics

表 2 实时音视频 QoS 指标示例

QoS 指标	示例
时间戳	1 652 544 000
重传率	0.018 488 881
最小 RTT/ μ s	49 166.052 23
最大 RTT/ μ s	267 747.368 9
最大发送窗口/byte	340 696.303 6
发生丢包的次数	2.944 504 897
.....

3.2 总体设计

为了解决单端无法实时检测 QoE 指标变动情况以及无法利用 QoE 指标衡量优化传输算法的作用这 2 个问题,本文设计了包含 QoE 异常检测和 QoS 因果推断 2 个部分的传输控制流程。

Table 1 Example of real-time audio/video QoE metrics

表 1 实时音视频 QoE 指标示例

时间戳	端到端时延/ms	首帧时延/ms	拉流成功率/%	百秒卡顿次数/ms	百秒卡顿时长/ms	视频渲染百秒卡顿次数
1652544000	6 164	514	98.31	0.56	1139	11.97
1652544300	6 169	541	97.66	0.80	1216	12.55
1652544600	6 329	534	97.73	0.74	1238	11.56
1652544900	6 162	558	97.64	1.02	1576	12.82

传输控制的功能模块关系如图 2 所示,整体由 3 个模块构成:训练模块、异常检测模块、因果分析模块。具体的控制流程是:(1)CDN 供应商在本地利用历史 QoS 和 QoE 数据进行分类和回归模型的训练,并将训练好的模型集成在视频数据发送端;(2)在传输过程中,CDN 供应商实时收集网络的 QoS 数据,并输入分类模型中,得到对对应时刻的 QoE 是否异常的判断;(3)当判断结果为异常时,将此组 QoS 数据传入因果分析模块,在这个模块中使用回归模型分析最有可能引起 QoE 指标异常的 3 个根因 QoS 指标。

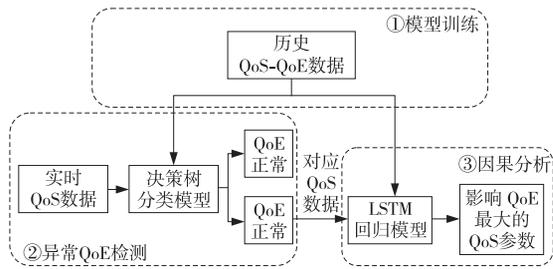


Figure 2 Overall design modules

图 2 总体设计模块图

在本节接下来的小节中,将具体描述 QoE 异常检测和因果分析 2 个关键模块的工作。

3.3 基于 QoS-QoE 预测的性能异常检测

3.3.1 性能异常的定义

在对异常 QoE 指标的定义环节,本文利用了正态分布 sigma 原则^[9]和假设检验中的“小概率事件”的基本思想,同时考虑到对于视频应用,QoE 指标通常是极大值或极小值为异常,而非极大值和极小值都是异常(比如,本文关注的百秒卡顿时长、渲染时长和首帧时延这 3 个指标均是过大为异常,而拉流成功率则是过小为异常),将异常 QoE 指标定义为超过(或低于) $\mu + 2\sigma$ (或 $\mu - 2\sigma$),其中 μ 表示该 QoE 指标的均值, σ 表示该 QoE 指标的方差。

当某一时刻的多个 QoE 指标有任意一个被认为是异常 QoE 指标时,该时刻就被定义为异常时刻。

QoE 指标异常与否是一个二分类问题,因此在下一小节中主要讨论选择哪种分类算法最适合在实时音视频传输端进行 QoE 指标预测。

3.3.2 分类算法选择

分类算法是机器学习中最常见的一种算法,它可以将数据集中的样本分成不同的类别。常见的分类算法包括 K 近邻算法、朴素贝叶斯、逻辑回归和决策树算法等。下面依次分析各个算法是否适

合应用于 QoS-QoE 映射。

K 近邻算法需要计算每个测试点与训练集的距离,当训练集较大时,计算量相当大,时间复杂度高,而且存在样本不平衡问题,对稀有类别的预测准确度低。由于在视频数据传输过程中,会持续产生大量的 QoS 数据,并且 QoE 异常在现实世界发生的概率远低于 QoE 正常的情况,因此 K 近邻算法不适合应用在实时视频传输的 QoS-QoE 预测上。

朴素贝叶斯算法的一个前提是各个特征相互独立,这与实际的 QoS 数据特征是不符的,比如 QoS 指标中的发送速率(Pacing Rate)是由平滑往返时延 SRTT(Smooth Round Trip Time)和拥塞窗口 CWND(Congestion Window)计算得到的。

逻辑回归分类算法的特点是模型简单和可解释性强,它通过拟合一系列参数来建立模型,其本质上是一个线性分类器,所以不能应对较为复杂的数据情况。

决策树分类算法包括特征选择、决策树生成和决策树剪枝 3 个步骤,在训练后可以生成一个树状模型,每个分支节点代表一个属性的判断,每个叶节点代表一个分类结果。决策树算法的优势有可以可视化分析、容易提取出规则,同时在进行测试时运行速度较快,不会对发送端造成过多资源占用。决策树算法在进行属性划分的时候,不同的判定准则会带来不同的属性选择倾向,因此需要依据实际情况选择判断准则。

结合上述各个算法的优点和缺点,本文选择决策树分类算法作为预测 QoE 指标是否发生异常的分类算法。

3.4 基于因果分析的性能瓶颈定位

检测到 QoE 指标异常后,本文希望更进一步,发掘是哪一个或哪几个 QoS 指标引起的异常,即进行因果分析。采取的方法是利用长短期记忆 LSTM(Long Short-term Memory)^[10] 回归模型,通过每次改变一个 QoS 指标到均值,输入回归模型后比较改变不同 QoS 指标对 QoE 指标的影响多少,认为改变前后影响最大的 QoS 指标即为影响性能的瓶颈指标。

3.4.1 基于 LSTM 的回归算法

LSTM 的核心思想是通过记忆单元来存储输入的时序信息,而记忆单元则使用“门”结构来筛选存储至记忆单元的信息。这种算法能够结合历史数据对当前数据进行预测,QoS-QoE 数据也是依据时间序产生的。在一段时间内,QoS 数据的变

化会影响接下来的 QoE 指标。因此,使用 LSTM 算法可以更好地捕获 QoS 数据的动态变化,并将其应用于 QoE 指标的预测和分析中。

LSTM 的训练环节也是利用历史 QoS-QoE 数据进行离线训练,得到一个由 QoS 到 QoE 的预测模型,训练后的 LSTM 模型会被集成到发送端。

只在因果分析模块使用 LSTM 模型而不在异常时刻检测时也使用 LSTM 模型的原因是,异常时刻分类环节需要所有 QoS 数据都进入模型进行预测,因此需要一个计算量小、轻量级的模型,LSTM 模型并不能较好地处理较大数量级序列的任务,因此选择决策树分类算法;因果分析只在分类预测结果为异常时才被调用,而异常是小概率事件,所以并不会对总体计算量产生过大负担;此外,回归模型比分类模型更能体现 QoS 变化对 QoE 指标的影响大小。

3.4.2 因果分析

因果分析的步骤如图 3 所示,原始 QoS 组合经过 LSTM 回归得到,依次替换每一个 QoS 指标为其均值,将新的 QoS 组合经 LSTM 回归得到新的 QoE 数值,比较其与原始 QoE 的差异值,认为得到差异值最大的 QoS 值为影响 QoE 指标的最大原因,即传输瓶颈所在。

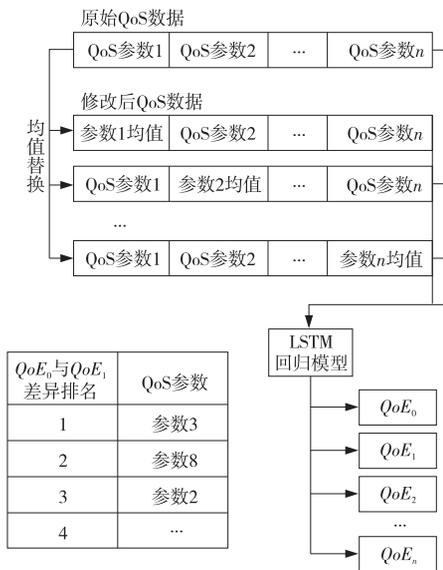


Figure 3 Schematic of causal analysis

图 3 因果分析示意图

4 实验评估

本节实验数据来源于腾讯直播业务三周的 QoS 和 QoE 数据,其中 QoS 数据每分钟收集一次,QoE 数据每 5 分钟收集一次。在数据预处理

环节,删除 QoS-QoE 时间戳数据列,并将 QoS 数据以 5 分钟为单位各指标取平均值,得到 QoS-QoE 一一对应的数据条目。

4.1 异常 QoE 检测

4.1.1 训练集和测试集划分

经过数据预处理后的 QoS-QoE 条目共有 6 042 条,由于数据集中样本数量不平衡,QoE 正常时刻占总体数据比例约为 87%,因此将样本按照类别以 7 : 3 的比例随机划分为训练集和测试集。

4.1.2 不平衡决策分类树

对于分类模型而言,样本量越大的标签可以学习的信息越多,算法就会更加依赖于从多数类中学到的信息来进行判断。并且,在样本中多数类占比高达 87% 的情况下,即使分类模型将测试集所有数据都分为多数类,模型准确率也有 87%,然而这对于异常 QoE 检测是毫无帮助的。

所以,本文通过改变决策树中调节样本均衡的参数 `class_weight` 和 `sample_weight`,来让算法意识到样本是不平衡的。在剪枝环节,也要改变 `min_weight_fraction_leaf` 这个基于权重的剪枝参数。

4.1.3 结果评估

本节分别训练了普通决策树和不平衡决策树^[11],并在测试集进行测试,最后计算 QoE 异常的 *Recall* 值(召回率),召回率计算公式如式(1)所示:

$$Recall_{anomaly} = \frac{TN}{TN + FP} \quad (1)$$

其中,*TN* 表示样本为 QoE 异常,预测结果为 QoE 异常;*FP* 表示样本为 QoE 异常,预测结果为 QoE 正常,召回率的大小表示了模型对异常情况的识别能力,召回率越大,说明异常情况被检测到的越多,模型的异常识别能力越强。

普通决策树和不平衡决策树的召回率如表 3 所示:

Table 3 Comparison of recall of decision trees and unbalanced decision trees

表 3 决策树和不平衡决策树召回率对比

学习算法	召回率
决策树	0.17
不平衡决策树	0.80

根据召回率结果,可以发现普通决策树几乎无法识别到 QoE 异常的分类,这是因为在训练集中 QoE 异常的分类占比很小,决策树无法学习到足

够的特征;而使用不平衡决策树,QoE异常的召回率可以达到0.80,即测试集中QoE异常情况有80%都能够被不平衡决策树预测到,满足QoE异常检测的要求。

4.2 传输瓶颈定位

4.2.1 训练集和测试集划分

LSTM回归模型以某一QoE指标为回归标准,并且因为LSTM模型是时序模型,所以划分训练集和测试集时,需要连续划分,本文将所有样本按顺序以7:3的比例划分为训练集和测试集。本节以视频渲染百秒卡顿时长为回归指标进行训练和测试。图4为以视频渲染百秒卡顿时长为目标的回归效果实例,从图4中可以看出预测值能够学习到QoE指标的变化趋势。

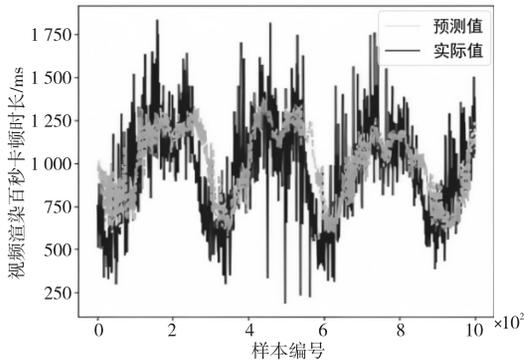


Figure 4 Video rendering 100 second lag time prediction vs. ground truth

图4 视频渲染百秒卡顿时长预测值与真实值对比

4.2.2 回归效果评估

为了对LSTM模型在QoE数据预测中的性能进行评价,本文采用平均绝对百分比误差MAPE(Mean Absolute Percentage Error)作为评价指标。平均绝对百分比误差的计算公式如式(2)所示:

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (2)$$

其中 \hat{y}_i 表示*i*时刻的预测值; y_i 表示*i*时刻的真实值; n 代表预测的总数量。

可以看出,MAPE是一个百分比值,主要反映的是误差占真实值的百分比,进而反映出预测的误差水平及可信程度,MAPE值越接近0,表示模型的拟合程度越好,MAPE值越接近100%,表示模型的拟合效果越差。本实验对各QoE指标分别进行LSTM模型拟合,并计算得到各QoE指标对应的MAPE值,如表4所示。

Table 4 MAPE values for each metric of LSTM models

QoE指标	MAPE值/%
端到端时延/ms	5.13
首帧时延/ms	9.74
拉流成功率	0.46
百秒卡顿次数	12.01
百秒卡顿时长/ms	18.73
视频渲染百秒卡顿次数	8.62
视频渲染百秒卡顿时长/ms	10.39

其中,拉流成功率模型拟合效果最好,MAPE值仅为0.46%。

4.2.3 因果分析示例

以时间戳为“1652916600”这条QoE指标异常数据为例,将其对应的QoS指标分别替换为整体均值,得到25条新QoS数据,经过回归模型预测后,得到QoE预测值和实际值差异排名(差异越大排名越前)如表5所示。

Table 5 Ranking of differences between QoE predictions and actual values
表5 QoE预测值和实际值差异排名

修改的QoS指标	差异值排名
retran_ratio	1
retran_bytes_ss	2
total_retran_pkts	3
srtt	4
max_rtt	5
app_limit_times	6
min_pacing_rate	7
.....

根据表5内容,本文认为导致时间“1652916600”QoE指标异常的QoS指标按影响大小前3个分别是“retran_ratio_avg”“retran_bytes_ss”和“total_retran_pkts”,都是和重传相关的QoS指标,可以推测此时网络状况较差,产生了较大的丢包,丢包率成为网络的瓶颈。

在对数据集中792条异常时刻全部进行因果分析后,发现出现频率最高的3条QoS指标分别是“retran_ratio”“retran_bytes_ss”和“max_rtt”,说明对直播业务来说,丢包和时延是影响用户体验的重要原因。

5 未来研究方向

针对分析出的引发QoE指标恶化的QoS参

数后,可以优化与其有关的现有传输策略。优化后对该 QoS 指标进行监控,如果 QoS 指标出现负向变化,则要及时回退传输策略,避免进一步引发 QoE 指标恶化。下面举例说明可以采取的优化方法:

(1)使用前向纠错减少丢包的影响。

前向纠错 FEC(Forward Error Correction)^[12]是一种通过增加冗余数据对丢失的数据包进行恢复的信道编码算法,如果在传输的全过程中开启 FEC,则会对带宽造成较大的浪费,但是不开启 FEC,则会在丢包率大的情况下无法快速恢复丢失数据,造成较大的传输延迟。理想的情况是在网络状况良好的时候关闭 FEC,在网络中丢包率高的情况下开启 FEC,而重传数据这一 QoS 指标可以反映实时的网络丢包情况,因此可以在重传数据这一 QoS 指标成为根因的时刻开启 FEC 算法,从而更好地应对丢包网络。

(2)优化 ARQ 以减少丢包的影响。

自动重传请求 ARQ (Automatic Repeat-Request)是在检测到丢包后重传丢失包的传输策略,然而当网络状态较差时,重传包也有很大可能丢失,导致较差的用户体验,现在有一些工作研究对 ARQ 的优化,比如 OR3^[13]通过增加冗余包数量应对重传包丢包,但是缺少自适应应对网络的功能;ART^[14]通过自适应副本数的方法解决重传丢包问题。未来可以将本文工作与优化 ARQ 相结合,在丢包率为瓶颈指标时增加重传冗余包数量,从而得到更好的 QoE 结果。

(3)切换拥塞控制算法以优化传输。

当瓶颈定位的结果是与拥塞控制算法中慢启动有关的指标时,可能是 Cubic 的某些特点影响了传输效果,比如对丢包敏感,此时可以切换拥塞控制算法^[15],使用 BBR(Bottleneck Bandwidth and RTT)或者 TACK(Tame ACK)^[16]等拥塞控制算法,提高传输速率。

在对传输策略进行优化后,需要监控根因 QoS 指标是否发生变化,以及预测的 QoE 指标是否依旧异常,如果根因 QoS 指标没有发生变化且预测的 QoE 指标依旧是异常,则需要尽快回退传输策略到改动之前。设计回退机制的目的是为了避免预测 QoE 是否异常和根因分析环节出现错误,造成网络状态恶化。

6 结束语

本文针对实时音视频业务的 QoS-QoE 指标

进行预测和因果分析,提出一种基于机器学习的瓶颈检测模型。模型使用不平衡决策树进行 QoS-QoE 预测,实现 QoE 异常检测;使用 LSTM 回归模型进行因果分析,实现瓶颈定位。实验结果表明,与普通决策树相比,不平衡样本的决策树可以在识别小比例样本时获得更好的效果;在因果分析和瓶颈定位环节,本文发现对直播业务来说,丢包和时延是影响用户体验的重要原因,下一步的工作将考虑针对分析得到的瓶颈 QoS 指标进行传输策略优化,从而提高用户体验。

参考文献:

- [1] Wu B, Li T, Luo C, et al. AutoPlex: Inter-session multiplexing congestion control for large-scale live video services[C]//Proc of the ACM SIGCOMM Workshop on Network-Application Integration, 2022: 1-6.
- [2] Zheng Z L, Ma Y F, Liu Y M, et al. XLINK: QoE-driven multi-path QUIC transport in large-scale video services[C]//Proc of ACM SIGCOMM Conference, 2021: 418-432.
- [3] 倪萍, 廖建新, 朱晓民, 等. 一种基于 QoS 的 QoE 到 SLA 映射方法[J]. 电子与信息学报, 2010, 32(6): 1463-1468.
Ni Ping, Liao Jian-xin, Zhu Xiao-min, et al. A method QoE mapped to SLA based on QoS[J]. Journal of Electronics & Information Technology, 2010, 32(6): 1463-1468.
- [4] 吴玉峰, 李建, 阮雅端, 等. 基于 BP 神经网络 QoS 到 QoE 映射模型[J]. 电子测量技术, 2016, 39(1): 84-87.
Wu Yu-feng, Li Jian, Ruan Ya-duan, et al. QoS-QoE mapping model based on back propagation neural network[J]. Electronic Measurement Technology, 2016, 36(1): 84-87.
- [5] 姚会刚, 黄永峰. 基于 BP 神经网络的网络视频 QoE 评估[J]. 计算机工程与设计, 2017, 38(1): 1-6.
Yao Hui-gang, Huang Yong-feng. BP-based estimate on network video QoE[J]. Computer Engineering and Design, 2017, 38(1): 1-6.
- [6] Orosz P, Skopkó T, Nagy Z, et al. A case study on correlating video QoS and QoE[C]//Proc of 2014 IEEE Network Operations and Management Symposium, 2014: 1-5.
- [7] Chen Y J, Wu K S, Zhang Q. From QoS to QoE: A tutorial on video quality assessment[J]. IEEE Communications Surveys & Tutorials, 2015, 17(2): 1126-1165.
- [8] 李彤, 郑凯, 徐格. 传输控制中的确认机制研究[J]. 软件学报, 2024, 35(4): 1993-2021.
Li Tong, Zheng Kai, Xu Ke. Research on acknowledgment mechanisms of transmission control[J]. Journal of Software, 2024, 35(4): 1993-2021.
- [9] Tariq A, Nourelfath M, Hassan J. Six sigma performance for non-normal processes[C]//[J]. European Journal of Operational Research, 2015, 247(3): 968-977.
- [10] Gers F A, Schmidhuber J, Cummins F. Learning to forget:

Continual prediction with LSTM[C]//Proc of 1999 9th International Conference on Artificial Neural Networks,2000: 850-855.

- [11] Cieslak D A, Chawla N V. Learning decision trees for unbalanced data[C]//Proc of European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases,2008:241-256.
- [12] Bolot J-C, Fosse-Parisis S, Towsley D. Adaptive FEC-based error control for Internet telephony[C]//Proc of IEEE INFOCOM'99. Conference on Computer Communications. Proceedings. 18th Annual Joint Conference of the IEEE Computer and Communications Societies. The Future is Now,1999:1453-1460.
- [13] Xie H, Li T. Revisiting loss recovery for high-speed transmission[C]//Proc of 2022 IEEE Wireless Communications and Networking Conference,2022:1987-1992.
- [14] Li T, Liu W, Ma X Y, et al. ART: Adaptive retransmission for wide-area loss recovery in the wild[C]//Proc of 2023 IEEE 31st International Conference on Network Protocols, 2023:1-11.
- [15] Zhou J, Qiu X Y, Li Z Y, et al. Antelope: A framework for dynamic selection of congestion control algorithms[C]//Proc of 2021 IEEE 29th International Conference on Network Protocols,2021:1-11.
- [16] Li T, Zheng K, Xu K, et al. TACK: Improving wireless transport performance by taming acknowledgments[C]//Proc of the Annual Conference of the ACM Special Interest Group on Data Communication on the Applications, Technologies, Architectures, and Protocols for Computer Communication,2020:15-30.

作者简介:



马心宇(2000-),女,河北石家庄人,硕士生,研究方向为计算机网络。**E-mail:** 2022104205@ruc.edu.cn

MA Xin-yu, born in 2000, MS candidate, her research interest includes computer network.



李彤(1989-),男,博士,副教授,CCF会员(50505M),研究方向为新一代互联网体系结构。**E-mail:** tong.li@ruc.edu.cn

LI Tong, born in 1989, PhD, associate professor, CCF member (50505M), his research interest includes the next-generation Internet architecture.



曹景堃(2003-),女,河北石家庄人,研究方向为计算机网络。**E-mail:** 2021201503@ruc.edu.cn

CAO Jing-kun, born in 2003, her research interest includes computer network.



吴波(1990-),男,博士,高级工程师,研究方向为计算机网络、网络传输优化和网络安全。**E-mail:** wub14@tsinghua.org.cn

WU Bo, born in 1990, PhD, senior engineer, his research interests include computer network, network transmission optimization, and network security.



孙永谦(1988-),男,博士,副教授,CCF高级会员(69960S),研究方向为服务异常检测、根因分析和故障诊断。**E-mail:** sunyongqian@nankai.edu.cn

SUN Yong-qian, born in 1988, PhD, associate professor, CCF senior member (69960S), his research interests include anomaly detection, root cause analysis, and failure diagnosis.



赵乙(1994-),男,博士,助理教授,CCF会员(50399M),研究方向为网络空间安全、人工智能安全和下一代互联网体系结构。**E-mail:** zhaoyi@bit.edu.cn

ZHAO Yi, born in 1994, PhD, assistant professor, CCF member (50399M), his research interests include cyberspace security, AI security, and next-generation internet architecture.