# Mitigating the Effect of Networks on Mobile Video Quality of Experience

**Tahir Nawaz Minhas**

Department of Technology and Aesthetics
Blekinge Institute of Technology

Doctoral Dissertation Series no. 2025:06

Blekinge Institute of Technology
Doctoral Dissertation Series No. 2025:06
ISSN 1653-2090
ISBN 978-91-7295-502-8

# Mitigating the Effect of Networks on Mobile Video Quality of Experience

**Tahir Nawaz Minhas**

DOCTORAL DISSERTATION
for the degree of Doctor of Philosophy at Blekinge Institute of Technology to be publicly defended on [2025-06-13] at [10:15] in [J1630 Hall, Karlskrona]

Supervisors
Markus Fiedler, Professor,
Hans-Jürgen Zepernick, Professor

Faculty Opponent
Dr. Kjell Brunnström, Research Institute of Sweden (RISE), Sweden
Prof. Narciso Garcia, Madrid Polytechnic University, Madrid, Spain
Prof. Helmut Hlavacs, University of Vienna Österrike, Austria
Associate Prof. Dr. Truong Thu Huong, Hanoi University of Sceience and Technology, Vietnam

# Abstract

The rapid growth in mobile video consumption, driven by advancements in mobile devices and network infrastructure, has raised user expectations for seamless video Quality of Experience (QoE) despite improvements in video streaming, network impairments like packet loss, delay, jitter, and outages. For instance, outages can cause visual artifacts like freezing, jumping, and missing frames, which negatively affect user perception. Understanding the relationship between network performance and QoE is crucial for improving user satisfaction.

This thesis investigates the impact of network performance on mobile video QoE and proposes strategies to mitigate these effects. The objectives include: (1) understanding TCP/IP's role in influencing QoE, (2) exploring the effects of Quality of Service (QoS) parameters such as delay, jitter, and packet loss on video quality, (3) analyzing the impact of network outages on QoE, and (4) developing a buffer-based solution to mitigate network disruptions.

The research employs theoretical modeling, controlled emulation experiments, and subjective assessments to evaluate QoE. The QoE Hourglass Model links network-layer parameters to user-perceived quality. Subjective tests, guided by ITU-T recommendations, use the Absolute Category Rating (ACR) method and Mean Opinion Scores (MOS) to assess video quality under various conditions. Additionally, the effectiveness of a sender buffer mechanism is tested through statistical analyses and user evaluations.

The findings reveal that network impairments, especially packet loss and delay variation, significantly degrade QoE. The QoE Hourglass Model provides a structured framework for understanding these effects. Experimental results show that higher frame rates and proactive buffering improve user perception. Perceptual Evaluation of Video Quality (PEVQ) and Temporal Quality Metric (TQM) measurements correlate with user ratings but are less accurate in predicting video freezes. The sender buffer mechanism effectively reduces freeze durations and enhances QoE during network outages.

This research emphasizes the impact of network impairments on video QoE and offers practical solutions, such as the sender buffer mechanism, to mitigate disruptions and enhance user satisfaction in video streaming.



**Keywords:** QoE, Quality of Experience, QoS, Quality of Service, Mobile Video, Live Video, Multimedia Streaming

Blekinge Institute of Technology
Doctoral Dissertation Series No. 2025:06

# Mitigating the Effect of Networks on Mobile Video Quality of Experience

**Tahir Nawaz Minhas**

Doctoral Dissertation in Telecommunication Systems

Department of Technology and Aesthetics
Blekinge Institute of Technology
SWEDEN

Blekinge Institute of Technology
Department of Technology and Aesthetics

Blekinge Institute of Technology Doctoral Dissertation Series No. 2025:06
ISBN 978-91-7295-502-8
ISSN 1653-2090
urn:nbn:se:bth-27779

Printed in Sweden by Media-Tryck, Lund University, Lund 2025

*Dedication*

*To my Parents and Family*

# Acknowledgements

All praise is due to Allah, the Most Gracious, the Most Merciful, whose guidance, blessings, and support have been the true source of all my achievements and success in life.

I am deeply indebted to my primary supervisor, Professor Markus Fiedler, whose unwavering support, invaluable encouragement, and expert guidance have been instrumental in shaping my understanding of research. His mentorship has not only enriched my academic growth but also inspired me to approach teaching with dedication and passion, a lesson I will carry with me throughout my career. I am equally grateful to my co-supervisor, Professor Hans-Jürgen Zepernick, for his constant supervision, insightful advice, and profound expertise, which have been pivotal in the progress of my research. His patience, support, and encouragement, especially during the extended period of my studies, have been a source of strength.

I would also like to extend my sincere thanks to my colleagues, especially Patrik Arlos, and fellow Ph.D. students at BTH, whose companionship, collaboration, and encouragement have enriched my academic experience.

I am profoundly grateful to my parents and other family members for their continuous prayers, unwavering moral support, and encouragement, which have been my pillar of strength throughout this journey. Lastly, I would like to express my heartfelt appreciation to all those who have prayed for me, offered their moral support, and encouraged me during this endeavor.

**Tahir Nawaz Minhas**

# List of Papers

The following is a list of papers included in this thesis.

## Paper I

T. N. Minhas and M. Fiedler, "Quality of experience hourglass model," In Proceedings of the *IEEE International Conference on Computing, Management and Telecommunications (ComManTel)*, Ho Chi Minh City, Vietnam, 2013, pp. 87-92, DOI: 10.1109/ComManTel.2013.6482371.

## Paper II

T. N. Minhas, O. G. Lagunas, P. Arlos and M. Fiedler, "Mobile video sensitivity to packet loss and packet delay variation in terms of QoE," In Proceedings of the *19th IEEE International Packet Video Workshop (PV)*, Munich-Garching, Germany, 2012, pp. 83-88, DOI: 10.1109/PV.2012.6229747.

## Paper III

T. N. Minhas and M. Fiedler. Impact of disturbance locations on video quality of experience. In Proceedings of the *EuroITV 2011 Workshop: Quality of Experience for Multimedia Content Sharing*, Lisbon, Portugal 2011.

## Paper IV

T. N. Minhas, M. Shahid, B. Lövström, A. Rossholm, H.-J. Zepernick, and M. Fiedler. "QoE rating performance evaluation of ITU-T recommended video quality metrics in the context of video freezes," In *Australian Journal of Electrical and Electronics Engineering*, pp. 122–131, 2016. DOI: 10.1080/1448837X.2015.1094855.

## Paper V

T. N. Minhas and M. Fiedler. "Mitigation of the Effects of Network Outage on Video QoE Using a Sender Buffer," In *MDPI Open Access Journal Electronics: Advances in Network*, 2021. DOI:10.3390/electronics10101209

List of research papers not included in the thesis.

## Paper VI

T. N. Minhas, O. Nawaz, M. Fiedler and S. Khatibi, “The effects of additional factors on subjective quality assessments,” In Proceedings of the *2nd IEEE International Conference on Advancements in Computational Sciences (ICACS)*, Lahore, Pakistan, 2019, pp. 1-5, DOI: 10.23919/ICACS.2019.8689138.

## Paper VII

O. Nawaz, T. N. Minhas and M. Fiedler, “QoE based comparison of H.264/AVC and WebM/VP8 in an error-prone wireless network,” In Proceedings of the *2017 IFIP/IEEE Symposium on Integrated Network and Service Management (IM)*, Lisbon, Portugal, 2017, pp. 1005-1010, DOI: 10.23919/INM.2017.7987426.

## Paper VIII

O. Nawaz, T. N. Minnas and M. Fiedler, “Optimal MTU for realtime video broadcast with packet loss-A QoE perspective,” In Proceedings of the *9th IEEE International Conference for Internet Technology and Secured Transactions (ICITST-2014)*, London, UK, 2014, pp. 396-401, DOI: 10.1109/ICITST.2014.7038844.

## Paper IX

T. N. Minhas, M. Shahid, A. Rossholm, B. Lövström, H.-J. Zepernick and M. Fiedler, “Assessment of the rating performance of ITU-T recommended video quality metrics in the context of video freezes,” In Proceedings of the *IEEE Australasian Telecommunication Networks and Applications Conference (ATNAC)*, Christchurch, New Zealand, 2013, pp. 207-212, DOI: 10.1109/ATNAC.2013.6705382.

## Paper X

T. N. Minhas, M. Fiedler, J. Shaikh and P. Arlos, “Evaluation of throughput performance of traffic shapers,” In Proceedings of the *7th IEEE International Wireless Communications and Mobile Computing Conference*, Istanbul, Turkey, 2011, pp. 1596-1600, DOI: 10.1109/IWCMC.2011.5982669.

## Paper XI

J. Shaikh, M. Fiedler, P. Arlos, T. Minhas and D. Collange, "Classification of TCP connection termination behaviors for mobile Web." In Proceedings of the *IEEE GLOBECOM Workshops (GC Wkshps)*, Houston, TX, USA, 2011, pp. 1111-1115, DOI: 10.1109/GLOCOMW.2011.6162351.

## Paper XII

T. N. Minhas, M. Fiedler, and P. Arlos. "Quantification of packet delay variation through the coefficient of throughput variation." In Proceedings of the *6th ACM International Wireless Communications and Mobile Computing Conference (IWCMC '10)*, Caen, Farance, 2010, pp. 336–340. DOI: 10.1145/1815396.1815474.

## Paper XIII

J. Shaikh, M. Fiedler, T. N. Minhas, P. Arlos, and Denis Collange. "Passive methods for the assessment of user-perceived quality of delivery." In Proceedings of the *7th Swedish National Computer Networking Workshop (SNCNW)*, Linköping, Sweden, 2011.

## Paper XIV

J. Shaikh, M. Fiedler, T. N. Minhas, P. Arlos, and Denis Collange. "Inferring user-perceived performance of network by monitoring TCP interruptions." In *Network Protocols and Algorithms*, 4(2): pp. 49-67, 2012.

## Paper XV

J. Shaikh, T. N. Minhas, P. Arlos and M. Fiedler, "Evaluation of delay performance of traffic shapers." In Proceedings of the *2nd IEEE International Workshop on Security and Communication Networks (IWSCN)*, Karlstad, Sweden, 2010, pp. 1-8, DOI: 10.1109/IWSCN.2010.5497994.

## Paper XVI

T. Zinner, T. Hossfeld, T. N. Minhas, and M. Fiedler. "Controlled vs. uncontrolled degradations of QoE: The provisioning-delivery hysteresis in case of video."

In Proceedings of the *EuroITV 2010 Workshop: Quality of Experience for Multimedia Content Sharing*, Tampere, Finland, 2010.

# Author's contribution to the papers

## Paper I

Conceptualization, T.N.M., and M.F.; Methodology, T.N.M., and M.F.; Results compilation, T.N.M.; Validation, M.F.; Writing-original draft, T.N.M.; Writing-review & editing, T.N.M., and M.F.; Supervision, M.F.

## Paper II

Conceptualization, T.N.M., and O.L.; Methodology, T.N.M., P.A., and O.L.; Results compilation, T.N.M.; Validation, M.F.; Writing-original draft, T.N.M.; Writing-review & editing, T.N.M., and M.F.; Supervision, M.F., and P.A.

## Paper III

Conceptualization, T.N.M., and M.F.; Methodology, T.N.M., and M.F.; Results compilation, T.N.M.; Validation, M.F.; Writing-original draft, T.N.M.; Writing-review & editing, T.N.M., and M.F.; Supervision, M.F.

## Paper VI

Conceptualization, T.N.M., and M.S.; Methodology, T.N.M., M.S., and A.R.; Results compilation, T.N.M., and M. S.; Validation, B.L., H.J.Z., and M.F.; Writing-original draft, T.N.M.; Writing-review & editing, T.N.M., and M.F.; Supervision, B.L., H.J.Z., and M.F.

## Paper V

Conceptualization, T.N.M., and M.F.; Methodology, T.N.M., and M.F.; Results compilation, T.N.M.; Validation, M.F.; Writing-original draft, T.N.M.; Writing-review & editing, T.N.M., and M.F.; Supervision, M.F.

# Abbreviations

| | |
|---|---|
| ACR | Absolute Category Rating |
| AVC | Advanced Video Coding |
| BP | H.264 Baseline Profile |
| CDN | Content Delivery Network |
| CIF | Common Intermediate Format |
| CTA | Consumer Technology Association |
| DPMI | Distributed Passive Measurement Infrastructure |
| fps | frames per second |
| HVS | Human Visual System |
| ITU | International Telecommunication Union |
| MOS | Mean Opinion Score |
| MSE | Mean Squared Error |
| NAK | Negative Acknowledgment |
| OWD | One-Way Delay |
| PDAs | Personal Digital Assistants |
| PDV | Packet Delay Variation |
| PEVQ | Perceptual Evaluation of Video Quality |
| PGC | Professionaly Generated Content |
| PSNR | Peak Signal-to-Noise Ratio |
| QCIF | Quarter Common Intermediate Format |
| QoD | Quality of Delivery |
| QoE | Quality of Experience |
| QoP | Quality of Presentation |
| QoS | Quality of Service |
| QVGA | Quarter Video Graphics Array |
| RERES | Reliability-based Real-time Streaming |
| SLA | Service Level Agreement |
| SNR | Signal-to-Noise Ratio |
| SSIM | Structural Similarity Index |
| TQM | Temporal Quality Metric |
| UGC | User-Generated Content |
| VGA | Video Graphics Array |
| VLC | VideoLAN Client |

| | |
|---|---|
| VoD | Video on Demand |
| VQA | Video Quality Assessment |
| VQM | Video Quality Metric |

# Table of Contents

# 1 Introduction

## 1.1 Overview

The surge in popularity of mobile Internet and streaming multimedia content has been remarkable in recent years. This shift is primarily driven by advancements in mobile device technology. These improvements have significantly enhanced capabilities such as processing power, memory capacity, display quality, audio performance, and various other functionalities. As a result, end-users are increasingly engaging not only in consuming online video content but also in creating their own video content for streaming purposes. This trend has contributed to a considerable increase in the overall consumption of online video. Live streaming, in particular, has become a favored method for sharing real-time experiences.

According to a study published in 2023 [1], video content constitutes an astounding 65.93% of total Internet traffic, compared to a mere 4.63% for web browsing activities; the same fact has also been reported in [2]. This statistic highlights the transformative impact video streaming has had on user interactions with digital media. Over the past decade, the rapid growth of video streaming platforms and applications has led to a significant increase in the share of Internet traffic attributed to these services.

Moreover, today's consumers are no longer passive recipients of media. They are now active participants in content creation. User-generated video content has emerged as one of the most prominent applications on the Internet, reflecting a fundamental shift in user engagement with media. Notably, research from the Consumer Technology Association (CTA) indicates that user-created content accounts for 39% of all media consumption, while professionally produced and sourced content represents the remaining 61% [3].

Additionally, historical data reveals that mobile video has demonstrated the highest growth rate in the digital landscape, with an impressive increase of 52.9% recorded in 2011. By 2015, mobile video had come to dominate mobile data consumption, accounting for approximately two-thirds of all mobile data traffic. In their 2018 annual Internet report, CISCO indicated that video streaming, particularly multimedia streaming, would make up more than two-thirds of global Internet traffic by 2019 [4]. Furthermore, in the white paper of the annual Internet report (2018-2023), CISCO predicted the growth of online video up to 82% [5].

Over the past five years, there has been a remarkable growth in mobile video consumption, with usage rates doubling annually. A significant portion of mobile data traffic is now attributed to streaming videos, indicating a shift toward more media-rich communication on handheld devices [6]. Among various devices, smartphones have emerged as the preferred platform for video consumption, showcasing their dominance in the digital media landscape [7].

As consumers engage more with video content, their expectations regarding quality have grown. This increased focus on quality, particularly their viewing experience, has made users more quality-conscious [8]. Research highlights an important aspect of user experience when individuals encounter poor service, they tend to share their negative experiences with an average of 13 potential users [9]. Such dissatisfaction can lead to user churn, a critical concern for service providers striving to retain their customer base. Minimizing churn is one of the foremost goals for these providers.

To address this challenge, recent studies have presented algorithms designed to reduce the likelihood of users discontinuing their services by taking into account Quality of Experience (QoE) and Quality of Service (QoS) factors [10]. Furthermore, the connection between user churn and QoE ratings has been explored in detail, shedding light on how perceived quality directly influences user retention [11].

Keeping user satisfaction as the highest priority is the key to winning user loyalty. In this regard, QoE is a primary concept that enables service providers, application developers, and other stakeholders to understand the experiences and perceptions of users. An exact assessment of video QoE is crucial for fulfilling the Service Level Agreement (SLA) commitments, which can help maintain or even enhance a positive user experience [12]. Recent advancements in QoE research highlight its significant role in driving user retention and establishing a competitive edge in the market [13].

In environments where user experience is central, the success of a service hinges on the QoE. Traditionally, QoE evaluation has been largely subjective, ideally relying on real-user testing to gather insights. However, this subjective approach can be resource-intensive, demanding significant investments in time, effort, and cost. Various authors have shifted their focus toward developing objective methods that can reliably and efficiently measure QoE, e.g., [14, 15]. These advancements suggest a more streamlined approach to understanding user satisfaction while alleviating some burdens associated with subjective evaluations.

QoS parameters such as packet loss, latency, delay variation, and bit rate are used to measure network performance. Despite offering valuable insight into the technical quality of the service, these metrics do not fully capture the end-user experience. In contrast, QoE directly reflects the perception of the service quality from the users' perspective. Estimating QoS parameters at the network and application layers is a common approach to monitor QoE [16]. The contribution of all TCP/IP layers to QoE is analyzed in [17], which highlights the holistic nature of network performance on perceived quality.

Multimedia applications have a significant and immediate impact on QoE as users are highly sensitive to even minor interruptions. It is therefore crucial to study QoE to determine the optimal QoS settings and other parameters that ensure excellent QoE for users. For instance, [18] proposes an exponential relationship between QoE and QoS, establishing a theoretical link between these two concepts. Similarly, [8, 19] investigate the influence of specific QoS parameters on video QoE, demonstrating how technical impairments can degrade user-perceived quality.

This thesis focuses on mobile video QoE, a critical area due to the increasing consumption of streaming multimedia on mobile devices. The QoE of streaming applications depends on a variety of indicators that influence user perception. Some of these indicators are directly tied to network performance, for example, poor network performance can cause visual artifacts such as freezing, jumping, jerking, or other disruptions, all of which adversely affect the perceived quality. Moreover, multiple simultaneous artifacts can amplify the degradation in QoE due to poor network conditions. One brief network outage can significantly affect user satisfaction.

Furthermore, the objective of this thesis is to explore how QoS parameters, such as packet loss, delay, and delay variation, impact QoE. Among additional objectives, the aim is to develop predictive models for mobile video QoE based on measurable QoS parameters. This includes identifying thresholds for network impairments that lead to significant degradation in QoE. We aim to answer questions such as: What levels of packet loss or latency variability result in a noticeable deterioration of QoE? How do specific network impairments impact the perceived quality of video content?

Similarly, another objective is to examine the effects of network outages on video QoE and to suggest mitigation strategies. The study will assess the severity of QoE degradation caused by outages and investigate techniques to minimize their impact. By understanding and addressing these issues, we aim to improve the overall user experience for mobile video streaming, ensuring that service providers can meet the growing demand for high-quality multimedia services. To study the mobile video QoE in relation to network performance parameters or QoS parameters, there are three possible options:

1. Real-time study;

2. Simulation;

3. Emulation.

An overview of the experimental setup is shown in Figure 1.1. In our research, we prioritize using real equipment whenever feasible to effectively capture the perception of a real system. However, replicating Internet behavior in a controlled real-time environment presents significant challenges and incurs high costs. To address this complexity, we opted for a traffic shaper from well-established options, such as NetEm, NISTnet, or KauNet, which are widely recognized and utilized by researchers in the field [20–23].

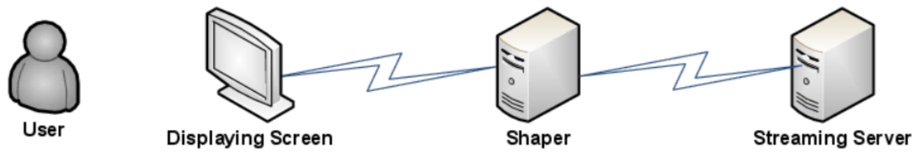



**Figure 1.1:** Overview of Experimental Setup to Study the Video QoE.

Our primary focus was to shape QoS parameters that influence user experience, including delay, packet delay variation, bitrate, and packet loss. Each of the traffic shapers we considered possesses the ability to modify these essential parameters; however, the accuracy of the timing output provided by each shaper can vary and may not always meet our specific requirements. If the performance of the selected shaper fails to align with the established criteria, the results generated could be deemed unreliable and unrealistic, undermining the integrity of our findings.

To ensure the validity of our QoE experiments, we undertook a thorough evaluation of the performance of the chosen traffic shapers. This process involved assessing their capabilities to guarantee their appropriateness for our study, thereby ensuring that the data collected would accurately reflect the intended QoS impacts on user experience.

## 1.2 Background

This section presents a brief description of the basic concepts that have been used in this thesis.

### 1.2.1 Mobile Video

Mobile video refers to the viewing of video content on mobile devices such as smartphones and tablets using wireless networks. Creating or capturing video content using mobile devices is also popular these days. Watching video on a mobile device can be real-time or Video on Demand (VoD) content from streaming platforms such as YouTube, Facebook, TikTok, etc. The convenience and portability of mobile devices have made mobile video a current way to consume. End-users can watch the media anywhere on a mobile device with access to a wireless network.

The history of mobile video, mobile networks, and devices is closely linked. In the late nineties, video recording was introduced on mobile devices. Later, 3G networks made it possible to send and receive the video content of mobile devices [24]. In 2007, a touchscreen phone with better Internet connectivity was introduced, which started the rise of smartphones and mobile video. At that time, mobile users were complaining about low video quality and poor viewing experience [25]. Later, 4G and now 5G and beyond have brought significant improvements in video streaming. The latest network developments have enabled high-definition and 4K video streams. Even augmented reality and virtual reality are on the way. All of this has made the use of mobile video very popular and trendy. Video streaming can be classified into two main groups: real-time video streaming and VoD.

#### *1.2.1.1 Online Video/Video on Demand*

Online video, or VoD, requires recording, storing it on media, and playing it when needed. There is freedom for the end-user to watch the video at one's convenience rather than follow a broadcast schedule. Figure 1.2 describes the steps involved in online video.

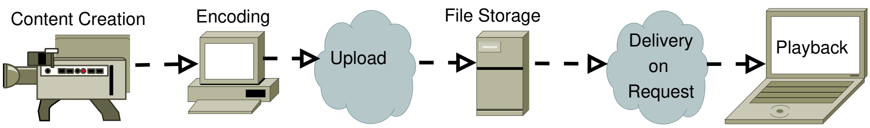



**Figure 1.2:** Online Video/Video on Demand.

#### *Content Creation*

Content creation is the first step and can be divided into two types, User-Generated Content (UGC) and Professionaly Generated Content (PGC). To begin with, a video-capturing device or a camera is used to capture or record video. Video content ranges from television programs and movies to educational videos. The content is typically recorded using cameras and other recording equipment and later edited to be of high quality.

### Encoding and Compression

After content creation, video files are encoded. During encoding, a raw video is converted into a digital format. The video is encoded by codecs such as H.264, H.265, or VP9 to allow for efficient broadcasting. Encoding is the conversion of raw video files into a digital format that can be streamed over the internet easily. Encoding also compresses video. The video file is encoded to reduce its size, but by keeping the similar quality, a smaller size is needed for streaming.

### File Storage and Distribution

Encoded videos are stored on the server, and the server sends the file to the Content Delivery Network (CDN). CDN is utilized to scatter the video file at different locations, which helps improve playback quality.

### Delivery on Request and Playback

When a user requests to watch a video, the VoD platform delivers the content using a CDN. The CDN ensures that the requested video is delivered from the server closest to the user. The video player receives the video stream from the nearest server and plays it back on the user's device.

### 1.2.1.2 Real-time video

Real-time video is delivered from the source to the viewer without any delay. The definition of real-time in the perceptual sense is given in [26] as *"The result of processing is effectively displayed 'immediately' (usually in the perceptual sense) once the input is available"*. How a real-time stream works is shown in Figure 1.3.

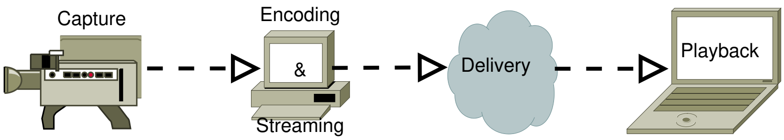



**Figure 1.3:** Real-time/Live Stream.

The following are the key characteristics of real-time streaming:

- **Immediacy:** In live streaming, the content is delivered immediately, with a sense of presence that it is happening now.
- **Data Continuity:** Data stream continuously to ensure a smooth and enjoyable viewing experience.
- **Low Latency:** The delay between the live event and its appearance on the viewer's screen is minimal.

### Content Creation

In live/real-time video streaming, content is created in real time. This can be anything from broadcasting an event to live gaming sessions, webinars, or even social media live streams. The content is captured using video cameras, mobile phone cameras, and other recording devices. Unlike VoD, live streaming does not involve editing before broadcasting, so whatever is captured is broadcast as it is.

### Encoding and Streaming

Capture devices typically capture video in raw format. A video encoder converts the raw video into various codecs and resolutions. The video encoding process also compresses the video to reduce its size while maintaining its quality.

Once the encoding is complete, the streaming server distributes the video content to playback devices over the Internet using CDN.

### Playback

The video player on the playback device decodes the stream and plays it. In contrast to VoD, there is no option to re-request the lost data in real-time streaming. It may result in low-quality video, which may be impaired with different artifacts.

## 1.2.2 Streaming Video Artifacts

There are also several video artifacts, and artifacts can potentially lower the video viewing quality, which impacts the QoE. Video streaming artifacts are defects or undesired features produced in the course of video content streaming. Artifacts may be due to many reasons, such as low bandwidth, poor network conditions, low-bitrate video streams, or high-bitrate video streams. Table 1.1 summarizes the artifacts, their descriptions, and potential reasons [27–30].

## 1.2.3 Video Quality Assessment

Video Quality Assessment (VQA) assesses the overall quality of a video throughout its journey from source to viewer. This assessment includes compression, transmission, and processing that affect the video's appearance and user experience. Video quality assesses the degradation or distortion of the video due to streaming, compression, or other factors involved in video transmission. Video quality is typically measured relative to the original video. To reduce the file size of the video, lossy compression is used, and during the quantization process, video quality may degrade. Similarly, impairments such as packet loss or packet delay may occur during video transmission over a wireless or mobile network, which can degrade video quality. Therefore, video quality assessment is important for tuning and optimizing streaming networks, compression, and processing tools, which helps improve the video experi-

**Table 1.1:** Real-Time Streaming Video Artifacts and Their Causes.

| Artifact | Description | Causes |
|---|---|---|
| Blocking/Pixelation | Video appears divided into small blocks or squares, losing detail. | Low bitrate, aggressive compression, or insufficient bandwidth. |
| Blurring | Loss of fine details, making the image appear soft or out of focus. | Over-compression, low resolution, or improper upscaling. |
| Color Banding | Visible as abrupt changes in color gradients instead of smooth transitions. | Insufficient color depth, low bitrate, or poor compression settings. |
| Mosquito Noise | Flickering or shimmering around edges, especially in high-motion scenes. | High compression, complex scenes, or inefficient encoding. |
| Ringing/Edge Artifacts | Halos or ghosting around sharp edges. | Aggressive compression algorithms or improper sharpening during encoding. |
| Stuttering/Frame Drops | Jerky or uneven playback due to missing or delayed frames. | Network congestion, insufficient buffering, or hardware limitations. |
| Buffering | Playback pauses to load more data. | Low bandwidth, network congestion, or insufficient buffering. |
| Freezing | Video temporarily stops playing. | Severe network congestion, packet loss, or hardware/software issues. |
| Missing of Video | Loss of video content. | Network outage, massive packet loss |
| Jumping | A sudden shift in the video's timeline. Usually occurs with missing. | Network outage, massive packet loss |
| Ghosting | A faint duplicate image slightly offset from the main image. | Packet loss, transmission errors, or improper deinterlacing. |
| Aliasing | Jagged or stair-stepped edges in the video. | Low resolution, improper scaling, or lack of anti-aliasing during encoding. |
| Combing Artifacts | Horizontal lines or distortions in moving scenes. | Interlaced video not properly deinterlaced before streaming. |
| Posterization | Loss of color detail making the image appear flat or cartoonish. | Excessive compression or insufficient color depth. |
| Noise/Grain | Random visual distortions or grain-like patterns in the video. | Low light conditions, high compression, or poor camera quality. |
| Tiling | Large blocks of the video appear distorted or misplaced. | Severe packet loss or corruption during transmission. |

ence for the end user. The end user, who experiences the video quality, is human, so the assessment of the video quality should be aligned with the user's perception. This evaluation method is called the subjective video quality assessment. Another evaluation method of video quality based on mathematical algorithms is objective video quality assessment. Both objective and subjective quality assessments are described in the following sections.

### 1.2.4 Objective Video Quality Assessment

Instead of humans, assessment algorithms or mathematical models are used to assess the video quality. The purpose of this assessment is to forecast the perceived quality of video without having the user tests. According to [31], the basic approach is to apply the still image quality on video frames, but a more promising approach is to consider the temporal aspects of the Human Visual System (HVS). Commonly used metrics are Signal-to-Noise Ratio (SNR), Peak Signal-to-Noise Ratio (PSNR), and Mean Squared Error (MSE) for objective video quality assessment [31, 32]. These methods are simple to use and compute. However, these methods are unable to consider the viewing conditions and characteristics of human visual perceptions [33].

Therefore SNR, PSNR, and MSE results are not mapped accurately to user perceived quality but can be used to evaluate the quality difference among videos [34]. There are some other tools used to measure the perceived video quality objectively. The Video Quality Metric (VQM) is a tool based on a general purpose video quality model used to measure perceptual quality objectively [35]. It has been adopted by ANSI (ANSI T1.801.03-2003) and ITU-T (ITU-R BT. 1683 and ITU-T J.144). The Structural Similarity Index (SSIM) index is an objective method for the assessment of video quality based on the degradation of structural information by considering the assumption that human visual perception is highly adopted for extracting structural information from a scene [36]. Another tool recommended by the ITU-T (ITU-T Rec-J.247) for objective perceptual multimedia video quality assessment is Perceptual Evaluation of Video Quality (PEVQ) in the presence of a full reference. It measures degradations due to a network by analyzing the degraded video pixel-by-pixel after temporal alignment of corresponding frames of reference and test sequence. This model measures QoE based on modeling the behavior of the human visual system [37]. Moreover, the objective assessment is based on the reference video and can be classified into full-reference (FR) method, reduce-reference (RR) method, and no-reference (NR) method [38, 39]. These methods are discussed below and are shown in Fig. 1.4.

#### *1.2.4.1 Full-Reference*

A full-reference method is used when the original reference sequence is available for comparison. Every pixel of reference- and test-sequence is compared to estimate the quality of test sequence. Metrics such as PSNR and SSIM are included in this classification. The accuracy of this method is very high, but it is suitable only in a non-real-time scenario, typically a lab environment.

#### *1.2.4.2 Reduced-Reference*

Instead of using the original reference, in this method, reduced information of the reference sequence is used for the estimation of test video quality. Reduced information is extracted at the sender side from the reference sequence and sent over the reliable channel to the receiver for comparison. The accuracy of results depends on the reduced information of reference and test sequence.

#### *1.2.4.3 No-Reference*

The no-reference method uses only the test sequence without any information about the original sequence. It is also known as the reference-free method or blind video quality assessment, and its accuracy can be very low because it is very difficult to assess the quality of a video without a reference. Reference-free techniques use machine learning algorithms or heuristic approaches to identify and measure artifacts such as blurring, blocking, or noise. These methods are especially useful for real-time

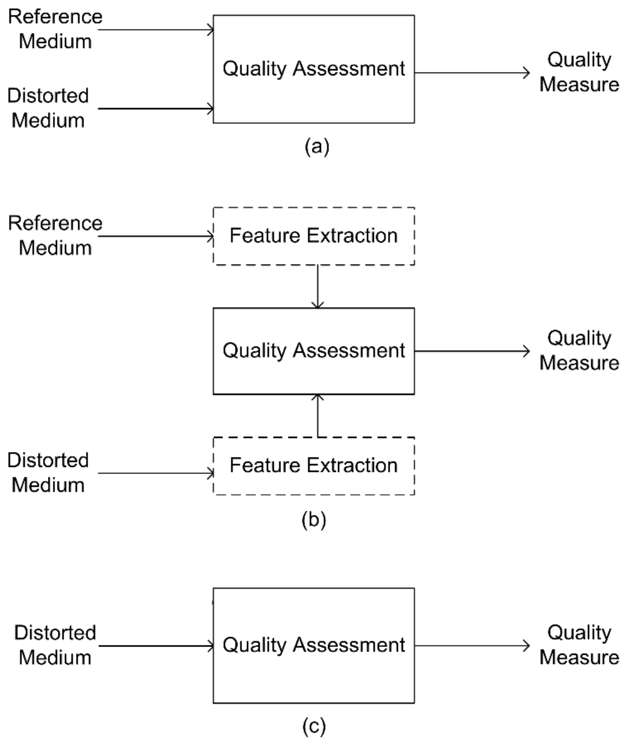



**Figure 1.4:** Video quality assessment methods: (a) Full-reference method, (b) Reduced-reference method, (c) No-reference method [39].

streaming scenarios where reference videos are not available. Furthermore, for real-time quality assessment by the receiver without any information about the original sequence, this method is one of the choices.

### 1.2.5 Subjective Video Quality Assessment

Subjective video quality assessment (SVQA) seeks to measure the quality of video as perceived by end-users. Unlike objective metrics that depend on mathematical computations, SVQA focuses on the direct evaluation of human perception. This assessment is concerned with the perception of a human and can be different from user to user [40]. When it comes to the human perceived quality, then various other factors play their role in user rating, for example, human psychology, the user interface of a device or application and its usability, viewing condition, and the content of video [17, 41, 42]. It acts as the benchmark for evaluating video quality, directing the advancement, and enhancement of video encoding, transmission, and display technologies. In ITU-R Rec. BT.500 and ITU-R Rec. P.910, the ITU gave detailed video quality assessment recommendations for television and multimedia applications, respectively. In [40, 43] different subjective measurement methods are discussed, two commonly used and recommended methods for multimedia applications are as follows:

#### 1.2.5.1 Single Stimulus or Absolute Category Rating (ACR)

In this category, test sequences are presented and rated separately. After each presentation, the subjects are asked to evaluate the quality of the video. This is a very efficient method, easy and fast to implement, and tests can be done in a shorter time as compared to other methods. It is well-suited for qualification tests and applied in several studies [43, 44]. To assess the video quality, the end-user/evaluator uses a predefined rating scale, which may consist of a 5-point or 9-point scale. The rating scales include a range of classifications, such as ”excellent,” ”good,” ”fair,” ”poor,” and ”bad.” In this work, we use the Absolute Category Rating (ACR).

#### 1.2.5.2 Double Stimulus or Degradation Category Rating (DCR)

In this category, test sequences are presented in pairs (reference and distorted sequences), and subjects are asked to evaluate the quality of the corrupted video compared to the reference video. Furthermore, the reference and impaired sequences can be presented on the display simultaneously if possible. The end-user/evaluator assesses the quality of the distorted video compared to the reference video. The rating scale ranges from “imperceptible”, “perceptible but not annoying”, “slightly annoying”, “annoying”, and “very annoying”.

#### 1.2.5.3 Differential Mean Opinion Score (DMOS)

In this category, reference, and test sequences are presented in pairs. Unlike the DCR, subjects are asked to evaluate the test video quality compared to the reference video and to assess the perceived quality difference between the two videos [45]. The rating scale ranges from -3 (significantly worse) to +3 (significantly better), whereas 0 indicates no perceptible difference.

#### 1.2.5.4 Pair Comparison Method

In this category, test sequences are presented in pairs. For each pair, the subject is asked to determine the better-quality video. No absolute quality rating is assigned. All potential combinations of pairs are used to assess the video quality. For example, in the case of two sequences A and B, possible pairs will be AB and BA.

### 1.2.6 Quality of Experience of Mobile Video

Audio is an essential part of mobile video. Distorted videos may contain distorted audio, which can affect user perception. However, audio is a separate entity and is inaudible to human ears, while video is a visual product. The focus of this thesis is on video for mobile devices. Different factors influence the mobile video QoE, which includes the non-technical and technical influencing factors [46]. This thesis investigates influencing factors like video codecs, frame rate, video resolution, display screen, packet loss, packet delay, jitter, and network outage to determine how they impact the QoE of mobile video. QoE measures the user's experience with a service such as a video/mobile video streaming service. There are various definitions of QoE, however, the definition that aligns well with IP networks or multimedia streaming is by ITU-T, which is:

*"The overall acceptability of an application or service, as perceived subjectively by the end user.*

- *Quality of experience includes the full end-to-end system effects (client, terminal, network, services infrastructure, etc.).*

- *Overall acceptability can be influenced by user expectations and context" [47].*

Another definition of QoE is
*"It is a measurement used to determine the extent to which a network is meeting the user's needs" [48].*

### 1.2.7 Mean Opinion Score

The Mean Opinion Score (MOS) is one of the most frequently employed subjective assessments. MOS is a numerical measure of video or image quality based on subjective evaluations by human viewers. In a typical MOS test, participants rate the quality of video samples on a predefined scale, such as 1 (bad) to 5 (excellent). The average of these ratings across participants provides the MOS. This score reflects the perceived quality of the video and is widely used for benchmarking and validating objective quality metrics. MOS is particularly valuable because it captures the viewer's perspective. It is a key measure for services/applications focused on user satisfaction. A similar method has been in use for decades to evaluate the telephony network quality experienced by the end-user and is referred to as Mean Opinion Score (MOS) (ITU-T Rec P.800). It is recommended by ITU-T and used for the assessment of telephone, IPTV, and multimedia applications [8, 40, 43, 49, 50]. Different rating scales are used for MOS; however, five- and nine-point scales are typically used. The nine-point scale is used if a higher discriminative power is required. For low-bitrate videos, a nine-point rating scale is recommended by ITU-T [43]. The rating of MOS is distributed between "Excellent" to "Bad" as shown in Figures 1.5 and 1.6.

A five-level scale, illustrated in Figure 1.7, is employed to assess the degree of impairment of the distorted sequence relative to the reference sequence.

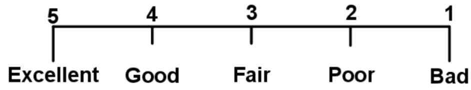



**Figure 1.5:** Five-Point Scale Mean Opinion Score [43].

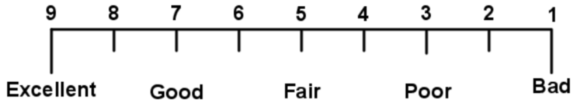



**Figure 1.6:** Nine-Point Scale Mean Opinion Score [43].

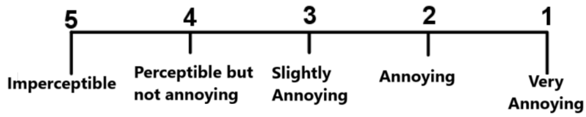



**Figure 1.7:** Five-Point Scale Degradation Category Rating [43].

### 1.2.8 Relationship between QoE and QoS

QoE is also referred to as users' perceived quality of service [51, 52] and captures the users' experience more clearly from end-to-end QoS [53]. For acceptable user-perceived QoS, different applications demand the distinctive optimization of parameters like jitter, delay, packet loss, and throughput. Best-effort IP networks do not provide any guarantee that the data is delivered to the user according to its requirements. Its possibility of being unreliable and unpredictable increases in wireless and mobile networks compared to wired ones. Network performance is measured by using QoS parameters such as bandwidth, jitter, packet delay, and packet loss. QoS is critically important in wireless video delivery as QoS parameters significantly impact the user's viewing experience [54]. Network transport QoS is used to prioritize and control network traffic at the transport level for optimal performance of time-sensitive applications. According to [55], for multimedia streaming, it is important to monitor the network transport QoS through QoS parameters, such as delay, delay variation, and packet loss. Real-time or multimedia applications are very sensitive to network performance parameters. Packet delays exceeding the playout time, or a lost packet may degrade the video quality. Poor performance of a network, especially fluctuating QoS, can degrade QoE. For mobile video/live streaming, QoS is a foundation for QoE, as QoS measures the network's quality of delivery. Real-time applications or multimedia streaming may suffer from degraded quality due to packet loss, delay, and delay variation [50]. Different applications require different network performance parameters tuning for good quality of experience. The study of QoE and QoS relationships can help to identify the threshold of network performance parameters to maintain the user-perceived quality at a satisfactory level. Delay can affect the user perception in many ways, e.g., long initial waiting time to establish the service can upset the user. The delay variation is more critical for multimedia streaming. It can cause packet re-ordering and degradation of the perceptual quality. A jitter buffer is used to reduce the impact of packet delay variation. Packet loss is another QoS parameter, and in the case of real-time applications, packet loss can result in information loss. If a keyframe is lost, the video may not be viewable, or it may play with a major interruption until the next frame arrives. Different studies have been conducted and have shown that network performance fluctuation affects the video quality [8, 19, 56, 57].

The relationship between QoE and QoS has drawn significant attention from researchers and service providers. Understanding this relationship helps identify thresholds for QoS parameters to ensure satisfactory user-perceived quality. In competitive environments, assessing user experience is crucial for service providers. Developing models that correlate QoE with QoS can enhance multimedia application efficiency and improve overall user satisfaction.

### 1.2.9 Network Emulation

For performance evaluation and testing of real-time applications and services, simulation, live testing, or emulation are used. Simulation is the most commonly used method for performance evaluation of the real-time application and is also used to test and verify the proof-of-concept of a newly designed application or protocol. A simulator is an application or piece of code running on a computer, based on a traffic model, which represents the behavior of the network and returns the performance results. However, it cannot provide the environment of real-time operation.

Another option to evaluate an application or service is to use a real environment, which is perhaps the best way to experiment. However, this method is hard to use and control because of cost and lack of easy access to equipment. The trade-off between simulation and live testing is emulation. Emulation is best described as an imitation and is used to imitate the internal design and functionality of the target platform. Network emulation mimics the target network and is used for performance analysis of a service, protocol, or application. It takes the input from the network or traffic source, applies defined rules, alters the packet flow, and injects it back into the network or connected computer. Emulation uses real equipment to a possible extent along with the simulated part of network attributes. Furthermore, it is flexible, easy to use, and gives better control in experiments with reproducibility options. The network emulator can be a computer running emulation software or a dedicated device, attached to computers and behaving like a real network. It is also referred to as a traffic generator or traffic shaper. Traffic shapers are used to study the behavior of networks, especially with user-defined traffic shaping parameters, e.g., delay, packet delay variation, throughput, and packet loss. Also, traffic shapers are used for the enforcement of SLA, e.g., for limiting throughput. Traffic shapers are used in research to emulate the behavior of targeted networks. In our experimental setup, we use emulation to study the impact of QoS parameters on mobile video QoE. To emulate the network behaviors, different shapers are in use by the research community, we selected three shapers NetEm, NISTnet, and KauNet to evaluate their performance for delay and throughput shaping, the results show, that the performance of NetEm is better than other two. Therefore, we selected NetEm for further experiments [58, 59].

## 1.3 Structure of the Thesis

The remainder of this thesis is organized as follows: Chapter 2 presents the research objectives, research questions, and methods employed in the papers included in this thesis. Chapter 3 provides an overview of the papers, along with discussions and conclusions. Chapter 4 is based on Paper I *"Quality of Experience Hourglass Model,"* while Chapter 5 discusses Paper II *"Mobile Video Sensitivity to Packet Loss and Packet Delay Variation in Terms of QoE."* Chapter 6 focuses on Paper III *"Impact of Disturbance Locations on Video Quality of Experience."* Lastly, Chapter 7 is based on Paper IV *"QoE Rating Performance Evaluation of ITU-T Recommended Video Quality Metrics in the Context of Video Freezes,"* and Chapter 8 covers Paper V *"Mitigation of the Effects of Network Outage on Video QoE Using a Sender Buffer."*

# 2 Methodology

The thesis investigates video QoE when the video is streamed and viewed on a mobile device. Various factors can affect the video quality of the experience, including video resolution, frame rate, color depth, compression, bit rate, codec, network conditions, etc.

This research focuses on the impact of network performance on the quality of transmitted video. Furthermore, it examines other factors of video quality and their resilience in the presence of poor network performance. This work uses subjective tests to evaluate the quality of all video streaming scenarios. These results are compared with algorithm-based applications like PEVQ and Temporal Quality Metric (TQM) [60, 61]. Besides network performance and its artifacts on video, this thesis explores the network outage and video freeze and proposes a sender buffer-based mitigation solution for network outages.

The subsequent sections present research objectives, research questions, and research methods employed in the papers included in this thesis.

## 2.1 Research Objectives

*Research Objective I: To understand the role of all TCP/IP model layers in video QoE.*

The purpose is to understand the role of each TCP/IP model layer towards QoE, especially from the network layer to the application layer. To achieve this goal, a theoretical QoE model is discussed, which is inspired by the TCP/IP hourglass model, and the role of each TCP/IP layer towards QoE is explained. Subjective tests were conducted on impaired mobile videos due to network performance parameters. Furthermore, user feedback was collected on laptops and mobile devices to see the impact of the Quality of Presentation (QoP) on QoE. *Paper I* is an umbrella for this objective, and *Paper II* contributes subjective test results toward this objective.

*Research Objective II: To study the role of QoS parameters in degrading video QoE.*

Objective II is closely related to the network layer of Objective I. However, the second objective of this research is to understand the relationship between quality of service and QoE. In particular, the impact of packet delay, packet delay variation, and packet loss on mobile video QoE. *Paper II - Paper IV* contribute the subjective results towards this objective.

*Research Objective III: Monitoring the impact of network outages on video QoE.*

This objective focuses on special artifacts like video freeze, miss, and jump that result from a network outage. This also addresses the user reaction when the same artifact occurs at a different timeline of a video sequence. Location-based user ratings are used to understand the limitations of PEVQ and TQM ratings. The results strengthened the understanding of user ratings of network outages and the location of artifacts results of network outages. *Paper III* and *Paper IV* contribute to the results and shed light on the impact of network outages.

*Research Objective IV: To investigate and propose a solution to mitigate the impact of network outages on the video QoE.*

Video coding and encoding can correct errors to some extent. However, video freezing and missing data correction/recovery due to network outages in live transmission are difficult. In transmission, network outages are obvious, either due to technical glitches or even due to handover. They can drastically reduce user pleasure, and user ratings go low. In addition to the other requirements, mitigation of end-user annoyance is an important factor. *Paper V* suggests a buffer-based solution to mitigate the impact of network outages on video QoE.

## 2.2 Research Questions

**RQ1:** What is the role of IP/Network and the above layers towards QoE in mobile video streaming?

**Motivation:** In mobile video streaming, user satisfaction is highly dependent on network performance parameters such as latency, jitter, and packet loss. Although these parameters directly affect QoS, the impact of these parameters on QoE is not straightforward and easy. This research question aims to fill the gap in how network layer parameters translate into QoE. Understanding this relationship is crucial for designing adaptive streaming solutions and optimizing multimedia delivery systems to ensure a smooth viewing experience under different network conditions.

**RQ2:** How do QoS parameters, especially packet delay, packet delay variation, and loss, affect mobile video QoE, and how does the playback device affect user ratings?

**Motivation:** Packet loss and delay variation are inevitable in real-world networks. Their impact on video quality can vary depending on the video encoding profiles and playback devices. With the widespread use of the H.264 Baseline Profile in mobile video streaming, it is critical to understand how these impairments affect QoE. Additionally, while mobile devices and laptops are both common playback platforms, it remains unclear whether device-specific factors, such as screen size and resolution, significantly influence user perception. This research question aims to provide insight into the QoS parameters of the H.264 baseline profile in relation to video QoE and to examine the role of playback devices on QoE.

**RQ3:** How do freezing artifacts, introduced at different locations in videos with varying resolutions and frame rates affect QoE ratings, and how accurately do the objective metrics proposed by ITU-T estimate it?

**Motivation:** Freezing artifacts caused by network issues are a common challenge in video streaming, directly affecting user experience. While objective metrics like PEVQ and TQM are used to estimate QoE, their accuracy and correlation with subjective user ratings, particularly under diverse video conditions, remain uncertain. This research aims to evaluate the effectiveness of these metrics, helping service providers better predict and mitigate user dissatisfaction.

**RQ4:** How does the introduction of a sender buffer mechanism mitigate the impact of network outages on video playback, and what is its effect on the QoE?

**Motivation:** Network outages disrupt video streaming by causing freezes and data loss, leading to degraded QoE. Existing buffering techniques often fail to fully address these issues in real-time streaming scenarios. By exploring the sender buffer's ability to recover from outages and improve QoE, this study provides insights into practical solutions for enhancing video streaming performance under adverse network conditions.

## 2.3 Research Methods

The research investigates the impact of network impairments, such as delay, jitter, and packet loss, on video streaming quality through a series of experiments. The experiments were conducted using a carefully designed emulation setup that allowed precise control over network conditions. The study used subjective and objective video quality assessments. For subjective tests besides user ratings, the ITU-T recommended tool PEVQ is used and evaluated [61]. We followed the ACR method for subjective tests. Three primary elements formed the experimental setup: video streaming devices, a network emulator, and video players. The experiments were carried out using standard H.264 encoded video sequences. The methodology emphasized real-world scenarios by simulating network impairments and recording their effects on video playback. The following subsections provide details of the experimental setup carried out in different studies conducted as part of this research. Table 2.1 provides an overview of the thesis, illustrating the relationships between the objectives, research questions, methods, and the corresponding papers included in the thesis. The details of the methods employed in each paper are presented in the subsequent section.

### 2.3.1 Approach followed in Study 1

The QoE Hourglass Model is introduced in this study, drawing inspiration from the Internet Hourglass Model, to conceptualize QoE across layers, including QoS, Quality of Delivery (QoD), Quality of Presentation (QoP), and overall QoE. The model establishes mathematical relationships between these layers, linking QoE to network parameters such as throughput, delay, jitter, and packet loss. Subjective evaluations using MOS were conducted to assess device-specific QoE variations across laptops and mobile devices, focusing on the impact of QoP. Matched-sample t-tests were employed to analyze the significance of these variations. Additionally, the study evaluated the performance of different video codecs and profiles, such as H.264, to determine their error resilience and influence on QoE under similar network conditions.

### 2.3.2 Approach followed in Study 2

The study investigates the sensitivity of mobile video quality to packet loss and packet delay variation in terms of QoE. It utilizes an experimental setup involving a network emulator to simulate packet loss and delay variation, streaming H.264 Baseline Profile videos over a 100 Mbps full-duplex link using VideoLAN Client (VLC) media player. Subjective assessments were conducted following ITU-T recommendations, where 34 participants rated video sequences displayed on both laptops and mobile devices. The study used ACR for user evaluations and MOS to quantify the perceived video quality. The effects of Packet loss ratios (0.4% to 7%) and delay variations (±2 ms to ±16 ms) were analyzed. A matched-sample t-test was applied to compare user perceptions across devices. The findings highlight that packet loss and delay variation significantly impact video QoE. In contrast, the type of device does

**Table 2.1:** Mapping of Objectives, Research Questions, Methods, and Papers.

| Objective | Related Research Questions (RQs) | Methods | Addressed in |
|---|---|---|---|
| **Objective 1:** Understand TCP/IP model layers' role in QoE | **RQ1**: What is the role of IP/Network and the above layers toward QoE in mobile video streaming? | Theoretical modeling: QoE Hourglass Model; Subjective evaluations using MOS on mobile and laptop devices | Paper I |
| **Objective 2:** Study QoS parameters' role in video QoE degradation | **RQ2**: How do QoS parameters (delay, latency, loss) affect QoE?<br>**RQ3**: Does the type of device impact QoE at the same resolution? | Network emulation for packet loss and delay variation; Subjective evaluations using ACR (Absolute Category Rating) scale; Comparison of mobile and laptop user ratings | Paper II<br>Paper III |
| **Objective 3:** Monitor network outages' impact on QoE | **RQ3**: How do freeze artifacts at various points affect QoE?<br>**RQ4**: How accurately do PEVQ and TQM metrics estimate QoE? | Experimental setup with controlled freezes and jumps; Subjective tests in a lab setting; Objective metrics evaluation using PEVQ and TQM; Comparison with subjective user ratings | Paper III<br>Paper IV |
| **Objective 4:** Propose mitigation solutions for network outages | **RQ5**: How does a sender buffer mitigate network outages on QoE?<br>**RQ6**: What is the optimal sender buffer size to minimize freeze time and data loss? | Theoretical modeling of sender buffer performance; Simulated network outages; Subjective evaluations using ACR and MOS; Statistical analysis of freeze durations and buffer configurations | Paper V |

not substantially influence QoE ratings when the resolution of the video is the same.

### 2.3.3 Approach followed in Study 3

This study used an experimental setup to investigate the effects of network impairments, such as latency, jitter, and packet loss, on video QoE. FFmpeg [62] was used to simulate video freezes. The setup included a video streamer, video player, shaper, and network measurement point, with VLC used for streaming and playback, and NetEm for generating latency and delay variance. Video traffic was captured before and after the shaper using a Distributed Passive Measurement Infrastructure (DPMI) [63]. The latency (D) and delay variance ($\Delta D$) were implemented as $D \pm \Delta D$ = 100 ms ± 0 ms, 2 ms, 4 ms, 6 ms, 8 ms, 10 ms, 12 ms, 14 ms, 16 ms. The PEVQ was used to measure objective QoE, while user perception tests were conducted following ITU-R BT 500-11 and ITU-T P.910 recommendations using the ACR method. Users rated video sequences with interruptions at different locations, such as freeze and jump effects. The results were analyzed to compare objective PEVQ ratings with user perceptions, highlighting variations in perceived QoE based on the type and location of interruptions.

### 2.3.4 Approach followed in Study 4

The study employed a comprehensive methodology to evaluate the impact of video freezes on QoE. Subjective assessments were conducted in the BTH perception laboratory following ITU-T P.910 guidelines. Six videos with varying motion intensities were encoded at multiple resolutions, i.e., Quarter Common Intermediate Format (QCIF), Common Intermediate Format (CIF), Quarter Video Graphics Array (QVGA), Video Graphics Array (VGA), and frame rates (6.25 to 30 frames per second (fps)). One-second freezing artifacts were introduced at different locations using MATLAB, with the last displayed frame repeated during freezes. Twenty-five participants rated the videos using the single stimulus method, with randomized presentation orders to avoid bias, and session durations limited to 30 minutes to prevent fatigue. Also, the ratings were measured using algorithm-based PEVQ and TQM. The results were compared with user ratings. PEVQ estimates video quality based on human visual system modeling, while TQM is a no-reference quality metric.

### 2.3.5 Approach followed in Study 5

The study investigated the impact of network outages on video QoE by employing a controlled experimental setup. The proposed methodology utilized sender and jitter buffers to mitigate the effects of outages. This study comprises two main components: the first involves theoretical modeling, analysis, and discussion, while the second involves simulated videos based on the theoretical results. Standard video sequences were streamed with controlled network outages, and the sender buffer stored video data during interruptions, transmitting it at a higher rate upon recovery to quickly refill the jitter buffer. The experiments simulated different network conditions, varying the durations and locations of network outages as well as buffer sizes. Both subjective and objective QoE assessments were performed. Subjective evaluations involved user ratings using the ACR method in a controlled perception lab following ITU-R BT.500 and ITU-T P.910 recommendations. The MOS were analyzed to assess the effectiveness of the sender buffer, and statistical methods, including Student's t-tests, were used to validate the results. The study demonstrated that the sender buffer significantly improved QoE by reducing video freeze durations, data loss, and latency.

# 3 Thesis Overview

## 3.1 Overview and Contributions

This thesis explores various aspects of video QoE, focusing on the interplay between network performance, video impairments, and user perception. The five papers comprising this thesis collectively contribute to understanding and improving video streaming under diverse conditions, ranging from conceptual modeling to experimental analysis and practical mitigation strategies. Paper I discusses the conceptual model of QoE in parallel to the classical Internet architecture hourglass model. Paper II investigates the impact of packet loss and packet delay variation on mobile video in terms of QoE. Paper III examines the relationship between freeze and its location for mobile video in terms of QoE. Paper IV addresses the impact of the freeze on real-time streaming and evaluates the performance of PEVQ and TQM compared to users' ratings. Paper V proposes a buffer-based solution to mitigate the impact of network outages on video QoE. Figure 3.1 summarizes the overall contribution of the thesis. Details of each paper is presented in the subsequent sections.

## Paper I: Quality of Experience Hourglass Model

The paper introduces the QoE Hourglass Model, a conceptual framework inspired by the classical Internet architecture hourglass model, designed to bridge the gap between technical network metrics and user-perceived video quality. The model comprises four layers: QoS, QoD, QoP, and QoE. Each layer uniquely contributes to the end-user's perceived quality, representing interactions from the network to the application level.

The study examines how technical metrics such as packet loss, delay, and jitter propagate through these layers to impact user perception. Subjective evaluations emphasize the significant role of QoP in shaping overall QoE. In addition, it identifies conditions in which QoE can be accurately predicted using QoS metrics alone, without additional factors.

This model offers a systematic framework for service providers, operators, and researchers to analyze and improve multimedia QoE. It links measurable network pa-

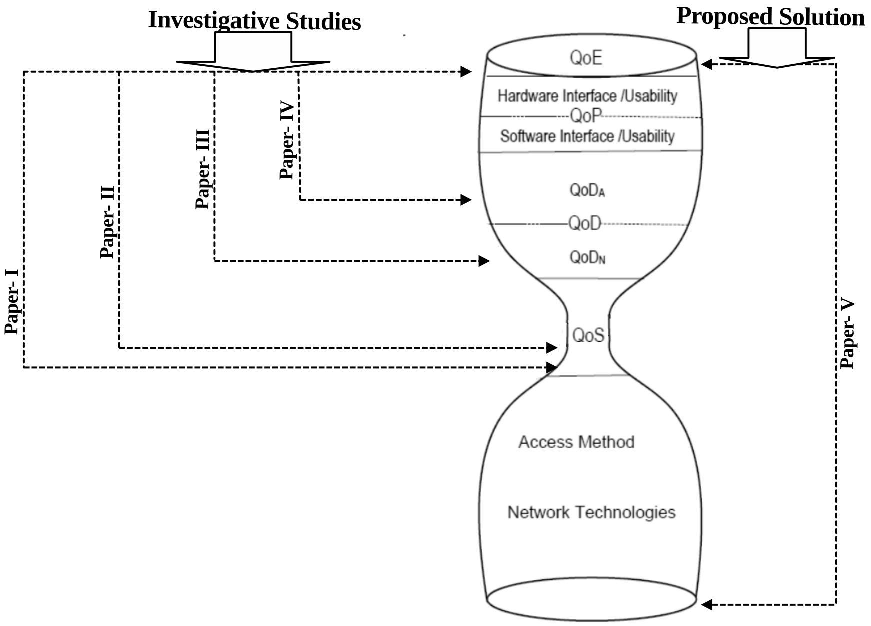



**Figure 3.1:** Thesis Contributions Along TCP/IP Model.

rameters to user perceptions and highlights the interrelationship between technical and perceptual factors.

**Key Questions Addressed:**

1. How can video streaming QoE be effectively modeled through a layered approach inspired by the classical Internet hourglass model?

2. How can QoE be objectively assessed using QoS parameters?

3. How do QoD and QoP influence overall QoE?

## Paper II: Mobile Video Sensitivity to Packet Loss and Packet Delay Variation in Terms of QoE

This paper examines the effects of packet loss and Packet Delay Variation (PDV) on the QoE in mobile video streaming. The focus is on videos encoded with the widely used H.264 Baseline Profile codec, commonly used for low-power processing devices. User perception was evaluated through subjective tests conducted under controlled network conditions using mobile devices and laptops.

The findings reveal that even minor network impairments can significantly degrade QoE, with PDV exerting a more pronounced effect than packet loss. Surprisingly, the results show that the type of device (mobile vs. laptop) has minimal influence on QoE when the video resolution remains the same. This challenges common assumptions about device-specific differences in user perception and highlights the critical need to adopt a mechanism to address network impairments to prevent a bad user experience.

By linking network parameters to user perception, this study provides valuable insights for service providers and developers aiming to optimize video streaming quality under varying network conditions.

**Key Questions Addressed:**

1. What is the effect of packet loss on the QoE of mobile video encoded with the H.264 Baseline Profile?

2. How does PDV affect mobile video QoE?

3. Does device type (mobile vs. laptop) significantly influence QoE when viewing videos at the same resolution?

## PART III: Impact of Disturbance Locations on Video Quality of Experience

This paper explores the effects of freeze-and-jump artifacts and delay variation on QoE in video streaming, with a focus on the role of disturbance location within a video sequence. Using subjective tests, the study compares user ratings with objective evaluations provided by the PEVQ model.

The results reveal that user perception is significantly influenced by both the type and location of disturbances. For instance, freezes during critical or engaging moments in the video receive more negative ratings than those occurring in less impactful segments. While PEVQ closely aligns with subjective ratings for delay variation, it consistently overestimates QoE in freeze-and-jump scenarios.

This study underscores the limitations of existing objective metrics in capturing the nuanced effects of complex impairments on user perception. It emphasizes the need for user-centric QoE assessments that account for the dynamic and context-sensitive nature of video streaming experiences.

**Key Questions Addressed:**

1. How do freeze-and-jump artifacts and delay variation affect the perceived quality of video streaming?

2. How does the location of disturbances within a video sequence influence user perception?

3. How well do objective metrics like PEVQ align with subjective user ratings under different impairment scenarios?

## PART IV: QoE rating Performance evaluation of ITU-T recommended Video Quality Metrics in the context of Video freezes

This paper evaluates the ability of PEVQ and TQM to predict QoE in scenarios involving video freezes. The study investigates the effects of varying video resolutions, frame rates, and freeze locations on the alignment between these metrics and subjective user ratings.

Findings reveal that both metrics have limitations in accurately predicting QoE during repeated freeze events. PEVQ tends to overestimate user ratings, particularly under complex freeze scenarios, while TQM demonstrates better alignment with subjective evaluations, especially at higher frame rates. These results highlight the need for improved objective metrics that consider the temporal and spatial characteristics of video impairments to enhance the accuracy of QoE predictions.

**Key Questions Addressed:**

1. How do video freezes impact QoE during real-time streaming from the user's perspective?

2. To what extent do PEVQ and TQM effectively represent user-perceived QoE for various resolutions and frame rates during freezes?

3. What are the limitations of PEVQ and TQM in predicting QoE under freeze-induced interruptions?

## PART V: Mitigation of the Effects of Network Outage on Video QoE Using a Sender Buffer

This paper proposes a sender buffer mechanism to mitigate the negative impact of network outages on video streaming QoE. The sender buffer stores video data during outages and transmits it at a higher rate upon recovery to minimize freeze durations and improve playback continuity.

The study derives theoretical models to quantify data loss and freeze times during outages and validates the findings through subjective tests using simulated network conditions. Results indicate that the sender buffer significantly enhances QoE, particularly when its size matches the jitter buffer, optimizing performance by reducing data loss and freeze durations. The study highlights the importance of buffer configurations and the capacity factor (i.e., the rate of buffered data transmission post-outage) in enhancing user satisfaction.

**Key Questions Addressed:**

1. How do network outages affect video QoE?
2. Can a sender buffer reduce the negative effects of network outages on streaming quality?
3. What is the optimal size of the sender buffer relative to the jitter buffer for minimizing freeze durations and data loss?
4. How does the capacity factor influence the effectiveness of the sender buffer in improving QoE?
5. How do user ratings and MOS compare between sender-buffer-based and traditional streaming methods?

## 3.2 Discussion

This thesis contributes to understanding mobile video QoE by systematically examining how video degradation, caused by network performance parameters such as delay, jitter, and packet loss, impacts user perception. It further explores mitigation strategies, such as buffering techniques, designed to address specific impairments, caused by network outages. By integrating theoretical modeling, experimental validation, and practical solutions, this work provides a comprehensive solution for enhancing QoE in real-time video streaming applications. The studies presented in this thesis are interconnected, with each building upon the insights of the previous, creating a cohesive narrative that advances both scientific knowledge and practical approaches to enhancing QoE. The subsequent paragraphs provide a brief discussion of each paper included in this thesis, followed by subsections on research implications and limitations. Implications are discussed from both research and application perspectives.

**Paper I** invites a deeper question about the tension between technical and human factors in user experience analysis. While it effectively structures the interdependence of QoS, QoD, and QoP with QoE, it also quietly disagrees with the traditional engineering mindset that prioritizes computable network metrics over subjective perception. So the question arises: to what extent, if any, can QoE ever be measured quantitatively without subjecting it to whimsical human perception? Furthermore, the model suggests that controlling QoE through QoS is insufficient–should future work be on adaptive, AI-driven systems that continuously respond to live user input? In practical terms, this could mean moving beyond hard-wired QoS guarantees to more comprehensive, context-aware service designs that take into account device diversity, user desires, and even emotional states. This note flags a broader trend in network services: from performance optimization to personalized experience.

**Paper II** extends this exploration by narrowing its focus to packet loss and delay variation as specific network impairments and their impact on mobile video QoE. While technological improvements in mobile devices enable them to handle a higher quality of video, they do not make up for the overarching issue: network instability causes degradation in QoE. Also, the results contradict the usual presumption that tge type of device plays an important role in perception–hypothesizing that resolution and encoding settings may be more important than the medium of display. This prompts a shift in focus: should research continue refining network performance metrics, or should it pivot toward more user-centric solutions that integrate perceptual resilience into video streaming technologies? Such considerations highlight a broader challenge in multimedia delivery–ensuring that user expectations align with what is technically achievable in real-world conditions.

**Paper III** explores the effects of delay, jitter, delay variation, and freeze and jump artifacts on QoE. A key finding is that the timing of freezes significantly influences user perception. Freezes early in a video have a lesser impact on QoE compared to those in the middle or end, where user involvement is higher. This reflects a user tolerance for disruptions early on but increased sensitivity as engagement deepens. The study also highlights that content complexity affects user sensitivity, with fast-moving scenes being less affected by delay variations than slower ones. In addition, the research validates PEVQ as a reliable metric for assessing QoE under network impairments. These insights emphasize the need for content-aware network optimizations to enhance user satisfaction and mitigate impairments dynamically.

**Paper IV** addresses the issue of freeze artifacts, which are one of the most perceptible forms of video degradation in streaming applications. The findings reveal that higher frame rates generally lead to better user ratings, as they mask the negative effects of freezes to some extent. However, the presence of repeated freezes significantly worsens the decline in perceived quality, regardless of frame rate or resolution. While the study compares subjective user ratings with objective metrics such as PEVQ and TQM, the results emphasize the limitations of these tools in fully capturing the details of user-perceived quality, particularly in scenarios involving complex temporal impairments.

Finally, **Paper V** proposes a sender buffer mechanism to mitigate the effects of network outages on video playback. By storing video data during interruptions and transmitting it at higher rates during recovery, the mechanism ensures smoother playback and significantly reduces freeze durations. Subjective assessments confirm that the buffer enhances QoE, demonstrating its practical potential for real-world applications. This study contributes to the body of work on proactive mitigation strategies and provides a robust foundation for further exploration into adaptive buffering techniques for video streaming.

### 3.2.1 Implications

**Implications for Research:** The findings of this thesis highlight several avenues for advancing video streaming research:

1. **Validation in Real-World Networks:** Papers I and II establish a theoretical foundation for linking network parameters to QoE, but the experiments were conducted in a controlled environment. This suggests a need for future research to validate these models in real-world, diverse environments featuring varying network conditions and device types.

2. **Refinement of QoE Metrics:** Papers III and IV expose the limitations of existing algorithms like PEVQ and TQM in capturing detailed user perceptions. In future research, developing hybrid frameworks that integrate subjective and objective methods will provide more accurate QoE assessments.

3. **Integration with Adaptive Protocols:** In Paper V the introduction of sender buffer mechanism shows promise for mitigating playback disruptions. Future work can explore its integration with adaptive streaming protocols and its performance across diverse scenarios.

4. **Content-Specific Insights:** Papers III and IV underscore the role of video content characteristics, such as motion intensity and scene complexity, in shaping

QoE. Researchers can focus on tailoring streaming solutions to specific content types.

5. **Inclusivity in QoE Models:** To ensure generalizability, future studies must account for broader user demographics, codec diversity, and emerging technologies like 5G and next-generation video codecs.

**Implications for Application:** The research offers actionable insights for improving video streaming services:

1. **QoE Monitoring Framework:** The QoE Hourglass Model (Paper I) provides a comprehensive framework for monitoring and optimizing video streaming services by addressing factors across QoS, QoD, and QoP.

2. **Error-Resilience Mechanisms:** Findings from Papers III and IV highlight the need for adaptive buffering strategies and refined evaluation tools to better align with user perceptions.

3. **Mitigation of Playback Disruptions:** Paper V's sender buffer mechanism demonstrates a practical solution for addressing network-induced playback disruptions. This approach can be integrated into real-time streaming platforms to enhance user satisfaction.

4. **Content-Aware Optimization:** Insights from Papers III and IV suggest developing streaming algorithms that dynamically adapt to video content characteristics, such as motion and complexity, to improve QoE.

5. **Service Robustness:** By integrating these findings, service providers can enhance user satisfaction and deliver consistent video quality, even in challenging network environments.

### 3.2.2 Limitations of the Research

While the studies collectively offer significant insights into QoE, certain limitations must be acknowledged. The research focuses primarily on videos encoded with the H.264 Baseline Profile, leaving room for exploration of codec diversity and its impact on QoE. The studies could benefit from broader demographic representation to ensure inclusivity and applicability across diverse user groups.

Further, this research provides valuable insights into the effects of delay and delay variation but does not address scenarios involving simultaneous impairments, such as the interplay between delay and packet loss. Concerning the evaluation of freeze artifacts, the research provides comprehensive mechanisms, but it does not fully explore the interaction between freeze frequency, freeze location, and user perception, which could further enrich the understanding of temporal impairments. Finally, this research primarily focuses on simulated network outages to demonstrate the usefulness of the sender buffer, however, it needs to be tested in real-world scenarios.

## 3.3 Conclusions

This thesis advances the understanding of video QoE by systematically investigating the impact of network impairments, video artifacts, and mitigation strategies on user perception. The QoE Hourglass Model provides a structured framework for linking measurable network parameters with user-perceived QoE, while findings across all studies demonstrate the significant influence of factors such as packet loss, delay variation, and freeze artifacts. Despite the promising alignment of objective metrics like PEVQ with user ratings, their limitations in detailed scenarios reinforce the indispensability of subjective evaluations. The sender buffer mechanism proposed in this work offers a practical solution for mitigating playback disruptions, showcasing the potential of proactive buffering strategies to enhance QoE in real-world applications.

### 3.3.1 Future Work

Building on the findings of this thesis, future research should address several key areas to further bridge the gap between technical optimizations and user-centric design:

*Refinement of QoE Metrics:* Develop hybrid evaluation frameworks that assess subjective assessments based on objective metrics under diverse network conditions and content types by using machine learning and artificial intelligence.

*Codec-Specific Optimizations:* Investigate the performance of other codecs under varying impairments to enhance error resilience and adaptive streaming mechanisms, to figure out the best suitable codec for better QoE.

*Real-World Network Environments:* Expand research to study the impact of physical networks on QoE to ensure the generalizability of findings across diverse scenarios.

*Content-Aware Strategies:* Explore the interplay between video characteristics (e.g., motion intensity, scene complexity) and user perception to design QoE models tailored to specific use cases, such as gaming and live sports streaming.

*Advanced Mitigation Techniques:* Investigating the role of a backup buffer parallel to a jitter buffer to diminish the freeze time due to network outage.

*Inclusive User Studies:* Broaden demographic representation in subjective evaluations to ensure QoE models are inclusive and reflective of diverse user groups and accessibility needs.

By addressing these areas, future research can contribute to the development of more robust, user-centric QoE frameworks that enhance video streaming experiences in increasingly complex and dynamic network environments.

# Bibliography


[1] *2023 Global Internet Phenomena Report*. SANDVINE, The App QoE Company, 2000.

[2] A. Saha, S. Pentapati, Z. Shang, R. Pahwa, B. Chen, H. Gedik, S. Mishra, and A. Bovik. "Perceptual video quality assessment: the journey continues!" In: *Frontiers in Signal Processing* 3 (June 2023). DOI: 10.3389/frsip.2023.1193523.

[3] *Dive Brief: User-created-content especially video gains on traditional media consumption.* [accessed on Dec 2024]. URL: https://www.marketingdive.com/news/user-created-content-especially-video-gains-on-traditional-media-consum/616745/.

[4] Cisco. *Cisco Annual Internet Report - Cisco Annual Internet Report (2019)*. en. Tech. rep. Library Catalog: www.cisco.com. 2019. URL: https://www.cisco.com/c/dam/en_us/about/annual-report/cisco-annual-report-2019.pdf.

[5] Cisco. *Cisco Annual Internet Report - Cisco Annual Internet Report (2018–2023) White Paper*. en. Tech. rep. Library Catalog: www.cisco.com. 2024. URL: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html.

[6] *Mobile traffic surges with video consumption*. [accessed on December 2024]. URL: https://csimagazine.com/csi/Mobile-traffic-surges-with-video-consumption.php.

[7] *Verizon. Consumer Connections Report.* [accessed on December 2024]. URL: https://www.verizon.com/about/connections.

[8] T. N. Minhas, O. Gonzalez Lagunas, P. Arlos, and M. Fiedler. "Mobile video sensitivity to packet loss and packet delay variation in terms of QoE". In: *2012 19th International Packet Video Workshop (PV)*. 2012, pp. 83–88. DOI: 10.1109/PV.2012.6229747.

[9] Nokia. *Quality of Experience (QoE) of mobile services:Can it be measured and improved*. [accessed on November 2011]. 2011. URL: http://www.nokia.com/NOKIA_COM_1/About_Nokia/Press/White_Papers/pdf_files/whitepaper_qoe_net.pdf.

[10] I. Mazur, J. Rak, and K. Nowicki. "Minimising the Churn Out of the Service by Using a Fairness Mechanism". In: *Computer Networks*. Ed. by P. Gaj, W. Gumiński, and A. Kwiecień. Cham: Springer International Publishing, 2020, pp. 117–137. ISBN: 978-3-030-50719-0.

[11] M. Fiedler, K. De Moor, H. Ravuri, P. Tanneedi, and M. Chandiri. "Users on the Move: On Relationships Between QoE Ratings, Data Volumes and Intentions to Churn". In: *2017 IEEE 42nd Conference on Local Computer Networks Workshops (LCN Workshops)*. 2017, pp. 97–102. DOI: 10.1109/LCN.Workshops.2017.70.

[12] L. J. Karam, T. Ebrahimi, S. S. Hemami, T. N. Pappas, R. J. Safranek, Z. Wang, and A. B. Watson. "Introduction to the Issue on Visual Media Quality Assessment". In: *Selected Topics in Signal Processing, IEEE Journal of* 3.2 (Apr. 2009), pp. 189–192. ISSN: 1932-4553. DOI: 10.1109/JSTSP.2009.2015485.

[13] J. Doe, A. Smith, and R. Brown. "Recent advancements in objective QoE measurement techniques". In: *IEEE Transactions on Multimedia* 25.3 (Mar. 2023), pp. 456–467.

[14] T. Hoßfeld and P. Pérez. "A theoretical framework for provider's QoE assessment using individual and objective QoE monitoring". In: *2024 16th International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE. 2024, pp. 235–241.

[15] I. Orsolic and L. Skorin-Kapov. "A Framework for in-Network QoE Monitoring of Encrypted Video Streaming". In: *IEEE Access* 8 (2020), pp. 74691–74706. DOI: 10.1109/ACCESS.2020.2988735.

[16] H. J. Kim, D. H. Lee, J. M. Lee, K. H. Lee, W. Lyu, and S. G. Choi. "The QoE Evaluation Method through the QoS-QoE Correlation Model". In: *2008 Fourth International Conference on Networked Computing and Advanced Information Management*. Vol. 2. 2008, pp. 719–725. DOI: 10.1109/NCM.2008.202.

[17] T. N. Minhas and M. Fiedler. "Quality of experience hourglass model". In: *2013 International Conference on Computing, Management and Telecommunications (ComManTel)*. 2013, pp. 87–92. DOI: 10.1109/ComManTel.2013.6482371.

[18] M. Fiedler, T. Hossfeld, and P. Tran-Gia. “A generic quantitative relationship between quality of experience and quality of service”. In: *Network, IEEE* 24.2 (Apr. 2010), pp. 36–41. ISSN: 0890-8044. DOI: 10.1109/MNET.2010.5430142.

[19] C.-H. Lin, C.-H. Ke, C.-K. Shieh, and N. Chilamkurti. “The Packet Loss Effect on MPEG Video Transmission in Wireless Networks”. In: *Advanced Information Networking and Applications, 2006. AINA 2006. 20th International Conference on*. Vol. 1. Apr. 2006, pp. 565–572. DOI: 10.1109/AINA.2006.325.

[20] M. Carson and D. Santay. “NIST Net: a Linux-based network emulation tool”. In: *SIGCOMM Comput. Commun. Rev.* 33.3 (July 2003), pp. 111–126. ISSN: 0146-4833. DOI: 10.1145/956993.957007. URL: https://doi.org/10.1145/956993.957007.

[21] P. Vicat-Blanc Primet, R. Takano, Y. Kodama, T. Kudoh, O. Gluck, and C. Otal. “Large Scale Gigabit Emulated Testbed for Grid Transport Evaluation”. In: *In Proceedings of The Fourth International Workshop on Protocols for Fast Long-Distance Networks, PFLDnet’2006*. 2006.

[22] The Linux Foundation. [Online] http://www.linuxfoundation.org/collaborate/workgroups/networking/netem. [accessed on June 2010]. 2010.

[23] J. Garcia, P. Hurtig, and A. Brunström. *KauNet: Design and Usage*. Tech. rep. MSU-CSE-99-39. Karlstad, Sweden: Karlstad University, Faculty of Economic Sciences, Communication and IT, Aug. 2008. URL: http://urn.kb.se/resolve?urn=urn:nbn:se:kau:diva-3189.

[24] A. Kukushkin. *Introduction to mobile network engineering: Gsm, 3g-wcdma, lte and the road to 5g*. John Wiley & Sons, 2018.

[25] H. Knoche and J. McCarthy. “Mobile users’ needs and expectations of future multimedia services”. In: *Proceedings of the WWRF12* (2004). URL: https://infoscience.epfl.ch/handle/20.500.14299/76001.

[26] N. Kehtarnavaz and M. N. Gamadia. *Real-time image and video processing: from research to reality*. Vol. 5. Morgan & Claypool Publishers, 2006.

[27] K. Zeng, T. Zhao, A. Rehman, and Z. Wang. “Characterizing Perceptual Artifacts in Compressed Video Streams”. In: *Proceedings of SPIE - The International Society for Optical Engineering* 9014 (Jan. 2014). DOI: 10.1117/12.2043128.

[28] C. Feng, D. Danier, F. Zhang, A. Mackin, A. Collins, and D. Bull. “BVI-Artefact: An artefact detection benchmark dataset for streamed videos”. English. In: *2024 Picture Coding Symposium (PCS)*. Picture Coding Symposium (PCS). 2024 Picture Coding Symposium (PCS), PCS 2024 ; Conference date: 12-06-2024 Through 14-06-2024. United States: Institute of Electrical and

Electronics Engineers, pp. 1–5. ISBN: 9798350358490. DOI: 10.48550/arXiv.2312.08859. URL: https://2024.picturecodingsymposium.org/.

[29] L. Lin, S. Yu, L. Zhou, W. Chen, T. Zhao, and Z. Wang. "PEA265: Perceptual Assessment of Video Compression Artifacts". In: *IEEE Transactions on Circuits and Systems for Video Technology* 30.11 (2020), pp. 3898–3910. DOI: 10.1109/TCSVT.2020.2980571.

[30] T. N. Minhas and M. Fiedler. "Mitigation of the Effects of Network Outage on Video QoE Using a Sender Buffer". In: *Electronics* 10.10 (2021). ISSN: 2079-9292. DOI: 10.3390/electronics10101209. URL: https://www.mdpi.com/2079-9292/10/10/1209.

[31] Z. Wang, H. R. Sheikh, and A. C. Bovik. "Objective video quality assessment". In: *In the handbook of video databases: Design and Applications*. CRC Press, 2003, pp. 1041–1078.

[32] V. Vassiliou, P. Antoniou, I. Giannakou, and A. Pitsillides. "Requirements for the Transmission of Streaming Video in Mobile Wireless Networks". In: *Artificial Neural Networks ICANN 2006*. Vol. 4132. Lecture Notes in Computer Science. 2006, pp. 528–537.

[33] L. Guo and Y. Meng. "What is Wrong and Right with MSE?" In: *Signal and Image Processing* (2006), pp. 212–215.

[34] Q. Huynh-Thu and M. Ghanbari. "Scope of validity of PSNR in image/video quality assessment". In: *Electronics Letters* 44.13 (2008), pp. 800–801. ISSN: 0013-5194. DOI: 10.1049/el:20080522.

[35] M. Pinson and S. Wolf. "A new standardized method for objectively measuring video quality". In: *Broadcasting, IEEE Transactions on* 50.3 (Sept. 2004), pp. 312–322. ISSN: 0018-9316. DOI: 10.1109/TBC.2004.834028.

[36] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. "Image quality assessment: from error visibility to structural similarity". In: *Image Processing, IEEE Transactions on* 13.4 (Apr. 2004), pp. 600–612. ISSN: 1057-7149. DOI: 10.1109/TIP.2003.819861.

[37] OPTICOM. *Perceptual Evaluation of Video Quality*. http://www.pevq.org/. [accessed on July 2010]. 2010.

[38] W. Stefen. *Digital Video Image Quality and Perceptual Coding*. Ed. by H. Wu and K. Rao. CRC Press, 2005. Chap. Perceptual video Quality metrics - a review.

[39] U. Engelke and H.-J. Zepernick. "Perceptual-based Quality Metrics for Image and Video Services: A Survey". In: *Next Generation Internet Networks, 3rd EuroNGI Conference on*. May 2007, pp. 190–197. DOI: 10.1109/NGI.2007.371215.

[40] ITU-R BT.500-11. *Methodology for the subjective assessment of the quality of television pictures*. International Telecommunications Union Radiocommunication Sector. 2002.

[41] VQEG. *Video Quality Experts Group*. http://www.its.bldrdoc.gov/vqeg/. [accessed on July 2010]. 2010.

[42] *Human Factors (HF);Quality of Experience (QoE) requirements for real-time communication services*. Tech. rep. European Telecommunications Standards Institute, Jan. 2010. URL: http://www.etsi.org/deliver/etsi_tr/102600_102699/102643/01.00.02_60/tr_102643v010002p.pdf.

[43] ITU-T P.910. *Subjective video quality assessment methods for multimedia applications.* International Telecommunications Union Telecommunication Sector. 1999.

[44] M. Claypool and J. Tanner. "The effects of jitter on the peceptual quality of video". In: *MULTIMEDIA '99: Proceedings of the seventh ACM international conference on Multimedia (Part 2)*. Orlando, Florida, United States: ACM, 1999, pp. 115–118. ISBN: 1-58113-239-5. DOI: http://doi.acm.org/10.1145/319878.319909.

[45] ITU-R BT.500-14. *Methodologies for the subjective assessment of the quality of television images*. International Telecommunications Union Radiocommunication Sector. [accessed on December 2024]. URL: https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.500-14-201910-S!!PDF-E.pdf.

[46] D.-H. Shin. "Conceptualizing and measuring quality of experience of the internet of things: Exploring how quality is perceived by users". In: *Information & Management* 54.8 (2017), pp. 998–1011. ISSN: 0378-7206. DOI: https://doi.org/10.1016/j.im.2017.02.006. URL: https://www.sciencedirect.com/science/article/pii/S0378720617300952.

[47] ITU-T FG IPTV. *Focus Group on IPTV Standardization: Definition of Quality of Experience (QoE)*. International Telecommunications Union Telecommunication Sector Publications. Feb. 2007. URL: http://ties.itu.int/ftp/public/itu-t/fgiptv/readonly/Previous_Meetings/20070122_MountainView/il/T05-FG.IPTV-IL-0050-E.htm.

[48] Webopedia. *QoE*. URL: http://www.webopedia.com/TERM/Q/QoE.html.

[49] ITU-T P.8000. *Methods for subjective determination of transmission quality.* International Telecommunications Union Telecommunication Sector. Aug. 1996.

[50] T. N. Minhas and M. Fiedler. "Impact of Disturbance Locations on Video Quality of Experience". In: *Quality of Experience for Multimedia Content Sharing, EuroITV2011*. Lisbon, Portugal, June 2011.

[51] A. Khan, L. Sun, E. Jammeh, and E. Ifeachor. "Quality of experience-driven adaptation scheme for video applications over wireless networks". In: *In Special Issue on Video Communication over Wireless Networks* (2009).

[52] T. N. Minhas. *Network Impact on Quality of Experience of Mobile Video*. Licentiate dissertation. Karlskrona, Sweden, 2012.

[53] H. A. Tran and A. Mellouk. "QoE Model Driven for Network Services". In: *Wired/Wireless Internet Communications*. Ed. by E. Osipov, A. Kassler, T. M. Bohnert, and X. Masip-Bruin. Vol. 6074. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 264–277. URL: `http://www.springerlink.com/content/7159q510q378m234/`.

[54] J. Jiang and B. Girod. "Review of cross-layer approaches for quality-optimized wireless video delivery". In: *Proceedings of the IEEE*. Vol. 97. 1. IEEE. 2009, pp. 186–223.

[55] G. Van der Auwera and M. Reisslein. "Implications of Smoothing on Statistical Multiplexing of H.264/AVC and SVC Video Streams". In: *Broadcasting, IEEE Transactions on* 55.3 (Sept. 2009), pp. 541–558. ISSN: 0018-9316. DOI: `10.1109/TBC.2009.2027399`.

[56] D. Loguinov and H. Radha. "End-to-end Internet video traffic dynamics: statistical study and analysis". In: 2 (2002), 723–732 vol.2. DOI: `10.1109/INFCOM.2002.1019318`.

[57] M. Li, M. Claypool, and R. Kinicki. "MediaPlayer™ versus RealPlayer™: a comparison of network turbulence". In: *Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurment*. IMW '02. Marseille, France: ACM, 2002, pp. 131–136. ISBN: 1-58113-603-X. DOI: `http://doi.acm.org/10.1145/637201.637221`.

[58] J. Shaikh, T. N. Minhas, P. Arlos, and M. Fiedler. "Evaluation of delay performance of traffic shapers". In: *Security and Communication Networks (IWSCN), 2010 2nd International Workshop on*. IEEE. 2010, pp. 1–8.

[59] T. N. Minhas, M. Fiedler, J. Shaikh, and P. Arlos. "Evaluation of throughput performance of traffic shapers". In: *Wireless Communications and Mobile Computing Conference (IWCMC), 2011 7th International*. IEEE. 2011, pp. 1596–1600.

[60] OPTICOM. *Perceptual Evaluation of Video Quality*. http://www.pevq.org/. accessed July 2010.

[61] ITU-T REC-J.247. *Objective perceptual multimedia video quality measurement in the presence of a full reference.* International Telecommunications Union Telecommunication Sector. Aug. 2008.

[62] FFMPEG. [Online] http://www.ffmpeg.org. accessed June 2010.

[63] P. Arlos, M. Fiedler, and Arne A. Nilsson. “A Distributed Passive Measurement Infrastructure”. In: *Proceedings of Passive and Active Measurement Workshop*. Boston, US, 2005, pp. 215–227.

# Paper I

## Quality of Experience Hourglass Model

*Tahir Nawaz Minhas and Markus Fiedler. Quality of experience hourglass model. In Proceedings of International Conference on Computing, Management and Telecommunications (ComManTel), pages 87–92. IEEE, 2013.*

## 1 Abstract

Recently, the user perception got the attention of research community as well as the service providers. Both are interested to know how the users perceive quality of multimedia streaming service. To avoid the fatigue of subjective testing, professional are interested to model the QoE based on QoS and other measurable parameters. However these parameters to estimate the QoE may differ per service.

In this paper, we propose a conceptual model of QoE in parallel to the classical internet architecture hourglass model. Moreover conditions are highlighted under which QoE can be evaluated objectively from QoS without entailing other factors. The main goal was to understand the factors that are playing the role on QoE. We also discuss the role of application on QoE.

## 2 Introduction

With the growth of mobile internet, the number of users of multimedia applications are also growing rapidly. Moreover, new mobile devices are well equipped to handle the multimedia applications, which made them a popular platform for watching video content. The end-user, i.e. the viewer, has certain expectations about the video quality and wants to have good quality of experience. These user expectations trigger the interest of service providers, operators, manufacturers and researchers in QoE. On the basis of her experience and satisfication, a user may decide to continue with the service provider or not. Nowadays the end-user has more than one option to choose the service provider. So to keep the user loyal with a service and its provider and to prevent churn is a big challenge for service providers. In [1], churn is considered as a key performance measure for mobile operators. Among other factors, customer experience also influence churn. The satisfied user recommends the service, while dissatisfied users shared their experience within the circle of friends. According to [2] 82% of users leave the service due to frustration, which is result of bad experience with the service and one abandonment user shares her experience with 13 other people. To know the perception and experience of the user about the network or network service is important to improve or tune the network in a way that the users face good QoE.

QoS is a well-known and established concept, especially for technical personnel and researchers. The service providers do their best to tune the network to achieve good quality of service. A technically well-established network with the optimal QoS gives the feedback of technical excellence but still can't estimate the end-user experience or perception. The estimation of user experience and perception comes under the umbrella of QoE, an entirely different measure than QoS. Different methods and terminologies are used to elaborate the user experience. For example, in [3] the authors used the term user-oriented QoS and measured the video quality using quality rating scaled from 1 to 5 (bad, poor, fair, good, excellent). Similarly, in [4] the authors define the video quality assessment in terms of perceived QoS and measure the video quality based on spatial and temporal activity using the 0-to-100 scale. Other terms used in similar concepts are end-to-end QoS, user-perceived QoS and QoS experienced [5–7]. The study of QoE is of utmost important for service providers because it helps them to prevent churn or failure of the service as well as to optimize their services for their survival in face of ever-growing competition. Most QoE-oriented researchers and professionals are working to define new reliable as well as easily computable assessment methodologies and quality prediction models. In [8] the authors proposed a pentagram model to measure the QoE, based on five factors: integrality, retainability, availability, usability and instantaneousness. The QoS and QoE correlation model and QoE evaluation using QoS parameters is discussed in [9].

In this paper, we propose the QoE model in parallel to the internet architecture, in which we consider and formalise the different factors affecting QoE, from IP layer to application layer.

The remainder of this paper is organized as follows. Section 3 defines the QoE and Section 4 describes the QoE hourglass model. Section 6 and 7 discuss the role of presentation devices and of application on QoE, respectively, whereas Section 8 concludes the paper.

# 3 Quality of Experience

QoE is an emerging thesis in the area of communications and networks, but the existence of this concept is not very new. The quality of experience is a well known concept in different subjects, for example, in philosophy and psychology [10, 11].

Moreover, the term quality of experience was used in 1938 by Jhon Dewey in his book "Experience and Education" [12].
There are a lot of definitions for QoE. We select the definition that fits to the IP networks or multimedia streaming. The definition of QoE by ITU-T is
*"The overall acceptability of an application or service, as perceived subjectively by the end-user.*

- *Quality of Experience includes the complete end-to-end system effects (client, terminal, network, services infrastructure, etc).*
- *The overall acceptability may be influenced by user expectations and context." [13]*

Another definition of QoE is *"It is a measurement used to determine how well that network is satisfying the end user's requirements" [14].*

# 4 The QoE Hourglass Model

The proposed QoE hourglass model is inspired by the classical internet hourglass model [15], which can be divided into the five layers. Table 1 describes the layers of the classical internet hourglass model with the corresponding OSI model's layers.
The QoE hourglass model and the classical internet hourglass model are shown side-by-side in Fig 1. There are ample applications and application layers protocols in use. New applications and application layer protocols are introduced on as-needed basis. Above the transport and internet layers, different applications and application layer protocols are working next to each other. Similarly below the internet layer different technologies, access methods and protocols are in use. New entities can be introduced on data link layer or physical layers, as desired. However, network layer and transport layers remain unchanged. Particularly, at network layer IPv4 or IPv6 rule the internet communication. Above and below the IP (Internet Protocol) or the network layer different kind of protocols exist and new ones can be added. Due to this addition bottom and top layers of internet architecture are expanding from the very first day of the internet due to the addition of new applications, protocols, technologies, and access methods. As a result of this addition internet architecture took the shape of an hourglass, with narrow neck at the network layer. In the internet hourglass model, IP finds itself at the narrow neck of an hourglass in a way that all applications and application layer protocols are above the IP layer, while all data link layer and physical layer protocols are below the IP [16, 17]. In this model, IP is playing the main role in communication and is independent of network technologies underlaying the IP. Similarly, any protocol can be used above IP. However, the network layer contains only the internet protocol that is IPv4 or IPv6. The internet protocol (IP) treats all networks equally [18].

| | Internet Hourglass Model's Layers | OSI Model's Layers |
|---|---|---|
| 1 | Application Layer | 5-7 |
| 2 | Transport Layer | 4 |
| 3 | Network Layer | 3 |
| 4 | Data Link Layer | 2 |
| 5 | Physical Layer | 1 |

**Table 1:** The classical Internet Hourglass Model with corresponding OSI Model's Layers.

The most common understanding of QoS relates to the IP level [19]. The QoE Hourglass Model has Quality of Service (QoS) at the narrow neck of the hourglass. As IP is independent of data-link layer and physical layer protocols, the QoE hourglass model bottom layer is Quality of Service corresponding to IP layer. The layers of QoE hourglass model are

1. Quality of Experience (QoE)
2. Quality of Presentation (QoP)
3. Quality of Delivery (QoD)
4. Quality of Service (QoS)

In case of network service or video streaming QoE is a comprehensive experience of end-user that depends on QoP, QoD and QoS. According to the QoE hour

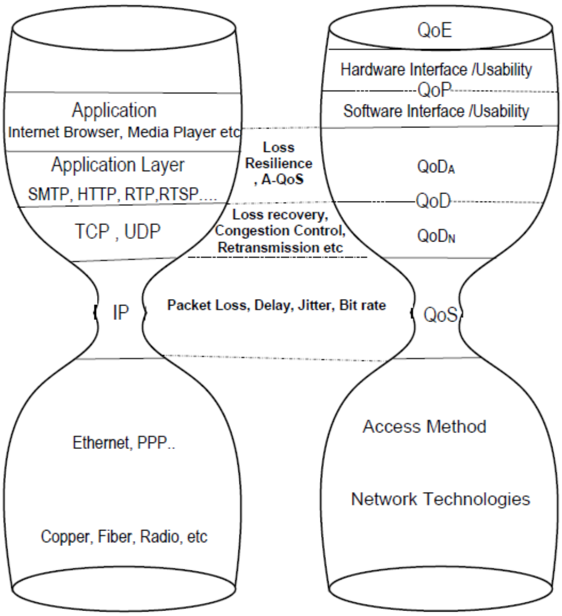



**Figure 1:** The QoE Hourglass Model

glass model, QoE depends on QoP, where the end-user has a direct interaction and experience with the displaying devices and interface of applications. Any faulty or unexpected behavior of device or application can damage the user perception even with excellent network performance. Along with QoP contribution towards the user perception, it depends on QoD. QoD concerns with the quality of data delivered. The quality of data delivery is two fold, (i) Quality of data delivered by the application and (ii) Quality of data delivered by the Network.

The quality of data delivered by the application ($\mathrm{QoD_A}$) is a result of the treatment of data by the application to enhance the end-to-end performance and also referred as application-based QoS [20]. The quality of data delivered by the network ($\mathrm{QoD_N}$) is cumulative effect of network performance parameters and data handling by transport layer protocols. The ($\mathrm{QoD_N}$) depends on quality of server (QoS), which finds itself on the narrow neck of the QoE Hourglass Model.

## 4.1 Quality of Service (QoS)

The QoS is corresponding to the network layer as shown in Fig. 2. The network layer manages the delivery of packets to and from the underlying networks and transport layer. QoS is a well established concept among the service providers and the research community. The most commonly used parameters to measure the QoS or network performance are delay, delay variation, throughput, response time and loss. The fundamental parameters associated with video streaming are throughput, packet delay, delay variation and packet loss. On the best effort based service network these parameters are changing with respect to time, but are measurable at particular instants. As these parameters are uncontrollable with best effort service, it is very hard to guarantee the QoS. For the safeguard of user QoE, it is very important to model the relation of QoS-QoE, which will help in improving and designing the applications or services. In the QoE hourglass model, QoS is a foundation layer of the model, which is measurable. Throughput ($R$) is measured as bit per second at destination, whereas packet delay ($D$) is defined as end-to-end transmission time. In case of packet delay variation ($\Delta D$), packets arrive at destination with different delays, usually both are measured in milliseconds. Packet loss ($L$) denotes the share of undelivered packets.Thus

$$QoS = f(L, D, \Delta D, R). \tag{I.1}$$

The effect of this bottom layer traverses towards the top layer of QoE through QoD.

## 4.2 Quality of Delivery (QoD)

QoD is concerned with the quality of the data delivery process. It maps the transport layer, application layer and partially the application to the classical internet hourglass model. Both transport and application layers have different mechanism to handle the data delivery process. At transport layer the delivery process of TCP and UDP are entirely different from each other. The data delivered by the transport layer is processed and presented by an application to the end-user. If this application comes with error recovery procedures, then the packet loss or any recoverable error of data has a nominal effect on quality, otherwise this error may damage the quality to a significant extent. QoD can be divided into two parts.

- Quality of Delivery by Network ($\mathrm{QoD_N}$)
- Quality of Delivery by Application ($\mathrm{QoD_A}$)

and is expressed as

$$QoD = f(QoD_N, QoD_A). \tag{I.2}$$

### 4.2.1 Quality of Delivery by Network ($\mathrm{QoD_N}$)

Data delivered by the network is the resulting effect of QoS and behavior of transport layers protocols. The quality of data delivered by the network is related to the data delivery process of transport layer. At this layer TCP and UDP are involved in process-to-process data delivery mechanism and both have different characteristics. The key features of TCP are retransmissions of lost packets, reliability, flow control and congestion control, while UDP doesn't have such features. In case of packet loss recovery process or in case of congestion control TCP reduces the throughput of stream. Similarly long packet delays (approximately 200 ms) can invoke the retransmission process of TCP. However, UDP doesn't care for anything and continues sending the data. Moreover, a data stream simultaneously can have packet delay variation and packet loss; both are measured and treated differently. However, in this case, $\mathrm{QoD_N}$ is a collective impact of both parameters (packet delay variation and packet loss) of QoS on the data stream.

$$QoD_N = g(QoS). \tag{I.3}$$

### 4.2.2 Quality of Delivery by Application ($\mathrm{QoD_A}$)

The transport layer is responsible for providing the data delivery services to the application layer, and the application layer provides the data to the application, which then presents the data to the user. To this end, the transport layer uses what it gets from the delivery process of underneath layers. $\mathrm{QoD_A}$ captures the improvement of data quality by the application to enhance the end-to-end performance. Any data processing-application (e.g. a of video codec) may have the ability to rectify the impact of QoS parameters. For example, to reduce the effect of packet delay variation, jitter buffer is used; similarly, video applications have error correction functionality to recover lost data. So the data delivered by the application may have different quality as compared to the data delivered by the network. In case of mobile video, $\mathrm{QoD_A}$

based on error recovery functionality of the codec and may vary from application to application. For a detail example and discussion of $\text{QoD}_\text{A}$ and the corresponding role of the application in QoE, see section 7. Using Equation I.3, Equation I.2 becomes

$$QoD = f(g(QoS), QoD_A). \tag{I.4}$$

As $\text{QoD}_\text{A}$ depends on the application, it is not directly related to the network data delivery process. So for, a particular application, this factor can be considered as invariant:

$$QoD = f(g(QoS)|_{QoD_A}). \tag{I.5}$$

## 4.3 Quality of Presentation (QoP)

QoP addresses the second layer of the QoE hourglass model. This is the layer where the end-user has her experience with displaying devices and applications. This layer mostly deals with the software and hardware interface and their usability along with the delivered content of video. Even with excellent QoD, poor presentation or bad interface may degrade the video quality of experience. For example, poor resolution of the displaying device or unbalanced settings of display brightness and contrast can damage the user experience. Similarly, a high quality video on a "dead pixels" screen will be perceived as a poor quality. In [21] the authors compared the video and image quality assessment on LCD and CRT displays and found that CRT has better performance for video while LCD was preferred for still images. QoP depends on the QoD and $\text{QoP}_\text{o}$, whereas QoD and $\text{QoP}_\text{o}$ are independent and $\text{QoP}_\text{o}$ depends on hardware and software interface and usability.

$$QoP = h(QoD, QoP_o). \tag{I.6}$$

$\text{QoP}_\text{o}$ is the contribution of this layer and does not depend on the data delivery process. By freezing this independent factor Equation I.6 becomes

$$QoP = h(QoD|_{QoP_o}). \tag{I.7}$$

## 4.4 Quality of Experience (QoE)

In the QoE hourglass model, QoE is the resultant of all quality factors contributed by the underlying layers as shown in figure 1. It is the feel or experience of the user residing at the top, which measures the true success of a service. The end-user experiences the QoE by interacting with presenting equipment and application, described by is QoP. Thus

$$QoE = k(QoP). \tag{I.8}$$

# 5 Estimation of QoE on the Basis on QoS Measurement

IP forwards the data to the transport layer, and then this data is handled by the application and presented to the user. In the QoE hourglass model, we map these layers with their contribution to QoE. Each layer is contributing towards the final assessment of QoE, independently or depending on previous layer input, or with both independent and dependent factors. For the simplification of this model, independent contributions of each layer are considered as constant e.g $\mathrm{QoD_A}$ and $\mathrm{QoP_o}$. From Equation (I.8) using the QoP from Equation (I.7), we arrive at

$$QoE = k(h(QoD|_{QoP_o})). \tag{I.9}$$

Using QoD definition with frozen $\mathrm{QoD_A}$, given by Equation (I.5), we obtain

$$QoE = k(h(f(g(QoS)|_{QoD_A})|_{QoP_o})). \tag{I.10}$$

Hence the QoE estimation can be modelled from the QoS parameters, with frozen or known factors of $\mathrm{QoD_A}$ and $\mathrm{QoP_o}$. Both $\mathrm{QoD_A}$ and $\mathrm{QoP_o}$ are contribution of presenting application and equipment, and independent of the network data delivery process. The role of $\mathrm{QoP_o}$ and $\mathrm{QoD_A}$ is discussed with examples in the following sections 6 and 7, respectively.

# 6 Role of $QoP_o$ in QoE

Usability and user interface are a well studied subjects in Computer Science. The $QoP_o$ comes under the influence of Human-Computer Interaction (HCI). In [22] the authors compare the subjective video quality results conducted with professional CRT and consumer LCD and conclude that under the tested resolution LCD and CRT can achieve very similar results. In [23] the authors compared the subjective quality results of videos shown on the laptop and on the mobile phone. The selected videos are impaired with packet loss and packet delay variation. Videos are encoded with H.264 baseline profile with 30 fps and have a resolution of $320 \times 240$.

Figures 2 and 3 show the MOS for packet loss and delay variation for laptop and mobile, respectively. For both graphs we can notice that the results are almost the same for both of the devices except for the packet loss of $0.4\%$.

To identify if these variations are significant, a matched-sample *t*-test was used. Table 2 presents the results for the calculations for each of the video pairs, referring to the same video sequence presented on the laptop and on the mobile phone. $D \pm \Delta D$ stands for delay variation in the table. The comparison of $t_{calc}$ with $t_{table}$ reveals that for most of the video sequences, there is no evidence that the different devices affect the perception of the shown video. For a detailed description see the paper [23]. We

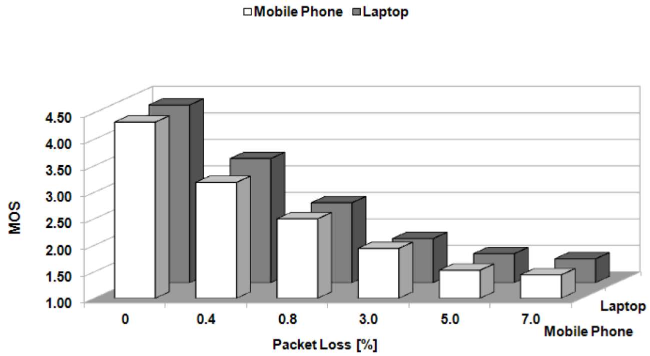



**Figure 2:** MOS as function of packet loss for videos displayed on laptop and mobile phone [23]

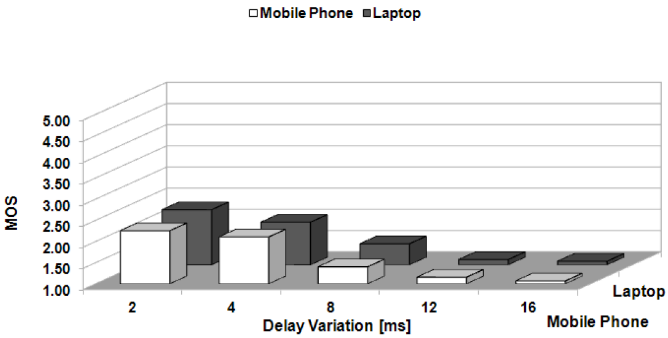



**Figure 3:** MOS as function of delay variation for videos displayed on the laptop and mobile phone [23]

can conclude that there is not a significant difference between doing the experiment in the mobile phone and in the laptop displaying the same resolution. Overall very similar rating behavior was observed for both devices as shown in Table 2. The hypothesis is rejected only for 0.4% packet loss of all videos and for $D \pm \Delta D = \{150 \pm 2, 8\}$ ms in the Football case.

**Table 2:** Matched-sample *t* test results, with $t_{calc} = 2.042$ [23]

| Foreman | | | Hall-Monitor | | | Football | | |
|---|---|---|---|---|---|---|---|---|
| $L$ | $t_{calc}$ | Hypothesis | $L$ | $t_{calc}$ | Hypothesis | $L$ | $t_{calc}$ | Hypothesis |
| 0.0% | −0.63 | Fail to reject | 0.0% | 0.90 | Fail to reject | 0.0% | 0.42 | Fail to reject |
| 0.4% | −1.14 | Fail to reject | 0.4% | 2.50 | Reject | 0.4% | 3.20 | Reject |
| 0.8% | 0.42 | Fail to reject | 0.8% | 1.36 | Fail to reject | 0.8% | −2.03 | Fail to reject |
| 3.0% | −2.05 | Fail to reject | 3.0% | −0.44 | Fail to reject | 3.0% | 0.25 | Fail to reject |
| 5.0% | −0.25 | Fail to reject | 5.0% | −0.27 | Fail to reject | 5.0% | 1.16 | Fail to reject |
| 7.0% | 0.00 | Fail to reject | 7.0% | 0.83 | Fail to reject | 7.0% | −0.63 | Fail to reject |
| $D \pm \Delta D$ | $t_{calc}$ | Hypothesis | $D \pm \Delta D$ | $t_{calc}$ | Hypothesis | $D \pm \Delta D$ | $t_{calc}$ | Hypothesis |
| $150 \pm 2$ ms | −1.36 | Fail to reject | $150 \pm 2$ ms | 0.00 | Fail to reject | $150 \pm 2$ ms | 2.46 | Reject |
| $150 \pm 4$ ms | 0.70 | Fail to reject | $150 \pm 4$ ms | −1.28 | Fail to reject | $150 \pm 4$ ms | −0.89 | Fail to reject |
| $150 \pm 8$ ms | 0.44 | Fail to reject | $150 \pm 8$ ms | −0.37 | Fail to reject | $150 \pm 8$ ms | 2.73 | Reject |
| $150 \pm 12$ ms | 0.44 | Fail to reject | $150 \pm 12$ ms | 0.00 | Fail to reject | $150 \pm 12$ ms | −1.68 | Fail to reject |
| $150 \pm 16$ ms | 0.00 | Fail to reject | $150 \pm 16$ ms | 0.00 | Fail to reject | $150 \pm 16$ ms | 0.81 | Fail to reject |

# 7 Role of $QoD_A$ in QoE

To enhance the video performance in case of errors, different error control techniques such as forward error correction (FEC), retransmission, error resilience, and error concealment, can be used. Different video codecs use different error control techniques, which results in different video quality under similar impairments. In [3] the authors discuss the loss-distortion model that accounts for the impact of network losses on video quality. For experimentation they select the H.264 and MPEG-2 codecs and found that under similar environments and packet loss conditions, the average distortion of video encoded with H.264 and MPEG-2 is different.

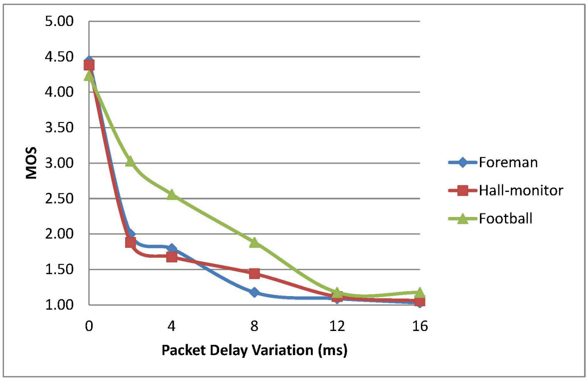



**Figure 4:** Mean Opinion Scores of videos encoded with H.264 Baseline Profile [23]

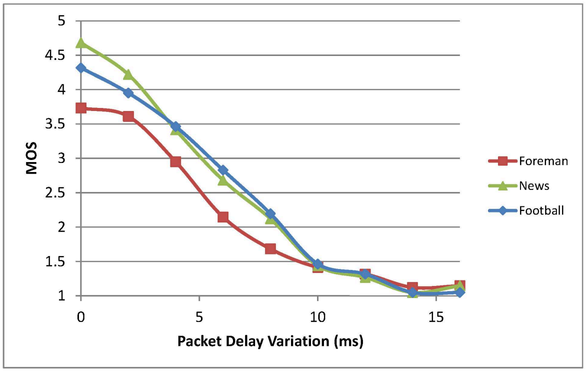



**Figure 5:** Mean Opinion Scores of videos encoded with H.264 Main Profile [24].

Figure 4 and Figure 5 shows the MOS of the three different video sequences encoded with H.264 main profile and baseline profile. MOS is given along the y-axis, while the x-axis shows the packet delay variation. To see the effect of packet delay variation, the player's jitter buffer is set to zero. NetEm [25] was used to shape the delay variation and by default it re-orders the packets while shaping the variable delay. From Figure 5 and 4 we can see that for 2 ms of packet delay variation, the

MOS rating of the Foreman video encoded with main profile is higher as compared to the same video encoded with baseline profile of H.264. In short, we saw that under similar network performance conditions, both profiles of H.264 show different MOS ratings, which highlights the role of $QoD_A$ in QoE. In [26] subjective assessment of codecs performace is performed for videos compressed with DivX 6.0 (DivX, Inc.), XviD (Open Source Project), WMV (Microsoft Corporation) and x264(Open Source Project) with the two bitrates 690 kbps and 1024 kbps. More than 50 participants took part in the assessment of these videos using the MOS 0-9 scale . Figure 6 shows the average results of MOS for all videos and bitrates. Despite similar resolutions and bitrates, their quality differs according to the use of different applications (codec) for compression. All these examples higlight the pronounced role of $QoD_A$ on user ratings.

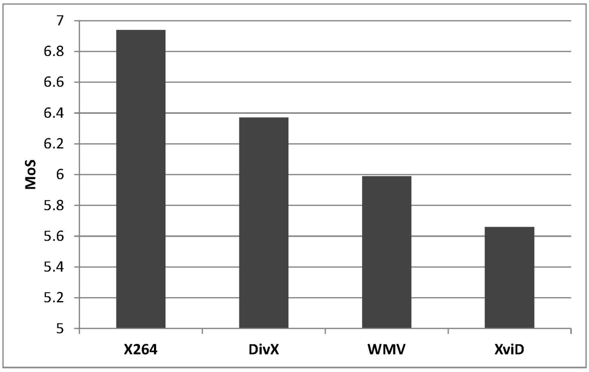



**Figure 6:** Mean Opinion Scores of videos encoded with x264, DivX WMV and Xvid [26].

# 8 Conclusion

In this paper, we proposed the QoE hourglass model, in which we addressed the streaming video quality of experience and its factors. This model has four quality layers corresponding to layers in the classical Internet hourglass model, and each layer has a role towards the user perceived quality. According to our argument, QoE of streaming videos cannot be evaluated only with QoS parameters in general because other factors are also involved. However, we showed that the QoE can be evaluated from QoS parameters by freezing the role of $QoD_A$ (or for particular Codec) and $QoP_o$. We also discussed the role of the application on QoE by presenting the results of video quality for different codecs as well as for the main profile and baseline profile of H.264.

# References


[1] W. I. Briefing. *Mobile Churn: The impact of churn on mobile subscriber profitability*. accessed November 2011. URL: %5BOnline%5D%20http://www.wds.co/enlightened/churn/WDS_Churn_2011.pdf.

[2] Nokia. *Quality of Experience (QoE) of mobile services:Can it be measured and improved*. accessed November 2011. URL: %5BOnline%5D%20http://www.nokia.com/NOKIA_COM_1/About_Nokia/Press/White_Papers/pdf_files/whitepaper_qoe_net.pdf.

[3] O. Verscheure, P. Frossard, and M. Hamdi. "User-Oriented QoS Analysis in MPEG-2 Video Delivery". In: *Real-Time Imaging* 5.5 (1999), pp. 305–314. ISSN: 1077-2014. DOI: 10.1006/rtim.1999.0175.

[4] H. Koumaras, A. Kourtis, D. Martakos, and J. Lauterjung. "Quantified PQoS assessment based on fast estimation of the spatial and temporal activity level". In: *Multimedia Tools and Applications* 34 (Mar. 2007), pp. 355–374. ISSN: 1380-7501, 1573-7721. DOI: 10.1007/s11042-007-0111-1. URL: http://www.springerlink.com/content/f8v2p4r852266415/.

[5] P. Casas, P. Belzarena, and S. Vaton. "End-2-End evaluation of IP multimedia services, a user perceived quality of service approach". In: *ITC Specialist Seminar on Quality of Experience, 18th. Proceedings. Karlskrona, Sweden*. 2008, pp. 1–11. URL: http://iie.fing.edu.uy/publicaciones/2008/CBV08.

[6] N. Bhatti, A. Bouch, and A. Kuchinsky. *Integrating User-Perceived Quality into Web Server Design*. Jan. 2000. URL: http://adrianmarriott.net/logosroot/papers/websiteDesign.pdf.

[7] A. Perkis, S. Munkeby, and O. I. Hillestad. "A model for measuring Quality of Experience". In: *Signal Processing Symposium, 2006. NORSIG 2006. Proceedings of the 7th Nordic*. IEEE, June 2006, pp. 198–201. ISBN: 1-4244-0412-6. DOI: 10.1109/NORSIG.2006.275209.

[8] Y. Gong, F. Yang, L. Huang, and S. Su. "Model-Based Approach to Measuring Quality of Experience". In: *Emerging Network Intelligence, 2009 First International Conference on*. Oct. 2009, pp. 29–32. DOI: 10.1109/EMERGING.2009.17.

[9] H. J. Kim, D. H. Lee, J. M. Lee, K. H. Lee, W. Lyu, and S. G. Choi. "The QoE Evaluation Method through the QoS-QoE Correlation Model". In: *Networked Computing and Advanced Information Management, 2008. NCM '08. Fourth International Conference on*. Vol. 2. Sept. 2008, pp. 719–725. DOI: 10.1109/NCM.2008.202.

[10] G. Harman. “The Intrinsic Quality of Experience”. In: *Philosophical Perspectives* 4 (Jan. 1990). ArticleType: research-article / Issue Title: Action Theory and Philosophy of Mind / Full publication date: 1990 / Copyright © 1990 Ridgeview Publishing Company, pp. 31–52. ISSN: 1520-8583. DOI: 10.2307/2214186. URL: http://www.jstor.org/stable/2214186.

[11] M. Csikszentmihalyi and J. LeFevre. “Optimal experience in work and leisure”. In: *Journal of Personality and Social Psychology* 56.5 (May 1989). PMID: 2724069, pp. 815–822. ISSN: 0022-3514. URL: http://www.ncbi.nlm.nih.gov/pubmed/2724069.

[12] J. Dewey. *Experience and Education*. Collier Macmillan Publishers, 1938.

[13] ITU-T FG IPTV. *Focus Group on IPTV Standardization: Definition of Quality of Experience (QoE)*. International Telecommunications Union Telecommunication Sector Publications. Feb. 2007. URL: http://ties.itu.int/ftp/public/itu-t/fgiptv/readonly/Previous_Meetings/20070122_MountainView/il/T05-FG.IPTV-IL-0050-E.htm.

[14] Webopedia. *QoE*. URL: http://www.webopedia.com/TERM/Q/QoE.html.

[15] N. Committee. *Realizing the Information Future:The Internet and Beyond*. National Academy Press.Washington, D.C. USA: Computer Science and Telecommunications Board, 1994.

[16] S. W. Brim and B. E. Carpenter. *Middleboxes: Taxonomy and Issues*. [Online] http://tools.ietf.org/html/rfc3234. URL: http://tools.ietf.org/html/rfc3234.

[17] S. Akhshabi and C. Dovrolis. “The evolution of layered protocol stacks leads to an hourglass-shaped architecture”. In: *Proceedings of the ACM SIGCOMM 2011 Conference*. SIGCOMM ’11. Toronto, Ontario, Canada: Association for Computing Machinery, 2011, pp. 206–217. ISBN: 9781450307970. DOI: 10.1145/2018436.2018460. URL: https://doi.org/10.1145/2018436.2018460.

[18] C. Douglas. *Internetworking With TCP/IP Volume 1: Principles Protocols, and Architecture*. 5th ed. Prentice Hall, 2006. ISBN: 0-13-187671-6.

[19] ITU-T Y.1541. *Network performance objectives for IP-based services*. International Telecommunications Union Radiocommunication Sector. 2011-12.

[20] T. M. ONeil. *Quality of Experience and Quality of Service*. URL: http://hive2.hive.packetizer.com/users/h323forum/papers/polycom/QualityOfExperience+ServiceForIPVideo.pdf.

[21] S. Tourancheau, P. Le Callet, and D. Barba. “Image and Video Quality Assessment Using LCD: Comparisons with CRT Conditions”. In: *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.* E91-A.6 (June 2008), 1383□1391. ISSN: 0916-8508. DOI: 10.1093/ietfec/e91-a.6.1383.

[22] P. Margaret H. and W. Stephen. *The Impact of Monitor Resolution and Type on Subjective Video Quality Testing*. U.S. DEPARTMENT OF COMMERCE. Mar. 2004. URL: http://www.its.bldrdoc.gov/pub/ntia-rpt/04-412/ntia412.pdf.

[23] T. N. Minhas, O. Gonzalez Lagunas, P. Arlos, and M. Fiedler. "Mobile video sensitivity to packet loss and packet delay variation in terms of QoE". In: *2012 19th International Packet Video Workshop (PV)*. 2012, pp. 83–88. DOI: 10.1109/PV.2012.6229747.

[24] T. N. Minhas and M. Fiedler. "Impact of Disturbance Locations on Video Quality of Experience". In: *Quality of Experience for Multimedia Content Sharing, EuroITV2011*. Lisbon, Portugal, June 2011.

[25] The Linux Foundation. [Online] http://www.linuxfoundation.org/collaborate/workgroups/networking/netem. accessed June 2010.

[26] D. Vatolin, A. Parshin, and O. Petrov. *MSU Subjective Comparison of Modern Video Codecs*. Jan. 2006. URL: http://www.compression.ru/video/codec_comparison/pdf/msu_subjective_codecs_comparison_en.pdf.

# Paper II

## Mobile Video Sensitivity to Packet Loss and Packet Delay Variation in Terms of QoE

*Tahir Nawaz Minhas, Oziel Gonzalez Lagunas, Patrik Arlos, and Markus Fiedler. Mobile video sensitivity to packet loss and packet delay variation in terms of QoE. In Proceedings of 19th International Packet Video Workshop (PV), pages 83–88. IEEE, 2012.*

## 1 Abstract

With the growth of the mobile internet, the popularity of multimedia services and applications have increased rapidly. As a result, the end-users become quality-conscious. To fulfill the users' expectations, the study of quality of experience (QoE) has become very important for both researchers and service providers. This paper analyses the impact on perceived quality of received videos encoded with the H.264 baseline profile, which is suitable for mobile video. These encoded videos are streamed through an emulated network with packet loss and packet delay variation. To evaluate the video QoE, tests are conducted on a mobile device and on a laptop. The users' responses show that baseline profile of H.264 is very sensitive to packet loss and packet delay variation. Moreover, there is no considerable impact on users' perception depending on whether the test is conducted on the mobile device or on the laptop, playing videos with the same resolution.

**Categories and Subject Descriptors**
**{C.2.3}**Computer Communication NetworksNetwork Operations
**{H.5.1}**Multimedia Information SystemEvaluation/methodology, Video

**General Terms**
Human Factors, Measurement, Performance
**Key Words**
Quality of Experience, User Perception,Packet loss, Packet delay variation, H.264 baseline profile

## 2 Introduction

The use of applications on mobile phones and particularly on smart phones are increasing with rapidly growing mobile internet. Among other services the popularity of multimedia services and applications are increasing rapidly. Moreover, the capabilities of mobile phones in terms of processing power, memory, color display and other functionalities are improving continuously. As a result of this growth and exposure, the user expects good quality of experience. Real-time applications and multimedia services can effect user perception very quickly, as compared to other applications on Internet. Due to this consciousness of the users, service providers, application developers and research community are focusing on user quality of experience(QoE). Therefore to send the video across the network, several techniques for video compressions are in use [1], one of the most widely used codecs is H.264/ Advanced Video Coding (AVC) (Advance Video Coding). It is used with different applications and each has its own specific requirement, so it is divided into different profiles with respect to their usage [2]. For example, the mobile applications use the Baseline Profile (H.264 Baseline Profile (BP)) of H.264.

Now-a-days the mobile phones are used for many tasks beyond making the phone call [3]. Especially mobile videos are getting popular. Mobile video quality can be degraded over the internet due to network disturbances, such as packet loss and packet delay variation [4].

In this paper we study the QoE of mobile video for defined values of packet loss and delay variation for H.264 baseline profile. Moreover, we also investigate the role of video playing device (in this case mobile phone and laptop) on QoE for a fixed resolution video.

In order to achieve these research objectives, an experimental set-up was implemented to emulate the characteristics of a network, over which videos were transmitted. Subjective assessments were realised following the recommendations from the International Telecommunications Union (International Telecommunication Union (ITU)) [5, 6]. The results are calculated and presented using Mean Opinion Score (MOS) and statistical methods.

The remainder of this paper is organized as follows. Section 3 covers the related work and Section 4 describe the experiment setup and assessment methodology. The results are presented and discussed in Section 5 and finally, Section 8 concludes the paper.

## 3 Related Work

Studies have addressed the video quality perception for video and different codecs; Calyam et al. [7] made a comparative study of subjective and objective video quality for the codec H.323. They added disturbances of loss, delay and jitter to the same video sequence and found that jitter has the biggest effect. Claypool et al. [8] realised a subjective perception study for the codec MPEG-1 making variations of jitter and packet loss.

Previous work has also addressed studies employing the codec H.264, using subjective and objective methods. Jusmisko et al. [9] worked in a subjective analysis comparing different codecs H.263, H.264 and XviD for mobile devices. Lin et al. [10] presented a model for packet prioritization for different GOP structures of H.264 and MPEG-2. The article by Loguinov and Radha [11] analysed the behaviour of different parameters in video coded with MPEG-4. The Simone et al. [12] offers a database of H.264 videos for research community and made experiments with packet loss. Some studies that have addressed the quality of video for H.264, mention the profile used, for example [10] and [12] used the high profile, [13] used the main profile, Ries et al. [14] uses the baseline profile for quality estimation based in motion characteristics.

Furthermore, studies of video quality perception for mobile devices in some cases are carried out using a monitor or big LCD display instead of a mobile phone [15, 16]. Further, in the study done by Jumisko et al. [9], mobile phones were used but the phones were attached to a fixed stand, so the user was not able to manipulate the phone as it would be the case in real life scenarios.

People have different perceptions of quality for different media devices. Television and computers are usually at a fixed distance in contrast with hand-held devices on which distance and angle can be adjusted easily by the user. The perception changes according to the image size. Previous TV experiments carried out found that the general rule to be the bigger the image the better quality perceived [17].

Consequently, a study that tests the user perception of video quality on different network conditions, for packet loss and delay variation encoded with H.264 and considers characteristics for mobile devices, will contribute to understanding the expectations of the users with mobile devices. Further, an understanding of the role the device plays in the user perception of packet loss and delay variation is required.

# 4 Experiment Setup

An emulation setup shown in Figure 1 is used to carry out the experiments. The streaming Server (S) is responsible to send the original encoded video sequences through Shaper (TS) to the Player (P), using UDP protocol. A full duplex link of bandwidth 100 Mbps is used from S to D. For streaming, VLC is used while, for delay variation and packet loss shaping the Network Emulator (NetEm) is used, as it provides the best delay shaping as compared to other shapers [18].

Computer S is a Toshiba laptop with an Intel Core Duo T2450 2.00GHz processor architecture running Ubuntu 10.04 LTS Linux, while computer P is a MacBook Pro with an Intel CPU Core 2 Duo 2.40 GHz, running MacOS X 10.5. Both computers have the VLC media player version 1.1.3. The shaper computer has an Intel CPU 2.66 GHz with Linux kernel 2.6.23.9. It is important to mention that after the version 2.6.15 of the Linux kernel, NetEm does re-ordering of packets if a lot of jitter is emulated. In this experiment, we leave packets re-ordering without change as configured by default.

We captured the video traffic before and after the shaper using the Distributed Passive Measurement Infrastructure [19] to verify the shaped delay, delay variation and packet loss.

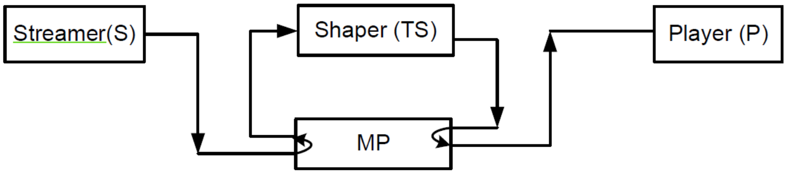



**Figure 1:** Experiment Setup

The setting used for the shaping of delay ($D$) and varying delay ($\Delta D$) were (150 $\pm$ $\{0, 2, 4, 8, 12, 16\}$) ms with normal distribution. The ITU standard G.114, one-way transmission time, states that the delay should not be more than 150 ms and jitter should not be more than 20 ms to 50 ms [20]. For this study, the maximum packet loss value is $7\%$, as samples taken in the laboratory showed that for packet loss greater than this levels, the video content was lost.

Video quality assessments can be objective and subjective. In subjective quality assessment, the description of quality is based on human perception and can be different between subjects [7]. Human perception involves various aspects of human psychology and viewing conditions; such as vision ability, lighting conditions, preference for content, displaying devices, understanding of the rating criteria [21] and Absolute Category Rating (ACR) is consider to be well-suited for qualification tests [5].

Another option considered was the software called *Perceptual Evaluation Video Quality* (PEVQ) [22]. It is a software approved by the VQEG and ITU-T that makes an objective analysis of the video and gives a correlation with the subjective MOS. But in [23] the authors questioned the performance of PEVQ for particular cases as compared to the user's rating.

ACR is the method recommended for qualification tests and applied in several studies [12, 13, 24].

The videos were taken from a commonly used repository for the purpose of video quality assessment studies [25] as suggested by The Simone et al. [12]. The selected videos used for experiments are Football, Foreman and Hall-monitor.

The length of each video sequences is between eight and ten seconds. This is the length suggested by the ITU-T [5].

The video sequences used for this study are encoded with H.264 Baseline Profile, level 1.3 because it is recommended for mobile videos. Each sequence has frame rate 30 fps and bit-rate 768 kbps. Previous video studies for mobile have used the QCIF ($176 \times 144$) resolution [9, 14, 15]; in some cases using mobile devices and some others displaying the video on a computer monitor displaying the QCIF resolution. For this study was decided to use higher resolution Quarter Video Graphics Array (QVGA) ($320 \times 240$) to cover the new segment of modern phones.

*Assessment Methodology*

The assessment sessions were held in rooms conforming to the specifications of the ITU-T [5] at the premises of two academic institutions in Kristianstad and Karlskrona in Sweden.

Two mobile devices were used for the assessment. A laptop Toshiba Satellite with an Intel Core Duo T2450 2.00GHz with a LCD display of 15.4 inches (diagonal) widescreen TruBrite TFT, resolution $(1280 \times 800)$, and an iPhone 3G. The media player VLC version 1.1.3 were used on the laptop and the standard media player from Apple Inc on the iPhone Mobile Phone.

For the mobile phone the user was able to choose the angle and distance to watch the videos. This imitates what happens in a real-world environment where the user can move the device as desired to achieve a comfortable viewing position.

The video sequences shown on the laptop complied with the recommendations made by the ITU-T [5]. The videos were played in VLC using the original resolution of the video sequence $(320 \times 240)$, in the center of the screen, with the background set to fifty percent grey.

# 5 Results

This section presents the results collected from the assessment sessions. During the assessment sessions, 34 volunteers graded the 33 videos sequences using the scale Excellent (5), Good (4), Fair (3), Poor (2) and Bad (1). Each video was shown on both a laptop and on a mobile phone. In total 66 video sequences were shown to the subjects.

## 5.1 Packet Loss

Figures 2 and 3 show the graphs of the MOS for the packet loss for the laptop and the mobile phone, respectively. The graphs illustrate that the smaller the values of loss the higher MOS, thus, better user perception of the video. As packet loss increases, the MOS drops steeply, but it remains “fair” until values of packet loss between $0.4\%$ and $0.8\%$. From $5\%$ packet loss the value of the MOS is close to 1 (bad). This indicates that at and above this value the user perception quality makes them reject the video. This behaviour suggest that packet loss is an important factor to consider. If this video is provided as a service will lead the user to stop watching the video.

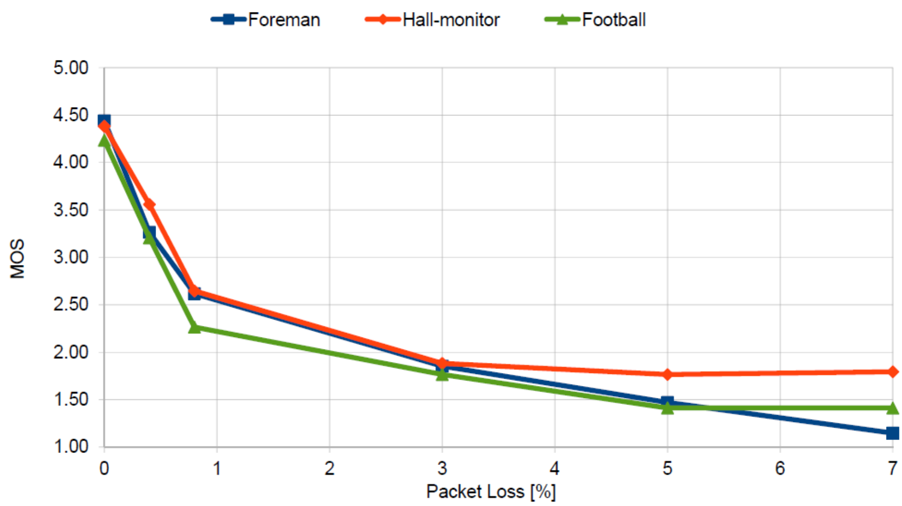



**Figure 2:** MOS for packet loss displayed on the laptop

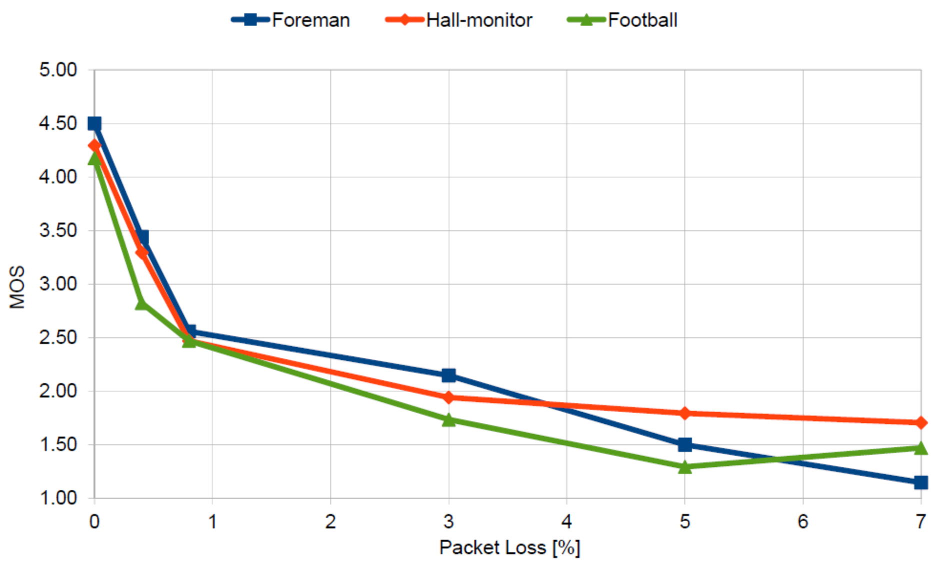



**Figure 3:** MOS for packet loss displayed on the mobile phone

The video sequence encoded with H.264 and without distortions was graded with a MOS that lies between 4.50 and 4.18 although there are no disturbances. This shows that the viewers in some cases are reluctant to score the videos as "excellent." This behaviour was also found in other studies [13].

For the different videos affected with packet loss it is noticed that the video with the lowest MOS, in both cases, is the Football video. For the video Hall-monitor in both devices the MOS decline to poor (2) and then it seems that there is not a big variation between 3% and 7%. The video sequence Foreman is the video with the lower MOS at 7%.

The values calculated for the confidence interval for packet loss are shown in the Figure 4. The x-axis shows the video sequence (FM, Foreman; HM, Hall-monitor; FT, Football), the device that was used to present the video sequence to the observer (L, Laptop; P, Mobile phone) and the percentage of packet loss emulated. The y-axis refers to the MOS. We can see that there is not a big difference within the three different video sequences for each of the packet loss values. The tendency shows that for bigger values of packet loss the rate degrades for all the video sequences. In general, for values above 5% of packet loss the observers rate is close to 1 (bad). Slightly higher MOS values are shown for the Hall-monitor video sequence in both devices for values of packet loss of 5% and 7%. However, the assessment value is close to 1 (bad).

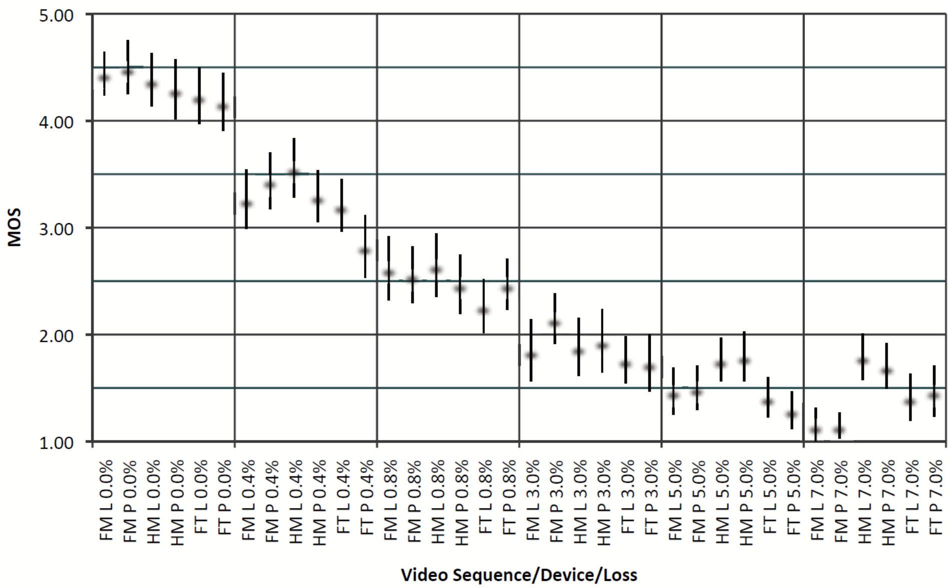



**Figure 4:** Confidence interval (95%) for packet loss

## 5.2 Packet Delay Variation

The delay variation results were plotted in Figures 5 and 6. It is shown that for a small increase delay variation the MOS decreases dramatically. For a variation of only $\pm 2$ ms the MOS for the foreman and hall-monitor videos decays to a value close to poor. The video sequence football is less affected by the delay variation for values between $\pm 2$ ms and $\pm 8$ ms. However, for bigger variations the MOS is similar for the three video sequences. Delay variation above $\pm 8$ ms for all the video sequences leads to a MOS close to bad (1) indicating a reject of the video from the observers. In the Figure 6 we can see that for the football sequence the MOS does not vary from $\pm 2$ ms to $\pm 4$ ms.

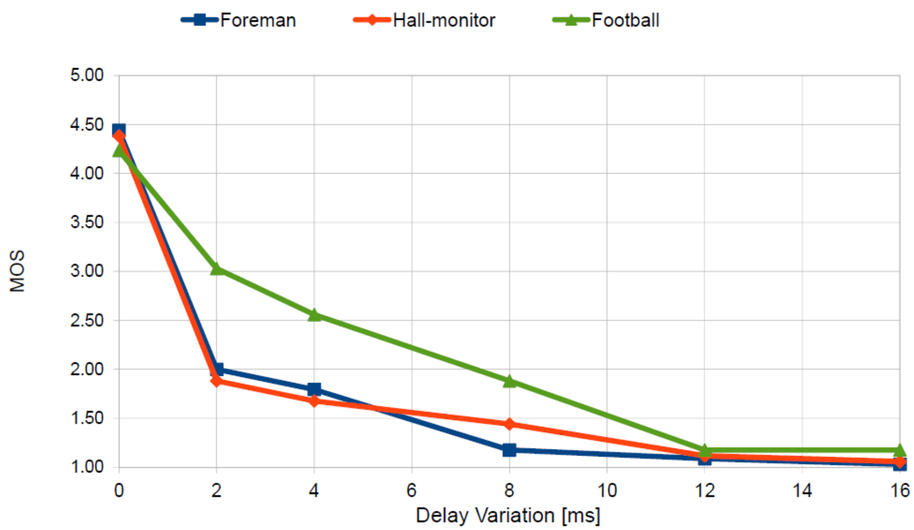



**Figure 5:** MOS for delay variation displayed on the laptop

The Figure 7 presents the confidence interval values for delay variation. The x-axis shows the video sequence, the device that was used to present the video sequence to the observer and the percentage of packet loss emulated, as explained with Figure 4. The y-axis refers to the MOS. It is appreciated that the football video sequence gets higher rates between $\pm 2$ ms and $\pm 8$ ms, the higher rates become less obvious for the values of $\pm 12$ ms and $\pm 16$ ms.

Table 1 shows the regression analysis of packet loss and packet delay variation shown in figures 2, 3, 5 and 6. The second-last column shows the power equation for every individual case, and the last column shows the R squared values. The power regression model fits for all cases as the R squared is alway greater the 0.9, it outperformed matching using linear, logarithmic and exponential models. The MOS values for zero loss ratio and zero packet delay variation could not be considered for

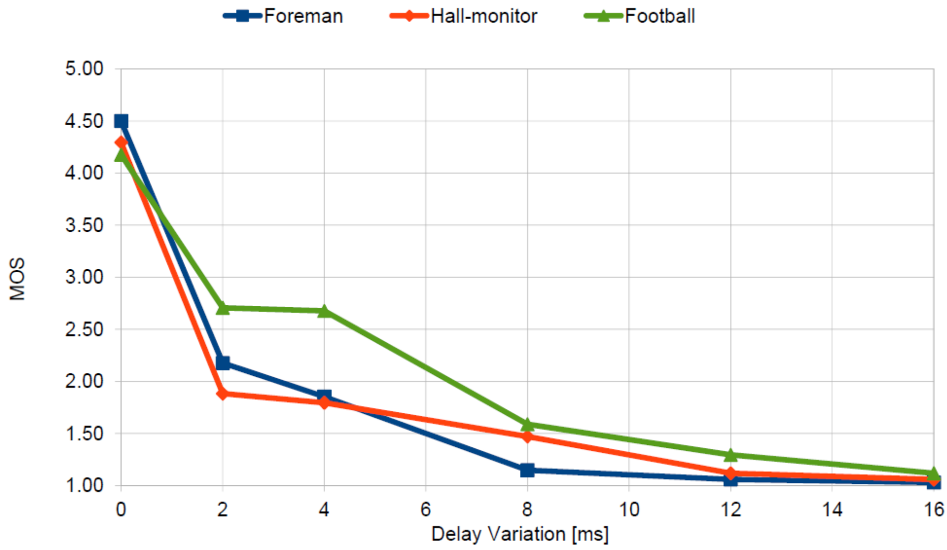



**Figure 6:** MOS for delay variation displayed on the mobile phone

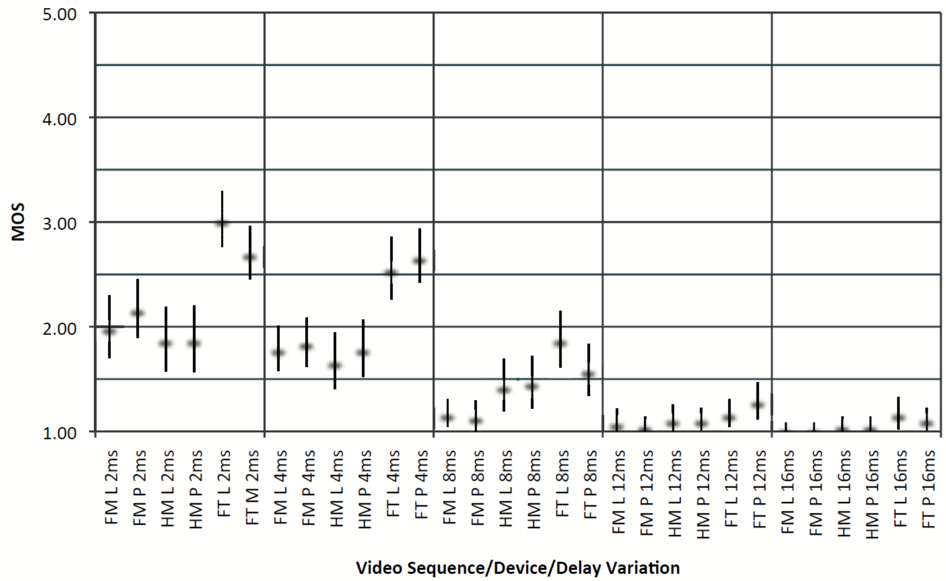



**Figure 7:** Confidence interval (95%) for delay variation

the power regression model.

Despite not adding both disturbances to the same video, packet loss and delay variation, we can notice that it is likely that the perception of quality in the video degrades more for small changes in the delay variation. This is similar to the findings

| Parameter | Video | Device | Equation | $R^2$ |
|---|---|---|---|---|
| Loss Ratio ($L$) | Foreman | Laptop | $0.511L^{-0.340}$ | 0.972 |
| | | Mobile | $0.536L^{-0.338}$ | 0.913 |
| | Hall-Monitor | Laptop | $0.863L^{-0.244}$ | 0.945 |
| | | Mobile | $0.926L^{-0.219}$ | 0.963 |
| | Football | Laptop | $0.643L^{-0.279}$ | 0.964 |
| | | Mobile | $0.668L^{-0.264}$ | 0.940 |
| Delay Variation ($J$) | Foreman | Laptop | $2.478J^{-0.342}$ | 0.938 |
| | | Mobile | $2.632J^{-0.268}$ | 0.923 |
| | Hall-Monitor | Laptop | $2.273J^{-0.282}$ | 0.944 |
| | | Mobile | $2.206J^{-0.265}$ | 0.924 |
| | Football | Laptop | $4.552J^{-0.504}$ | 0.943 |
| | | Mobile | $4.247J^{-0.481}$ | 0.931 |

**Table 1:** MOS Trend Fitting of Packet Loss and Delay Variation

of the study performed by Calyam et al. [7]. Although their study was for the codec H.263 and was adding different disturbances to the same video sequence.

## 5.3 Mobile Phone and Laptop

The MOS for the packet loss and delay variation was plotted for both devices, the laptop and mobile phone in Figure 8 and Figure 9 respetively. For both graphs we can notice that the results are almost the same for both of the devices. The biggest difference is found for packet loss of $0.4\%$ that in the laptop obtained a MOS of $3.19$ and in the mobile phone a slightly higher rate of $3.34$.

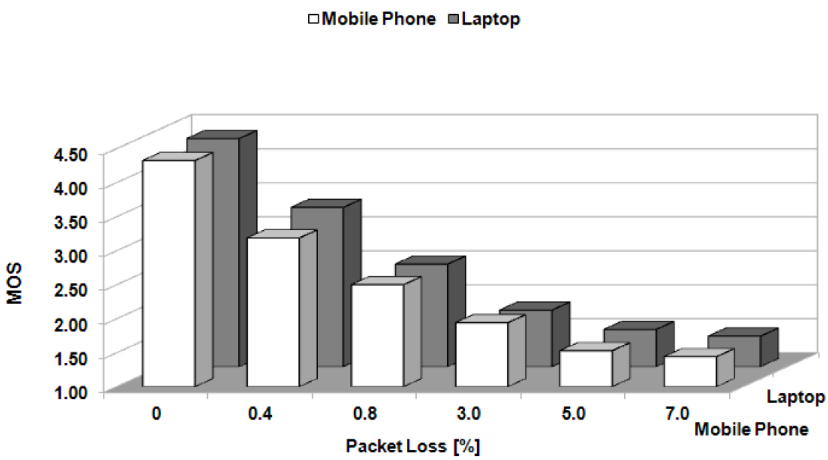



**Figure 8:** MOS packet loss displayed on the laptop and mobile phone

To identify if these variations are significant we applied matched-sample $t$-test. This test is used when we have two matched or correlated samples, and when we need to test the difference between two averages [26]. In this specific case we want to compare the mean between the assessment rates obtained for the laptop and the one obtained when the observers used the mobile phone as displaying device.

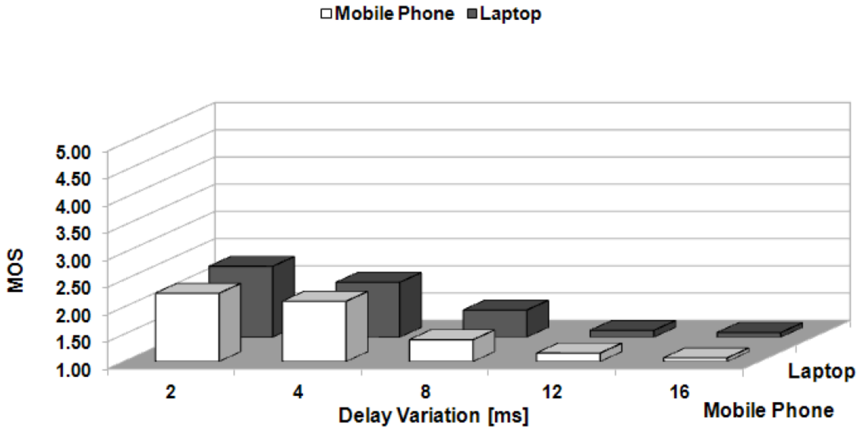



**Figure 9:** MOS delay variation displayed on the laptop and mobile phone

Let us consider that $\bar{u}_L$ is the average for the laptop, $\bar{u}_P$ is the average for the mobile phone. Thus the hypothesis is if $\bar{u}_L = \bar{u}_P$ is true, then there is no difference between doing the experiment on the mobile phone or on the laptop when using the same resolution in both devices.

If we consider the difference of the means $\bar{u}_D = \bar{u}_L - \bar{u}_P$, then the hypothesis can also be expressed as:

$$H : \bar{u}_D = \bar{u}_L - \bar{u}_P = 0 \tag{II.1}$$

This test needs the calculation of the $t_{calc}$ parameter and then compare it with the $t_{table}$ that is obtained from the Student-$t$ distribution table [26]. The $t_{table}$ value depends on the level of significance, in this case a $95\%$ confidence interval is used. Thus, the level of significance is $0.05$, the degrees of freedom (df) is defined as the number of observers minus one (*N-1*) is 33. The value from the table to compare in our case corresponds to $t_{table}$ = 2.042.

The value $t_{calc}$ is defined as:

$$t_{calc} = \frac{\bar{D} - \bar{u}_D}{\frac{s_D}{\sqrt{N}}} \tag{II.2}$$

where $\bar{D}$ and $s_D$ are the mean and standard deviation of the difference scores (difference between the data obtained for the laptop and mobile phone); $\bar{u}_D$ in this case is zero as our hypothesis and $N$ is the number of difference scores (number of pairs).

If we find that $t_{calc} < t_{table}$ we fail to reject the hypothesis, so there is no evidence to suggest that the different devices affect the perception of the shown video.

Table 2 presents the results for the calculations for each of the video pairs, referring to the same video sequence presented on the laptop and on the mobile phone. DV stands for delay variation in the table. The comparison of $t_{calc}$ with $t_{table}$ reveals that for the most of the video sequences, there is no evidence to suggest that the different devices affect the perception of the shown video.

| Foreman | | | Hall-Monitor | | | Football | | |
|---|---|---|---|---|---|---|---|---|
| $L$ | $t_{calc}$ | Hypothesis | $L$ | $t_{calc}$ | Hypothesis | $L$ | $t_{calc}$ | Hypothesis |
| $0.0\%$ | $-0.63$ | Fail to reject | $0.0\%$ | $0.90$ | Fail to reject | $0.0\%$ | $0.42$ | Fail to reject |
| $0.4\%$ | $-1.14$ | Fail to reject | $0.4\%$ | $2.50$ | Reject | $0.4\%$ | $3.20$ | Reject |
| $0.8\%$ | $0.42$ | Fail to reject | $0.8\%$ | $1.36$ | Fail to reject | $0.8\%$ | $-2.03$ | Fail to reject |
| $3.0\%$ | $-2.05$ | Fail to reject | $3.0\%$ | $-0.44$ | Fail to reject | $3.0\%$ | $0.25$ | Fail to reject |
| $5.0\%$ | $-0.25$ | Fail to reject | $5.0\%$ | $-0.27$ | Fail to reject | $5.0\%$ | $1.16$ | Fail to reject |
| $7.0\%$ | $0.00$ | Fail to reject | $7.0\%$ | $0.83$ | Fail to reject | $7.0\%$ | $-0.63$ | Fail to reject |
| $D \pm \Delta D$ | $t_{calc}$ | Hypothesis | $D \pm \Delta D$ | $t_{calc}$ | Hypothesis | $D \pm \Delta D$ | $t_{calc}$ | Hypothesis |
| $150 \pm 2$ ms | $-1.36$ | Fail to reject | $150 \pm 2$ ms | $0.00$ | Fail to reject | $150 \pm 2$ ms | $2.46$ | Reject |
| $150 \pm 4$ ms | $0.70$ | Fail to reject | $150 \pm 4$ ms | $-1.28$ | Fail to reject | $150 \pm 4$ ms | $-0.89$ | Fail to reject |
| $150 \pm 8$ ms | $0.44$ | Fail to reject | $150 \pm 8$ ms | $-0.37$ | Fail to reject | $150 \pm 8$ ms | $2.73$ | Reject |
| $150 \pm 12$ ms | $0.44$ | Fail to reject | $150 \pm 12$ ms | $0.00$ | Fail to reject | $150 \pm 12$ ms | $-1.68$ | Fail to reject |
| $150 \pm 16$ ms | $0.00$ | Fail to reject | $150 \pm 16$ ms | $0.00$ | Fail to reject | $150 \pm 16$ ms | $0.81$ | Fail to reject |

**Table 2:** Matched-sample *t* test results, with $t_{calc} = 2.042$

Therefore, we can conclude that there is not a significant difference between doing the experiment in the mobile phone and in the laptop displaying the same resolution. Overall very similar rating behavior was observed for both devices.

# 6 Conclusion

In this paper we investigated the mobile video quality of experience, encoded with the H.264 baseline profile. The network performance parameter plays an important role on the quality of streamed video. In this study we focused on the effect of packet loss and packet delay variation on mobile video quality of experience. We found that mobile video encoded with the H. 264 baseline profile is very sensitive to the network disturbance and user's rating dropped quickly with nominal packet loss and packet delay variation. Moreover the users rating survey is conducted on mobile phone as well as on the laptop. Results for both devices were compared using matched-sample t-test. We did not find any strong evidence to conclude that devices have any impact on user perception for video which has the same resolution.

# References


[1] D. Wu, Y. Hou, W. Zhu, Y.-Q. Zhang, and J. Peha. “Streaming video over the Internet: Approaches and directions”. In: *IEEE Transactions on Circuits and Systems for Video Technology* 11.3 (2001), pp. 282–300. ISSN: 1051-8215. DOI: 10.1109/76.911156.

[2] ITU-T REC-H.264. *H.264 : Advanced video coding for generic audiovisual services*. International Telecommunications Union Telecommunication Sector Publications. Mar. 2010. URL: http://www.itu.int/rec/T-REC-H.264-201003-I/en.

[3] L. Bouchard. “Multimedia Software for Mobile Phones”. In: *Software, IEEE* 27.3 (May 2010), pp. 8–10. ISSN: 0740-7459. DOI: 10.1109/MS.2010.78.

[4] J. G. Apostolopoulos, W.-t. Tan, and S. J. Wee. *Video Streaming: Concepts, Algorithms, and Systems*. Tech. rep. HP Laboratories, 2002, 2002.

[5] ITU-T P.910. *Subjective video quality assessment methods for multimedia applications*. International Telecommunications Union Telecommunication Sector. 1999.

[6] ITU-R BT.500-11. *Methodology for the subjective assessment of the quality of television pictures*. International Telecommunications Union Radiocommunication Sector. 2002.

[7] P. Calyam, M. Sridharan, W. Mandrawa, and P. Schopis. “Performance Measurement and Analysis of H.323 Traffic”. In: *Passive and Active Network Measurement*. Ed. by C. Barakat and I. Pratt. Vol. 3015. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2004, pp. 137–146.

[8] M. Claypool and J. Tanner. “The effects of jitter on the peceptual quality of video”. In: *MULTIMEDIA ’99: Proceedings of the seventh ACM international conference on Multimedia (Part 2)*. Orlando, Florida, United States: ACM, 1999, pp. 115–118. ISBN: 1-58113-239-5. DOI: http://doi.acm.org/10.1145/319878.319909.

[9] S. Jumisko-Pyykkö and J. Häkkinen. “Evaluation of subjective video quality of mobile devices”. In: *MULTIMEDIA ’05: Proceedings of the 13th annual ACM international conference on Multimedia*. Hilton, Singapore, 2005, pp. 535–538. ISBN: 1-59593-044-2. DOI: http://doi.acm.org/10.1145/1101149.1101270.

[10] T.-L. Lin, S. Kanumuri, Y. Zhi, D. Poole, P. Cosman, and A. Reibman. “A Versatile Model for Packet Loss Visibility and its Application to Packet Prioritization”. In: *IEEE Transactions on Image Processing* 19.3 (2010), pp. 722–735. ISSN: 1057-7149. DOI: 10.1109/TIP.2009.2038834.

[11] D. Loguinov and H. Radha. "Measurement study of low-bitrate internet video streaming". In: *Proceedings of the 1st ACM SIGCOMM Workshop on Internet Measurement*. San Francisco, California, USA, 2001, pp. 281–293. ISBN: 1-58113-435-5. DOI: `http://doi.acm.org/10.1145/505202.505238`.

[12] F. De Simone, M. Tagliasacchi, M. Naccari, S. Tubaro, and T. Ebrahimi. "A H.264/AVC video database for the evaluation of quality metrics". In: *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*. 2010, pp. 2430–2433. DOI: `10.1109/ICASSP.2010.5496296`.

[13] M. Pinson, S. Wolf, and G. Cermak. "HDTV Subjective Quality of H.264 vs. MPEG-2, With and Without Packet Loss". In: *IEEE Transactions on Broadcasting* 56.1 (2010), pp. 86–91. ISSN: 0018-9316. DOI: `10.1109/TBC.2009.2034511`.

[14] M. Ries, O. Nemethova, and M. Rupp. "Video quality estimation for mobile H. 264/AVC video streaming". In: *Journal of Communications* 3.1 (2008), pp. 41–50.

[15] V. Vassiliou, P. Antoniou, I. Giannakou, and A. Pitsillides. "Requirements for the Transmission of Streaming Video in Mobile Wireless Networks". In: *Artificial Neural Networks – ICANN 2006*. Ed. by S. Kollias, A. Stafylopatis, W. Duch, and E. Oja. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 528–537. ISBN: 978-3-540-38873-9.

[16] S. Winkler and F. Dufaux. "Video quality evaluation for mobile applications". In: *Proceedings of SPIE Conference on Visual Communications and Image Processing*. Vol. 5150. 2003, pp. 593–603.

[17] H. Knoche, J. D. McCarthy, and M. A. Sasse. "Can small be beautiful?: Assessing image resolution requirements for mobile TV". In: *MULTIMEDIA '05: Proceedings of the 13th annual ACM international conference on Multimedia*. Hilton, Singapore, 2005, pp. 829–838. ISBN: 1-59593-044-2. DOI: `http://doi.acm.org/10.1145/1101149.1101331`.

[18] J. Shaikh, T. Minhas, P. Arlos, and M. Fiedler. "Evaluation of delay performance of traffic shapers". In: *Second International Workshop on Security and Communication Networks*. 2010, pp. 1–8. DOI: `10.1109/IWSCN.2010.5497994`.

[19] P. Arlos, M. Fiedler, and Arne A. Nilsson. "A Distributed Passive Measurement Infrastructure". In: *Proceedings of Passive and Active Measurement Workshop*. Boston, US, 2005, pp. 215–227.

[20] B. Kelly. *Quality of Service In Internet Protocol (IP) Networks*. Wainhouse Research. 2002.

[21] Z. Wang, H. R. Sheikh, and A. C. Bovik. "Objective video quality assessment". In: *In the handbook of video databases: Design and Applications*. CRC Press, 2003, pp. 1041–1078.

[22] OPTICOM. *Perceptual Evaluation of Video Quality*. http://www.pevq.org/. accessed July 2010.

[23] T. N. Minhas and M. Fiedler. "Impact of Disturbance Locations on Video Quality of Experience". In: *Quality of Experience for Multimedia Content Sharing, EuroITV2011*. Lisbon, Portugal, June 2011.

[24] S. Spirou. "Packet Reordering Effects on the Subjective Quality of Broadband Digital Television". In: *2006 IEEE International Symposium on Consumer Electronics*. 2006, pp. 1–6. DOI: `10.1109/ISCE.2006.1689474`.

[25] *Xiph.org Test Media*. [Online] http://media.xiph.org/video/derf/. accessed July 2010.

[26] D. C. Howell. *Statistical Methods for Psychology*. 7th ed. Wadsworth, 2010.

# Paper III

## Impact of Disturbance Locations on Video Quality of Experience

*Tahir Nawaz Minhas and Markus Fiedler. Impact of disturbance locations on video quality of experience. In Proceeding of EuroITV 2011 Workshop: Quality of Experience for Multimedia Content Sharing, 2011.*

## 1 Abstract

Quality of experience is getting the attention of the research community as well as the industry. In case of real-time streaming, packet loss, delay and jitter degrades the video quality. The player buffer can be emptied due to long delay, which freezes video at playout, while resuming the streaming video content causes jumps to current location. There is the question, how user is going to react to such kind of artifacts with respect to location where they arise. We collected user ratings for videos showing the artifacts due to delay variation as well as freezes and jumps at different locations. We also verified these results with Perceptual Evaluation of Video Quality (PEVQ) application. For delay and delay variation case, the PEVQ results are aligning to the human rating, but both differ in freeze-and-jump case. The users' responses in case of freeze-and-jump shows interesting results with respect to location. **Categories and Subject Descriptors**
**{C.2.3}**Computer Communication NetworksNetwork Operations
**{H.5.1}**Multimedia Information SystemEvaluation/methodology, Video

**General Terms**
Human Factors, Measurement
**Key Words**
Quality of Experience, User Perception, PEVQ

## 2 Introduction

To win the user's loyalty by providing the good services, service provider use to measure the quality of services with different methods. Amongst others the concept of Quality of Experience (QoE) is emerging quickly and getting attention of service providers and the research community. In case of video streaming, QoE can be influenced by video transcoding and transmission. Real-time video transmission over the Internet may suffer with degraded quality due to packet loss, delay and delay variation. Moreover, if by any reason (due to long delay, packet loss or key frame loss) the receiver stops receiving the stream and the player empties the play-out buffer, then freezing will occur until the stream is resumed. For real-time streaming, the user will miss the content of video for that time period because the video will jump from stop position to resumed position. This will influence the video quality and thus also the user perception. The video quality is assessed by using either objective or subjective method. Subjective quality depends on various factors based on human psychology and viewing conditions, such as observer vision ability, translation of quality perception into ranking score, preference for content, adaptation, display devices, ambient light levels etc. [1]. Different studies have been undertaken for video quality of experience in packet networks [2] [3] [4]. In [5], authors investigated the effect of frozen and skipped frames on video quality. The mean opinion score (MOS) has been considered the most reliable subjective quality measurement method, however it is time consuming and inconvenient. Alternative algorithms have been implemented and approved by VQEG [6] and ITU-T. Basically these softwares analyze the objective metrics of video quality and correlate them with subjective quality MOS. The Perceptual Evaluation of Video Quality (PEVQ) [7] model is recommended by ITU-T (J.247) in the category of objective perceptual multimedia video quality measurement in the presence of a full reference [8]. In this work we evaluate the PEVQ with reference to user perception for different jitter conditions. Also, we test the PEVQ for a fixed duration of freeze-and-jump of a video sequence at different locations and compare the results with users ratings.

The remainder of this paper is organized as follows. Section 3 introduce the perceptual evaluation of video quality (PEVQ). In Section 4 we describe the experiment setup and we discuss freezes and jumps of video sequences with respect to different location of sequence. The results are presented and discussed in Section 5 and 6. In Section 7, we conclude the paper.

# 3 Perceptual Evaluation of Video Quality (PEVQ)

PEVQ is provided by OPTICOM, is part of the PEXQ software suite and recommended by ITU for perceptual multimedia video quality measurement. It measures degradations due to network by analyzing the degraded video. This model measures the Quality of Experience (QoE) based on modeling the behavior of the human [7]. It also quantifies the other video quality parameters like PSNR, distortion indicator and lip-sync delay. PEVQ is built on PVQM and designed for mobile applications and multimedia applications. For video quality detection, it is based on five indicators that are motivated by the human visual system (HVS). These indicators operate in temporal, spatial, luminance and chrominance domains [8]. The results of these indicators are incorporated and integrated in order to derive the MOS [7].

# 4 Experiments

To study the impact of delay and delay variation on video, the experimental setup shown in figure 1 is used. It consists of a video streamer, video player, shaper and measurement point (MP). For streaming and playing VLC [9] is used, while for delay and variable delay shaping the Network Emulator ”NetEm” [10] is used, as it provides the best delay shaping as compared to other shapers [11].We captured the video traffic before and behind the shaper using the Distributed Passive Measurement Infrastructure [12] based on DAG cards [13] to verify the shaped delay and delay variation. Streamer and player were installed on Microsoft Window XP while the shaper ran on Linux. The delay ($D$) and variable delay ($\Delta D$) settings used for these experiments are $D \pm \Delta D$=100 ms$\pm\{$0 ms, 2 ms, 4 ms, 6 ms, 8 ms, 10 ms, 12 ms, 14 ms, 16 ms$\}$.

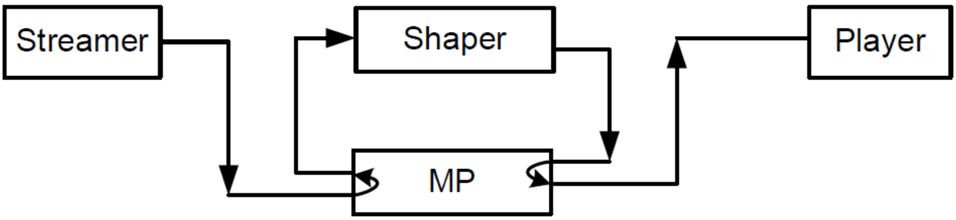



**Figure 1:** Experiment Setup

Figure 2 presents a view of the video sequences that were created to study the user perception with respect to a sequence of one-second freeze followed by a jump of video on different locations. The video sequences used for the experiments have 25 fps. The first test sequence starts with freezed frame number one for one second and then continues playing from 25$^{\text{th}}$ frame to the end of video. In the 2$^{\text{nd}}$ test sequence the video plays for one second, then the video freezes at 25$^{\text{th}}$ frame awaiting one second duration and then jumps to the 50$^{\text{th}}$ frame, from where it plays till the end. Similarly the process will continue till the last second of the video sequence, i.e we move the freeze-and-jump through the video second-by-second.

User perception tests were conducted on campus, most of the participants are undergraduate and graduate students. Moreover the tests were conducted according to the recommendation of ITU-R Rec. BT. 500-11 [14] and ITU-T P.910 [15] using absolute category rating (ACR).

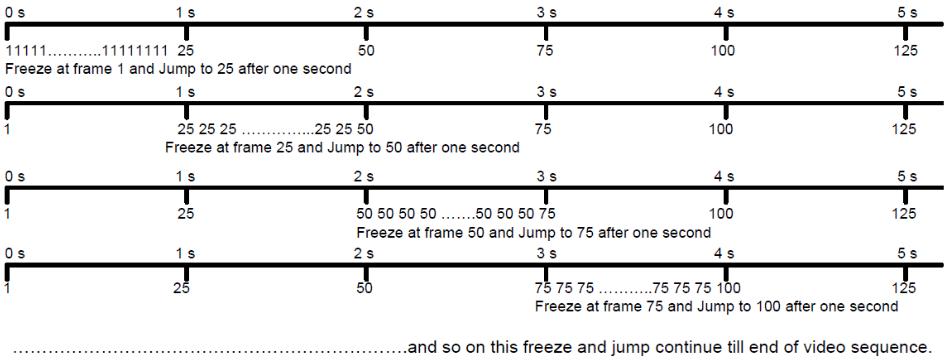



**Figure 2:** Freeze and Jump Variation on Video Sequence

# 5 Delay/Delay Variation

Figure 3 depicts the users' perception for "$Foreman$" video. The x-axis shows the delay variation $\Delta D$ in milliseconds around a fixed delay $D$=100 ms. The MOS is shown along the y-axis. On the x-axis zero means that video is received with a fixed delay of 100 ms, in this case MOS is 3.72$\pm$0.247. There is no remarkable change in MoS in case of 100 ms and 100$\pm$2 ms but after that with the increase of $\Delta D$, MOS decreases linearly till 100$\pm$8 ms and enters the BAD perception region at $\Delta D = 4$ ms. Similarly, Figure 4 shows the users' feedback for "$Football$" video. Here, in the first case with zero $\Delta D$ the MOS is higher as compared to $Foreman$ case. MOS is decreasing linearly with the rising delay variation till 100$\pm$12 ms and signals BAD perception for $\Delta D > 6$ ms. The $Football$ video shows more resistance to delay variation as compared to $Foreman$ because it is a fast moving video that makes it difficult for a user to notice quick changes and disturbance.

Those two video sequences are also analysed with PEVQ by OPTICOM [7]. The original video is used as reference for 100$\pm$0 ms case and later all other videos are ranked with reference to 100$\pm$0 ms video. Figures 3 and 4 shows the result along the results discussed above. From the figures, it is easy to see that the PEVQ results are very much in agreement with the user ratings.

As we saw in figures 3 and 4 the PEVQ and users' ratings are very much aligned with each other, which can also be verified numerically from the table 1. We can see that the difference between user rating and PEVQ is small. However after $\Delta D >$ 10 ms the PEVQ ratings remain greater than or equal to 1.5, while user ratings approach to one for $\Delta D \geq 14$ ms.

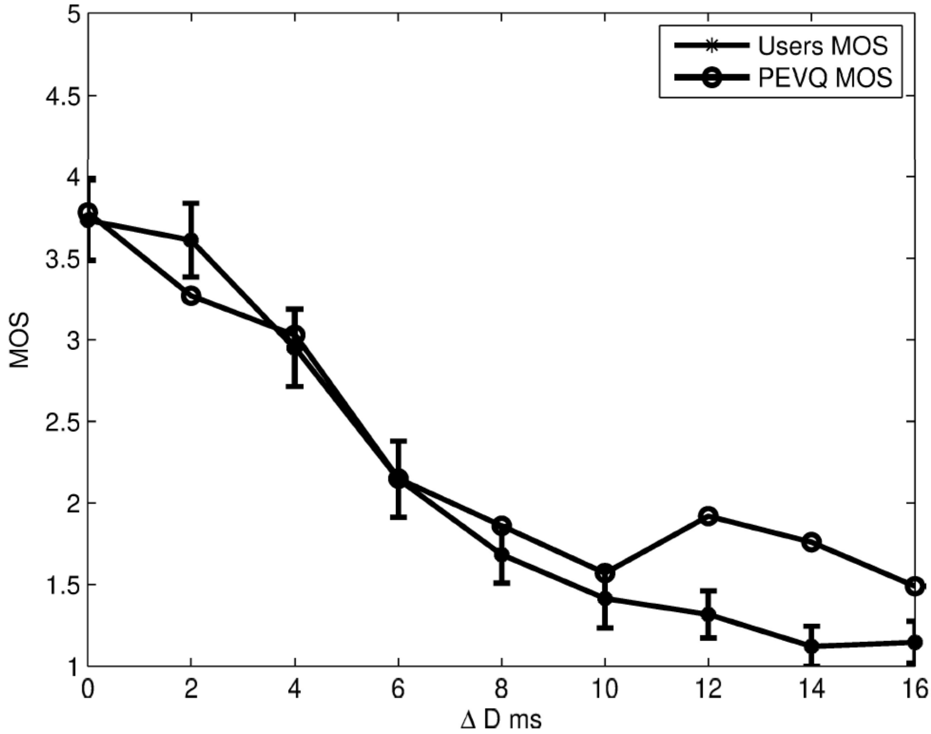



**Figure 3:** PEVQ MOS and Users'MOS rating for $Foreman$ Video

**Table 1:** PEVQ MOS and Users' MOS for Foreman and Football Video

| Delay | Foreman | | Football | |
|---|---|---|---|---|
| $D\pm\Delta D$ | PEVQ | User MOS | PEVQ | User MOS |
| 100± 0 | 3.78 | 3.73± 0.247 | 5.00 | 4.71± 0.186 |
| 100± 2 | 3.27 | 3.61± 0.226 | 4.04 | 4.25± 0.177 |
| 100± 4 | 3.03 | 2.95± 0.237 | 3.89 | 3.79± 0.204 |
| 100± 6 | 2.15 | 2.15± 0.233 | 3.16 | 3.25± 0.177 |
| 100± 8 | 1.86 | 1.68± 0.174 | 2.22 | 2.50± 0.264 |
| 100± 10 | 1.57 | 1.41± 0.181 | 1.69 | 1.67± 0.226 |
| 100± 12 | 1.92 | 1.32± 0.144 | 1.56 | 1.50± 0.264 |
| 100± 14 | 1.76 | 1.12± 0.122 | 1.70 | 1.08± 0.113 |
| 100± 16 | 1.49 | 1.15± 0.129 | 1.61 | 1.08± 0.113 |

# 6 Freeze-and-Jump

As shown in Figure 2 we focus on special cases, in which a one-second freeze is followed by a jump of video moves along the whole sequence. For this test we selected the $Foreman$, $News$ and $HallMonitor$ videos. The corresponding users' perceptions are shown in figures 5, 6 and 7. The MOS is shown along the y-axis, whereas frame numbers are shown along the x-axis. The data shown for $x = 0$ corresponds to the original videos of 10 seconds length, whereas all other videos have nine seconds actual video plus one second freeze. This one second freeze-and-

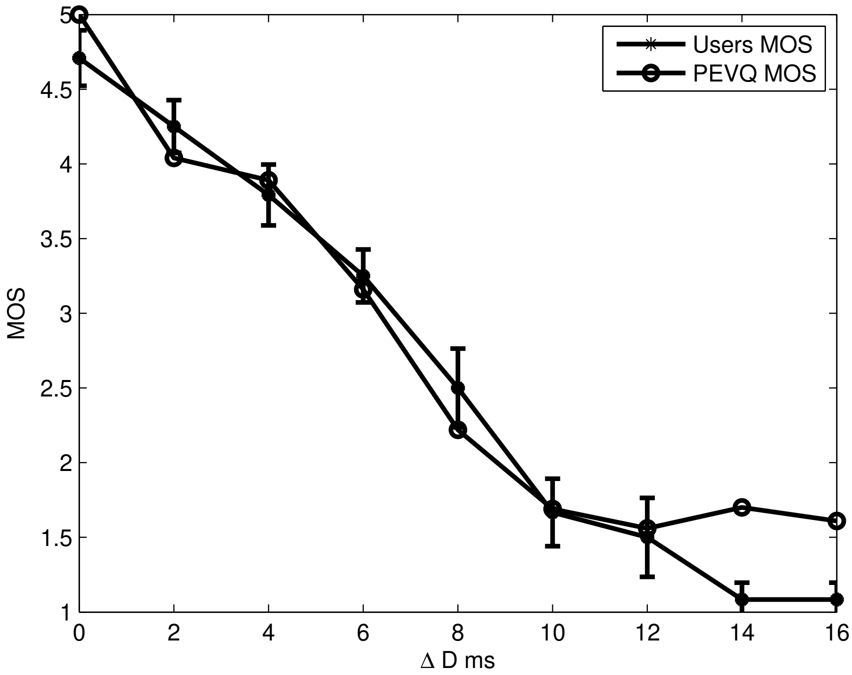



**Figure 4:** PEVQ MOS and Users'MOS rating for $Football$ Video

jump moves from first second to last second of video. In both figures we can see that the users' perception varies with respect to the location where this one-second freeze-and-jump occurs. For example in figure 5 when the freeze occurs at frame number 25 and jumps the frames from 25 to 50, the MOS is 3.63 whereas when freeze and jump cover frames 100 to 125, the MOS drops to 2.63. The corresponding results of PEVQ are also given in figures 5, 6 and 7 along with the users' results. There is a considerable gap between the users and PEVQ rating, but the graphs are similar in shape. For the $Foreman$ and $News$ videos we can see the ratings variation by users as well as PEVQ, however in case of $HallMonitor$ the PEVQ rating remains approximately constant except for the first two and the last values.

Table 2 shows the results of moving freeze and jump along the videos of $Foreman$, $News$ and $HallMonitor$.

From the table we can see that the rating of PEVQ and Users differ from each other. For all videos the rating of PVEQ is greater than four, while Users' rating is varying between two and four, which indicates that user is pickier with respect to the location where freeze and jump occur, whereas the PEVQ just compares the frozen frame with the reference video. Obviously this yields a better rating as compared to user rating. PEVQ gives approximately the same rating for all $HallMonitor$ videos, except for first and last video. It has an unchanged background of hall with two

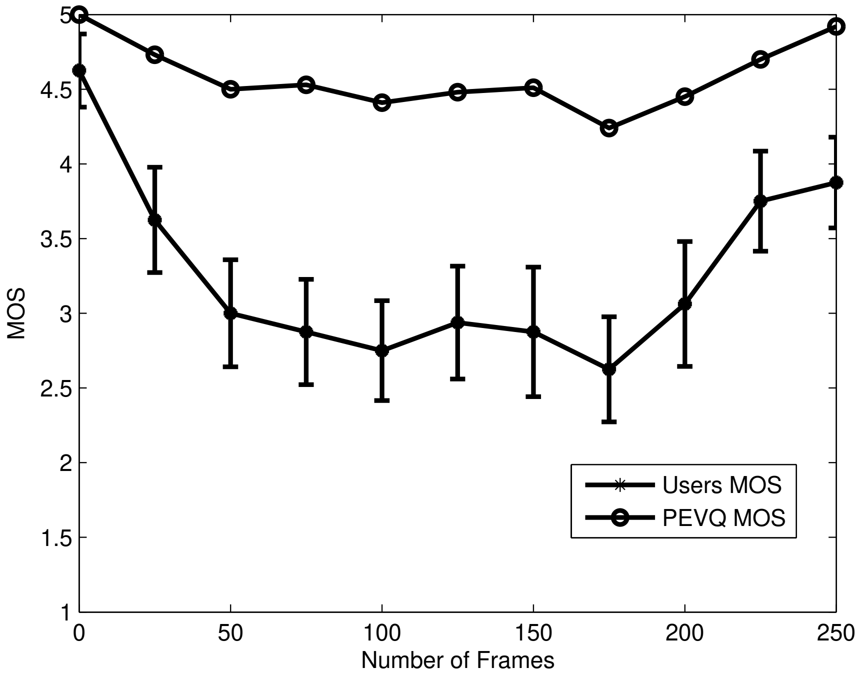



**Figure 5:** PEVQ MOS and Users'MOS rating for $Foreman$ Video with Freeze and Jump

moving persons, so PEVQ obviously could not figure out any big difference between the freeze-and-jump video and the reference video while comparing the frames with each other. From table 2 we can see that PEVQ rating for these videos is around 4.5 except for first and last video, whereas the user's ratings are significantly lower than the PEVQ ratings and varying with respect to the location of the disturbance.

| Frame | | Foreman | | News | | Hall Monitor | |
|---|---|---|---|---|---|---|---|
| Freeze | Jumped | PEVQ | User MOS | PEVQ | User MOS | PEVQ | User MOS |
| 0 | 0 | 5.00 | 4.63± 0.245 | 4.84 | 4.69± 0.261 | 5.00 | 4.56± 0.344 |
| 1 | 1-25 | 4.73 | 3.63± 0.352 | 4.53 | 3.77± 0.394 | 4.77 | 3.67± 0.462 |
| 25 | 25-50 | 4.50 | 3.00± 0.358 | 4.47 | 3.00± 0.544 | 4.47 | 2.33± 0.566 |
| 50 | 50-75 | 4.53 | 2.88± 0.352 | 4.46 | 3.08± 0.564 | 4.48 | 2.44± 0.576 |
| 75 | 75-100 | 4.41 | 2.75± 0.335 | 4.05 | 2.85± 0.696 | 4.48 | 2.22± 0.544 |
| 100 | 100-125 | 4.48 | 2.94± 0.378 | 4.45 | 2.61± 0.522 | 4.54 | 2.22± 0.714 |
| 125 | 125-150 | 4.51 | 2.88± 0.434 | 4.25 | 2.92± 0.469 | 4.52 | 2.11± 0.393 |
| 150 | 150-175 | 4.24 | 2.63± 0.352 | 4.47 | 2.85± 0.660 | 4.55 | 2.33± 0.800 |
| 175 | 175-200 | 4.45 | 3.06± 0.418 | 4.46 | 3.08± 0.564 | 4.54 | 2.56± 0.808 |
| 200 | 200-225 | 4.70 | 3.75± 0.335 | 3.91 | 3.08± 0.564 | 4.52 | 2.44± 0.808 |
| 225 | 225-250 | 4.92 | 3.88± 0.303 | 4.49 | 3.92± 0.348 | 4.70 | 3.22± 0.635 |

**Table 2:** PEVQ MOS and Users' MOS for freeze and jump videos

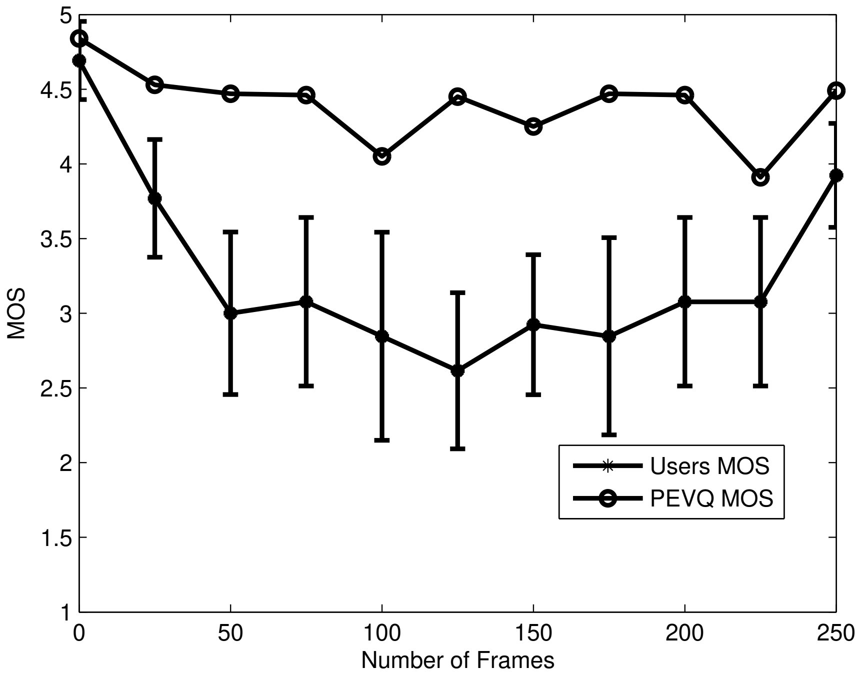



**Figure 6:** PEVQ MOS and Users'MOS rating for $News$ Video with Freeze and Jump

# 7 Conclusion

In this paper, we have presented the study of users' perception for delay variation and freeze-and-jump cases and also compare the user rating with PEVQ result. In case of delay and delay variation we observed that the MOS rating drops linearly towards the bad region due to increase in delay variation. In this case PEVQ shows the result that are well aligned to users' rating.

In the freeze-and-jump case, the user reacts differently with respect to the location where the problem happens. User ratings vary for the same kind of disturbance arising at different locations within the video, which indicates that the disturbance and its location have a combined impact on human perception. We observed similar behavior of user for three videos chosen for this study. In this case the results of PEVQ show variations similar to users' observations but the reduction of the MOS differs significantly in magnitude. In general the rating of PEVQ is between good and excellent, while the user rating is between poor and fair. On the basis of these results one can use PEVQ for MOS study in case of delay and delay variation and potentially for other typical network performance issues. But for freeze-and-jump and similar cases, one has to be careful as PEVQ overestimates user perception.

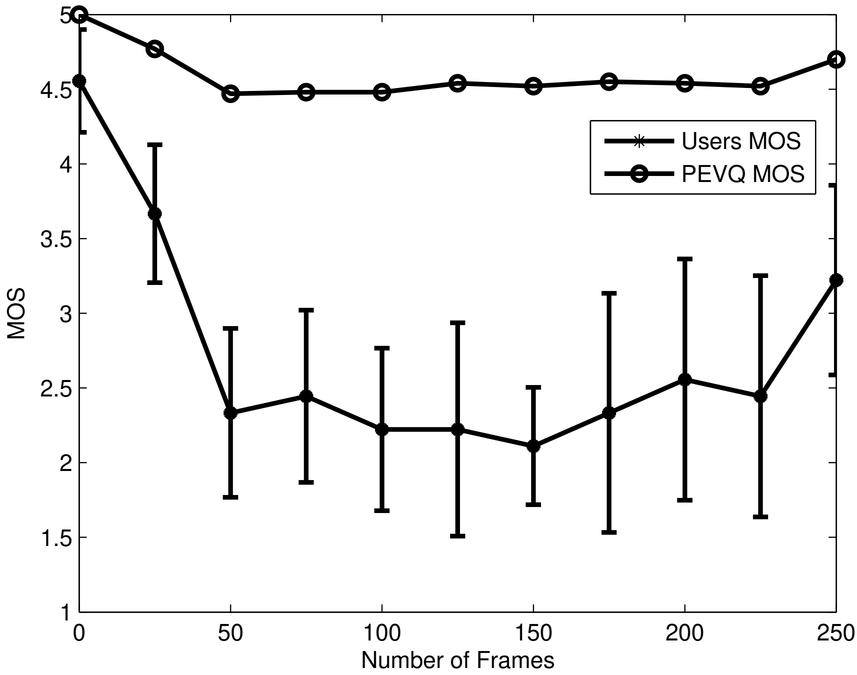



**Figure 7:** PEVQ MOS and Users'MOS rating for $HallMonitor$ Video with Freeze and Jump

## Acknowledgment

I would like to thanks Vasanthi D. Bhamidipati and Swetha Kilari for the QoE survey of delay/delay variation videos.

# References


[1] V. Vassiliou, P. Antoniou, I. Giannakou, and A. Pitsillides. “Requirements for the Transmission of Streaming Video in Mobile Wireless Networks”. In: *Artificial Neural Networks ICANN 2006*. Vol. 4132. Lecture Notes in Computer Science. 2006, pp. 528–537.

[2] P. Calyam, M. Sridharan, W. Mandrawa, and P. Schopis. “Performance Measurement and Analysis of H.323 Traffic”. In: *Passive and Active Network Measurement*. Ed. by C. Barakat and I. Pratt. Vol. 3015. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2004, pp. 137–146.

[3] M. Claypool and J. Tanner. “The effects of jitter on the peceptual quality of video”. In: *MULTIMEDIA ’99: Proceedings of the seventh ACM international conference on Multimedia (Part 2)*. Orlando, Florida, United States: ACM, 1999, pp. 115–118. ISBN: 1-58113-239-5. DOI: http://doi.acm.org/10.1145/319878.319909.

[4] Y. J. Liang, J. G. Apostolopoulos, and B. Girod. “Analysis Of Packet Loss For Compressed Video: Does Burst-Length Matter?” In: 5 (2003), pp. 684–687. URL: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.11.7729.

[5] Y. Qi and M. Dai. “The Effect of Frame Freezing and Frame Skipping on Video Quality”. In: *Intelligent Information Hiding and Multimedia Signal Processing, International Conference on* 0 (2006), pp. 423–426. DOI: http://doi.ieeecomputersociety.org/10.1109/IIH-MSP.2006.162.

[6] VQEG. *Video Quality Experts Group*. http://www.its.bldrdoc.gov/vqeg/. accessed July 2010.

[7] OPTICOM. *Perceptual Evaluation of Video Quality*. http://www.pevq.org/. accessed March 2011.

[8] ITU-T REC-J.247. *Objective perceptual multimedia video quality measurement in the presence of a full reference.* International Telecommunications Union Telecommunication Sector. Aug. 2008.

[9] VideoLAN. [Online] http://www.videolan.org/vlc/. accessed June 2010.

[10] S. Hemminger. “Network Emulation with NetEm”. In: *Linux Conf Au*. 2005.

[11] J. Shaikh, T. Minhas, P. Arlos, and M. Fiedler. “Evaluation of delay performance of traffic shapers”. In: *Second International Workshop on Security and Communication Networks*. 2010, pp. 1–8. DOI: 10.1109/IWSCN.2010.5497994.

[12] P. Arlos, M. Fiedler, and Arne A. Nilsson. “A Distributed Passive Measurement Infrastructure”. In: *Proceedings of Passive and Active Measurement Workshop*. Boston, US, 2005, pp. 215–227.

[13] Endace. *Endace Measurement Systems*. [Online] http://www.endace.com/. accessed March 2010.

[14] I.-R. R. BT.500-11. *Methodology for the subjective assessment of the quality of television pictures.* International Telecommunications Union Telecommunication Sector. Geneva, Switzerland, 2002.

[15] ITU-T P.910. *Subjective video quality assessment methods for multimedia applications.* International Telecommunications Union Telecommunication Sector. 1999.

# Paper IV

## QoE rating Performance evaluation of ITU-T recommended Video Quality Metrics in the context of Video freezes

*Tahir Nawaz Minhas, Muhammad Shahid, Benny Lövström, Andreas Rossholm, Hans-Jürgen Zepernick, and Markus Fiedler. Qoe rating performance evaluation of itu-t recommended video quality metrics in the context of video freezes. Australian Journal of Electrical and Electronics Engineering, 13(2):122–131, 2016.*

## 1 Abstract

In real-time video streaming, video quality can be degraded due to network performance issues. Among other artifacts, video freezing and video jumping are factors that influence user experience. Service providers, operators, and manufacturers are interested in evaluating the Quality of Experience (QoE) objectively because subjective assessment of QoE is expensive and in many user cases subjective assessment is not possible to perform. Different algorithms have been proposed and implemented in this regard. Some of them are in the recommendation list of the ITU Telecommunication Standardization Sector (ITU-T). In this paper, we study the effect of the freezing artifact on user experience and compare the mean opinion score (MOS) of these videos with the results of two algorithms, the Perceptual Evaluation of Video Quality (PEVQ) and Temporal Quality Metric (TQM). Both metrics are part of the ITU-T Recommendation J.247 Annex B and C. PEVQ is a full-reference video quality metric, whereas TQM is a no-reference quality metric. Another contribution of this paper is the study of the impact of different resolutions and frame rates on user experience and how accurately PEVQ and TQM measure varying frame rates. uality of Experience, Video Quality, Objective Video Quality Assessment, Subjective Video Quality Assessment, Perceptual Evaluation of Video Quality, Temporal Quality Metric

# 2 Introduction

The usage of video streaming is growing quickly with the rapid development of the mobile Internet. Watching videos over mobile devices, like smart phones, tablets and laptops, is getting more popular day by day. At the end of 2013, mobile video traffic has been estimated for 53% of the total mobile traffic [1]. With this new development and trend, the users become experienced and quality-conscious. This attracts the attention of researchers, manufacturers, service providers, and network providers towards the Quality of Experience (QoE). QoE is an influencing factor of churn along with other factors. Churn is considered a key performance measure for network providers or service providers. To maintain an acceptable level of the quality for user satisfaction, service providers are interested to measure QoE to take necessary action in time. The ITU-T defines QoE as follows [2]:

*"The overall acceptability of an application or service, as perceived subjectively by the end-user."*

To assess the user experience, either subjective or objective methods are used. Subjective assessment of streaming video is based on different factors such as data transmission rate, packet loss, delay and jitter [3, 4]. Moreover, subjective assessment of video quality is also influenced by seven factors; colour, brightness, background stability, speed in image reassembling, outline definition, 'dirty window', and the mosaic/blocking effect [5]. Due to the involvement of human beings in subjective experiments, it is an expensive and time consuming method. Therefore, service providers, manufacturers, and researchers are interested to objectively measure QoE.

In video streaming, a video encoder, a network, a decoder, and a video player are involved. During the encoding process, a video is compressed spatially and temporally. When the video is transmitted over a mobile communication network or wireless network, it may get impaired by spatial or temporal artifacts. Among other artifacts, frame freeze can occur due to packet loss, jitter, long delay, and low bitrate. Moreover, an error-prone channel can also be a reason of frame freeze. In case of real-time streaming, users can miss the content of a video during the freeze time. This frame freezing and skipping may lead to perceptual degradation. Besides the subjective experiments performed by the Video Quality Experts Group (VQEG), independent efforts have been made to share the results of subjective experiments, e.g. [6–8].

For judging user perception objectively, approaches have been reported that work in the spatial domain [9] or temporal domain, e.g. [10–12]. Similarly, several authors have investigated the effect of network performance parameters on video QoE, e.g. [13–15]. Moreover, in [4], the authors proposed the QoE hourglass model, which describes the factors affecting the QoE from network layer (IP) to application layer. Other efforts have been also made to assess the user perception using different network parameters [16, 17]. In [12], the authors studied the effect of packet loss on decoded video quality. Apart from the video QoE, others have investigated the user behavior on the Web to assess the user experience [18, 19].

In [20], the author presented a no-reference and reduced-reference metric to detect the dropped video frames. This approach is based on an algorithm that quantifies the interruptions to flow of motion in the video scene. Moreover, in [21], the author proposed a model of jerkiness for temporal impairments in video transmission, which can be used to predict jerkiness from small to large resolution. For a more detailed treatise on no-reference methods of image and video quality assessment, including temporal impairments, see [22]. The ITU-T has a list of recommended models or algorithms for perceptual video quality measurement in its recommendation J.247 [23]. It is important to evaluate the accuracy and performance of these algorithms for further use. Therefore, we evaluate two of them with respect to user ratings of the freeze artifact.

In this paper, our first contribution is the subjective tests for freeze frame videos. We consider six videos encoded with two resolutions, three frame rates, and four freeze constellations each which altogether make 144 test samples. This survey enables us to compare the ratings of users for different frame rates as well as resolutions. As a second contribution, we evaluate the performance of ITU-T recommended Perceptual Evaluation of Video Quality (PEVQ) and Temporal Quality Metric (TQM) being part of ITU-T Recommendation J.247 Annex C. PEVQ is also evaluated by [14] with user ratings for network performance parameters and frame freezes. Moreover, in [24], the authors found that the temporal model specified for freezing with skipping in ITU-T J.247 is not applicable to stalling, neglects the influence of video duration, and fails to predict the mean opinion score (MOS) for shorter stalling events. However, our results are more extensive than the existing studies for the frame freeze scenario and differ in resolution and frame rate scenario as well as the comparison with user ratings. Hence, the comparison of PEVQ and TQM results with user ratings for six videos with four resolutions and three frame rates gives us insights into the performance of these models. This article is an extended version of the paper [25], containing more results and their descriptions. In particular, we have added the user ratings with 95% confidence intervals, and compared these with PEVQ ratings and TQM ratings.

The remainder of the paper is organised as follows. Section 3 describes the measurement methodology, which covers the information about the videos, and how the freezing is introduced and tests are conducted. Section 4 provides and discusses the results. Finally, Section 5 concludes this work.

# 3 QoE Assessment Experiment

## 3.1 Test Videos

Six videos “Akiyo”, “Foreman”, “Football”, “Pedestrian”, “Crew”, and “Soccer” are selected from xiph.org video test media. Table 1 details the videos used in the subjective experiments. The six videos are divided into two groups; each group consists of three videos (slow, medium, and fast) based on their temporal spectral information (TI) as per ITU-T recommendation P.910 [26]. The first group contains the resolutions quarter common intermediate format (QCIF) and common intermediate format (CIF). The second group comprises the resolutions of quarter video graphics array (QVGA) and video graphics array (VGA). CIF ($352 \times 288$) and QCIF ($178 \times 144$) video formats are part of the ITU H.261 videoconferencing standard and serve as standard resolutions for video conferences. On the other hand, VGA ($640 \times 480$) and QVGA ($320 \times 240$) are well-known resolutions associated with computer displays. Mobile phones, personal digital assistants (Personal Digital Assistants (PDAs)), and other handheld devices support the QVGA format. Previous video QoE studies used these video formats [7].

Freezing was introduced by using Matlab at different locations within the videos. The length of each freeze was set to one second. After each freeze, frames were skipped. In case of a freeze, the last displayed frame was repeated during one second, thereby replacing the original frames of the video. As a result, the length of the test videos remained unchanged. This freeze introduction process imitates real-time streaming where the last successfully received frame is repeated until the correct frame is received. Figure 1 explains the number of freezes and their locations in the video sequences.

**Table 1:** Video properties used for subjective analysis

| Video (Motion) | Frame Rate | Resolution | Video [s] | No. of freezes for each Video | Freeze Location for each Video | Freeze Length [s] |
|---|---|---|---|---|---|---|
| Akiyo (Slow) | 7.5, 15, and 30 fps | QCIF and CIF | 10 | 1 | 1st quarter | 1 |
| Foreman (Medium) | " | " | 10 | 2 | 1st and 2nd quarter | 1 each |
| Football (Fast) | " | " | 8 | 3 | 1st, 2nd, and 3rd quarter | 1 each |
| Pedestrian (Slow) | 6.25, 12.5 and 25 fps | QVGA and VGA | 15 | 1 | 1st quarter | 1 |
| Crew (Medium) | 7.5, 15, and 30 fps | " | 10 | 2 | 1st and 2nd quarter | 1 each |
| Soccer (Fast) | " | " | 8 | 3 | 1st, 2nd, and 3rd quarter | 1 each |

The fifth column of Table 1 shows the number of freezes for each video and corresponding explanations in the sixth column:

- “1” represents a single freeze event occurring in the first quarter of the video,
- “2” indicates a freeze in the first and the second quarter of the video.
- “3” describes three freeze events, i.e., in the first, second, and third quarter of the video.

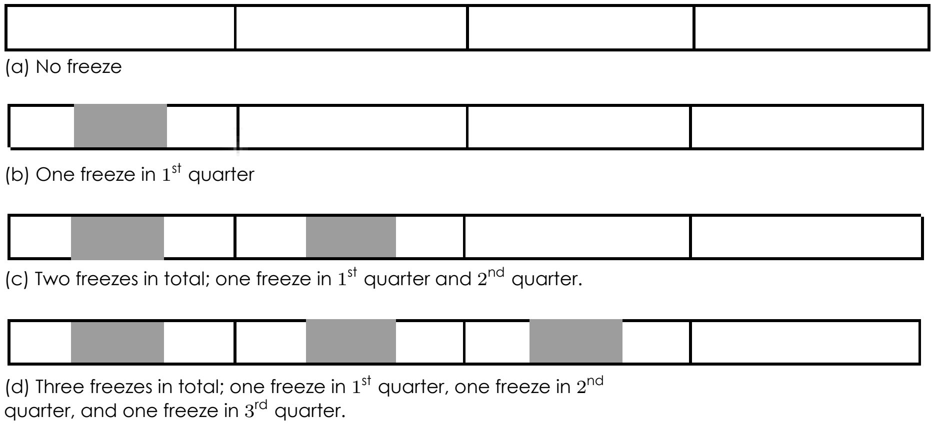



**Figure 1:** Number of freezes and their location in the video sequence: (a) No freeze, (b) One freeze, (c) Two freezes, (d) Three freezes.

## 3.2 Test Methodology

The perceived qualities of the test videos were determined from subjective tests, conducted in the perception laboratory of the Blekinge Institute of Technology (BTH). The BTH perception laboratory fulfils the requirements for subjective tests as prescribed by ITU-T P.910. The tests were carried out by using the single stimulus absolute category rating. Each video sequence is presented one at a time and rated. After watching each video, a viewer has ten seconds to judge the quality of the video. During these ten seconds of voting time, the user has the freedom to change the rating of the video. Most of the 25 non-expert participants who volunteered to take part in the experiments were students. According to the ITU-T recommendation, the minimum number of observers for subjective tests is fifteen [27]. The subjects were trained for the experiment with written instructions and verbal discussions. Users are allowed to adjust their viewing distance according to their comfort level. However, they are instructed to keep their back with the chair. Videos are played in the centre of the screen with grey background. Viewing sessions were kept approximately half an hour each to avoid subject fatigue. To unbias the ratings, the presentation orders of the videos were randomized in a way such that each of the participants faces a different order of presentation of the videos. To relax and reduce the fatigue, refreshments were served during the break between the three sessions.

## 3.3 Objective Quality Metrics

Besides subjective assessment of the quality of the various frame freeze conditions, as detailed in Section 3.1, two objective metrics have been used. The TQM proposed in [28] is centered on measuring the annoyance of frame freeze duration and can be employed in no-reference conditions. This metric uses the mean squared difference value to mark freeze events and builds a mapping function based on such durations of freeze to estimate the subjective quality of a video. The metric is a part of ITU-T Recommendation J.247 Annex C [23] for the objective perceptual quality measurement of videos. TQM shows promising correlation with the subjective scores. The second objective metric is a full-reference commercially available tool called PEVQ, which is a standardized end-to-end measurement algorithm that estimates the frame quality of videos on a 5-point MOS scale. After successful benchmarking by VQEG, PEVQ has become part of ITU-T Recommendation J.247 [23]. PEVQ measures the video quality based on modelling the behaviour of the human visual system (HVS) and estimates the MOS based on indicators of the HVS [29].

# 4 Results

This section describes the results obtained from our experiments. Section 4.1 describes the user ratings, PEVQ ratings, and TQM ratings for QCIF and CIF video resolutions. Section 4.2 discusses the user ratings, PEVQ ratings, and TQM ratings for the video sequences encoded with QVGA and VGA resolutions. In Section 4.3 and 4.4, we compare the user ratings with the PEVQ and TQM ratings, respectively.

A total of 30 participants took part in the subjective assessment. However, five outliers were filtered out. The outliers are based on the user ratings which have either some un-rated videos or a set of videos assessed with a fixed opinion score.

## 4.1 Comparison of QCIF and CIF Video Ratings

Table 2 shows the results for the videos 'Akiyo', 'Football', and 'Foreman'. Each video was encoded at QCIF and CIF resolutions with frame rates of 30 fps, 15 fps, and 7.5 fps. The third column of the table shows the number of freezes, as explained in Table 2, where zero means no freeze, i.e. original video clip. The fourth, fifth, and sixth column show the results for videos endcoded with QCIF resolution. The fourth column of the table, provides the user ratings of the test videos. The fifth column $\Delta$PM is the difference between PEVQ ratings and MOS.

The sixth column $\Delta$TM shows the difference between the results of TQM and MOS.

Similarly, the seventh, eighth and ninth column are based on the results for CIF resolution videos. In the fifth, sixth, eighth, and ninth column, a positive number indicates that PEVQ or TQM rating is higher than the MOS. A negative number in these columns tells that the MOS rating is higher than the PEVQ or TQM rating. Note that $\Delta$PM is always positive, while $\Delta$TM is flucuating arround zero. The rating of PEVQ and TQM for the original video (with zero freezes) is 5 irrespective of the frame rate. However, for other videos both ratings differ. The value of $\Delta$PM and $\Delta$TM is increasing with the decrease in the frame rate because the user ratings decreased with reduced frame rate. For video clip ‘Akiyo’, except the original video, the value of $\Delta$PM is greater than approximately 0.6 and $\Delta$TM is approximately equal to -0.6 for 30 fps. This implies that the PEVQ rating is higher, while TQM is lower than the user rating. However, for 15 fps and 7.5 fps, $\Delta$TM is close to zero whereas $\Delta$PM is near to 1.5, which is a visible difference. Moreover, for the other two videos, PEVQ rating is higher than the user rating. A similar behavior of PEVQ rating in case of frame freeze is reported in [14].

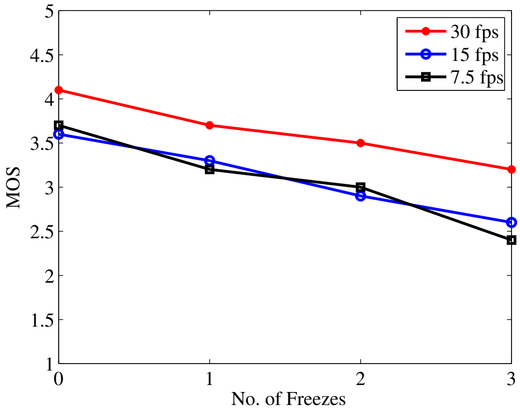
To illustrate and discuss the user ratings, PEVQ ratings, and TQM ratings, we selected one video for each resolution out of the three videos. Figures 2 and 3 describe the ratings for ‘Akiyo’ with QCIF resolution and ‘Football’ with CIF resolution, respectively. Figures 2a,b,c and 3a, b, c, respectively, show the user ratings, PEVQ rating and TQM rating, for 30 fps, 15 fps and 7.5 fps. MOS ratings from 1 to 5 are given along the y-axis, while the x-axis shows the number of freezes. Figures 2a and 3a illustrate that the user ratings are fluctuating with the change of frame rate. Similar trends can be seen for PEVQ in Figures 2b and 3b. However, there is no remarkable change in TQM rating with the change in frame rate. By comparing the user, PEVQ, and TQM ratings, we can see the difference between them as illustrated in Table 2.

User ratings vary with respect to the frame rate as illustrated by Figures 2a and 3a. The rating for 30 fps is better than for the other two frame rates, which is intuitive, because the higher frame rate ensures good video quality. However, by reducing the frame rate by 50%, the decrease in the user rating is negligible. Therefore, with respect to the network condition, e.g., low bandwidth, the frame rate can be adopted accordingly. The MOS, PEVQ, and TQM ratings for the other videos show a similar behavior as we have seen and discussed in Figures 2 and 3 for ‘Akiyo’ and ‘Football’ respectively.

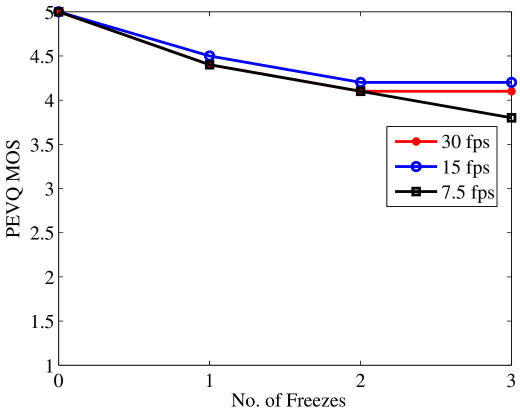
## 4.2 Comparison of QVGA and VGA Video Ratings

Table 3 shows the results of the videos encoded with QVGA and VGA resolution. Two of them have frame rates 30 fps, 15 fps, and 7.5 fps, while one of them has frame rates 25 fps, 12.5 fps, and 6.25 fps. Table 3 has a similar skeleton as Table 2, which is described in Section 4.1. For both resolutions, the values of $\Delta$PM are mostly greater than or equal to one, whereas the values of $\Delta$TM are less than or equal to 0.5, except for the original videos. This indicates that the TQM and user ratings are closer to each other as compared to PEVQ.

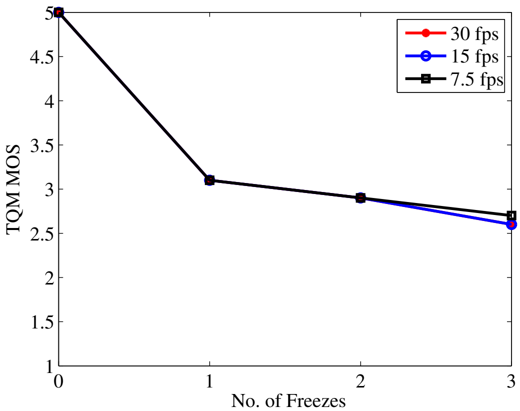
Like the QCIF and CIF videos, we selected two videos, one with QVGA and one with VGA resolution, out of six videos to discuss the user ratings, PEVQ ratings, and TQM ratings. Figure 4 illustrates the result of ‘Soccer’ with QVGA resolution and Figure 5 shows the result of ‘Crew’ with VGA resolution. Each figure consists of three subfigures a, b, and c, which are describing the user rating, PEVQ rating, and TQM rating, for 30 fps, 15 fps and 7.5 fps, respectively. The x-axis shows the number




(a) User ratings

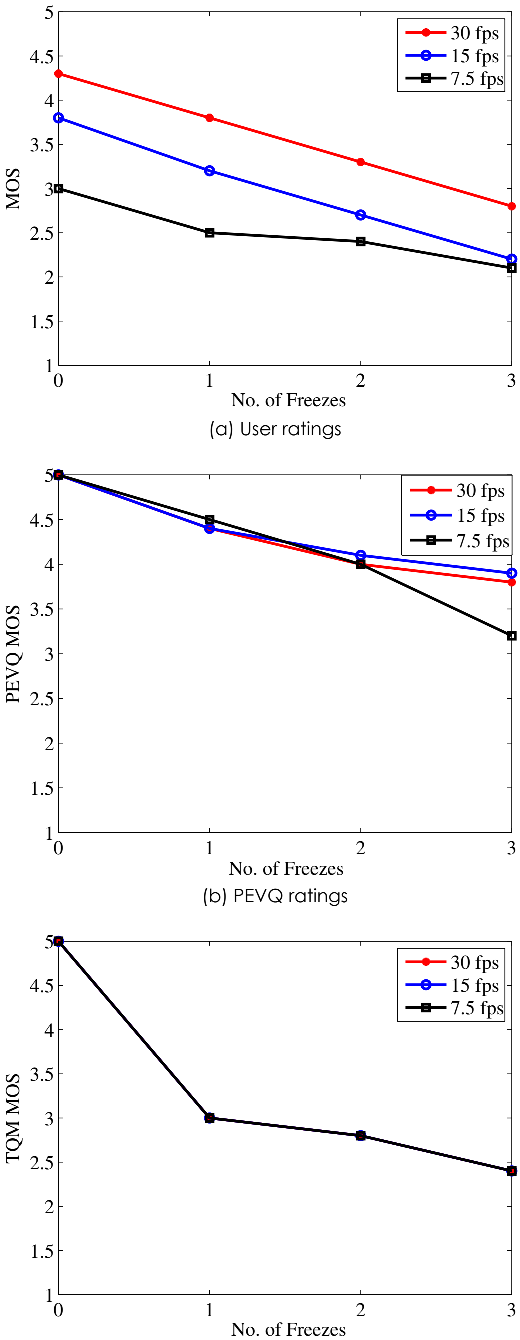



(b) PEVQ ratings




(c) TQM ratings

**Figure 2:** Rating of `Akiyo' with QCIF resolution: (a) User ratings, (b) PEVQ ratings, (c) TQM ratings.




(a) User ratings

(b) PEVQ ratings

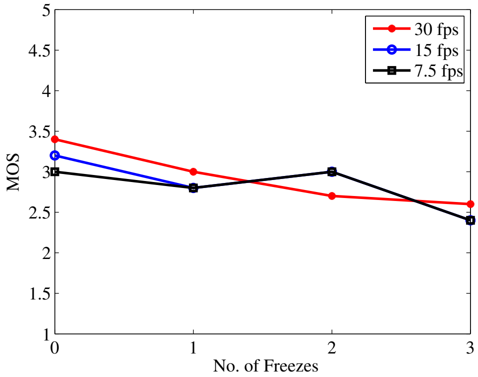
(c) TQM ratings

**Figure 3:** Ratings of `Football' with CIF resolution: (a) User ratings, (b) PEVQ ratings, (c) TQM ratings.

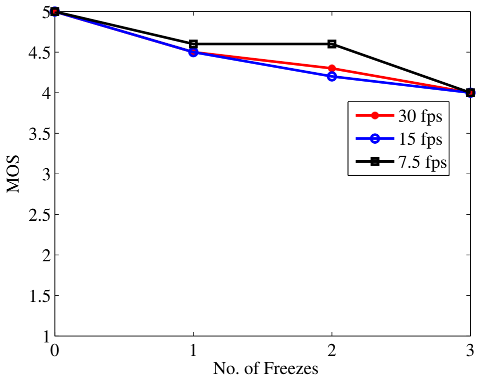
**Table 2:** Comparison of QCIF and CIF video ratings

| Video | fps | No of Freezes | QCIF | | | CIF | | |
|---|---|---|---|---|---|---|---|---|
| | | | MOS | ΔPM | ΔTM | MOS | ΔPM | ΔTM |
| AKIYO | 30 | 0 | 4.1 | 0.9 | 0.9 | 4.0 | 0.9 | 1.0 |
| | 30 | 1 | 3.7 | 0.7 | -0.6 | 3.5 | 0.8 | -0.4 |
| | 30 | 2 | 3.5 | 0.6 | -0.6 | 3.1 | 1.1 | -0.1 |
| | 30 | 3 | 3.2 | 0.9 | -0.6 | 2.6 | 1.4 | 0.0 |
| | 15 | 0 | 3.6 | 1.4 | 1.4 | 3.4 | 1.5 | 1.6 |
| | 15 | 1 | 3.3 | 1.2 | -0.2 | 3.1 | 1.3 | 0.1 |
| | 15 | 2 | 2.9 | 1.3 | 0.0 | 2.6 | 1.5 | 0.4 |
| | 15 | 3 | 2.6 | 1.6 | 0.0 | 2.0 | 2.0 | 0.6 |
| | 7.5 | 0 | 3.7 | 1.3 | 1.3 | 2.9 | 2.0 | 2.1 |
| | 7.5 | 1 | 3.2 | 1.2 | -0.1 | 2.4 | 2.0 | 0.7 |
| | 7.5 | 2 | 3.0 | 1.1 | -0.1 | 2.2 | 1.9 | 0.7 |
| | 7.5 | 3 | 2.4 | 1.4 | 0.3 | 2.1 | 1.7 | 0.6 |
| FOOTBALL | 30 | 0 | 4.4 | 0.6 | 0.6 | 4.3 | 0.8 | 0.8 |
| | 30 | 1 | 4.1 | 0.3 | -1.1 | 3.8 | 0.6 | -0.7 |
| | 30 | 2 | 3.6 | 0.4 | -0.8 | 3.3 | 0.7 | -0.5 |
| | 30 | 3 | 3.2 | 0.0 | -0.8 | 2.8 | 1.0 | -0.4 |
| | 15 | 0 | 4.1 | 0.9 | 0.9 | 3.8 | 1.2 | 1.2 |
| | 15 | 1 | 3.5 | 1.0 | -0.5 | 3.2 | 1.2 | -0.2 |
| | 15 | 2 | 3.2 | 1.0 | -0.4 | 2.7 | 1.4 | 0.1 |
| | 15 | 3 | 2.7 | 1.3 | -0.3 | 2.2 | 1.7 | 0.2 |
| | 7.5 | 0 | 3.8 | 1.2 | 1.2 | 3.0 | 2.0 | 2.0 |
| | 7.5 | 1 | 2.7 | 1.8 | 0.3 | 2.5 | 2.0 | 0.5 |
| | 7.5 | 2 | 2.3 | 1.7 | 0.5 | 2.4 | 1.6 | 0.4 |
| | 7.5 | 3 | 2.1 | 2.1 | 0.4 | 2.1 | 1.1 | 0.3 |
| FOREMAN | 30 | 0 | 4.2 | 0.9 | 0.9 | 4.4 | 0.6 | 0.6 |
| | 30 | 1 | 3.7 | 0.8 | -0.6 | 4.1 | 0.3 | -1.0 |
| | 30 | 2 | 3.4 | 0.8 | -0.5 | 3.6 | 0.5 | -0.7 |
| | 30 | 3 | 3.0 | 0.8 | -0.4 | 3.2 | 0.6 | -0.5 |
| | 15 | 0 | 3.7 | 1.3 | 1.3 | 4.0 | 1.0 | 1.0 |
| | 15 | 1 | 3.4 | 1.2 | -0.3 | 3.6 | 1.0 | -0.5 |
| | 15 | 2 | 3.3 | 1.0 | -0.4 | 3.1 | 1.3 | -0.2 |
| | 15 | 3 | 3.0 | 1.1 | -0.4 | 2.6 | 1.4 | 0.0 |
| | 7.5 | 0 | 3.5 | 1.5 | 1.5 | 3.0 | 2.0 | 2.0 |
| | 7.5 | 1 | 3.1 | 1.4 | 0.1 | 2.6 | 2.1 | 0.6 |
| | 7.5 | 2 | 2.8 | 1.3 | 0.1 | 2.3 | 2.2 | 0.6 |
| | 7.5 | 3 | 2.7 | 1.2 | 0.0 | 2.2 | 1.8 | 0.5 |

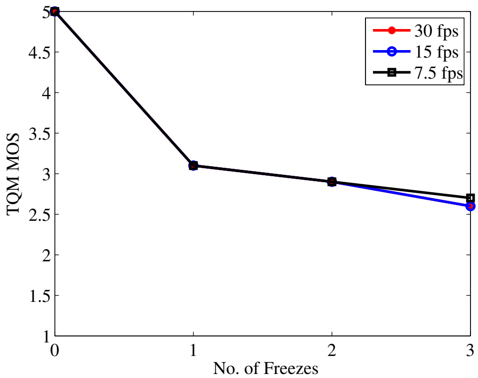
of freezes and the y-axis represents the MOS rating from 1 to 5. In Figures 4a and 5a, the user ratings are not varying much with the change in frame rate. However, we have seen that the user rating varies for QCIF and CIF videos in Figures 2a and 3a, respectively. From Figures 4b and 5b, we can see that there is a nominal variation in the ratings of PEVQ with respect to the change in the frame rate. There is no significant change in TQM ratings with the variation of frame rate, see Figures 4c and 5c. PEVQ and TQM ratings show consistency in their behavior as we can observe in Figures 2b, 2c, and 3b, 3c. As can be determined from Table 3, ΔTM of video 'Soccer' and 'Football' is closer to zero for the frame freeze case. This shows that TQM ratings are closer to the user ratings. Furthermore, the value of ΔPM is greater than 0, which shows that the PEVQ ratings are greater as compared to the user ratings. In both groups of videos, we have seen that user ratings and PEVQ ratings vary with frame rates. However, the TQM ratings remain similar for all frame rates.




(a) User ratings

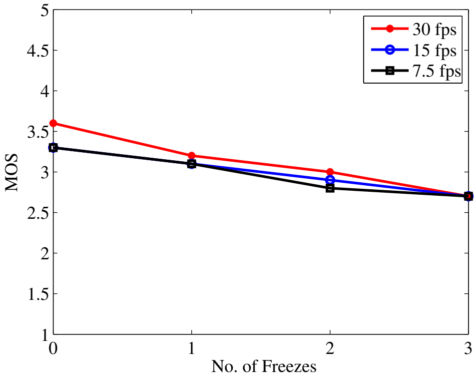



(b) PEVQ ratings

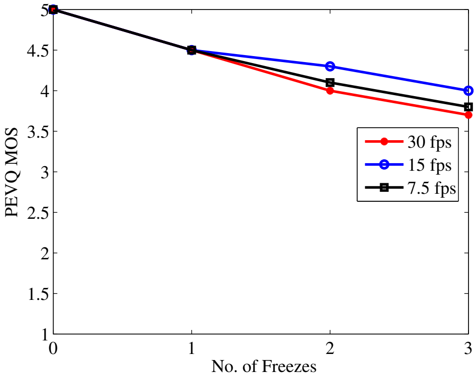



(c) TQM ratings

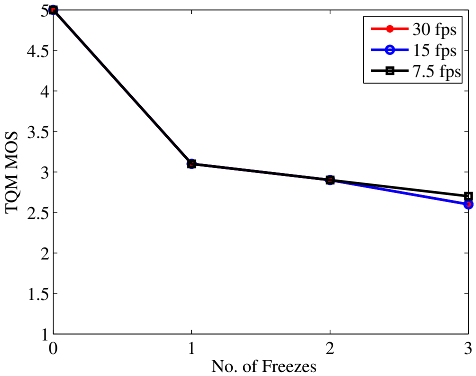
**Figure 4:** Ratings of `Soccer' with QVGA resolution: (a) User ratings, (b) PEVQ ratings, (c) TQM ratings.




(a) User ratings




(b) PEVQ ratings




(c) TQM ratings

**Figure 5:** Ratings of `Crew' with VGA resolution: (a) User ratings, (b) PEVQ ratings, (c) TQM ratings.

**Table 3:** Comparison of QVGA and VGA video ratings

| Video | fps | No of Freezes | QVGA | | | VGA | | |
|---|---|---|---|---|---|---|---|---|
| | | | MOS | ΔPM | ΔTM | MOS | ΔPM | ΔTM |
| CREW | 30 | 0 | 3.8 | 1.2 | 1.2 | 3.6 | 1.4 | 1.4 |
| | 30 | 1 | 3.4 | 1.0 | -0.3 | 3.2 | 1.2 | -0.1 |
| | 30 | 2 | 3.0 | 1.0 | -0.1 | 3.0 | 1.0 | -0.1 |
| | 30 | 3 | 2.9 | 1.0 | -0.2 | 2.7 | 1.0 | 0.0 |
| | 15 | 0 | 3.5 | 1.5 | 1.5 | 3.3 | 1.7 | 1.7 |
| | 15 | 1 | 3.3 | 1.2 | -0.1 | 3.1 | 1.4 | 0.0 |
| | 15 | 2 | 2.9 | 1.1 | 0.0 | 2.9 | 1.4 | 0.0 |
| | 15 | 3 | 2.6 | 1.0 | 0.0 | 2.7 | 1.2 | -0.1 |
| | 7.5 | 0 | 3.0 | 2.0 | 2.0 | 3.3 | 1.7 | 1.7 |
| | 7.5 | 1 | 2.7 | 1.8 | 0.5 | 3.1 | 1.4 | 0.0 |
| | 7.5 | 2 | 2.4 | 1.7 | 0.6 | 2.8 | 1.3 | 0.1 |
| | 7.5 | 3 | 2.3 | 1.7 | 0.4 | 2.7 | 1.1 | -0.1 |
| SOCCER | 30 | 0 | 3.4 | 1.7 | 1.7 | 3.4 | 2.1 | 2.1 |
| | 30 | 1 | 3.0 | 1.5 | 0.1 | 3.1 | 1.8 | 0.5 |
| | 30 | 2 | 2.7 | 1.6 | 0.2 | 3.1 | 1.7 | 0.4 |
| | 30 | 3 | 2.6 | 1.4 | 0.0 | 2.9 | 1.4 | 0.1 |
| | 15 | 0 | 3.2 | 1.8 | 1.8 | 3.3 | 2.3 | 2.3 |
| | 15 | 1 | 2.8 | 1.7 | 0.3 | 3.2 | 1.9 | 0.5 |
| | 15 | 2 | 3.0 | 1.2 | -0.1 | 3.0 | 1.5 | 0.3 |
| | 15 | 3 | 2.4 | 1.6 | 0.2 | 2.9 | 1.4 | 0.1 |
| | 7.5 | 0 | 3.0 | 2.1 | 2.1 | 3.2 | 2.4 | 2.4 |
| | 7.5 | 1 | 2.8 | 1.9 | 0.4 | 3.1 | 2.1 | 0.6 |
| | 7.5 | 2 | 3.0 | 1.6 | 0.0 | 3.0 | 2.0 | 0.5 |
| | 7.5 | 3 | 2.4 | 1.6 | 0.3 | 2.8 | 1.7 | 0.4 |
| PEDESTRIAN | 25 | 0 | 3.4 | 1.7 | 1.7 | 3.4 | 1.6 | 1.6 |
| | 25 | 1 | 3.2 | 1.2 | -0.1 | 3.3 | 1.1 | -0.2 |
| | 25 | 2 | 3.1 | 1.1 | -0.2 | 3.0 | 1.2 | 0.0 |
| | 25 | 3 | 3.0 | 0.8 | -0.3 | 2.7 | 1.1 | -0.1 |
| | 12.5 | 0 | 3.3 | 1.7 | 1.7 | 3.4 | 1.6 | 1.6 |
| | 12.5 | 1 | 3.2 | 1.3 | 0.0 | 3.2 | 1.3 | -0.1 |
| | 12.5 | 2 | 3.1 | 1.2 | -0.1 | 3.1 | 1.1 | -0.1 |
| | 12.5 | 3 | 2.9 | 1.0 | -0.2 | 2.8 | 1.1 | -0.2 |
| | 6.25 | 0 | 3.3 | 1.8 | 1.8 | 3.3 | 1.7 | 1.7 |
| | 6.25 | 1 | 3.2 | 1.2 | 0.0 | 3.2 | 0.9 | 0.0 |
| | 6.25 | 2 | 2.9 | 1.2 | 0.1 | 2.9 | 0.9 | 0.0 |
| | 6.25 | 3 | 2.7 | 1.1 | 0.0 | 2.7 | 0.9 | 0.0 |

## 4.3 Comparison of User Ratings and PEVQ Ratings

From Tables 2 and 3, we can see that ΔPM is always positive and in most of the cases, it is greater than or equal to one. This indicates that both ratings differ from each other, and PEVQ ratings are always greater than the user ratings.

To compare the user ratings and PEVQ ratings, we selected one video from QCIF resolution, and one from VGA resolution. Figures 6a, b, c and 7a, b, c, present the user ratings with 95% confidence intervals, and PEVQ ratings, for 30 fps, 15 fps and 7.5 fps for ‘Akiyo’ and ‘Crew’. The ratings are shown along the y-axis, while the x-axis shows the number of freezes. PEVQ has always rated the original videos ‘5’ regardless of the frame rate, whereas the user ratings vary for the original video with respect to the frame rates and consistently remain below ‘5’. In both figures clearly, PEVQ ratings are more eminent than the upper bound of the 95% confidence interval of user ratings. The user ratings are changing with respect to the number of freezes. An approximately similar variation can be seen in PEVQ ratings. This indicates that

PEVQ successfully captures the variation in the freeze arifacts. Besides the higher ratings of PEVQ, the curve of PEVQ shows that it is capturing the response trend for the freeze artifacts parallel to the user ratings.

### 4.4 Comparison of User Ratings and TQM Ratings

Tables 2 and 3 show that $\Delta$TM is very low and is varying between 0.7 to -0.7 for all freeze cases. In most of the cases, $\Delta$TM is approximately zero, which shows that both ratings are close to each other. Yet, for the no freeze case, TQM has always rated the video as '5'.

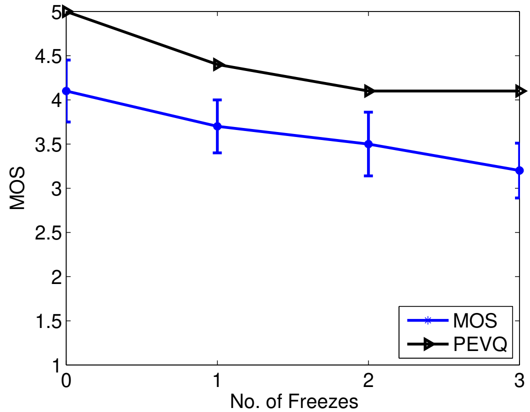
Results of all video sequences depict a similar behaviour in comparison of user rating versus TQM ratings. To illustrate the difference between user ratings and TQM ratings, we selected the two video sequences 'Akiyo' and 'Crew'. Figures 8a, b, c and 9a, b, c show the results of user ratings and TQM ratings for 30 fps, 15 fps, and 7.5 fps, respectively. From higher to lower frame rate, we can see that the gap between both ratings decreases. The reason of this variation is that user ratings vary with respect to the frame rate. However, TQM is unable to capture frame rate variation and rates the video based on the freeze artifact. Figures 2c, 3c, 4c, and 5c confirm this behaviour of TQM. The user ratings and the TQM ratings are very much aligned with each other for all freeze cases.

## 5 Conclusions

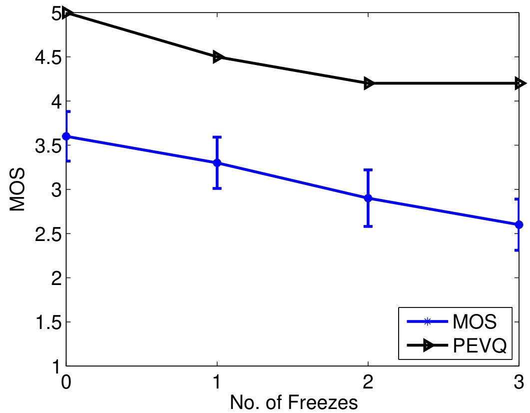
In this paper, we have assessed the subjective quality of videos with four resolutions and three frame rates. Furthermore, video quality has also been evaluated by PEVQ and TQM. Then, user ratings have been compared with the results of PEVQ and TQM.

Users are immediately reacting to the freeze artifact and rate the video as fair on the quality scale. However, on the repetition of the freeze artifact, users drop the rating of the video between fair and poor. Moreover, the user ratings varied due to the change in the video frame rate. Similarly, a small change is noticed in ratings due to change in resolution of video from CIF to QCIF and VGA to QVGA. At the same time, both variations in user rating are minimal. This indicates that the frame rate and resolution of a video can be adopted according to network conditions for better user perception.

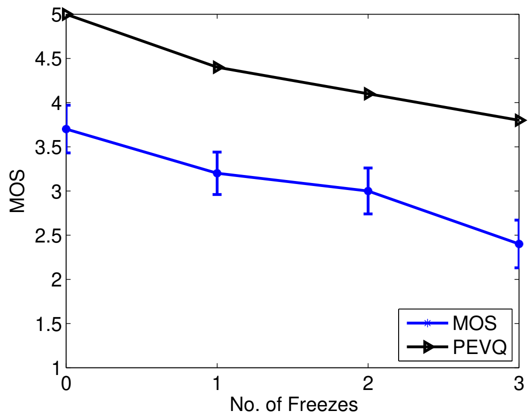
PEVQ ratings and TQM ratings differ from the user ratings. TQM seems to be rather conservative, while PEVQ is optimistic because results of TQM are more inclined to the users' ratings as compared to PEVQ. PEVQ ratings remain between good and excellent while the ratings of TQM and users are between poor and fair. In contrast to the user ratings, the ratings of TQM for no-freeze videos are excellent. Like the user ratings, PEVQ ratings vary due to the change of video frame rate, which



(a) 30 fps



(b) 15 fps



(c) 7.5 fps

**Figure 6:** Rating of `Akiyo' with QCIF resolution: (a) 30 fps, (b) 15 fps, (c) 7.5 fps for user ratings (MOS) and PEVQ ratings.

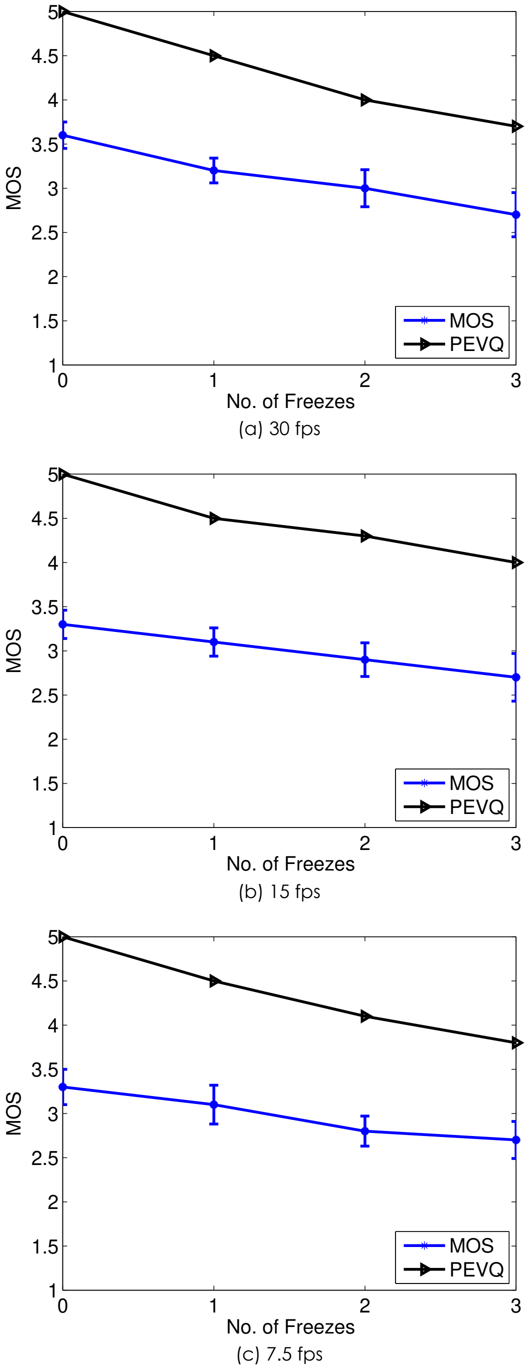
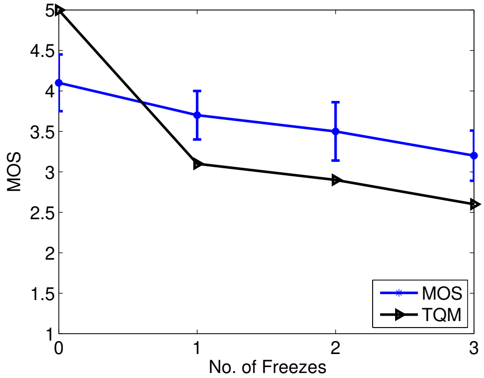
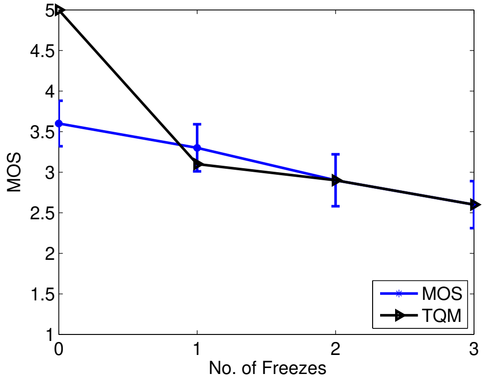


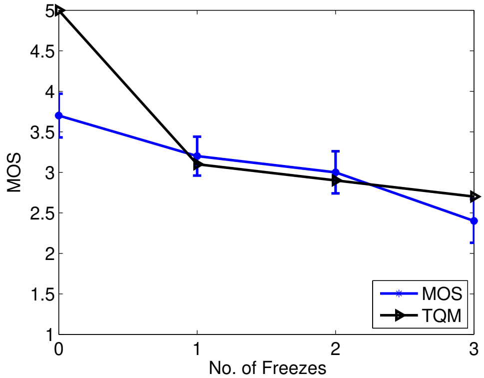

**Figure 7:** Rating of `Crew' with VGA resolution: (a) 30 fps, (b) 15 fps, (c) 7.5 fps for user ratings (MOS) and PEVQ ratings.




(a) 30 fps

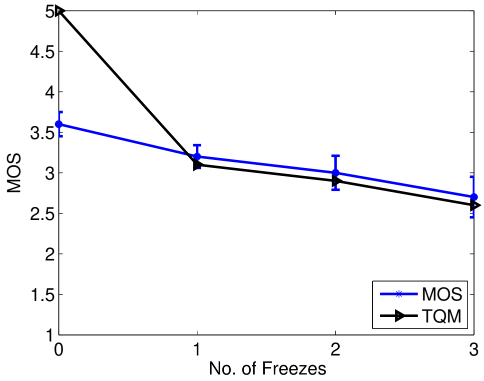



(b) 15 fps

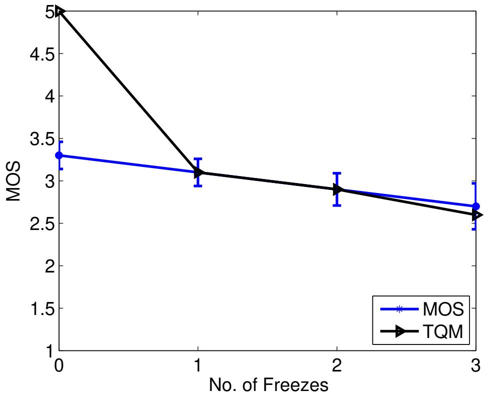



(c) 7.5 fps

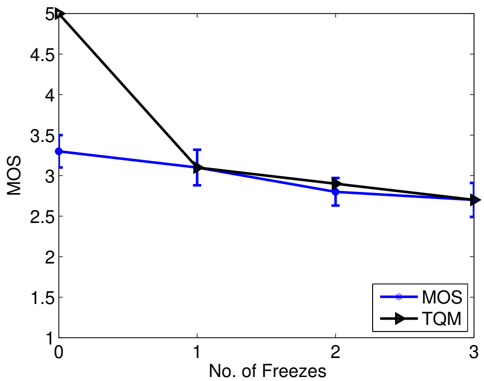
**Figure 8:** Rating of `Akiyo' with QCIF resolution: (a) 30 fps, (b) 15 fps, (c) 7.5 fps for user ratings (MOS) and TQM ratings.




(a) 30 fps




(b) 15 fps




(c) 7.5 fps

**Figure 9:** Rating of `Crew' with VGA resolution: (a) 30 fps, (b) 15 fps, (c) 7.5 fps for user ratings (MOS) and TQM ratings.

shows that PEVQ detects the change in video frame rate and evaluates the video quality accordingly. However, for user ratings and PEVQ ratings, this variation is negligible. On the other hand, the TQM ratings remain unchanged for all frame rates.

Overall, for the set of videos and artifacts that we considered, the visual quality changes due to frame freezes are better estimated by TQM, being closer to the subjective perception as compared to PEVQ.

# 6 Future Work

Future work is to perform similar tests on high resolution videos and to compare other quality assessment models for different frame rates and frame freezes. Moreover, the performance of other no-reference video quality metrics for frame freeze and jerkiness are under investigation in comparison to user ratings.


## Aknowlegment

The authors are very grateful to Bharat Reddy Ramancha for his help in conducting the subjective experiments.


# References


[1] CiscoSystem. *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2013–2018*. 2014.

[2] ITU-T FG IPTV. *Definition of Quality of Experience (QoE)*. Focus Group on IPTV, ITU-T. 2007.

[3] V. Vassiliou, P. Antoniou, I. Giannakou, and A. Pitsillides. “Requirements for the Transmission of Streaming Video in Mobile Wireless Networks”. In: *Artificial Neural Networks □ ICANN 2006*. Vol. 4132. Lecture Notes in Computer Science. 2006, pp. 528–537.

[4] T. N. Minhas and M. Fiedler. “Quality of experience hourglass model”. In: *2013 International Conference on Computing, Management and Telecommunications (ComManTel)*. 2013, pp. 87–92. DOI: `10.1109/ComManTel.2013.6482371`.

[5] M. J. Gili, E. L. Janez, L. M. Hernandez, and M. Szymanski. “Subjective image quality assessment and prediction in digital video communications”. In: COST212 HUFIS Report. 1991, pp. 65–79.

[6] M. Shahid, A. K. Singam, A. Rossholm, and B. Lövström. "Subjective quality assessment of H.264/AVC encoded low resolution videos". In: *Fifth International Congress on Image and Signal Processing*. Chongqing, China, Oct. 2012, pp. 63–67.

[7] T. N. Minhas, O. Gonzalez Lagunas, P. Arlos, and M. Fiedler. "Mobile video sensitivity to packet loss and packet delay variation in terms of QoE". In: *2012 19th International Packet Video Workshop (PV)*. 2012, pp. 83–88. DOI: 10.1109/PV.2012.6229747.

[8] R. Mok, E. Chan, and R. Chang. "Measuring the quality of experience of HTTP video streaming". In: *IFIP/IEEE International Symposium on Integrated Network Management*. Dublin, Ireland, 2011, pp. 485–492. DOI: 10.1109/INM.2011.5990550.

[9] M. Shahid, A. Rossholm, and B. Lövström. "A reduced complexity no-reference artificial neural network based video quality predictor". In: *Fourth International Congress on Image and Signal Processing*. Shanghai, China, Oct. 2011, pp. 517–521.

[10] R. R. Pastrana-Vidal, J. C. Gicquel, J. L. Blin, and H. Cherifi. "Predicting subjective video quality from separated spatial and temporal assessment". In: *Proceedings of SPIE*. Vol. 6057. San Jose, CA, USA, 2006, pp. 276–286.

[11] S. Winkler. *Digital video quality: Vision models and metrics*. New York, USA: John Wiley & Sons, 2005.

[12] Q. Dai and R. Lehnert. "Impact of Packet Loss on the Perceived Video Quality". In: *Second International Conference on Evolving Internet (INTERNET),* Valcencia, Spain, 2010, pp. 206–209. DOI: 10.1109/INTERNET.2010.51.

[13] M. Claypool and J. Tanner. "The Effects of Jitter on the Peceptual Quality of Video". In: *Proceedings of the Seventh ACM International Conference on Multimedia (Part 2)*. MULTIMEDIA. Orlando, Florida, USA: ACM, 1999, pp. 115–118. ISBN: 1-58113-239-5. DOI: 10.1145/319878.319909.

[14] T. N. Minhas and M. Fiedler. "Impact of Disturbance Locations on Video Quality of Experience". In: *Quality of Experience for Multimedia Content Sharing, EuroITV2011*. Lisbon, Portugal, June 2011.

[15] T. Zinner, T. Hossfeld, T. N. Minhas, and M. Fiedler. "Controlled vs. Uncontrolled Degradations of QoE - The Provisioning-Delivery Hysteresis in Case of Video". In: *Quality of Experience for Multimedia Content Sharing*. Tampere, Finland, June 2010.

[16] M. Mu, R. Gostner, A. Mauthe, G. Tyson, and F. Garcia. “Visibility of individual packet loss on H.264 encoded video stream – A user study on the impact of packet loss on perceived video quality”. In: *Proceedings of SPIE*. Ed. by R. Rejaie and K. D. Mayer-Patel. Vol. 7253. San Jose, CA, USA, Jan. 2009, pp. 02–12. DOI: 10.1117/12.815538.

[17] M. Esteve, C. E. Palau, J. Martínez-Nohales, and B. Molina. “A Video Streaming Application for Urban Traffic Management”. In: *J. Netw. Comput. Appl.* 30.2 (Apr. 2007), pp. 479–498. ISSN: 1084-8045. DOI: 10.1016/j.jnca.2006.06.001.

[18] J. Shaikh, M. Fiedler, D. Collange, P. Arlos, and T. N. Minhas. “Inferring User-Perceived Performance of Network by Monitoring TCP Interruptions”. In: *Network Protocols and Algorithms* 4.2 (2012), pp. 49–67.

[19] J. Shaikh, M. Fiedler, P. Arlos, and D. Collange. “Modeling and analysis of web usage and experience based on link-level measurements”. In: *24th IEEE International Teletraffic Congress*. Krakow, Poland, 2012, pp. 1–8.

[20] S. Wolf. “A no reference (NR) and reduced reference (RR) metric for detecting dropped video frames”. In: *Fourth International Workshop on Video Processing and Quality Metrics for Consumer Electronics*. Scottsdale, USA, Jan. 2009.

[21] S. Borer. “A model of jerkiness for temporal impairments in video transmission”. In: *Second International Workshop on Quality of Multimedia Experience (QoMEX)*. Trondheim, Norway, June 2010, pp. 218–223. DOI: 10.1109/QOMEX.2010.5516155.

[22] M. Shahid, A. Rossholm, B. Lövström, and H.-J. Zepernick. “No-reference image and video quality assessment: A classification and review of recent approaches”. In: *EURASIP Journal on Image and Video Processing* (2014). ISSN: 1687-5281. DOI: 10.1186/1687-5281-2014-40.

[23] ITU-T Rec. J.247. *Objective perceptual multimedia video quality measurement in the presence of a full reference*. ITU-T. 2008.

[24] T. Hossfeld, S. Egger, R. Schatz, M. Fiedler, K. Masuch, and C. Lorentzen. “Initial delay vs. interruptions: Between the devil and the deep blue sea”. In: *Fourth International Workshop on Quality of Multimedia Experience*. Yarra Valley, Australia, July 2012, pp. 1–6. DOI: 10.1109/QoMEX.2012.6263849.

[25] T. N. Minhas, M. Shahid, A. Rossholm, B. Lövström, H.-J. Zepernick, and M. Fiedler. “Assessment of the rating performance of ITU-T recommended video quality metrics in the context of video freezes”. In: *2013 Australasian Telecommunication Networks and Applications Conference (ATNAC)*. 2013, pp. 207–212. DOI: 10.1109/ATNAC.2013.6705382.

[26] ITU-T Rec. P.910. *Subjective video quality assessment methods for multimedia applications*. ITU-T. Apr. 2008.

[27] ITU-R Rec. BT.500-13. *Methodology for the subjective assessment of the quality of television pictures*. ITU-T. 2012.

[28] Q. Huynh-Thu and M. Ghanbari. “No-reference temporal quality metric for video impaired by frame freezing artefacts”. In: *16th IEEE International Conference on Image Processing*. Cario, Egypt, Nov. 2009, pp. 2221–2224. DOI: 10.1109/ICIP.2009.5413894.

[29] OPTICOM. *Perceptual Evaluation of Video Quality*. http://www.pevq.org/. accessed July 2010.

# Paper V

## Mitigation of the Effects of Network Outage on Video QoE Using a Sender Buffer

*Tahir Nawaz Minhas and Markus Fiedler.Mitigation of the Effects of Network Outage on Video QoE Using a Sender Buffer. MDPI Open Access Journal Electronics: Advances in Network, 2021.*

## 1 Abstract

–With the growth of multimedia applications and the mobile Internet, quality sense and quality expectation of the end-user are rising rapidly. A small notable distortion in the multimedia applications may degrade the degree of delight of the user, who is very considerate of the video Quality of Experience (QoE). During live streaming, a network outage may result in video freezes and video jumps. To dampen the impact of a network outage on the video QoE, we propose the use of a well-sized sender buffer. We present the concept, derive key analytical relations, and perform a set of subjective tests. Based on those, we report a significant enhancement of user ratings due to the proposed sender buffer in the presence of network outages.

*keyword*–Quality of Experience; network outage; video streaming; sender buffer; user experiments

## 2 Introduction

The use of multimedia applications is tremendously increasing with the rapid growth of wireless and mobile wireless Internet. Mobile devices, such as smartphones, pads, and laptops, are capable of connecting to the Internet via mobile (e.g., 3G, 4G, or 5G) and wireless (Wi-Fi) interfaces. Moreover, mobile devices are continuously improving in terms of computing capabilities and power, data speed, and screen resolution. This exposure strengthens the quality sense of the user, and she expects a good Quality of Experience (QoE), describing the degree of delight or annoyance of a user's experience with a service [1]. However, multimedia applications are sensitive to jitter, bandwidth limitation, and delay. The disruption of the network is more noticeable in multimedia applications than in other applications. As glitches and artifacts are more perceivable in multimedia applications, they can easily affect users' QoE. Live video transmission may undergo QoE degradation due to interruptions such as video freezes, jumps, and misses.

The use of Internet services increased by factor 2.5 as compared to COVID-19 pre-lockdown [2]. Due to the COVID-19 pandemic, many businesses and other activities have moved online. This increases the use of live multimedia applications, e.g., online meetings, lectures, discussions, etc. For example, during an online lecture, when a teacher is teaching, the students may miss key information and lecture material due to a network problem. Similarly, watching a cricket match online and missing a big shot of six of one's favorite team can dramatically lower QoE ratings. We will see subsequently that buffering some content on the sender-side can help with mitigating the impact of network problems on QoE and its ratings, especially when buffered data are sent to the recipient at a higher rate than the required data rate. The theoretical download speed of 5G ranges from 1 Gbps to 10 Gbps, and the average download speed of 5G is 50 Mbps and above [3]. Therefore, getting high data rates in 5G is not a big challenge.

Nowadays, the user has more options to select both network provider and service provider. One may leave or join a network or a service provider based on personal or cumulative QoE ratings. Thus, in this user-centric era, the satisfaction and good perception of users are vital for the network providers and service providers to survive in the competitive market. Studying QoE and related user perception has been attracting the attention of researchers, manufacturers, network, and service providers for about two decades. In Reference [4], the authors discuss the QoE management challenge for multimedia applications and also identify future research directions. In References [5–8], the authors have studied video QoE related to network performance parameters, while others also have looked into web QoE [9, 10]. In these studies, the parameters of interest are network performance parameters and their impact on QoE. Despite the advances of networks and mobile devices, maintaining acceptable QoE under the Internet's best-effort networking paradigm is a challenge. The transmitted data suffer from impairments through both transmissions over congested links and error-prone channels. This results in a degradation of user ratings. The most frequent network-induced issues that arise with video transmission are throughput variability, packet loss, packet delay, and packet delay variation. The receiver buffer is used to accommodate throughput variability and jitter. It stores the data temporarily to smooth the throughput variation and the jitter, and provides data with the required data rate to the client. The receiver buffer is also referred to as the jitter buffer in this article.

The packet delay variation in the Internet is very small, and is typically measured in milliseconds. However, in the case of mobile Internet, it may reach the order of seconds, which is jitter of a completely different order of magnitude that has to be taken care of by a properly dimensioned jitter buffer. In Reference [11], the authors measured the one-way delay on the down-link in mobile networks and realised spans of one-way delays (One-Way Delay (OWD)) between less than 50 ms to 2.1 s and beyond. Similar results were reported in [12], in which the authors performed active delay measurements in the 3G HSPA network on the application level. Long delays in the down-link at the application layer in the live stream may empty the jitter buffer. This results in video freezes, which may hamper the user's QoE. Moreover, it may happen that the receiver is not receiving any valid data during periods due to other reasons such as bit errors corrections, handoffs, congestion, etc. In Reference [13], the authors built an argument based upon previous studies that handover delay and packet loss due to path change and binding updates may degrade the QoS. In Reference [14], the authors reported that, from server to UE, some pings experienced long RTTs (>500 ms) as compared to the expected RTTs of 35 ms with a 100% success rate on handover. The consequence of such communication break is a paused play-out until new data are available. During the live streaming, when the sender is sending continuously, data may be lost during such a network outage. In this case with paused play-out, the user may miss multimedia content. This phenomenon is more likely to appear in mobile and wireless networks as compared to wired networks. In the case of video streaming, video freezes degrade the QoE of the user remarkably [15, 16]. Multiple freeze events have a severe impact on QoE, while the impact of data loss also depends on the video content [17].

Along with the evaluation of QoE, various efforts have been made to maintain an acceptable level of QoE. These efforts included the use of different codecs, error recovery schemes, recovery of the lost or dropped frames, jitter buffer, lost packet recovery, and so on. A dynamic transport architecture was proposed for next-generation mobile networks adapted to video service requirements and aiming at improving the QoE [18]. In Reference [19], the authors proposed an intelligent packet drop concept to improve the QoE in case of congestion.

In this article, we propose a solution by using a sender buffer for the live video stream to diminish the impact of the network outages, which may be the result of the delay, hand-off, or any other network interruption, on video QoE. The proposed approach can be implemented within the existing setup and within the available resources.

The first part of the article covers the theoretical modeling, analysis, discussion, and comparison of the proposed solution with existing streaming service configuration. We also discuss the proper dimensioning of the sender buffer size in relation to the jitter buffer.

The second part of the article discusses the experimental result of videos streamed without and with the proposed sender buffer solution, in the presence of network outages. The experimental data are statistically analyzed, in terms of Mean Opinion Scores (MOS) and MOS gains due to sender buffer as well as due to increased transmission rates between sender and receiver buffer.

Moreover, to the best of our knowledge, there is no specific work in which the sender buffer is used only on a need-basis to resolve the network outage problem, nor has there been any study that is addressing the dimensioning of the sender buffer in comparison to the jitter buffer for optimizing QoE, which are the main contributions of this article.

The remainder of the paper is organized as follows: Section 3 presents related work. Section 4 describes artifacts in video delivery and play-out, such as freezes, jumps, and misses. Section 5 elaborates on the video stream, while Section 6 discusses the proposed video stream model. Section 7 explains the role of the jitter buffer in multimedia streaming, and Section 8 figures out the role of the sender buffer in relation to the receiver buffer. Section 9 covers the discussion of the analytical part. Section 10 provides the user ratings of videos for different scenarios and compares the existing and proposed streaming approaches in terms of the user ratings. Finally, Section 11 concludes this article.

# 3 Related Work

The data packets that constitute a live stream over a network may experience varying OWD, which is usually referred to as packet delay variation or jitter [20]. There are many reasons for such kind of behavior: Different queues (at sender, receiver and in the network), network congestion, route changes, or mobile or wireless link layer impairments are major causes for varying OWD of packets. For instance, the error correction of the radio channels usually delivers error-free packets using the Automatic Repeat reQuest (ARQ) process at the price of jitter [21].

In multimedia applications, a stream whose packets are affected by jitter may create artifacts and degrade user perception, due to late arrival of essential information. In Reference [22], the authors used the measured Quality of Service (QoS) distribution to evaluate the distribution of QoE metrics. In Reference [23], the authors investigated the buffer size and video QoE for Netflix and found that improved buffer size can improve the video QoE. In addition, they found that TCP Reno with a smaller buffer is causing higher loss and has a negative impact on applications. In Reference [24], the authors examined latency and throughput in the face of heterogeneous buffer sizes of the network router for live video streaming application, and recommended the middle-of-the-road scope of buffer sizes. As many others, these rather recent publications focus on TCP-based content streaming, as opposed to UDP-based live streaming as addressed by our work.

In Reference [25], the authors studied the effect of jitter on the perceptual video quality and found that even low amounts of jitter or packet loss degrade the video quality severely as compared to the perceptual video quality of the reference video. To overcome the after effects of the jitter, and to smooth the incoming stream, jitter buffer is used. The jitter buffer plays a useful role in multimedia streaming by buffering the variable-rate data and maintains the required data rate for playback. However, the jitter buffer compensates the jitter at the expense of latency. Different efforts have been made to manage the jitter for multimedia applications [26, 27]. In Reference [28], the authors proposed a mechanism by distributing the buffer space requirements more uniformly over the route of an end-to-end path to control the jitter for packet-switching internetworks.

The size of the jitter buffer can vary according to the requirement of the application and the availability of resources. The data in the jitter buffer may shrink or expand. The size of the jitter buffer is a trade-off between the initial play-out delay and packet loss. To reduce the initial delay, a jitter threshold level is set for data in the jitter buffer, from where play-out begins. The jitter threshold level is smaller than the jitter buffer. The threshold level of the jitter buffer is very critical because the initial delay of the video depends on the threshold level. The threshold level of the jitter buffer should not be so small that the player consumes the buffer data during a very short period of time, which may anyway result in a video freeze. Moreover, a comparably high threshold level of jitter buffer may increase the initial delay which may degrade the user experience [29]. Similarly, after a network outage, it may take longer to restart playout, also resulting in a poor user experience.

To maintain the QoE and for smooth play-out of multimedia content on the Internet various efforts have been made besides the use of a jitter buffer. The use of the sender buffer is one of them. In Reference [30], the authors propose a buffer-driven adaptation streaming scheme for a stored video, which scales the quality of transmission, based on both receiver buffer occupancy and sender buffer occupancy, instead of bandwidth. In Reference [31], the authors investigate the double-buffer traffic shaper, implemented using Token and Leaky Bucket techniques, to adjust the video frame rate inflow into the TCP sender buffer of a multimedia application source across a slow-speed link. The authors of [32] introduced a scheduling algorithm based on the sender buffer backlog for real-time application. It schedules time slots based on the sender buffer backlog at the base station, which is believed to be correlated with a play-out buffer backlog at the receiver.

In Reference [33], the authors proposed enhanced transport named paceline to improve the performance and availability of streaming video and time-sensitive media by reducing the TCP queuing delay at the sender-side. To address the shortcoming of the existing error control solution for real-time streaming, the authors of [34] proposed packet reliability-based real-time streaming (Reliability-based Real-time Streaming (RERES)). In this solution, the authors came up with scheduling algorithms for reliable adaptation and buffer control. In Reference [35], the authors proposed to send all the packets via sender buffer using a UDP socket. On receiving negative acknowledgment packets (Negative Acknowledgment (NAK)), based on previous information and the sender buffer level, they decide whether to retransmit the video packet or not. The important factors used for the decision of retransmission of a video packet are the timeout counter and transmission rate.

# 4 Video Freeze, Jump, and Miss Artifacts

During live video transmission over an error-prone channel, data may be lost or delayed. As a result, the user may face temporal artifacts. Along with other temporal artifacts, video freezes, jumps, and misses are the common artifacts of live video transmission. Figure 1 illustrates the freeze, jump, and miss artifacts of video streaming.

**Perfect Video Streaming**: Figure 1a shows video streaming without any artifact like freeze-jump and miss. Initially, the player fills the player buffer before starting the play-out. After that, if the arrival rate of data is not less than the playing rate, the video will play smoothly.

**Video Freeze-and-Jump**: Video freeze-and-jump occurs due to excessive delay or data loss. On regaining the network, the on-demand video starts playing from the frame where it stalled. However, in the case of live transmission, video starts playing from the new frame that is transmitted at that instant. Figure 1b demonstrates the freeze-and-jump case. After receiving the video chunk $r_2$ at time $t_2$, the video stream stopped due to network outage and resumed back at time $t_6$. The player remains halted at the last displayed picture $r_2$ for the duration of $t_2$ to $t_6$. The *freeze-time* is the time duration for which the player remains paused at the last displayed frame, without losing the connection with the server.

In video streaming, there are several causes of data loss, such as intermediate buffer overflow, network outage, etc. As a result of this data loss, the player jumps from the last successfully played frame to the new correctly received frame. Figure 1b explains the jump of the video, after playing $r_5$, the player jumps to $r_9$ in the next time interval. The video *jump-length* is the duration of the skipped video due to data loss, which is a result of a network outage.

**Video Miss**: In live streaming, usually, the video freeze is followed by a video jump. However, when freeze-and-jump both take place together simultaneously, this implies a video miss, depicted in Figure 1c. The *miss-time* is the time interval for which video freeze and jump occur simultaneously.

# 5 Live Video Streaming

Video streaming requires a steady frame rate for smooth play-out. However, during the transmission of the data over the networks, packet arrival time may vary due to congestion, time drift, routing, etc. This irregularity in the arrival time of the packets is called jitter. To combat the jitter, a jitter buffer or receiving buffer is used at the receiving end. The presence of a jitter or application buffer is very common in multimedia applications [36]. However, the jitter buffer introduces some initial delay equal to the time length of the jitter in the play-out. Larger jitter buffers provide more tolerance to the streaming applications, as they have enough data to support the

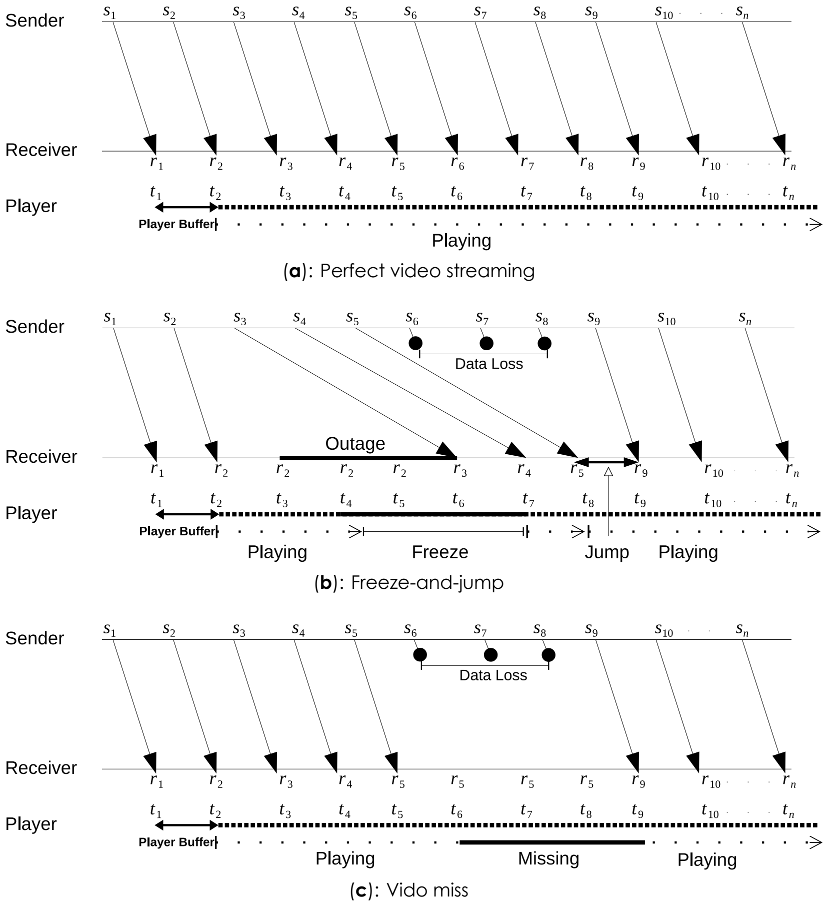



**Figure 1:** Description of (**a**): perfect video streaming; (**b**): video freeze-and-jump; (**c**): video miss.

play-out during the network disturbance. Nevertheless, a larger jitter buffer increases the initial delay, and the user needs to wait extra time during re-filling of the larger buffer.

A jitter buffer is measured in time units (typically milliseconds) and can be sized to the number of video frames. Let $\tau_{\mathrm{j}}$ represent the jitter buffer in time units. In the presence of $\tau_{\mathrm{j}}$, the end-user will face an initial delay $d_{\mathrm{i}}$ equal to $\tau_{\mathrm{j}}$, when starting the video play-out. The jitter buffer usually has a minimum threshold level, from which the play-out begins. However, for the sake of simplicity of the discussion, the size of jitter buffer is considered equal to the player buffer threshold to start the playing.

Figure 2 shows the video streaming model, which includes the camera attached to the sender and the jitter buffer players. Let us assume that the receiver is receiving the video from the sender with the data rate $R_{\mathrm{sr}}$, and the jitter buffer can be filled with this data rate in time $\tau_{\mathrm{j}}$. Then, the size of jitter buffer is given as

$$X_{\mathrm{sr}}^{\tau_{\mathrm{j}}} = R_{\mathrm{sr}}\tau_{\mathrm{j}} \tag{V.1}$$

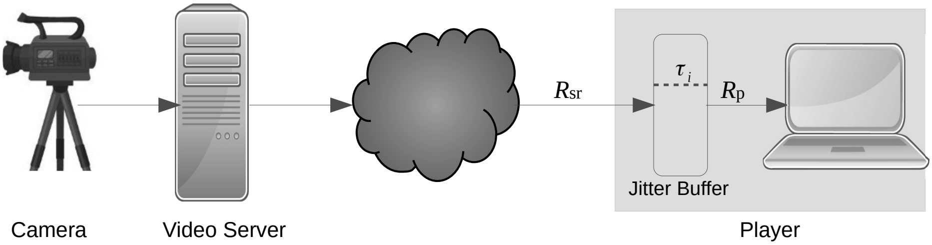



**Figure 2:** Video streaming model.

The player starts consuming the data from the jitter buffer after the initial delay $d_{\mathrm{i}}$. The initial delay is the time needed to fill the jitter buffer up to threshold level $\tau_{\mathrm{j}}$. Once the jitter buffer is full and ready, the player starts consuming the data. Let $R_{\mathrm{p}}$ be the video consumption rate, i.e., the decoding rate for smooth play-out. For smooth play-out, $R_{\mathrm{sr}}$ should not be smaller than $R_{\mathrm{p}}$; otherwise, the jitter buffer may underrun. Consequently, the video will freeze until re-buffering is over. In this case, play-out and freeze occur periodically.

Moreover, the jitter buffer may underrun due to *network outage*. The latter is a temporary downtime of network, during which the data flow is interrupted. It may occur due to network performance issues, such as loss of link, delay, handover, packet loss, etc. The *outage duration* $\tau_{\mathrm{o}}$ refers a time-period in which the network fails to deliver valid data to the receiver.

Due to a network outage and a subsequent underrun of the jitter buffer, the video will freeze at play-out. When a channel becomes available again after a network outage, the restart of the play-out depends on a number of factors, such as packet loss, loss recovery, and loss flexibility of the codec. The authors of [7] have discussed these factors as Quality of Delivery (QoD) and how the QoD has an impact on the QoE. The play-out may restart from the freeze frame or jump to the new location of video by skipping the video equal to the duration of the outage. If the retransmission is possible, then the video play-out restarts from the freezing point. Otherwise, play-out restarts from the newly arrived frame, i.e., there is no retransmission. For example, in live transmission, after the network outage, there is no option to retransmit the data transmitted during the network outage unless it was saved for retransmission at the sender. In case of live transmission, network outage may entail video freeze-and-jump. In this case, the *video freeze duration* $F_{\Delta t}$ and amount of *data loss* $L$ are directly proportional to the outage duration.

The video freeze-and-jump has an adverse impact on the user QoE [15, 16, 29]. In case of a video jump, users may feel to be deprived for missing video content of an important moment, e.g., a goal or an attempt for goal in a soccer match, and this may degrade the user QoE. The length of freeze-and-jump depends on the duration of outage. To reduce the duration of these artifacts and to improve the QoE of the user, we propose a buffer at the sender side to save the data during the network outage.

Before discussing the proposed model, we present the variables and related key relations used in this article in Table 1.

# 6 Proposed Video Streaming Model

Figure 3 illustrates the proposed video streaming model, which consists of a receiver and a proposed sender buffer. The sender sends the data directly to the receiver while the sender buffer is used under adverse conditions. In case of network outage, the sender's data are stored in the sender buffer. After the outage, when the channel gets available again, the sender buffer provides the saved data for quick rebuffering of the jitter buffer, which may be empty due to the outage. For quick rebuffering of the jitter buffer, it is assumed that the sender transmits the buffered data with maximum possible capacity

$$C = nR_{\text{sr}} \text{ where } n > 1 \tag{V.2}$$

If the sender fails to send the buffer data with higher speed, i.e., if $n = 1$, then the sender buffer may not be emptied at the end of the transmission. In this case, it is up to the user to play delayed video or skip the sender buffer content to watch the video in real time. Thus, a higher data speed $(n > 1)$ is required to empty the sender buffer and get back the video in real time.

**Table 1:** Terminology used.

| Variable | Description | Expression |
|---|---|---|
| $C$ | Capacity of the channel between sender and receiver | $C = nR_{\text{sr}}$ for $n > 1$ |
| $F_{\Delta t}$ | Video freeze duration due to outage | |
| $L$ | Data loss during the outage | |
| $n$ | Capacity factor | |
| $R_{\text{sr}}$ | Data rate from sender to receiver for smooth play-out | |
| $R_{\text{p}}$ | Player data consumption rate for smooth play-out | |
| $t_{\text{e}}(\tau_{\text{s}})$ | Time to empty the sender buffer at data rate $C$ | $t_{\text{e}}(\tau_{\text{s}}) = \frac{\tau_{\text{s}}}{n-1}$ for $n > 1$ |
| $t_{\text{f}}(\tau_{\text{j}})$ | Time to fill the jitter buffer at data rate $nR_{\text{sr}}$ | $t_{\text{f}}(\tau_{\text{j}}) = \frac{\tau_{\text{j}}}{n}$ |
| $\tau_{\text{o}}$ | Network outage duration | |
| $\tau_{\text{j}}$ | Time to fill the jitter buffer at data rate $R_{\text{sr}}$ | |
| $\tau_{\text{s}}$ | Time to fill the sender buffer at data rate $R_{\text{sr}}$ | |
| $X_{sr}(\tau_{\text{j}})$ | Data buffering capacity of the jitter buffer at data rate $R_{\text{sr}}$ | $X_{sr}(\tau_{\text{j}}) = R_{\text{sr}}\tau_{\text{j}}$ |
| $X_{sr}(\tau_{\text{s}})$ | Data buffering capacity of the sender buffer at data rate $R_{\text{sr}}$ | $X_{sr}(\tau_{\text{s}}) = R_{\text{sr}}\tau_{\text{s}}$ |
| $X_{\tau_{\text{s}}}(\tau_{\text{o}})$ | Data sent to the sender buffer during the network outage | $X_{\tau_{\text{s}}}(\tau_{\text{o}}) = R_{\text{sr}}\tau_{\text{o}}$ |

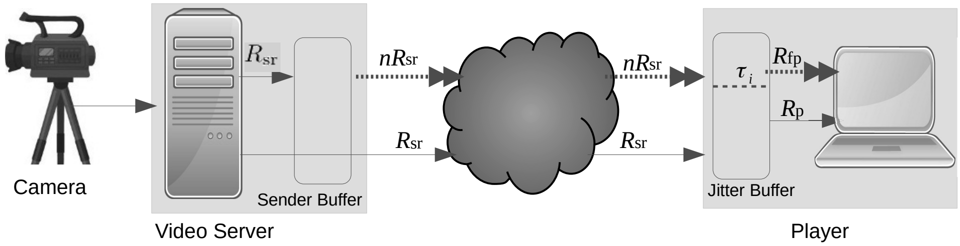



**Figure 3:** Proposed video streaming model with sender buffer.

The presence of the sender buffer helps to save the data during the outage and reduces the video freeze time, which is described through the following key formulae for *data loss* (V.3) and *freeze time* (V.4)

$$L = R_{\text{sr}}(\tau_{\text{o}} - \min\{\tau_{\text{s}}, \tau_{\text{j}}\}) \text{ where } \tau_{\text{o}} \geq \min\{\tau_{\text{s}}, \tau_{\text{j}}\} \tag{V.3}$$

$$F_{\Delta t} = \tau_{\text{o}} - \min\{\tau_{\text{s}}, \tau_{\text{j}}\}\Big(1 - \frac{1}{n}\Big) \text{ where } \tau_{\text{o}} \geq \min\{\tau_{\text{s}}, \tau_{\text{j}}\} \tag{V.4}$$

The derivation and discussion of Equations (V.3) and (V.4) will follow in Sections 8 and 9, respectively.

There will be no data loss if the outage duration $\tau_{\text{o}}$ is smaller than or equal to $\tau_{\text{s}}$. However, if $\tau_{\text{o}}$ is greater than $\tau_{\text{s}}$, then data will be lost due to overflow of the sender buffer. Moreover, any network outage $\tau_{\text{o}}$ greater than the jitter buffer $\tau_{\text{j}}$ results in a video freeze equal to $\tau_{\text{o}}$. The minimum duration of the video freeze the user has to face is the time to re-fill the empty jitter buffer to start the play-out. After the outage, this freeze duration can be reduced by flushing the sender buffer data with maximum channel capacity $C$. In this way, the sender buffer is used for saving the data during the outage and reducing the freeze duration after the outage.

In the following sections, we will quantify and discuss the usefulness of sender buffer, along with the jitter buffer, in detail. In Section 7, we will discuss the role of the jitter buffer in multimedia streaming without sender buffer using Equations (V.3) and (V.4). In addition, Sections 8 and 9 will estimate the size of the sender buffer corresponding to the size of the jitter buffer.

# 7 The Role of the Jitter Buffer in Live Multimedia Streaming

To overcome the after effects of the jitter, and to smooth the incoming stream, an application jitter buffer is used. The size of the jitter buffer can vary according to the requirement of the application and the availability of resources. The data in the jitter buffer may shrink or expand. As mentioned earlier in this article, the size of the jitter buffer is considered equal to the threshold of the jitter buffer from where the play-out starts.

In case there was only the jitter buffer, then $\tau_\mathrm{s} = 0$. If $\tau_\mathrm{o} < \tau_\mathrm{j}$, then, during the network outage, the jitter buffer will not become empty, and the user will not see any impact on play-out. On the other hand, if $\tau_\mathrm{o} \geq \tau_\mathrm{j}$, then, during the network outage, the jitter buffer will dry-out, and the user will face data loss and video freeze.

With $\tau_\mathrm{s} = 0$ in (V.3), we obtain

$$L = R_\mathrm{sr}(\tau_\mathrm{o} - \min\{0, \tau_\mathrm{j}\}) = R_\mathrm{sr}\tau_\mathrm{o} \text{ for } \tau_\mathrm{o} \geq \tau_\mathrm{j}, \text{i.e., } L \propto \tau_\mathrm{o} \qquad \text{(V.5)}$$

The data loss is proportional to the network outage. Thus, all data sent during the network outage will be lost.

Similarly, by using Equation (V.4), the video freeze duration can be calculated as

$$F_{\Delta t} = \tau_\mathrm{o} - \min\{0, \tau_\mathrm{j}\}\Big(1 - \frac{1}{n}\Big) = \tau_\mathrm{o} \text{ for } \tau_\mathrm{o} \geq \tau_\mathrm{j}, \text{i.e., } F_{\Delta t} \propto \tau_\mathrm{o} \qquad \text{(V.6)}$$

which shows that the video freeze duration will be equal to that of the network outage.

In the absence of the sender buffer, the network outage results in the jump-and-freeze artifact, the length of which is determined by the duration of the network outage.

# 8 The Role of the Proposed Sender Buffer in Live Multimedia Streaming

The proposed sender buffer at the sender provides the option to save the data during the network outage. During a normal flow of data, the sender sends the data directly to the receiver. However, during a network outage, the sender saves the data in sender buffer. After the network outage, the sender sends that data through the sender buffer unless the sender buffer empties. As a result, the user receives delayed data until the sender buffer is empty. It is important to empty the sender buffer as quickly as possible to get the live stream back.

To recover the live stream in a shorter time, it is necessary to flush the sender buffer data with maximum channel capacity. In this way, the sender buffer helps to save the data and reduce the video freeze duration by quickly re-filling the jitter buffer. Ultimately, it improves the user video QoE. Thus, the outlet data rate must be greater than the inlet data rate of the sender buffer. Otherwise, if the inlet and outlet data rates of the sender buffer are the same, then the user may not get the live data stream back. If the sender buffer holds data, then the sender sends data through the sender buffer. As a result, the sender buffer will not empty, and the user will face latency equal to the sender buffer time, i.e., $\tau_\mathrm{s}$.

If a higher data rate is not achievable, then the user may have to skip the sender buffer data to get the live stream. Hence, a higher bit rate for the sender buffer data is an important and required factor in this proposed solution. Otherwise, no timely recovery is possible.

In live transmission, the sender continuously transmits data regardless of the network status. The sender buffer saves the data during the network outage and reduces the freeze time after the network outage. During the network outage, the player continues the play-out unless the jitter buffer dries out. The corresponding video freeze starts at this moment and ends when the jitter buffer gets re-filled. The video freezes duration ($F_{\Delta t}$) depends on the network outage duration ($\tau_{\mathrm{o}}$), time length of jitter buffer ($\tau_{\mathrm{j}}$), and time to re-buffer the jitter buffer $t_{\mathrm{f}}(\tau_{\mathrm{j}})$:

$$F_{\Delta t} = \tau_{\mathrm{o}} - \tau_{\mathrm{j}} + t_{\mathrm{f}}(\tau_{\mathrm{j}}) \tag{V.7}$$

During the network outage, data may be lost at the sender side when the network outage duration is bigger than $\tau_{\mathrm{s}}$. Similarly, after the network outage, data may be lost at the receiver side, especially when $\tau_{\mathrm{j}} < \tau_{\mathrm{s}}$ and data are sent with high speed $C > R_{\mathrm{p}}$, i.e., the player consumption rate. The total data loss ($L$) is the sum of data loss at the sender ($L_{\mathrm{s}}$) and the data loss at the receiver ($L_{\mathrm{r}}$), given by

$$L = L_{\mathrm{s}} + L_{\mathrm{r}} \tag{V.8}$$

After the network outage, the amount of time it takes until the play-out restarts depends on the time to fill up the jitter buffer

$$t_{\mathrm{f}}(\tau_{\mathrm{j}}) = \frac{\tau_{\mathrm{j}}}{n} \tag{V.9}$$

Similarly, after the network outage, it is also important how fast the user gets to see the live stream. Obviously, this depends on the time it takes to empty the sender buffer

$$t_{\mathrm{e}}(\tau_{\mathrm{s}}) = \frac{\tau_{\mathrm{s}}}{n-1} \tag{V.10}$$

Let $\Delta t$ denote the time to get back the live stream after the play-out restarts. In the absence of the sender buffer, the time to start the smooth play-out is equal to $\tau_{\mathrm{j}}$. Thus, in this case,

$$\Delta t = \tau_{\mathrm{j}} \ \text{ when } \ \tau_{\mathrm{s}} = 0 \tag{V.11}$$

However, in the presence of the sender buffer, $\Delta t$ depends on the time to empty the sender buffer ($t_{\mathrm{e}}(\tau_{\mathrm{s}})$) and the time to re-fill the jitter buffer ($t_{\mathrm{f}}(\tau_{\mathrm{j}})$):

$$\Delta t = |t_{\mathrm{e}}(\tau_{\mathrm{s}}) - t_{\mathrm{f}}(\tau_{\mathrm{j}})| \tag{V.12}$$

The performance of the proposed sender buffer in the streaming setup can be evaluated in terms of the freeze-time ($F_{\Delta t}$), total data loss ($L$), and the time to get back the live stream ($\Delta t$).

As discussed, the network outage has a direct impact on the video jump and the video freeze. The presence of the sender buffer helps in reducing the length of the video jump and the video freeze. The duration of the network outage $\tau_{\mathrm{o}}$ may vary with respect to the size of the sender buffer $\tau_{\mathrm{s}}$ and the jitter buffer $\tau_{\mathrm{j}}$. However, the end-user will face the artifacts in the video only when $\tau_{\mathrm{o}}$ is greater than or equal to $\tau_j$. Along with the channel capacity, the performance of the proposed setup depends on the right combination of the sender buffer and the jitter buffer. With the assumption that $\tau_{\mathrm{o}} \geq \tau_{\mathrm{j}}$, the following three possible combinations are of interest:

1. The sender buffer is larger than the jitter buffer;
2. The jitter buffer is larger than the sender buffer;
3. Sender and jitter buffer are equally sized.

## 8.1 Sender Buffer Is Larger Than Jitter Buffer

With given $\tau_{\mathrm{s}} > \tau_{\mathrm{j}}$, there are two possibilities with respect to the network outage. Either, the network outage is greater than or equal to the sender buffer, i.e.,

$$\tau_{\mathrm{o}} \geq \tau_{\mathrm{s}} > \tau_{\mathrm{j}} \tag{V.13}$$

or the network outage is smaller than the sender buffer but greater than the jitter buffer, i.e.,

$$\tau_{\mathrm{s}} \geq \tau_{\mathrm{o}} > \tau_{\mathrm{j}} \tag{V.14}$$

Let us evaluate the performance of these cases in terms of video freeze duration, data loss, and time to get back the live stream.

**Video Freeze Duration ($F_{\Delta t}$)**: In both cases, either $\tau_{\mathrm{o}} \geq \tau_{\mathrm{s}}$ or $\tau_{\mathrm{o}} < \tau_{\mathrm{s}}$, the $F_{\Delta t}$ depends on $\tau_{\mathrm{o}}$, $\tau_{\mathrm{j}}$, and $t_{\mathrm{f}}(\tau_{\mathrm{j}})$ as given in Equation (V.7). The data capacity of the jitter buffer is given as

$$X_{\mathrm{sr}}(\tau_{\mathrm{j}}) = R_{\mathrm{sr}}\tau_{\mathrm{j}} \tag{V.15}$$

After the network outage, the jitter buffer fills with data rate $C$. Thus, $t_{\mathrm{f}}(\tau_{\mathrm{j}})$ is given as

$$t_{\mathrm{f}}(\tau_{\mathrm{j}}) = \frac{X_{\mathrm{sr}}(\tau_{\mathrm{j}})}{C} = \frac{R_{\mathrm{sr}}\tau_{\mathrm{j}}}{nR_{\mathrm{sr}}} = \frac{\tau_{\mathrm{j}}}{n}$$

Thus, Equation (V.7) becomes

$$F_{\Delta t} = \tau_{\mathrm{o}} - \tau_{\mathrm{j}} + \frac{\tau_{\mathrm{j}}}{n} = \tau_{\mathrm{o}} - \tau_{\mathrm{j}}\Big(1 - \frac{1}{n}\Big) \quad \text{where} \quad \tau_{\mathrm{s}} > \tau_{\mathrm{j}} \tag{V.16}$$

**Data Loss ($L$)**: If $\tau_{\mathrm{o}} > \tau_{\mathrm{s}}$, then the data loss on the sender side equals the difference between the data sent during the network outage and data saved in the sender buffer:

$$L_{\mathrm{s}} = R_{\mathrm{sr}}\tau_{\mathrm{o}} - R_{\mathrm{sr}}\tau_{\mathrm{s}} \quad \text{where} \quad \tau_{\mathrm{s}} > \tau_{\mathrm{j}} \quad \text{and} \quad \tau_{\mathrm{o}} > \tau_{\mathrm{s}} \tag{V.17}$$

After a network outage, the sender buffer data are sent to the jitter buffer with a higher data rate ($C$). The sender buffer data are bigger than the size of the jitter buffer, so the data loss at the jitter buffer is given as

$$L_{\mathrm{r}} = R_{\mathrm{sr}}\tau_{\mathrm{s}} - R_{\mathrm{sr}}\tau_{\mathrm{j}} \;\; \text{where} \;\; \tau_{\mathrm{s}} > \tau_{\mathrm{j}} \tag{V.18}$$

By using Equations (V.17) and (V.18), Equation (V.8) becomes

$$L = R_{\mathrm{sr}}(\tau_{\mathrm{o}} - \tau_{\mathrm{j}}) \tag{V.19}$$

Now, if $\tau_{\mathrm{o}} < \tau_{\mathrm{s}}$, then there will be no data loss at the sender as the data sent by the sender during the network outage is less than the size of the sender buffer. Thus,

$$L_{\mathrm{s}} = 0 \tag{V.20}$$

After the network outage, the data may be lost from the jitter buffer, as the data stored in the sender buffer are greater than the size of the jitter buffer:

$$L_{\mathrm{r}} = R_{\mathrm{sr}}\tau_{\mathrm{o}} - R_{\mathrm{sr}}\tau_{\mathrm{j}} \;\; \text{where} \;\; \tau_{\mathrm{s}} > \tau_{\mathrm{j}} \;\; \text{and} \;\; \tau_{\mathrm{s}} > \tau_{\mathrm{o}} \tag{V.21}$$

With Equations (V.20) and (V.21), Equation (V.8) becomes

$$L = R_{\mathrm{sr}}(\tau_{\mathrm{o}} - \tau_{\mathrm{j}}) \tag{V.22}$$

**Latency to the Live Stream ($\Delta t$)**: $\Delta t$ defines the time to get back the live stream after the play-out restarts and is defined in Equation (V.12). As the sender buffer is bigger than the jitter buffer,

$$\begin{aligned} \tau_{\mathrm{s}} > \tau_{\mathrm{j}} \Rightarrow \frac{\tau_{\mathrm{s}}}{n-1} > \frac{\tau_{\mathrm{j}}}{n}&, \text{then using Equations (V.10) and (V.9)} \\ \Rightarrow t_{\mathrm{e}}(\tau_{\mathrm{s}}) > t_{\mathrm{f}}(\tau_{\mathrm{j}})& \end{aligned} \tag{V.23}$$

A larger sender buffer takes more time to empty as compared to the re-buffering of the jitter buffer. As a result, the latter delays the live stream even beyond the end of the very outage.

As the sender buffer is larger than the jitter buffer, it saves more data during the network outage as compared to the size of the jitter buffer. However, after the network outage, when data are transmitted at the higher data rate, the jitter buffer may overflow. This data loss occurs after saving and transmitting the data, which is not an efficient approach.

## 8.2 Sender Buffer Is Smaller Than Jitter Buffer

In this section, we will analyze the performance of the sender buffer in the proposed video streaming scenario when the sender buffer is smaller than the jitter buffer, i.e., $\tau_{\mathrm{s}} < \tau_{\mathrm{j}}$. With respect to the network outage, there are two possibilities: Either the network outage is smaller than the jitter buffer and greater than or equal to the sender buffer, i.e.,

$$\tau_{\mathrm{s}} \leq \tau_{\mathrm{o}} < \tau_{\mathrm{j}} \tag{V.24}$$

or the network outage is greater than or equal to the jitter buffer, i.e.,

$$\tau_{\mathrm{s}} < \tau_{\mathrm{j}} \leq \tau_{\mathrm{o}} \tag{V.25}$$

**Video Freeze Duration ($F_{\Delta t}$)**: Let us first consider that the duration of the network outage $\tau_{\mathrm{o}}$ is smaller than the size of the jitter buffer and greater than or equal to the size of sender buffer. As $\tau_{\mathrm{o}} < \tau_{\mathrm{j}}$, the jitter buffer will not dry-out during the network outage, so there will be no video freeze.

Now, let us consider that the network outage is bigger than the sender buffer and the jitter buffer. In this case, the jitter buffer dries out during the network outage and the user faces a video freeze. As shown in Equation (V.7), the video freeze duration depends on $\tau_{\mathrm{o}}$, $\tau_{\mathrm{j}}$ and $t_{\mathrm{f}}(\tau_{\mathrm{j}})$. During the network outage, the data saved in the sender buffer are not sufficient to fill the jitter buffer, as $\tau_{\mathrm{s}} < \tau_{\mathrm{j}}$. Thus,

$$t_{\mathrm{f}}(\tau_{\mathrm{j}}) = \frac{\tau_{\mathrm{s}}}{n} + (\tau_{\mathrm{j}} - \tau_{\mathrm{s}}) \tag{V.26}$$

Equation (V.7) becomes

$$F_{\Delta t} = \tau_{\mathrm{o}} - \tau_{\mathrm{j}} + \frac{\tau_{\mathrm{s}}}{n} + (\tau_{\mathrm{j}} - \tau_{\mathrm{s}}) = \tau_{\mathrm{o}} - \tau_{\mathrm{s}}\Big(1 - \frac{1}{n}\Big) \text{ where } \tau_{\mathrm{s}} < \tau_{\mathrm{j}} \tag{V.27}$$

**Data Loss ($L$)**: If $\tau_{\mathrm{s}} \leq \tau_{\mathrm{o}} < \tau_{\mathrm{j}}$, then the data sent by the sender during the network outage exceeds the size of the sender buffer. Thus,

$$L_{\mathrm{s}} = R_{\mathrm{sr}}\tau_{\mathrm{o}} - R_{\mathrm{sr}}\tau_{\mathrm{s}} \text{ where } \tau_{\mathrm{j}} > \tau_{\mathrm{o}} \geq \tau_{\mathrm{s}} \tag{V.28}$$

However, there will be no loss at the receiver because the jitter buffer is bigger than the sender buffer, i.e.,

$$L_{\mathrm{r}} = 0 \tag{V.29}$$

Then, the total loss amounts to

$$L = R_{\mathrm{sr}}\tau_{\mathrm{o}} - R_{\mathrm{sr}}\tau_{\mathrm{s}} = R_{\mathrm{sr}}(\tau_{\mathrm{o}} - \tau_{\mathrm{s}}) \text{ where } \tau_{\mathrm{s}} < \tau_{\mathrm{j}} \tag{V.30}$$

Now, if the duration of the network outage is larger than both the sender buffer and the jitter buffer, then, during the network outage, there will be data loss at the sender, which is given as

$$L_{\mathrm{s}} = R_{\mathrm{sr}}\tau_{\mathrm{o}} - R_{\mathrm{sr}}\tau_{\mathrm{s}} \text{ where } \tau_{\mathrm{o}} \geq \tau_{\mathrm{j}} > \tau_{\mathrm{s}} \tag{V.31}$$

After the network outage, there will be no data loss at the receiver because the data stored in the sender buffer are less than the size of the jitter buffer:

$$L_{\mathrm{r}} = 0 \tag{V.32}$$

so the total loss amounts to

$$L = R_{\mathrm{sr}}(\tau_{\mathrm{o}} - \tau_{\mathrm{s}}) \text{ where } \tau_{\mathrm{s}} < \tau_{\mathrm{j}} \tag{V.33}$$

**Latency to the Live Stream (**$\Delta t$**)**: As the sender buffer is smaller than the jitter buffer,

$$\tau_{\mathrm{s}} < \tau_{\mathrm{j}} \Rightarrow \frac{\tau_{\mathrm{s}}}{n-1} < \frac{\tau_{\mathrm{j}}}{n} \cdot \frac{n}{n-1} \tag{V.34}$$

Using Equations (V.9) and (V.10), we obtain

$$t_{\mathrm{e}}(\tau_{\mathrm{s}}) < t_{\mathrm{f}}(\tau_{\mathrm{j}}) \left(\frac{n}{n-1}\right) \quad (n > 1) \tag{V.35}$$

The latter inequality indicates the condition under which a smaller sender buffer may take less time to empty as compared to the time needed for re-buffering. Consequently, the user may need to wait to restart the play-out even after receiving the data from the sender buffer. In any case, the user will get to see the live stream in less time as compared to the assumption discussed in Section 8.1.

As the sender buffer is smaller than the jitter buffer, the sender buffer may not hold enough data to re-buffer the jitter buffer after the network outage. Unlike the above case, after the network outage, no data loss occurs at the receiving side. However, during the network outage, data may be lost at the sender side. Indeed, longer freezes or waiting to restart the play-out may affect the user's QoE. Thus, a sender buffer that is smaller than the jitter buffer does not save enough data during the network outage that can help to resume the play-out faster after the network outage.

## 8.3 Sender Buffer Size Is Equal to Jitter Buffer Size

Besides the benefit of the sender buffer, we have seen the shortcomings of having it sized smaller or larger than the jitter buffer. Hence, based on the above discussion, it is more convincing to argue that, for the optimum results of the proposed streaming method, the size of both buffers should be equal, denoted by $\tau_{\mathrm{sj}}$.

In that particular case, the network outage will be either greater or smaller than these buffers. If the network outage is shorter than the corresponding buffer time $\tau_{\mathrm{sj}}$, then there will be no data loss or video freeze. However, if the network outage is greater than both jitter and sender buffer, then video freeze, data loss, and time to get back the live stream can be calculated as follows.

**Video Freeze Duration (**$F_{\Delta t}$**)**: As $\tau_{\mathrm{o}} > \tau_{\mathrm{sj}}$, the receiver buffer be will drained during the outage. As a result of this, the video will freeze and the freeze time can be calculated using Equation (V.4):

$$F_{\Delta t} = \begin{cases} \tau_{\mathrm{o}} - \tau_{\mathrm{sj}}\Big(1 - \dfrac{1}{n}\Big) & \text{if } \tau_{\mathrm{o}} > \tau_{\mathrm{sj}} \\ \dfrac{\tau_{\mathrm{sj}}}{n} & \text{if } \tau_{\mathrm{o}} = \tau_{\mathrm{sj}} \end{cases} \tag{V.36}$$

**Data Loss ($L$)**: As both the sender buffer and the jitter buffer are equal, we consider $\tau_{\mathrm{o}} \geq \tau_{\mathrm{sj}}$:

$$\Rightarrow R_{\mathrm{sr}}\tau_{\mathrm{o}} \geq R_{\mathrm{sr}}\tau_{\mathrm{sj}} \Rightarrow X_{\mathrm{sr}}^{\tau_{\mathrm{o}}} \geq X_{\mathrm{sr}}^{\tau_{\mathrm{sj}}} \tag{V.37}$$

The data sent by the sender during the network outage are greater than or equal to the sender buffer. As a result, the sender buffer may overflow. The data loss is the difference of the data sent during the network outage and data buffered by the sender buffer, and is given as

$$L_{\mathrm{s}} = X_{\mathrm{sr}}^{\tau_{\mathrm{o}}} - X_{\mathrm{sr}}^{\tau_{\mathrm{sj}}} = R_{\mathrm{sr}}(\tau_{\mathrm{o}} - \tau_{\mathrm{sj}}) \tag{V.38}$$

The data sent by the sender during after the network outage is equal to the size of the jitter buffer, so there is no more additional data loss at the receiver, i.e.,

$$L_{\mathrm{r}} = 0 \tag{V.39}$$

The total data loss is the sum of losses at the sender side and the jitter side.:

$$L = L_{\mathrm{s}} + L_{\mathrm{r}} = R_{\mathrm{sr}}(\tau_{\mathrm{o}} - \tau_{\mathrm{sj}}) \tag{V.40}$$

**Latency to the Live Stream ($\Delta t$)**: Equation (V.12) defines the time to get back the live stream ($\Delta t$) after the end of the outage. With the time to empty the sender buffer given by Equation (V.10) and the time to re-fill the jitter buffer given by Equation (V.9), we obtain

$$\Delta t = \frac{\tau_{\mathrm{sj}}}{n-1} - \frac{\tau_{\mathrm{sj}}}{n} = \frac{\tau_{\mathrm{sj}}}{n(n-1)} \tag{V.41}$$

Obviously, the data saved in the sender buffer helps to re-fill the jitter buffer faster, the larger that the capacity factor $n$ becomes. Consequently, the time required to get the live stream back ($\Delta t$) decreases quickly as $n$ grows.

# 9 Discussion

Section 8 discussed the role of the sender buffer in the proposed streaming setup. Clearly, the sender buffer helps to reduce the data loss ($L$), the video freeze duration ($F_{\Delta t}$) as well as the latency to the live stream ($\Delta t$) after the network outage ($\tau_{\mathrm{o}}$). To find out the balance between the size of the sending buffer and the size of the jitter buffer, in order to obtain the best results in terms of $L$, $F_{\Delta t}$, and $\Delta t$, we will analyze and compare the results discussed in Sections 8.1–8.3.

## 9.1 Data Loss

The equations that discuss the data loss under different conditions with respect to the sender buffer, the jitter buffer, and outage are given in Sections 8.1–8.3. These equations are

$$L = R_{\text{sr}}(\tau_{\text{o}} - \tau_{\text{j}}) \text{ where } \tau_{\text{o}} > \tau_{\text{s}} > \tau_{\text{j}} \tag{V.19}$$

$$L = R_{\text{sr}}(\tau_{\text{o}} - \tau_{\text{j}}) \text{ where } \tau_{\text{s}} > \tau_{\text{o}} > \tau_{\text{j}} \tag{V.22}$$

$$L = R_{\text{sr}}(\tau_{\text{o}} - \tau_{\text{s}}) \text{ where } \tau_{\text{s}} < \tau_{\text{j}} \tag{V.33}$$

$$L = R_{\text{sr}}(\tau_{\text{o}} - \tau_{\text{sj}}) \text{ where } \tau_{\text{s}} = \tau_{\text{j}} = \tau_{\text{sj}} \tag{V.40}$$

Equations (V.19), (V.22), and (V.33) show that the loss depends on the network outage duration $\tau_{\text{o}}$ and the smaller of either the sender or the jitter buffer Equations (V.19), (V.22), (V.33) and (V.40) can be summarized as

$$L = R_{\text{sr}}(\tau_{\text{o}} - \min\{\tau_{\text{s}}, \tau_{\text{j}}\}) \text{ where } \tau_{\text{o}} \geq \min\{\tau_{\text{s}}, \tau_{\text{j}}\} \tag{V.42}$$

which is the same as the proposed Equation (V.3). Considering equally sized buffers ($\tau_{\text{sj}}$), Equation (V.42) simplifies to

$$L = R_{\text{sr}}(\tau_{\text{o}} - \tau_{\text{sj}}) \text{ where } \tau_{\text{o}} \geq \tau_{\text{sj}} \tag{V.43}$$

Taking the partial derivative of Equation (V.43) w.r.t $\tau_{\text{sj}}$ yields

$$\frac{\partial L}{\partial \tau_{\text{sj}}} = -R_{\text{sr}} \tag{V.44}$$

Equation (V.44) shows the rate of change of loss with respect to the jitter buffer or the sender buffer. The negative sign of the data rate indicates that larger buffers result in less data loss and vice versa. Furthermore, taking the partial derivative of Equation (V.43) w.r.t $\tau_{\text{o}}$ yields

$$\frac{\partial L}{\partial \tau_{\text{o}}} = R_{\text{sr}} \tag{V.45}$$

Equation (V.45) shows the rate of change of data loss with respect to the network outage. The positive sign of the data rate shows that the data loss is directly proportional to the network outage, i.e., if $\tau_{\text{o}}$ increases, the loss will increase and vice versa.

## 9.2 Video Freeze Duration

In Sections 8.1–8.3, the video freeze duration is described by

$$F_{\Delta t} = \tau_{\text{o}} - \tau_{\text{j}}\Big(1 - \frac{1}{n}\Big) \text{ where } \tau_{\text{s}} > \tau_{\text{j}} \tag{V.16}$$

$$F_{\Delta t} = \tau_{\text{o}} - \tau_{\text{s}}\Big(1 - \frac{1}{n}\Big) \text{ where } \tau_{\text{s}} < \tau_{\text{j}} \tag{V.27}$$

$$F_{\Delta t} = \tau_{\text{o}} - \tau_{\text{j}}\Big(1 - \frac{1}{n}\Big) \text{ where } \tau_{\text{s}} = \tau_{\text{sj}} = \tau_{\text{sj}} \tag{V.36}$$

Equation (V.16) shows that, when the jitter buffer is smaller than the sender buffer, then the freeze time depends on the network outage duration, jitter butter size, and capacity factor $n$. In comparison to this, Equation (V.27) indicates that sender and jitter buffer have basically changed their roles, whereas, in Equation (V.36), any buffer size can be used. Obviously, the video freeze duration is determined by the the smaller one of the two buffers:

$$F_{\Delta t} = \tau_{\mathrm{o}} - \min(\tau_{\mathrm{s}}, \tau_{\mathrm{j}})\Big(1 - \frac{1}{n}\Big) \text{ where } \tau_{\mathrm{o}} \geq \min(\tau_{\mathrm{s}}, \tau_{\mathrm{j}}) \tag{V.46}$$

The latter equation is identical to the proposed Equation (V.4). With $\tau_{\mathrm{sj}}$ representing the same-sized sender and jitter buffers, the above Equation (V.46) simplifies to

$$F_{\Delta t} = \tau_{\mathrm{o}} - \tau_{\mathrm{sj}}\Big(1 - \frac{1}{n}\Big) \text{ where } \tau_{\mathrm{o}} \geq \tau_{\mathrm{sj}} \tag{V.47}$$

Taking the partial derivative of Equation (V.47) w.r.t $\tau_{\mathrm{sj}}$ yields

$$\frac{\partial F_{\Delta t}}{\partial \tau_{\mathrm{sj}}} = -\Big(1 - \frac{1}{n}\Big) \tag{V.48}$$

The rate at which the duration of video freeze changes as a function of the size of the jitter buffer is determined by the capacity factor $n$. When $\tau_{\mathrm{sj}}$ increases by one unit, then $F_{\Delta t}$ decreased by $(1 - \frac{1}{n})$ if $n > 1$; otherwise,

$$\frac{\partial F_{\Delta t}}{\partial \tau_{\mathrm{sj}}} = 0 \text{ for } n = 1 \tag{V.49}$$

Taking the partial derivative of Equation (V.46) w.r.t. $n$

$$\frac{\partial F_{\Delta t}}{\partial n} = -\frac{\tau_{\mathrm{sj}}}{n^2} \quad \text{where } n > 1 \tag{V.50}$$

The latter illustrates that, as $n$ increases, the freeze duration decreases quickly.

## 9.3 Latency to Live Stream

The receiver gets the back the live stream after the network outage when the sender buffer empties, which takes time according to Equation (V.10). Furthermore, after the network outage, the play-out starts when the jitter buffer gets re-filled, which takes time according to Equation (V.9).

The latency to live stream $\Delta t$ is discussed in Sections 8.1–8.3. The inequalities (V.23) and (V.35) together with Equation (V.41) describe $\Delta t$ under the different conditions:

$$t_{\mathrm{e}}(\tau_{\mathrm{s}}) > t_{\mathrm{f}}(\tau_{\mathrm{j}}) \quad \text{where } \tau_{\mathrm{s}} > \tau_{\mathrm{j}} \tag{V.23}$$

$$t_{\mathrm{e}}(\tau_{\mathrm{s}}) < t_{\mathrm{f}}(\tau_{\mathrm{j}}) \left(\frac{n}{n-1}\right) \quad \text{where } \tau_{\mathrm{s}} < \tau_{\mathrm{j}} \tag{V.35}$$

$$\Rightarrow \Delta t = |t_\mathrm{e}(\tau_\mathrm{s}) - t_\mathrm{f}(\tau_\mathrm{j})| \tag{V.51}$$

Obviously, the “slower” one of the two buffers governs the time the user has to wait for the live stream. In case both buffers are the same size, we obtain

$$\Delta t = \frac{\tau_\mathrm{sj}}{n(n-1)} \quad \text{where } \tau_\mathrm{s} = \tau_\mathrm{j} = \tau_\mathrm{sj} \tag{V.41}$$

The corresponding partial derivatives

$$\frac{\partial \Delta t}{\partial \tau_\mathrm{sj}} = \frac{1}{n(n-1)} \tag{V.52}$$

$$\frac{\partial \Delta t}{\partial n} = -\frac{2\tau_\mathrm{sj}}{n^2(n-1)} \tag{V.53}$$

underline the damping impact of the capacity factor $n$, which decreases quickly as $n$ grows beyond 1.

## 9.4 Limitations of the Model

Our model did not particularly include a mechanism to detect a network outage, inform the sender, and take countermeasures. While TCP appears to be a natural solution, it might cause additional jitter when mitigating loss [37]. Thus, UDP is the best choice to keep the real-time streaming character intact. Furthermore, control mechanisms on top of UDP can be used to identify and communicate the occurrence of network outages through (a) the exchange of regular “hello ”or echo messages, and detection of delivery issues of the same, and (b) the observation of the fill-level of the sender buffer. Indeed, once a network outage becomes effective, the sender buffer will start filling up without delay, which may trigger some countermeasure at the sender side immediately. For instance, a bit rate reduction might imply shorter outage durations [38]. Now that the basic model has been established, it would be interesting to study a corresponding extension of this work.

The effect of the outage mitigation depends on the capacity factor $n$, which is not necessarily known beforehand. In an outage-prone mobile context, the authors in [3] report an average download speed of at least 50 Mbps and peak values in the order of Gbps, while Zoom recommends 3.8 Mbps for HD live streaming [39]. These values imply $n > 10$. At this point, it is important (a) to realize the sinking impact of $n$ as it grows, cf. the partial derivative of the freeze duration (25) with respect to $n$ that reveals a quick decay, cf. Section 9.3; and (b) that capacity sharing will reduce the effective $n$ per user, which calls for further studies, using, for example, the stochastic fluid flow model with its dedicated focus on capacity ratios [40].

# 10 Subjective Results

We performed a subjective test to study the impact of the sender buffer on the video QoE in the presence of the network outages. In the experiment, we use the three video sequences “Foreman”, “Football,” and “News.” The resolution of all videos is QVGA and the frame rate is 30 fps. The QVGA videos are providing a baseline to study the artifact like jump, freeze, and miss. The temporal effects that we focus on are agnostic of the resolution, while single packet losses would yield different artifacts depending on resolution, codec, etc.

Each video was simulated with a single network outage, represented as L1 or L2 as shown in Figure 4. L1 means that the network outage occurs in the first half of the video after playing a one-second video, and L2 represents the outage location in the second half of the video after playing $t/2+1$ seconds of video. For each video, the length of network outage was either 1 s or 2 s. The videos and network outage details are given in Table 2.

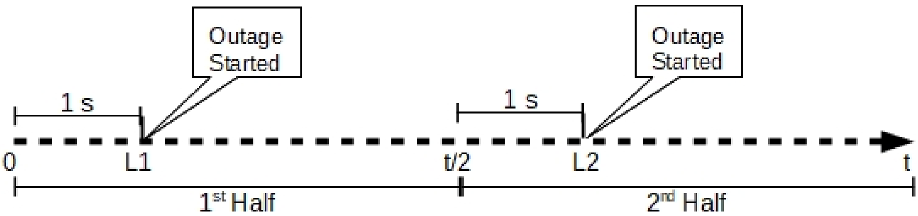



**Figure 4:** Network outage location in the video.

**Table 2:** Video used for subjective analysis.

| Video (Motion) | Video Length $t$ [s] | Outage Location | Outage Length [s] |
|---|---|---|---|
| Foreman (Medium) | 10 | L1 or L2 | 1 s or 2 s for L1 or L2 |
| Football (Fast) | 8 | L1 or L2 | 1 s or 2 s for L1 or L2 |
| News (Slow) | 10 | L1 or L2 | 1 s or 2 s for L1 or L2 |

In the presence and absence of a sending buffer, videos are simulated in two scenarios. Table 3 lists all simulation scenarios for all videos. We selected different simulation scenarios to test the performance of the proposed video streaming, without and with sender buffer. In Section 9, we concluded that both the sender buffer and the jitter buffer must be the same size to maximize the efficiency of the buffers, which we follow in the experiments. The first column and the second column represent the size of the sender buffer and jitter buffer in time units. The third column shows the duration of the network outage.

Obviously, when the network outage duration exceeds the jitter buffer, the user will experience video freeze and jump artifacts. However, when the network outage is shorter than the jitter buffer, the user does not see any artifacts. To reduce the complexity of discussion and experiments, we kept the network outage equal to the length of the jitter buffer and/or sender buffer. Cases 1a and 2a do not use a sender buffer. In these cases, even when the network outage is equal to the jitter buffer size, the user will face some video freeze and loss of video content. As discussed in Section 9, when a network outage is greater than the sender buffer, data loss occurs at the sender side. However, after the network outage, only the data stored in the sender buffer are of interest. Thus, to test the usefulness of the sender buffer, cases 1b, 1c, 2b, and 2c were chosen for the simulation, in which the network outage is equal to the sender buffer.

The capacity factor $n$ is another important multiplicative factor, which helps to reduce the freeze time duration and get back the live video in a faster way. Column 4 shows the capacity factor $n$. The last two columns represent the freeze time, cf. Equation (V.47), and the data loss duration according to Equation (V.43)

$$L' = \frac{L}{R_{\mathrm{sr}}} = \tau_{\mathrm{o}} - \tau_{\mathrm{sj}} \text{ where } \tau_{\mathrm{o}} \geq \tau_{\mathrm{sj}} \tag{V.54}$$

**Table 3:** Simulation parameters and their results in terms of video freeze and loss duration.

| **Case** | $\boldsymbol{\tau_{\mathrm{s}}}$ **[s]** | $\tau_{\mathrm{j}}$ [s] | $\tau_{\mathrm{o}}$ [s] | $\boldsymbol{n}$ | $\boldsymbol{F_{\Delta t}}$ [s] | $\boldsymbol{L'}$ [s] |
|---|---|---|---|---|---|---|
| 1a | 0 | 1 | 1 | n/a | 1 | 1 |
| 1b | 1 | 1 | 1 | 1 | 1 | 0 |
| 1c | 1 | 1 | 1 | 2 | 0.5 | 0 |
| 2a | 0 | 2 | 2 | n/a | 2 | 2 |
| 2b | 2 | 2 | 2 | 1 | 2 | 0 |
| 2c | 2 | 2 | 2 | 2 | 1 | 0 |

As described in Table 3, six different simulation scenarios are used to create test videos. For each selected simulation scenario, we apply the network outage in two different locations, one in the first half of the video and the other in the other half of the video. This results in a total of 12 videos for each video sequence and a total of 36 videos for subjective testing.

The subjective tests were conducted in the perception lab of Blekinge Institute of Technology, Sweden using absolute category rating (ACR). The perception lab is designed according to the recommendation of ITU-R BT. 500 [41] and ITU-T P.910 [42]. Before starting the user rating, verbal and written instructions were provided to each participant. All participants are given the instruction “In this experiment, you are going to rate the quality of the video regardless of the content of the video. Please be considerate about your judgment and remember there is not an exact score for a video. It is an opinion and can vary from person to person.” After viewing each video, participants are asked “Please rate the quality of this video.” A total of 46 subjects participated in this study and gave feedback on the video quality using the single stimulus method on the 5-point Absolute Category Rating (ACR) scale with categories 1 = *bad*, 2 = *poor*, 3 = *fair*, 4 = *good*, and 5 = *excellent*. All test videos have artifacts like freeze, jump, and miss, so a certain spread of the users’ ratings was expected, along with some rather low user ratings. However, the feedback of three participants was different. One of them had given score 5 to all videos. Most of the feedback of two other users fall outside the interval obtained from applying the Interquartile range (IQR), so they were considered as outliers. Thus, 43 participant ratings were analyzed further.

Figures 5 and 6 show the user ratings for the “Foreman” video simulated with and without the sender buffer, and the corresponding network outages. The abscissa indicates the outage location in the video, and the ordinate shows the Mean Opinion Score (MOS).

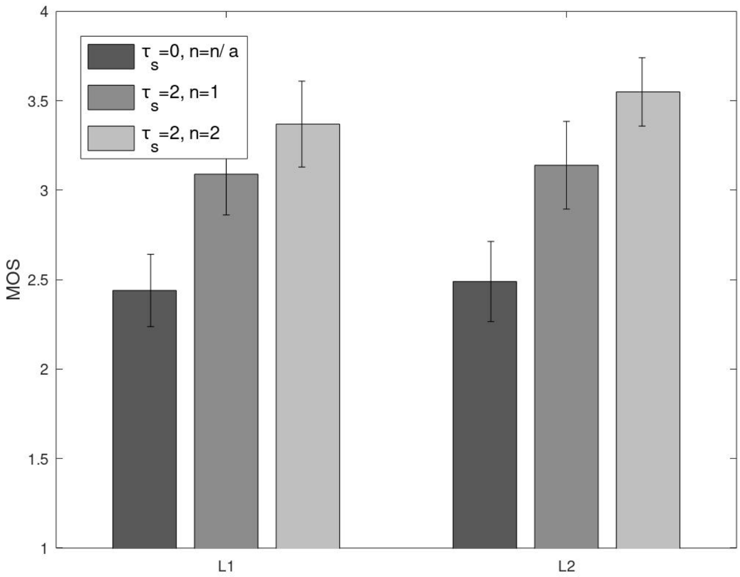



**Figure 5:** MOS with 95% confidence intervals of the ``Foreman'' video with two outage locations for $\tau_{\mathrm{j}} = 1$ and $\tau_{\mathrm{o}} = 1$.

In the absence of the sender buffer ($\tau_{\mathrm{s}} = 0$), the network outage results in video freeze-and-jump. From Table 3, we can see that, in the absence of the sender buffer, the video loss and video freeze are equal to the outage duration. Users react to these artifacts and rate the video on average below 2.5. This user rating applies to both $\tau_{\mathrm{o}} = 2$ and $\tau_{\mathrm{o}} = 1$. However, average user ratings for $\tau_{\mathrm{o}} = 2$ are lower than average users’ratings for $\tau_{\mathrm{o}} = 1$, which is obvious, as the users faced longer video losses and video freezes. Moreover, it is also observed that the user ratings vary with respect to the location of video under similar artifacts. Similar reactions are observed and reported by the authors in [15, 16]. The presence of the sender buffer helps to save the data during the network outage. As a result, it mitigates the video artifacts like freeze, jump, and miss as observed during the experiments and shown in Table 4 as well as in Figures 5 and 6. The sender buffer of length 1 s ($\tau_s = 1$) helps to improve MOS from 2.44 to 3.55, while an extended length to 2 s ($\tau_s = 2$) makes MOS improve

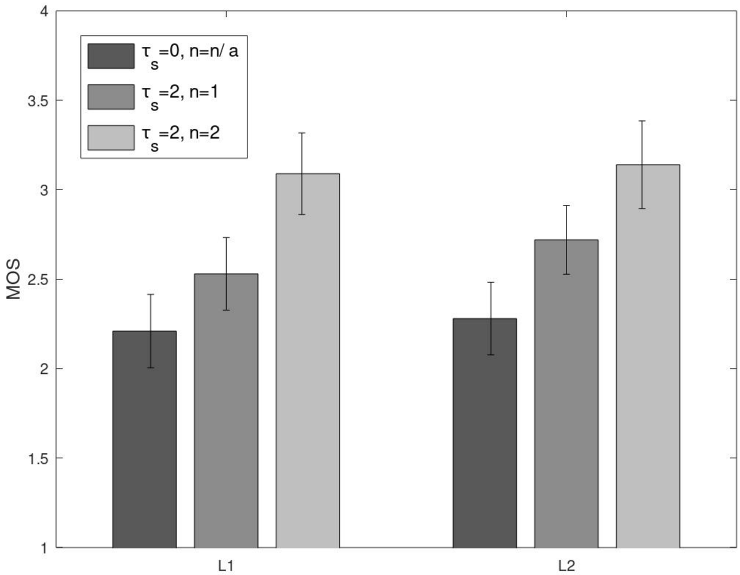



**Figure 6:** MOS with 95% confidence intervals of the ``Foreman'' video with two outage locations for $\tau_{\mathrm{j}} = 2$ and $\tau_{\mathrm{o}} = 2$.

from 2.21 to 2.72. Figures 5 and 6 illustrate that MOS values increase from 3.09 and 3.14 for $n = 1$ and from 3.37 and 3.55 for $n = 2$ when increasing $\tau_s = \tau_j = \tau_o = 1$ to $\tau_s = \tau_j = \tau_o = 2$. Obviously, an increase in the data rate after network outage in the presence of the sender buffer also improves the user ratings.

Figures 7–10 show the average user ratings for "Football" and "News," respectively. These bar charts for "Football" and "News" are depicting similar behaviors as the "Foreman" videos.

Table 4 shows the detailed MOS results of the "Foreman," "Football," and "News" videos.

The average user ratings of the original videos "Foreman", "Football", and "News" are 4.57, 4.30, and 4.35, respectively. This indicates that the users did not rank the original video as excellent, and the user ratings vary from video to video. This behavior was also reported in a previous study [8]. From the results, we can see that this random behavior is also found in the average user rating of all videos. From Table 4, we can realize that the network outage ($\tau_o = 1$) degrades the average user ratings from *good* to *poor*. However, for the same network outage settings, the sender buffer helps to improve the average user ratings from *poor* to *fair*.

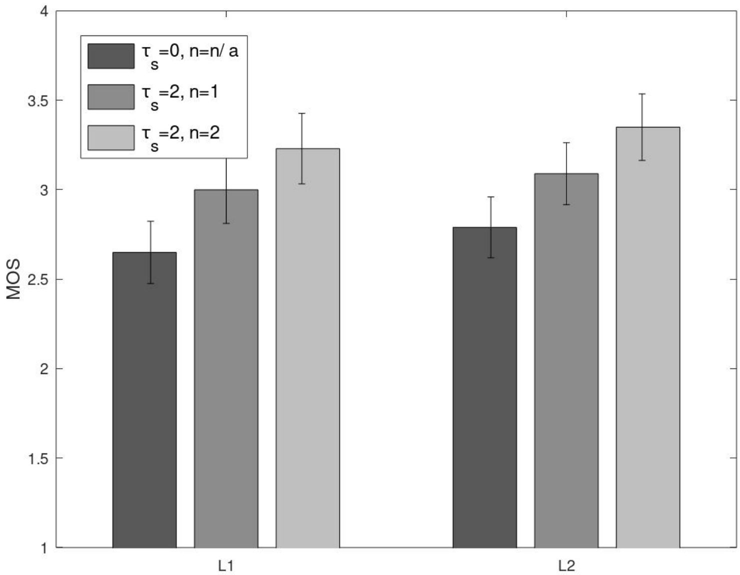



**Figure 7:** MOS with 95% confidence intervals of the ``Football'' video with two outage locations for $\tau_{\mathrm{j}} = 1$ and $\tau_{\mathrm{o}} = 1$.

## 10.1 Analysis of the Impact of the Sender Buffer

From the above results, the MOS of videos streamed with the sender buffer ($\mu_1$) is apparently better than the MOS of video streamed without the sender buffer ($\mu_2$) under otherwise similar conditions. To confirm that $\mu_1$ is significantly different and greater than $\mu_2$, Student's $t$-test was used (as implemented in R). The null hypothesis is defined as

$$H_0 : \mu_1 \leq \mu_2 \tag{V.55}$$

Moreover, to calculate how much MOS is augmented due to the sender buffer, we also computed the MOS gain

$$\Delta\mathrm{MOS} = \mu_1 - \mu_2 \tag{V.56}$$

Table 5 shows the Student's $t$-test results along with the MOS gain. Results are grouped for $\tau_{\mathrm{j}} = \tau_{\mathrm{o}} = 1$ and $\tau_{\mathrm{j}} = \tau_{\mathrm{o}} = 2$ for each video. Columns 5 and 9 show the MOS gain, whereas columns 7 and 11 are based on the outcome of the null hypothesis. The MOS gain is always greater than zero, which supports the assumption that the presence of a sender buffer helps to improve the MOS.

From Table 5, we can see that the null hypothesis is rejected in all cases. As in columns 6 and 10, we can see that all $p$-values are much smaller than $\alpha = 0.05$. Hence, the conclusion is that MOS with a sender buffer is greater than the MOS

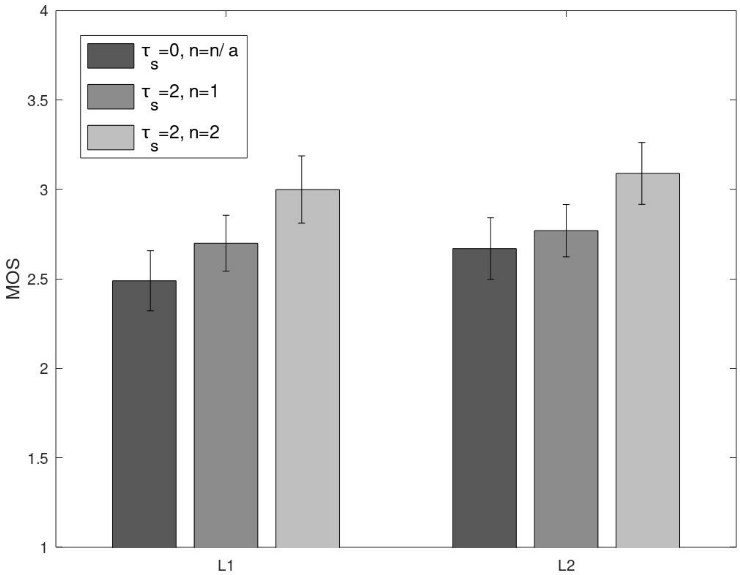



**Figure 8:** MOS with 95% confidence intervals of the ``Football'' video with two outage locations for $\tau_{\mathrm{j}} = 2$ and $\tau_{\mathrm{o}} = 2$.

without the sender buffer, which implies that the sender buffer minimizes the effects of network outages and improves the user experience.

## 10.2 Analysis of the Impact of the Capacity Factor $n$ on MOS Gain

To examine the role of the capacity factor ($n$) on MOS in the presence of sender buffer, we analyze MOS for $n = 2$ denoted by $\mu_1(n = 2)$ and MOS for $n = 1$ denoted by $\mu_2(n = 1)$ under otherwise similar conditions, to substantiate that the bigger capacity factor $n$ better mitigates the impact of the network outage. The Student's $t$-test was used with the following null hypothesis:

$$H_0 : \mu_1(n = 2) \leq \mu_2(n = 1) \tag{V.57}$$

In order to calculate how much MOS is increased due to the capacity factor ($n$) in the presence of the sender buffer, the MOS gain is calculated as

$$\Delta\mathrm{MOS} = \mu_1(n = 2) - \mu_2(n = 1) \tag{V.58}$$

In Table 6, the impact of the capacity factor $n$ is illustrated by the MOS gain in columns 4 and 9, all with positive values. Column 5 and 10 revealed the $p$-values

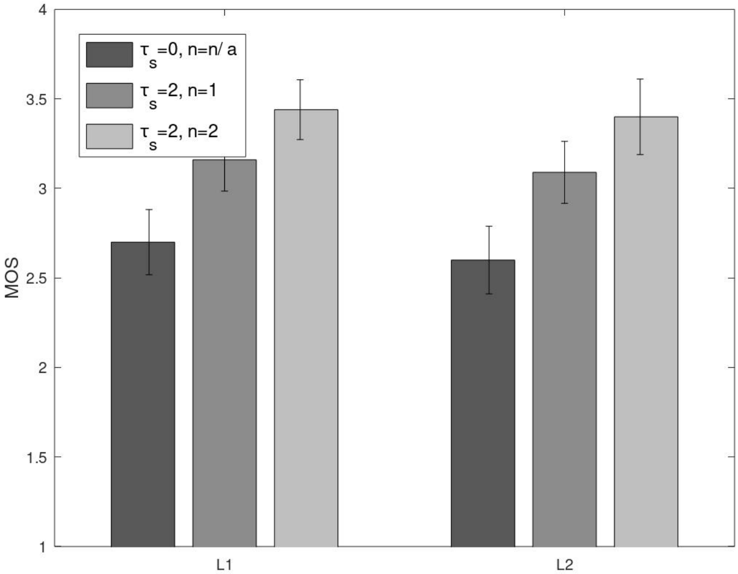



**Figure 9:** MOS with 95% confidence intervals of the ``News'' video with two outage locations for $\tau_{\mathrm{j}} = 1$ and $\tau_{\mathrm{o}} = 1$.

of the Student's $t$-test applied on the MOS for $n = 2$ and MOS for $n = 1$. As all $p$-values are less than 0.05, we can reject the null hypothesis. Both MOS gain and Student's $t$-test show that a bigger capacity factor $n$ in the presence of a sender buffer helps to mitigate the impact of the network outage in a better way through reduced video freeze durations, which obviously improves the QoE.

## 10.3 Limitations Regarding Resolutions

The experiments were performed using low-end QVGA resolution, which may appear outdated at least in the wealthy economies, while 3G is still standard in developing countries, or offered as low-cost alternative even in the developed countries.

The questions arise regarding what to expect from a higher resolution. At a higher resolution, an outage of a certain duration will naturally affect more packets. However, in our work, the *duration* of the outage, as compared to sender and jitter buffer, is the key parameter.

As expected, high-resolution videos yield better subjective ratings in terms of MOS [43] or objective Video Quality Metric (VQM) and Structural Similarity Index Measure (SSIM) values [44], as compared to low-resolution videos. In Reference [16], the authors studied the effects of video freezes with resolutions (QCIF, CIF, QVGA, and VGA), and compared MOS ratings with the Temporal Quality Metric (TQM) and Perceptual Evaluation of Video Quality (PEVQ) values. They found that

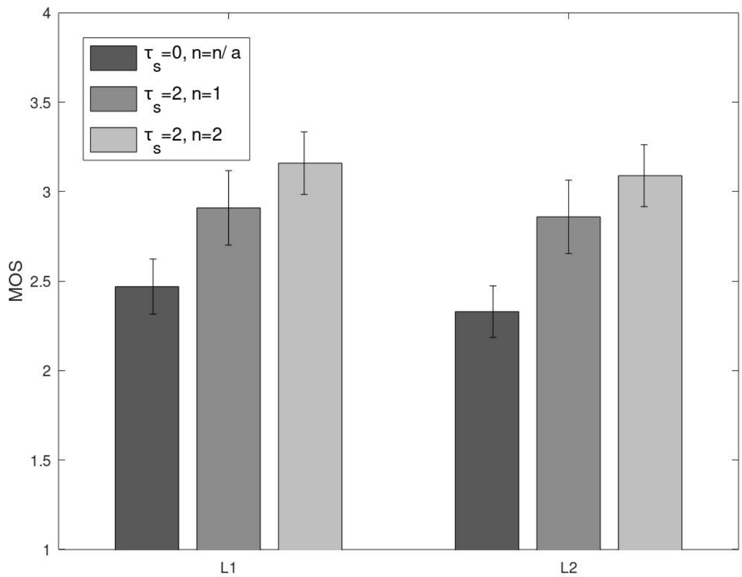



**Figure 10:** MOS with 95% confidence intervals of the ``News'' video with two outage locations for $\tau_{\mathrm{j}} = 2$ and $\tau_{\mathrm{o}} = 2$.

video resolutions played a minor role on both subjective and objective video quality ratings. However, it remains to be investigated to which extent a high streaming resolution (such as 4K) would affect the ratings of the temporal disturbances.

# 11 Conclusions

In live streaming, network outage may result in video freezes, jumps, or misses. Subjective studies showed that artifacts like freeze-and-jump degrade the video QoE. To mitigate the degradation of the QoE, a sender buffer-based solution was proposed.

Analytically, we have discussed that the sender buffer helps in reducing the data loss and freeze time duration during the network outage. The possibility to flush the sender buffer data with a higher data rate on the availability of the channel can reduce the freeze time duration even further. The relationship between the size of the sender buffer and receiver buffer was also analyzed in terms of various parameters. It was found that the size of both buffers should be equal for the optimal result.

Subjective tests were also conducted for the videos exposed to the network outage during the transmission. It is found that the user rating improves a lot in the presence of sender buffer, especially with a higher capacity factor ($n > 1$). Based on both our analytical and simulation-based user study, we may conclude that using a well-

**Table 4:** Simulation details and MOS for various videos (with * denoting n/a).

| Outage Postion | $\tau_s$ [s] | $\tau_j$ [s] | $\tau_o$ [s] | $n$ | $F_{\Delta t}$ [s] | $L$ [s] | Foreman MOS | Football MOS | News MOS |
|---|---|---|---|---|---|---|---|---|---|
| Original | | | | | | | 4.57 ± 0.187 | 4.30 ± 0.230 | 4.35 ± 0.256 |
| L1 | 0 | 1 | 1 | * | 1 | 1 | 2.44 ± 0.202 | 2.60 ± 0.164 | 2.70 ± 0.182 |
| L2 | 0 | 1 | 1 | * | 1 | 1 | 2.49 ± 0.224 | 2.67 ± 0.159 | 2.60 ± 0.189 |
| L1 | 0 | 2 | 2 | * | 2 | 2 | 2.21 ± 0.205 | 2.44 ± 0.153 | 2.47 ± 0.154 |
| L2 | 0 | 2 | 2 | * | 2 | 2 | 2.28 ± 0.203 | 2.56 ± 0.167 | 2.33 ± 0.144 |
| L1 | 1 | 1 | 1 | 1 | 1 | 0 | 3.09 ± 0.228 | 3.00 ± 0.188 | 3.16 ± 0.175 |
| L2 | 1 | 1 | 1 | 1 | 1 | 0 | 3.14 ± 0.245 | 3.09 ± 0.173 | 3.09 ± 0.173 |
| L1 | 1 | 1 | 1 | 2 | 0.5 | 0 | 3.37 ± 0.240 | 3.23 ± 0.197 | 3.44 ± 0.167 |
| L2 | 1 | 1 | 1 | 2 | 0.5 | 0 | 3.44 ± 0.153 | 3.35 ± 0.186 | 3.40 ± 0.211 |
| L1 | 2 | 2 | 2 | 1 | 2 | 0 | 2.53 ± 0.203 | 2.72 ± 0.153 | 2.91 ± 0.208 |
| L2 | 2 | 2 | 2 | 1 | 2 | 0 | 2.72 ± 0.192 | 2.81 ± 0.152 | 2.86 ± 0.205 |
| L1 | 2 | 2 | 2 | 2 | 1 | 0 | 3.09 ± 0.228 | 3.00 ± 0.188 | 3.16 ± 0.175 |
| L2 | 2 | 2 | 2 | 2 | 1 | 0 | 3.14 ± 0.245 | 3.09 ± 0.173 | 3.09 ± 0.173 |

dimensioned sender buffer in combination with a faster-than-needed channel helps to mitigate the effect of the network outages on the video QoE.

In our future work, we plan to investigate the role of a backup buffer at the receiver, in parallel to the jitter buffer. During our work, so far, we have observed that it is very difficult to minimize the freeze time so that the user cannot feel it. In order to remove this minimum freeze time from the user display, we are investigating the concept of a backup buffer in parallel to the jitter buffer at the receiving side, which is be used to replay contents when the jitter buffer is empty in order to keep the user involved.

We are also considering ways to detect and inform the sender about the network outage, as quickly as possible. Furthermore, we are also interested in optimising the usage of the available link capacity after the network outage in the presence of other streams, and for various combinations of resolutions and mobile network connectivity.

# References


[1] P. Le Callet, S. Möller, A. Perkis, et al. "Qualinet white paper on definitions of quality of experience". In: *European network on quality of experience in multimedia systems and services (COST Action IC 1003)* 3.2012 (2012).

[2] R. De', N. Pandey, and A. Pal. "Impact of digital surge during Covid-19 pandemic: A viewpoint on research and practice". In: *International Journal of*

**Table 5:** Analysis of the impact of the sender buffer using Student's $t$-test.

| | | $\tau_\text{j} = \tau_\text{o}$ = 1 | | | | | $\tau_\text{j} = \tau_\text{o}$ = 2 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Foreman** | | | | | | | | | | | |
| OL | $n$ | MOS $\tau_\text{s} = \tau_\text{j}$ | MOS $\tau_\text{s} = 0$ | ΔMOS | p-value $\alpha = 0.05$ | $H_0$ $\mu_1 \le \mu_2$ | MOS $\tau_\text{s} = \tau_\text{j}$ | MOS $\tau_\text{s} = 0$ | ΔMOS | p-value $\alpha = 0.05$ | $H_0$ $\mu_1 \le \mu_2$ |
| L1 | 1 | 3.09 | 2.44 | 0.65 | $1.2 \times 10^{-5}$ | Reject | 2.53 | 2.21 | 0.33 | $7.3 \times 10^{-3}$ | Reject |
| L2 | 1 | 3.14 | 2.49 | 0.65 | $3.8 \times 10^{-5}$ | Reject | 2.72 | 2.28 | 0.44 | $2.4 \times 10^{-4}$ | Reject |
| L1 | 2 | 3.37 | 2.44 | 0.93 | $1.1 \times 10^{-7}$ | Reject | 3.09 | 2.21 | 0.88 | $1.6 \times 10^{-7}$ | Reject |
| L2 | 2 | 3.44 | 2.49 | 0.95 | $3.2 \times 10^{-10}$ | Reject | 3.14 | 2.28 | 0.86 | $6.9 \times 10^{-9}$ | Reject |
| Football | | | | | | | | | | | |
| | | $\tau_\text{j} = \tau_\text{o}$ = 1 | | | | | $\tau_\text{j} = \tau_\text{o}$ = 2 | | | | |
| OL | $n$ | MOS $\tau_\text{s} = \tau_\text{j}$ | MOS $\tau_\text{s} = 0$ | ΔMOS | p-value $\alpha = 0.05$ | $\text{H}_0$ $\mu_1 \le \mu_2$ | MOS $\tau_\text{s} = \tau_\text{j}$ | MOS $\tau_\text{s} = 0$ | ΔMOS | p-value $\alpha = 0.05$ | $\text{H}_0$ $\mu_1 \le \mu_2$ |
| L1 | 1 | 3.00 | 2.60 | 0.40 | $1.9 \times 10^{-3}$ | Reject | 2.72 | 2.44 | 0.28 | $4.4 \times 10^{-3}$ | Reject |
| L2 | 1 | 3.09 | 2.67 | 0.42 | $8.2 \times 10^{-5}$ | Reject | 2.81 | 2.56 | 0.26 | $7.3 \times 10^{-3}$ | Reject |
| L1 | 2 | 3.35 | 2.60 | 0.74 | $6.4 \times 10^{-9}$ | Reject | 3.00 | 2.44 | 0.56 | $2.5 \times 10^{-6}$ | Reject |
| L2 | 2 | 3.23 | 2.67 | 0.56 | $2.5 \times 10^{-6}$ | Reject | 3.09 | 2.56 | 0.53 | $5.4 \times 10^{-6}$ | Reject |
| News | | | | | | | | | | | |
| | | $\tau_\text{j} = \tau_\text{o}$ = 1 | | | | | $\tau_\text{j} = \tau_\text{o}$ = 2 | | | | |
| OL | $n$ | MOS $\tau_\text{s} = \tau_\text{j}$ | MOS $\tau_\text{s} = 0$ | ΔMOS | p-value $\alpha = 0.05$ | $\text{H}_0$ $\mu_1 \le \mu_2$ | MOS $\tau_\text{s} = \tau_\text{j}$ | MOS $\tau_\text{s} = 0$ | ΔMOS | p-value $\alpha = 0.05$ | $\text{H}_0$ $\mu_1 \le \mu_2$ |
| L1 | 1 | 3.16 | 2.70 | 0.47 | $2.1 \times 10^{-5}$ | Reject | 2.91 | 2.47 | 0.44 | $1.5 \times 10^{-4}$ | Reject |
| L2 | 1 | 3.09 | 2.60 | 0.49 | $7.4 \times 10^{-5}$ | Reject | 2.86 | 2.33 | 0.53 | $9.2 \times 10^{-5}$ | Reject |
| L1 | 2 | 3.44 | 2.70 | 0.74 | $9.8 \times 10^{-12}$ | Reject | 3.16 | 2.47 | 0.70 | $4.1 \times 10^{-9}$ | Reject |
| L2 | 2 | 3.40 | 2.60 | 0.79 | $7.5 \times 10^{-9}$ | Reject | 3.09 | 2.33 | 0.77 | $2.1 \times 10^{-11}$ | Reject |

*Information Management* 55 (2020). Impact of COVID-19 Pandemic on Information Management Research and Practice: Editorial Perspectives, p. 102171. ISSN: 0268-4012. DOI: `https://doi.org/10.1016/j.ijinfomgt.2020.102171`. URL: `https://www.sciencedirect.com/science/article/pii/S0268401220309622`.

[3] C. de Looper and S. Hill. *Is 5G as fast as they're saying? We break down the speeds*. `https://www.digitaltrends.com/mobile/how-fast-is-5g/`. Accessed: 2021-04-25. Feb. 2021.

[4] A. A. Barakabitze, N. Barman, A. Ahmad, S. Zadtootaghaj, L. Sun, M. G. Martini, and L. Atzori. "QoE Management of Multimedia Streaming Services in Future Networks: A Tutorial and Survey". In: *IEEE Communications Surveys Tutorials* 22.1 (2020), pp. 526–565. DOI: `10.1109/COMST.2019.2958784`.

[5] T. N. Minhas, O. Nawaz, M. Fiedler, and S. Khatibi. "The Effects of Additional Factors on Subjective Quality Assessments". In: *2nd International Conference on Advancements in Computational Sciences (ICACS)*. IEEE. 2019, pp. 1–5.

**Table 6:** Analysis of the impact of the capacity factor $n$ on the MOS Gain using Student's *t*-test.

| **Foreman** | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\tau_j = \tau_s = \tau_o = 1$ | | | | | $\tau_j = \tau_s = \tau_o = 2$ | | | | |
| OL | MOS $n = 2$ | MOS $n = 1$ | $\Delta$MOS | p-value $\alpha = 0.05$ | $H_0$ $\mu_1 \leq \mu_2$ | MOS $n = 2$ | MOS $n = 1$ | $\Delta$MOS | p-value $\alpha = 0.05$ | $H_0$ $\mu_1 \leq \mu_2$ |
| L1 | 3.37 | 3.09 | 0.28 | $1.1 \times 10^{-2}$ | Reject | 3.09 | 2.53 | 0.56 | $5.2 \times 10^{-5}$ | Reject |
| L2 | 3.44 | 3.14 | 0.30 | $8.8 \times 10^{-3}$ | Reject | 3.14 | 2.72 | 0.42 | $1.2 \times 10^{-3}$ | Reject |
| Football | | | | | | | | | | |
| | $\tau_j = \tau_s = \tau_o = 1$ | | | | | $\tau_j = \tau_s = \tau_o = 2$ | | | | |
| OL | MOS $n = 2$ | MOS $n = 1$ | $\Delta$MOS | p-value $\alpha = 0.05$ | $H_0$ $\mu_1 \leq \mu_2$ | MOS $n = 2$ | MOS $n = 1$ | $\Delta$MOS | p-value $\alpha = 0.05$ | $H_0$ $\mu_1 \leq \mu_2$ |
| L1 | 3.23 | 3.00 | 0.23 | $4.3 \times 10^{-2}$ | Reject | 3.00 | 2.72 | 0.28 | $1.6 \times 10^{-2}$ | Reject |
| L2 | 3.35 | 3.09 | 0.26 | $2.0 \times 10^{-2}$ | Reject | 3.09 | 2.81 | 0.28 | $1.1 \times 10^{-2}$ | Reject |
| News | | | | | | | | | | |
| | $\tau_j = \tau_s = \tau_o = 1$ | | | | | $\tau_j = \tau_s = \tau_o = 2$ | | | | |
| OL | MOS $n = 2$ | MOS $n = 1$ | $\Delta$MOS | p-value $\alpha = 0.05$ | $H_0$ $\mu_1 \leq \mu_2$ | MOS $n = 2$ | MOS $n = 1$ | $\Delta$MOS | p-value $\alpha = 0.05$ | $H_0$ $\mu_1 \leq \mu_2$ |
| L1 | 3.44 | 3.16 | 0.28 | $4.4 \times 10^{-3}$ | Reject | 3.16 | 2.91 | 0.26 | $1.3 \times 10^{-2}$ | Reject |
| L2 | 3.40 | 3.09 | 0.30 | $1.1 \times 10^{-2}$ | Reject | 3.09 | 2.86 | 0.23 | $2.9 \times 10^{-2}$ | Reject |


[6] J. Shaikh, M. Fiedler, and D. Collange. “Quality of Experience from user and network perspectives”. In: *annals of telecommunications-annales des télécommunications* 65.1-2 (2010), pp. 47–57.

[7] T. N. Minhas and M. Fiedler. “Quality of experience hourglass model”. In: *2013 International Conference on Computing, Management and Telecommunications (ComManTel)*. 2013, pp. 87–92. DOI: 10.1109/ComManTel.2013.6482371.

[8] T. N. Minhas, O. Gonzalez Lagunas, P. Arlos, and M. Fiedler. “Mobile video sensitivity to packet loss and packet delay variation in terms of QoE”. In: *Packet Video Workshop (PV), 2012 19th International*. IEEE. 2012, pp. 83–88.

[9] A. S. Asrese, E. A. Walelgne, V. Bajpai, A. Lutu, Ö. Alay, and J. Ott. “Measuring Web Quality of Experience in Cellular Networks”. In: *Passive and Active Measurement*. Ed. by D. Choffnes and M. Barcellos. Cham: Springer International Publishing, 2019, pp. 18–33. ISBN: 978-3-030-15986-3.

[10] J. Shaikh, M. Fiedler, D. Collange, P. Arlos, and T. N. Minhas. “Inferring User-Perceived Performance of Network by Monitoring TCP Interruptions”. In: *Network Protocols and Algorithms* 4.2 (2012), pp. 49–67.

[11] P. Arlos and M. Fiedler. “Influence of the packet size on the one-way delay on the down-link in 3G networks”. In: *Wireless Pervasive Computing (ISWPC),*

*2010 5th IEEE International Symposium on*. 2010, pp. 573–578. DOI: 10.1109/ISWPC.2010.5483732.

[12] J. Fabini, W. Karner, L. Wallentin, and T. Baumgartner. “The Illusion of Being Deterministic – Application-Level Considerations on Delay in 3G HSPA Networks”. In: *NETWORKING 2009*. Ed. by L. Fratta, H. Schulzrinne, Y. Takahashi, and O. Spaniol. Lecture Notes in Computer Science 5550. Springer Berlin Heidelberg, Jan. 2009, pp. 301–312. ISBN: 978-3-642-01398-0, 978-3-642-01399-7. URL: http://link.springer.com/chapter/10.1007/978-3-642-01399-7_24.

[13] Y. Kyung and T.-K. Kim. “QoS-Aware Flexible Handover Management in Software-Defined Mobile Networks”. In: *Applied Sciences* 10.12 (June 2020), p. 4264. ISSN: 2076-3417. DOI: 10.3390/app10124264.

[14] M. P. Wylie-Green and T. Svensson. “Throughput, Capacity, Handover and Latency Performance in a 3GPP LTE FDD Field Trial”. In: *2010 IEEE Global Telecommunications Conference GLOBECOM 2010*. 2010, pp. 1–6. DOI: 10.1109/GLOCOM.2010.5683398.

[15] T. N. Minhas and M. Fiedler. “Impact of disturbance locations on video quality of experience”. In: *Quality of Experience for Multimedia Content Sharing, EuroITV2011* (2011).

[16] T. N. Minhas, M. Shahid, A. Rossholm, B. Lövström, H.-J. Zepernick, and M. Fiedler. “Assessment of the rating performance of ITU-T recommended video quality metrics in the context of video freezes”. In: *2013 Australasian Telecommunication Networks and Applications Conference (ATNAC)*. 2013, pp. 207–212. DOI: 10.1109/ATNAC.2013.6705382.

[17] Y. Qi and M. Dai. “The Effect of Frame Freezing and Frame Skipping on Video Quality”. In: *Proceedings of the 2006 International Conference on Intelligent Information Hiding and Multimedia*. IIH-MSP ’06. Washington, DC, USA: IEEE Computer Society, 2006, pp. 423–426. ISBN: 0-7695-2745-0. DOI: 10.1109/IIH-MSP.2006.162.

[18] N. Amram, B. Fu, G. Kunzmann, T. Melia, D. Munaretto, S. Randriamasy, B. Sayadi, J. Widmer, and M. Zorzi. “QoE-based transport optimization for video delivery over next generation cellular networks”. In: *2011 IEEE Symposium on Computers and Communications (ISCC)*. Kerkyra, Greece, 2011, pp. 19–24. DOI: 10.1109/ISCC.2011.5984019.

[19] T. N. M. V. Caenegem, K. O. Struyve, K. Laevens, D. D. Vleeschauwer, and R. Sharpe. “Maintaining video quality and optimizing video delivery over the bandwidth constrained DSL last mile through intelligent packet drop”. In: *Bell Labs Technical Journal* 13.1 (2008), pp. 53–68. ISSN: 1538-7305. DOI:

10.1002/bltj.20282. URL: http://onlinelibrary.wiley.com/doi/10.1002/bltj.20282/abstract.

[20] C. Demichelis and P. Chimento. *IP Packet Delay Variation Metric for IP Performance Metrics (IPPM)*. RFC 3393. IETF, Nov. 2002, pp. 1–20. URL: https://tools.ietf.org/html/rfc3393.

[21] H. Inamura, O. Takahashi, H. Nakano, T. Ishikawa, and H. Shigeno. "Impact of layer two ARQ on TCP performance in W-CDMA networks". In: *Distributed Computing Systems, 2004. Proceedings. 24th International Conference on*. 2004, pp. 284–291. DOI: 10.1109/ICDCS.2004.1281593.

[22] A. Wahab, N. Ahmad, and J. Schormans. "Direct propagation of network QoS distribution to subjective QoE for Video on Demand applications using VP9 codec". In: *2020 International Wireless Communications and Mobile Computing (IWCMC)*. 2020, pp. 929–933. DOI: 10.1109/IWCMC48107.2020.9148101.

[23] B. Spang, B. Walsh, T.-Y. Huang, T. Rusnock, J. Lawrence, and N. McKeown. "Buffer Sizing and Video QoE Measurements at Netflix". In: *Proceedings of the 2019 Workshop on Buffer Sizing*. BS '19. Palo Alto, CA, USA: Association for Computing Machinery, 2019. ISBN: 9781450377454. DOI: 10.1145/3375235.3375241. URL: https://doi.org/10.1145/3375235.3375241.

[24] S. A. Soomro, M. M. Shaikh, N. Nizamani, E. A. Buriro, and K. M. Zuhaib. "Heterogeneous Buffer Size Impact on UDP Performance for Real-Time Video Streaming Application". In: *International Journal of Advanced Computer Science and Applications* 9.6 (2018). DOI: 10.14569/IJACSA.2018.090638.

[25] M. Claypool and J. Tanner. "The Effects of Jitter on the Peceptual Quality of Video". In: *Proceedings of the Seventh ACM International Conference on Multimedia (Part 2)*. MULTIMEDIA '99. Orlando, Florida, USA: ACM, 1999, pp. 115–118. ISBN: 1-58113-239-5. DOI: 10.1145/319878.319909.

[26] R. Ramjee, J. Kurose, D. Towsley, and H. Schulzrinne. "Adaptive playout mechanisms for packetized audio applications in wide-area networks". In: *INFOCOM '94. Networking for Global Communications., 13th Proceedings IEEE*. June 1994, 680–688 vol.2. DOI: 10.1109/INFCOM.1994.337672.

[27] D. Stone and K. Jeffay. "An empirical study of delay jitter management policies". English. In: *Multimedia Systems* 2.6 (1995), pp. 267–279. ISSN: 0942-4962. DOI: 10.1007/BF01225244.

[28] D. Ferrari. "Delay Jitter Control Scheme for Packet-switching Internetworks". In: *Comput. Commun.* 15.6 (July 1992), pp. 367–373. ISSN: 0140-3664. DOI: 10.1016/0140-3664(92)90011-3.

[29] T. Hossfeld, S. Egger, R. Schatz, M. Fiedler, K. Masuch, and C. Lorentzen. "Initial delay vs. interruptions: Between the devil and the deep blue sea". In: *Fourth International Workshop on Quality of Multimedia Experience*. Yarra Valley, Australia, July 2012, pp. 1–6. DOI: 10.1109/QoMEX.2012.6263849.

[30] D. Ye, X. Wang, Z. Zhang, and Q. Wu. "A buffer-driven approach to adaptively stream stored video over Internet". In: *High Speed Networks and Multimedia Communications 5th IEEE International Conference on*. 2002, pp. 81–85. DOI: 10.1109/HSNMC.2002.1032552.

[31] O. O. Ayodeji, T. O. Damilola, E. O. Folake, A. A. Ganiyu, and A. A. Emmanuel. "Buffer Occupancy of Double-Buffer Traffic Shaper in Real-Time Multimedia Applications across Slow-Speed Links". In: *Communications and Network* 05.01 (2013), pp. 84–92. ISSN: 1949-2421, 1947-3826. DOI: 10.4236/cn.2013.51008. URL: http://www.scirp.org/journal/PaperDownload.aspx?DOI=10.4236/cn.2013.51008.

[32] H. Koto, M. Fukushima, S. Nomoto, and F. Takahata. "Scheduling algorithm based on sender buffer backlog for real-time application in mobile packet networks". In: *Wireless Communications and Networking Conference, 2005 IEEE*. Vol. 1. Mar. 2005, 151–157 Vol. 1. DOI: 10.1109/WCNC.2005.1424491.

[33] A. Erbad and C. Krasic. "Sender-side buffers and the case for multimedia adaptation". In: *Communications of the ACM* 55 (Dec. 2012). DOI: 10.1145/2380656.2380671.

[34] J. Wu, R. Tan, and M. Wang. "Streaming High-Definition Real-Time Video to Mobile Devices with Partially Reliable Transfer". In: *IEEE Transactions on Mobile Computing* 18.2 (2019), pp. 458–472. DOI: 10.1109/TMC.2018.2836914.

[35] J. W. Jiyan Wu Zhizhuan Wen and H. Pan. "Buffer-aware transmission rate control for real-time video streaming system". 010645448B2. May 2020. URL: https://www.google.com/patents/US010645448B2.

[36] J. F. Kurose and K. W. Ross. *Computer Networking: A Top-Down Approach*. 6th. Pearson, Mar. 2012. ISBN: 0132856204.

[37] A. Shafer. *UDP vs. TCP and Which One to Use for Video Streaming*. https://www.wowza.com/blog/udp-vs-tcp/. Accessed: 2021-05-08. Dec. 2020.

[38] M. Fiedler. "On the limited potential of buffers to improve quality of experience". In: *2014 IEEE International Conference on Pervasive Computing and Communication Workshops (PERCOM WORKSHOPS)*. 2014, pp. 419–424. DOI: 10.1109/PerComW.2014.6815243.

[39] Zoom. *System requirements for Windows, macOS, and Linux*. [Online] https://support.zoom.us/hc/en-us/articles/201362023-System-requirements-for-Windows-macOS-and-Linux. Accessed: 2021-05-10. Apr. 2021.

[40] D. Anick, D. Mitra, and M. M. Sondhi. "Stochastic theory of a data-handling system with multiple sources". In: *The Bell System Technical Journal* 61.8 (1982), pp. 1871–1894. DOI: 10.1002/j.1538-7305.1982.tb03089.x.

[41] I.-R. R. BT.500-11. *Methodology for the subjective assessment of the quality of television pictures.* International Telecommunications Union Telecommunication Sector. Geneva, Switzerland, 2002.

[42] ITU-T P.910. *Subjective video quality assessment methods for multimedia applications.* International Telecommunications Union Telecommunication Sector. 1999.

[43] A. A. Laghari, H. He, A. Khan, and S. Karim. "Impact of Video File Format on Quality of Experience (QoE) of Multimedia Content". In: *3D Res.* 9.3 (Sept. 2018). ISSN: 2092-6731. DOI: 10.1007/s13319-018-0191-x. URL: https://doi.org/10.1007/s13319-018-0191-x.

[44] T. Zinner, O. Hohlfeld, O. Abboud, and T. Hossfeld. "Impact of frame rate and resolution on objective QoE metrics". In: *2010 Second International Workshop on Quality of Multimedia Experience (QoMEX)*. 2010, pp. 29–34. DOI: 10.1109/QOMEX.2010.5518277.

The rapid growth in mobile video consumption, driven by advancements in mobile devices and network infrastructure, has raised user expectations for seamless video Quality of Experience (QoE) despite improvements in video streaming, network impairments like packet loss, delay, jitter, and outages. For instance, outages can cause visual artifacts like freezing, jumping, and missing frames, which negatively affect user perception. Understanding the relationship between network performance and QoE is crucial for improving user satisfaction.

This thesis investigates the impact of network performance on mobile video QoE and proposes strategies to mitigate these effects. The objectives include: (1) understanding TCP/IP's role in influencing QoE, (2) exploring the effects of Quality of Service (QoS) parameters such as delay, jitter, and packet loss on video quality, (3) analyzing the impact of network outages on QoE, and (4) developing a buffer-based solution to mitigate network disruptions.

The research employs theoretical modeling, controlled emulation experiments, and subjective assessments to evaluate QoE. The QoE Hourglass Model links networklayer parameters to user-perceived quality. Subjective tests, guided by ITU-T recommendations, use the Absolute Category Rating (ACR) method and Mean Opinion Scores (MOS) to assess video quality under various conditions. Additionally, the effectiveness of a sender buffer mechanism is tested through statistical analyses and user evaluations.

The findings reveal that network impairments, especially packet loss and delay variation, significantly degrade QoE. The QoE Hourglass Model provides a structured framework for understanding these effects. Experimental results show that higher frame rates and proactive buffering improve user perception. Perceptual Evaluation of Video Quality (PEVQ) and Temporal Quality Metric (TQM) measurements correlate with user ratings but are less accurate in predicting video freezes. The sender buffer mechanism effectively reduces freeze durations and enhances QoE during network outages.

This research emphasizes the impact of network impairments on video QoE and offers practical solutions, such as the sender buffer mechanism, to mitigate disruptions and enhance user satisfaction in video streaming.

Printed by Media-Tryck, Lund 2025 NORDIC SWAN ECOLABEL 3041 0903

Blekinge Institute of Technology
Doctoral Dissertation Series No. 2025:06
ISSN: 1653-2090
ISBN: 978-91-7295-502-8