



Addis Ababa University  
Addis Ababa Institute of Technology  
School of Electrical and Computer Engineering  
Telecommunication Engineering Graduate Program

**Correlation among Quality of Experience and Quality of  
Service Parameters for LTE Video Streaming Service: In  
case of Addis Ababa, Ethiopia**

By  
Alemayehu Tiruneh

Advisor  
Dr. Beneyam Berehanu

A Thesis Submitted to the School of Graduate Studies of Addis Ababa University  
in Partial Fulfillment of the Requirements for the Degree of Master of Science in  
Telecommunication Engineering.

October, 2021  
Addis Ababa, Ethiopia

Addis Ababa University  
Addis Ababa Institute of Technology  
School of Electrical and Computer Engineering  
Telecommunication Engineering Graduate Program

**Correlation among Quality of Experience and Quality of  
Service Parameters for LTE Video Streaming Service: In  
case of Addis Ababa, Ethiopia**

By  
Alemayehu Tiruneh

Signed by the Examining Committee:

1. Beneyam Berehanu (PhD) \_\_\_\_\_ \_\_\_\_/\_\_\_\_/2022  
Advisor Signature Date

2. Dereje Hailemariam (PhD) \_\_\_\_\_ \_\_\_\_/\_\_\_\_/2022  
Examiner Signature Date

3. Yalemzewd Negash (PhD) \_\_\_\_\_ \_\_\_\_/\_\_\_\_/2022  
Examiner Signature Date

---

Dean,  
(School of Electrical and Computer Engineering)

---

## Declaration

I, the undersigned, declare that this thesis is my original work, has not been presented for a degree in this or any other university, and all sources of materials used for the thesis have been fully acknowledged.

Name: -Alemayehu Tiruneh

Signature: \_\_\_\_\_

Place: Addis Ababa

Date of Submission: \_\_\_\_\_

This thesis has been submitted for examination with my approval as a university advisor.

Advisor's Name: Dr. Beneyam Berehanu

Signature: \_\_\_\_\_

# Acknowledgment

First and foremost, I would like to express my deepest gratitude to Almighty God for giving me the wisdom and skill to carry out my thesis work.

I would like to express my sincere gratitude to my esteemed advisor, Dr. Beneyam Berehanu; without his assistance, this work would not have progressed from its inception to the master thesis that has emerged.

I would also like to thank Dr.-Ing. Dereje Hailemariam and Dr. Yalemzewd Negash for their kind and responsible advice during the preparation of this thesis. Without their continuous support, all of this would not have been possible.

I must extend my gratitude to ethio-telecom employees, especially to Lake Gebrekidan and all Radio Access Network (RAN) optimization team members as they were very cooperative in providing the necessary data for this work.

I would also like to thank all of my family and friends for their inspiration and encouragement, particularly during the data collection process and encouragement to finish the thesis.

Last but not least, I would really like to express my gratitude to my lovely wife Meseret Chala and children(Nathan Alemayehu, Yemariam Alemayehu and Tomas Alemayehu) for their priceless care throughout my postgraduate program.

# Abstract

The growing popularity of multimedia video streaming services, combined with the increased prevalence of smartphones, has resulted in impressive demand growth and a significant need for quality of service (QoS). As a result, making video streaming services available to users in a cost-effective manner without sacrificing quality is a significant challenge. In Ethiopia, the sole telecom services provider, Ethio Telecom (ET), has launched several network expansion projects in recent years to provide quality services. However, there are still customer complaints from different parts of the city.

This thesis intends to study the spatio-temporal correlation analysis among Quality of Experience (QoE) and QoS parameters for LTE video streaming service and indicate measurement variations. The objective data collection has been collected using Nemo Handy, Opensignal, and Cloudflare speed tests. The subjective perceptions of video streaming users were gathered using a contextually formulated survey questionnaire via a field experiment campaign including 300 different sample users who watched a single YouTube video session. This study is conducted in six different locations considering the application, network, and physical layer quality metrics, and we analyze the data using Statistical Package for the Social Sciences (SPSS), Actix Analyzer, and Excel.

In all study locations, in particular during the evening and nighttime periods, the measured latency and jitter are not suitable for video streaming services. During these periods, the Mean Opinion Scores (MOS) are less than 3.3 and on average, a MOS result of 3.9 is obtained. According to the correlation analysis results, the highest correlation magnitude is achieved between MOS and playback with a strong positive and statistically significant value of ( $r = .746$ ,  $p < .001$ ). That is, playback is the main factor responsible for the MOS variation. The strongest correlation among application and network layer parameters is found between download throughput and playback, with values of ( $r = .711$ ,  $p < .001$ ). As a result, rather than latency and jitter, download throughput is the main reason for the variation of application

layer quality metrics. We found weaker correlation results between network and physical layer quality metrics at Mebrathail and Jemo areas, indicating that the network layer quality matrices are influenced by the increase in the number of users and other factors, rather than RSRP and SINR.

**Key words:-** LTE, QoE, QoS, Correlation, Subjective Measurement and Objective Measurement.

# Table of Contents

<b>Acknowledgment</b>	<b>iii</b>
<b>Abstract</b>	<b>iv</b>
<b>List of Tables</b>	<b>ix</b>
<b>List of Figures</b>	<b>xi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Background . . . . .	1
1.2 Statement of the Problem . . . . .	3
1.3 Objectives . . . . .	4
1.3.1 General Objective . . . . .	4
1.3.2 Specific Objectives . . . . .	4
1.4 Literature review . . . . .	4
1.5 Methodology . . . . .	7
1.6 Scope and Limitation . . . . .	7
1.6.1 Scope of the Thesis . . . . .	7
1.6.2 Limitations of the Thesis . . . . .	8
1.7 Contributions . . . . .	8
1.8 Thesis Organization . . . . .	8
<b>2 LTE Mobile Network and Video Streaming Service</b>	<b>10</b>
2.1 LTE Mobile Network . . . . .	10
2.1.1 Background of LTE Mobile Network . . . . .	10
2.1.2 LTE Network Architecture . . . . .	11
2.1.3 LTE Quality of Service . . . . .	13
2.2 Video Streaming Service . . . . .	15
2.2.1 Background of Video Streaming Service . . . . .	15

2.2.2	Video Streaming Architecture . . . . .	15
2.2.3	Video Streaming Quality Metrix . . . . .	16
<b>3</b>	<b>QoS Measurement and Analysis</b>	<b>18</b>
3.1	QoS Measurement . . . . .	18
3.1.1	Objective measurements . . . . .	18
3.1.2	Subjective measurements . . . . .	19
3.2	QoS Measurement Techniques . . . . .	20
3.2.1	Opensignal . . . . .	20
3.2.2	Nemo Handy . . . . .	21
3.2.3	Cloudflare Speed Test . . . . .	23
3.3	Correlation Coefficients . . . . .	23
3.3.1	Pearson’s correlation coefficient . . . . .	23
3.3.2	Spearman’s rank-order correlation coefficient . . . . .	24
<b>4</b>	<b>Data Collection and Analysis Methodologies</b>	<b>26</b>
4.1	Objective Data . . . . .	26
4.2	Subjective Data . . . . .	29
4.3	System Model . . . . .	34
4.4	Correlation Coefficient Selection . . . . .	35
<b>5</b>	<b>Analysis, Results and Interpretations</b>	<b>36</b>
5.1	Subjective Evaluation Results . . . . .	36
5.2	Objective Evaluation Results . . . . .	41
5.2.1	Application Layer Evaluation Results . . . . .	41
5.2.2	Network Layer Evaluation Results . . . . .	45
5.2.3	Physical Layer Evaluation Results . . . . .	49
5.3	Correlation Analysis Results . . . . .	52
5.3.1	Correlation among MOS and Application, Network and Physical Layer Quality Metrics. . . . .	52
5.3.2	Correlation between Application and Network Layer Quality Metrics. . . . .	55
5.3.3	Correlation between Network and Physical Layer Quality Metrics. . . . .	56
<b>6</b>	<b>Conclusion and Future Work</b>	<b>60</b>
6.1	Conclusion . . . . .	60

6.2 Future Work . . . . .	62
<b>Reference</b>	<b>63</b>
<b>A Sample SPSS Correlation results</b>	<b>68</b>
<b>B Subjective Metrics Questionnaire</b>	<b>70</b>

# List of Tables

3.1	A Conventional Approach to Interpreting a Correlation Coefficient [40]. . . . .	24
4.1	Overview of the subscribers' personal profiles. . . . .	30
4.2	Experiment Parts of the Day with Respective No. of Subjects Percentage. . . . .	31
4.3	Characteristics of videos used in subjective tests. . . . .	31
4.4	ITU-T 5-Point Scale – Absolute Category Rating. . . . .	32
4.5	Contents of Demographic Questions. . . . .	32
4.6	Contents of User Perception Survey Questions. . . . .	33
5.1	Results of Application Layer Quality Metrics. . . . .	42
5.2	Results of Physical Layer Quality Metrics. . . . .	52
5.3	Correlation results among MOS and Application, Network and Physical Layer Quality Metrics. . . . .	53
5.4	Correlation results among Application and Network Layer Quality Metrics. . . . .	55
5.5	Correlation results among Network and Physical Layer Quality Metrics. . . . .	57

# List of Figures

1.1	Mobile Data Demand Forecast in Addis Ababa (from 2017 – 2021) [5].	2
1.2	Different Factors influencing QoE [13]. . . . .	3
1.3	Summary of Methodology . . . . .	7
2.1	LTE Network Architecture [26]. . . . .	12
2.2	QoS Bearer [27]. . . . .	14
2.3	Video streaming Architecture [19]. . . . .	16
3.1	Opensignal’s Metrics [31]. . . . .	20
4.1	Opensignal Home Page and Sample Measured Outputs . . . . .	27
4.2	Sample Actix Analyzer route map . . . . .	28
4.3	Cloudflare speed test sample measured output . . . . .	29
4.4	System Model for LTE Video Streaming Correlation Analysis . . . . .	34
5.1	Average MOS per frequency of participants . . . . .	37
5.2	Result of Average MOS per Experiment Locations . . . . .	37
5.3	Indoor and Outdoor MOS Difference . . . . .	38
5.4	Variation of ”MOS” per Experiment Locations and Time . . . . .	39
5.5	Impact of Application KPIs on MOS . . . . .	40
5.6	MOS per Participant’s Age and Gender . . . . .	41
5.7	Average Startup Delay, Stalling Duration and Playback per Experimental Locations. . . . .	43
5.8	Average Startup Delay, Stalling Duration and Playback Results per Experiment Parts of the Day. . . . .	44
5.9	Application Layer KPIs per the Experimental Environment Context. . . . .	44
5.10	Geographical Locations of the Experiments. . . . .	45
5.11	Average Network Layer Quality Metrics per Experimental Locations. . . . .	46

5.12	Average Network Layer Quality Metrics per Experiment Parts of the Day. . . . .	47
5.13	Average Network Layer Quality Metrics per Experiment Environmental Context. . . . .	47
5.14	Evaluation Results of Average Download Throughput in Mbps. . . . .	48
5.15	Evaluation Results of Average Latency in ms. . . . .	49
5.16	Evaluation Results of Average Jitter in ms. . . . .	49
5.17	Average RSRP/SINR achieved per Experimental Locations. . . . .	50
5.18	Average RSRP/SINR Achieved per Environmental Context. . . . .	50
5.19	Average RSRP/SINR Achieved per Parts of the Day. . . . .	51
5.20	Scatter Plot Representing the Correlation between Playback and MOS. . . . .	54
5.21	Scatter Plot Representing the Correlation between Download throughput and Playback. . . . .	56
5.22	Scatter Plot Representing the Correlation among Download Throughput with RSRP and SINR. . . . .	58

# Acronyms

<b>2G</b>	Second-Generation
<b>5G</b>	Fifth Generation
<b>ACR</b>	Absolute Category Rating
<b>ACR-HR</b>	Absolute Category Rating-Hidden Reference
<b>ARP</b>	Allocation and Retention Priority
<b>AVC</b>	Advanced Video Codec
<b>CAAZ</b>	Central Addis Ababa Zone
<b>CDMA</b>	Code Division Multiple Access
<b>DSCQS</b>	Double Stimulus Continuous Quality Scale
<b>E-UTRAN</b>	Evolved UMTS Terrestrial Radio Access Network
<b>EAAZ</b>	East Addis Ababa Zone
<b>eNB</b>	Evolved Node B
<b>EPC</b>	Evolved Packet Core
<b>ET</b>	Ethio Telecom
<b>EVDO</b>	Evolution-Data Optimized
<b>FDMA</b>	Frequency-Division Multiple Access
<b>FPS</b>	Frames Per Second
<b>FR</b>	Full Reference
<b>FTP</b>	File Transfer Protocol
<b>GBR</b>	Guaranteed Bit Rate
<b>GSM</b>	Global System for Mobile Communications

<b>HetNet</b>	Heterogeneous Networks
<b>HSPA</b>	High Speed Packet Access
<b>HSPA+</b>	Evolved High Speed Packet Access
<b>HSS</b>	Home Subscriber Server
<b>HTML</b>	Hypertext Markup Language
<b>HTTP</b>	Hypertext Transfer Protocol
<b>ICT</b>	Information and Communication Technologies
<b>IMS</b>	Multimedia Subsystem
<b>ISP</b>	Internet Service Provider
<b>ITU</b>	International Telecommunication Union
<b>LTE-A</b>	LTE-Advanced
<b>MCM</b>	Multicarrier Modulation
<b>MIMO</b>	Multiple Input Multiple Output
<b>MME</b>	Mobility Management Entity
<b>MOS</b>	Mean Opinion Score
<b>NMS</b>	Network Management Systems
<b>Non-GBR</b>	Non-Guaranteed Bit Rate
<b>NR</b>	No Reference
<b>NR</b>	New Radio
<b>OFDMA</b>	Orthogonal Frequency Division Multiple Access
<b>PAR</b>	Peak to Average Ratio
<b>PC</b>	Personal Computer
<b>PCRF</b>	Policy and Charging Rules Function
<b>PSNR</b>	Peak Signal Noise Ratio
<b>QCI</b>	Quality of Service Class Identifier
<b>QoE</b>	Quality of Experience
<b>QoS</b>	Quality of Service
<b>RAN</b>	Radio Access Network

<b>RR</b>	Reduced Reference
<b>RSRQ</b>	Reference Signal Received Quality
<b>RTT</b>	Round-trip Time
<b>S-GW</b>	Serving Gateway
<b>SC-FDMA</b>	Single-Carrier Frequency Division Multiple Access
<b>SISO</b>	Single-input Single-output
<b>SPSS</b>	Statistical Package for the Social Sciences
<b>SSCQE</b>	Single Stimulus Continuous Quality Evaluation
<b>SSIM</b>	Structural Similarity
<b>SWAAZ</b>	South West Addis Ababa Zone
<b>SWAAZ</b>	South West Addis Ababa Zone
<b>TDMA</b>	Time-Division Multiple Access
<b>TEPs</b>	Telecom Expansion Projects
<b>TFT</b>	Traffic Flow Templates
<b>UE</b>	User Equipment
<b>VLoTE</b>	Voice over Long-Term Evolution
<b>VQM</b>	Video Quality Metric
<b>WCDMA</b>	Wideband Code Division Multiple Access
<b>WiFi</b>	Wireless Fidelity

# Chapter 1

## Introduction

### 1.1 Background

Information and communication technology (ICT) has become more valuable in all aspects of life [1]. Cellular mobile communication networks are one of the basic parts of ICT. In the past two decades, the number of mobile users has increased in a very rapid manner. It is also estimated that in 2023 there will be 5.7 billion mobile users [2] and they require access to a wide spectrum of various multimedia application services without being limited by constraints such as time, location, technology, device, and mobility restrictions.

Most of the connections to video platforms and social networking applications originate from wireless cellular mobile communication networks. This massive access from wireless networks introduces several issues. Following that, telecom service providers are increasing the access speed that allows a huge number of users and changing the way of managing the mobile communication network to guarantee QoS and Quality of Experience (QoE).

In the case of Ethiopia, under different Telecom Expansion Projects (TEPs), network capacity, coverage, and quality of various mobile data services are improved. The expansion work so far and the undertaking of additional network expansion projects have allowed ET to carry a large number of customers. As of January 2021, the total number of subscribers has reached more than 50.7 million. Of these, 23.5 million customers are data and internet users [3, 4], and the mobile data demand in Addis Ababa city is forecasted to be 20 PB per month in 2021, as shown in Figure 1.1 [5].

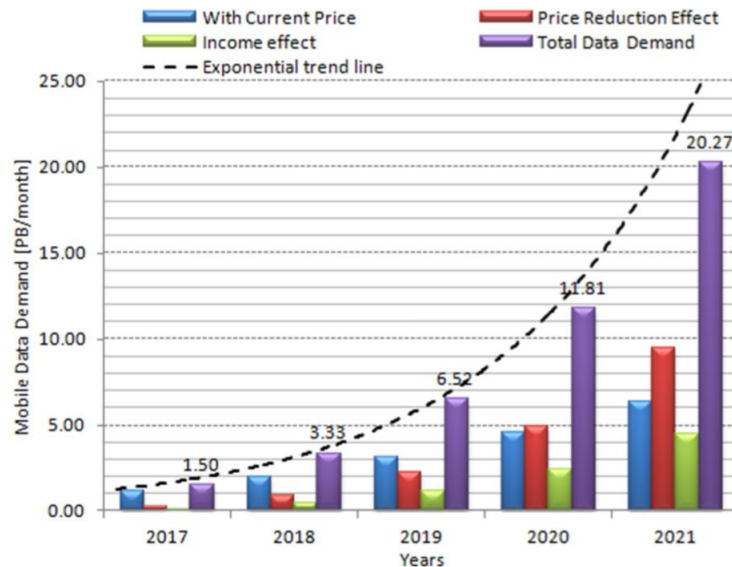


Figure 1.1: Mobile Data Demand Forecast in Addis Ababa (from 2017 – 2021) [5].

The 3GPP has defined the 4th mobile generation radio technology, named LTE or the Evolved Universal Terrestrial Access Network (E-UTRAN), to provide wireless broadband with respect to QoS support [6, 7].

The quality concept takes two forms, which are QoS and QoE. QoS is the network quality evaluated from a technical perspective to determine network performance, through measuring several factors (i.e., throughput, available bandwidth, delay, error probability, jitter, packet loss, etc.) According to the International Telecommunication Union (ITU) recommendation, the QoS was defined as a collective effect of service performance, which determines the degree of satisfaction of a user of the service. On the other hand the QoE is the degree of delight or annoyance of the user of an application or service as perceived subjectively includes the complete end-to-end system effects... maybe influenced by user state, content and context. [8, 9].

Many systems, such as video streaming QoS, are unable to capture the impact of network fluctuations on the user experience. To overcome this issue, the QoE concept was developed, which focuses purely on subjectively perceived quality. It results from the fulfillment of his or her expectations with respect to the utility and/or enjoyment of the application or service in the light of the user’s personality and current state [8, 10].

Assessing QoE is very challenging because the evaluation involves a number of subjective factors such as the mood of the user, responsiveness of the system, etc., as shown in Figure 1.2. Any feature of a user, system, service, application, or context

whose current state or setting may have an impact on the user's QoE [11, 12].

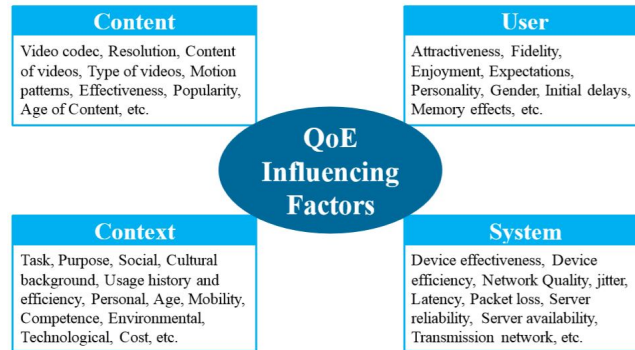


Figure 1.2: Different Factors influencing QoE [13].

Video streaming service is one of the dominant service around the world [13]. Similarly, a recent study in the Ethiopian capital, Addis Ababa, shows that video streaming service has the largest share of LTE services [14]. Although the service has a large number of users and generates better revenue, there are many complaints about the service [3, 15]. Hence, it is very important to investigate the current network performance to demonstrate the extent to which the areas surveyed are sufficient for different video quality and study correlation among QoE and QoS parameters to identify the most influential factors on the QoE of end users.

## 1.2 Statement of the Problem

Online streaming traffic is increasing and is estimated to increase in a very fast manner. Specifically, video streaming has become a dominant service with almost 58% of overall downstream traffic around the globe [16, 17]. This trend leads to an unprecedented boost in cellular data traffic and network congestion, which results in high latency and packet loss. At the end of the day, end users' overall service quality as well as operators' revenue will be affected [18].

Similarly, in Addis Ababa, where the case study was conducted, the data collected from October 2019 to April 2020 shows that the LTE customers' data usage has become more than tripled within a seven-month interval and the video streaming service has the largest share [14].

Although ET has deployed the LTE network and has quality assurance processes in order to make the network operation more reliable and improve the network service quality, data-related customer complaints are still continuing [3, 15]. Therefore,

it is essential to study the correlation between QoE and QoS, as well as spatial and temporal LTE network measurement variations in terms of video streaming service.

## 1.3 Objectives

### 1.3.1 General Objective

The main aim of this thesis is to study the correlation among QoE and QoS for LTE Video Streaming service to identify the QoS level of impact on QoE of video streaming users and to identify spatio-temporal measurement variations.

### 1.3.2 Specific Objectives

The specific objective of the thesis is summarized as follows:

- To identify relevant subjective and objective quality measuring metrics for video streaming services.
- To evaluate ET's LTE services quality monitoring process.
- To identify tools used to measure and analyze QoS performance metrics and QoE.
- To collect video streaming QoS and QoE data from Addis Ababa's LTE network and users.
- Review the literature on video streaming services to identify factors that influence users' QoE, types of correlations, and different layers' quality metrics.
- Indicate the current LTE network measurement variations in terms of location and time, in terms of the video streaming service quality requirements.
- Conducting the correlation study between QoE and QoS parameters to identify the most influential factors with respect to space-and-time manner.

## 1.4 Literature review

Recently, video streaming service has received a tremendous amount of research attention and become a point of interest in research [19].

In [20], the author studied how network path quality affects the QoE of HTTP video streaming. The study addressed the problem by dividing into two parts: measuring the correlation between the network QoS and application QoS and measuring the correlation between application QoS and QoE. In the first subproblem, the author proposed three application performance metrics for HTTP video streaming and used an analytical model and empirical evaluation to characterize their correlation. In the second subproblem, subjective assessments were used to correlate the MOS and the application QoS. As a network QoS, round-trip time (RTT), packet loss rate, and network bandwidth have been used. On the application side, initial buffering time, mean rebuffering duration, and rebuffering frequency are considered, and from the user side, MOS is used. Finally, the findings of this research showed that packet losses and RTT reduce network throughput and the rebuffering frequency to be the main factor responsible for the MOS variance.

V. D. Bhamidipati investigated the effect of delay and delay variation on the QoE in video streaming [21]. The perceived QoE for different streaming sessions has been collected using MOS and finally the author figures out the threshold level of delay that is acceptable by the users, the tolerance of users for different delay levels, and the effect of packet delay. The acceptable tolerance was found to be  $\pm 8$ ms and  $\pm 10$ ms for different video types. The quality of experience is degraded even for small variations ( $\pm 4$ ms) of delay, and the author concluded that when delay and delay variation increase above  $\pm 10$ ms the QoE decreases.

V. D. Bhamidipati investigated the effect of delay and delay variation on the QoE in video streaming [21]. The perceived QoE for different streaming sessions has been collected using MOS and analysis is done for all videos streamed with constant and varying delay. The author tried to address the following research questions: what is the influence of packet delay/delay variation on video QoE, what is the threshold level of delay/delay variation on video QoE from the user's perspective, and what is the relationship between MOS and delay/delay variation for the videos. The author figures out if the packet delay is kept constant, the user's QoE does not appear to change significantly, indicating that the variance of QoE is low. The quality of experience is degraded even for small variations ( $\pm 4$ ms) of delay, and the acceptable tolerance was found to be  $\pm 8$ ms and  $\pm 10$ ms for different video types, and for an increasing delay/delay variation, the mean opinion score (MOS) decreases and vice versa. Finally, concluded that as the delay and delay variation increase above  $\pm 10$ ms the QoE decreases.

In order to find the correlation between QoS and QoE, the author in [22] in-

investigated user perception of videos (MOS), objective video quality metrics (Video Quality Metric (VQM), Structural Similarity (SSIM), and Peak Signal Noise Ratio (PSNR)) and the network QoS (Packet loss, jitter, and reordering). This study considered three video streaming quality levels of 480p, 720p, and 1080p. The paper introduced a three-phase evaluation method, concluded that MOS is affected by the previous experiences of the watchers, linear combinations of some QoS metrics can correlate well with MOS values, and finally recommended that the method be fine-tuned through further research.

[23] presents a video streaming QoE evaluation over 3G networks for Android based smartphone users. This study tries to show the relationship between QoE and QoS. The author considered the application and network layer quality measures, including average video resolution watched, average initial latency, the average number of stalling events, average total stalling length, and average buffer level, as well as network level parameters such as average download throughput.

Crowd-sourcing measurement approaches, Network Management Systems (NMS), and questionnaires are used to collect objective and subjective data. As a result, the correlation between the average acceptability range and the Mean Opinion Score (MOS) has been found to be strong with a correlation coefficient and significant level values of ( $r=0.852$ ,  $p=0.01$ ). On the other hand, it was determined that the download throughput was lower than the recommended values, with an overall MOS score of 2.79. Finally, the author recommended that ET should give special attention to consumers' perception in addition to network capacity improvement in order to increase customer satisfaction.

As a summary, all the above studies focused on the relationships between different QoS parameters and the QoE of video streaming users, with different approaches in terms of data gathering, analysis, and methodologies. The first three papers are intended to investigate the correlation between network, application, and user-level quality metrics. They did not include spatial and temporal phenomena, as well as the correlation between physical, network, and application layer quality metrics. However, in this study, in addition to the correlation analysis among QoE and QoS parameters, the spatial and temporal aspects, as well as correlations among application, network, and physical layer quality metrics, were considered.

## 1.5 Methodology

To achieve the objectives and address the research questions, the following methodology is used.

The first is a literature review, in which various academic research has been reviewed regarding video streaming, QoS, QoE, objective and subjective quality metrics as well as basic knowledge of correlation coefficients and applications of performance assessment. Next, the quality metrics and correlation coefficients needed for the research are identified, then the objective metrics and user ratings are collected. For objective data measurement, Opensignal, Nemo Handy and Cloudflare measurement tools were employed, while a questionnaire for subjective QoE feedback was used. The collected data were then analyzed with SPSS, Actix Analyzer, and Excel to identify the measurement variations and correlation between QoE and QoS parameters for video streaming service users under the LTE network. Finally, the findings and results were then discussed, and recommendations were made.

The methodology, as shown in Figure 1.3, can be summarized as follows:

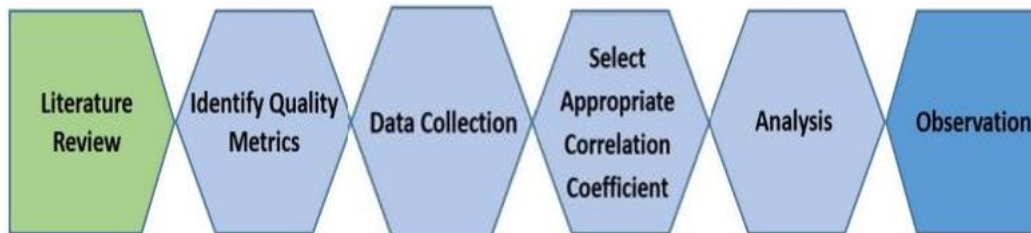


Figure 1.3: Summary of Methodology

## 1.6 Scope and Limitation

### 1.6.1 Scope of the Thesis

This thesis focuses on the correlation among QoE and QoS parameters, as well as indicating the spatiotemporal measurement variations limited to LTE network video streaming service users in Addis Abeba, Ethiopia, where the case study was conducted.

### 1.6.2 Limitations of the Thesis

The study was limited to Addis Ababa; other cities that have recently begun to use LTE services were not considered. This study used 640x480-pixel (480p) resolution videos, but other video quality types such as 720p, 1080p, and so on were not included. Although only eight quality metrics are considered in this paper, additional parameters can be included. The joint effects of quality metrics on QoE are not considered and because it was challenging to find volunteers for the evening and night sessions, the participant distribution for all periods was not proportional.

## 1.7 Contributions

The main contributions of this thesis are:

- The correlation analysis results between QoS factors and QoE show how different factors affect the perceived QoE of video streaming service users and identify the the most influential factors.
- The correlation analysis results among different layers of QoS metrics show how they affect each other and their impact on the upper layer quality metrics.
- The result of spatiotemporal measurement analysis indicate the measurement variations based on video streaming service requirements for different locations and time intervals.
- Overcome the shortcomings of the current ET's root cause analysis approaches and QoS/QoE measurement in terms of video streaming service.
- Consequently, it improves customer satisfaction, and in a future competitive environment, it helps to reduce customer migration threats.

## 1.8 Thesis Organization

This thesis work is organized into six chapters. Chapter 1 introduces a statement of the problem, objectives, methodologies, and scope. Chapter 2 provides technical background on the LTE Mobile Network and Video Streaming Service. Chapter 3 focuses on current QoS measurement and analysis, as well as correlation concepts and mechanisms for choosing correlation coefficients. The data collection and analysis methodologies used in this thesis are presented in Chapter 4. A summary of

the data analysis, results, and interpretations are covered in Chapters 5. Finally, conclusions and future work are included in Chapter 6.

# Chapter 2

## LTE Mobile Network and Video Streaming Service

### 2.1 LTE Mobile Network

#### 2.1.1 Background of LTE Mobile Network

Telephone technology began by providing voice, then data, and finally multimedia data, but this was insufficient for humans, who remained unsatisfied until now. They may demand high data traffic that varies in quantity and variety, such as video streaming, online gaming, video conferencing, etc. To accommodate the ever-increasing demand from users, in 2004, 3GPP developed a new system having different requirements such as [24]:

- Packet-switched domain optimization.
- Roundtrip time between server and user equipment (UE) must be below 30ms and access delay below 300ms.
- Uplink peak rate 75Mbps
- Downlink peak rate 300Mbps
- Improvement to mobility and security
- Terminal power efficiency improvements

Furthermore, due to the fact that LTE is a relatively new technology, it has the following characteristics: High data rate, improved latency, wide coverage, high

spectral efficiency, low cost, and a simplified structure. Therefore, in order to meet the aforementioned requirements, LTE network system mainly used three important technologies such as:

**Orthogonal Frequency Division Multiple Access (OFDMA):** OFDMA is a higher-quality air access method than its predecessor. The downlink of LTE uses OFDMA, whereas the uplink uses Single Carrier Frequency Division Multiple Access (SC-FDMA). This technology was chosen because it can provide the desired high data rates with a simple implementation that involves low-cost and energy-efficient hardware. OFDM is a type of Multicarrier Modulation (MCM), which is a parallel transmission technology that divides a radio frequency channel into many narrow-bandwidth subcarriers and then sends data on each subcarrier.

**Single-Carrier Frequency Division Multiple Access (SC-FDMA):** it is mainly used for uplink transmission and like other multiple access technologies (TDMA, FDMA, CDMA, OFDMA), it works by allocating a single communication channel to multiple users. In uplink, this technology was used to minimize the Peak to Average Ratio (PAR). The key advantages of this technology for mobile devices are improved transmission power efficiency and lower power amplifier costs.

**Multiple Input Multiple Output (MIMO):** This technology is primarily used to reduce noise impacts and improve link reliability. The core concept is that a transmitter transmits multiple streams on multiple transmit antennas, and each transmitted stream takes a distinct path to reach each receiver antenna. The same stream takes different paths to reach multiple receivers, allowing errors to be cancelled using superior signal processing technique.

### 2.1.2 LTE Network Architecture

LTE network consists of two key parts: The radio access network, known as Evolved UMTS Terrestrial Radio Access Network (E-UTRAN), and Evolved Packet Core (EPC), which is the packet core network.

**E-UTRAN:** it is the LTE radio access network which consists of User Equipment (UE) and evolved NodeBs (eNBs).

**UE:** Any device used by an end user to communicate with the base station.

**eNodeB:** is a major part of LTE network structure and unlike earlier generation systems, LTE integrates a radio different functionalities into the eNodeB such as: radio resource management, IP header compression and encryption of user data stream, scheduling and transmission of paging messages, selection of Mobility Management Entity (MME) at UE attachment, and the like. eNodeB uses two interfaces, the first is called S1 interface, and its purpose is to connect eNodeB with EPC which allows the UE to communicate with EPC. The other one is X2 interface mainly used to interconnect eNodeB and supports enhanced mobility and inter-cell interference management [25, 26, 27]. Figure 2.1 shows an overview of the LTE architecture and its basic components.

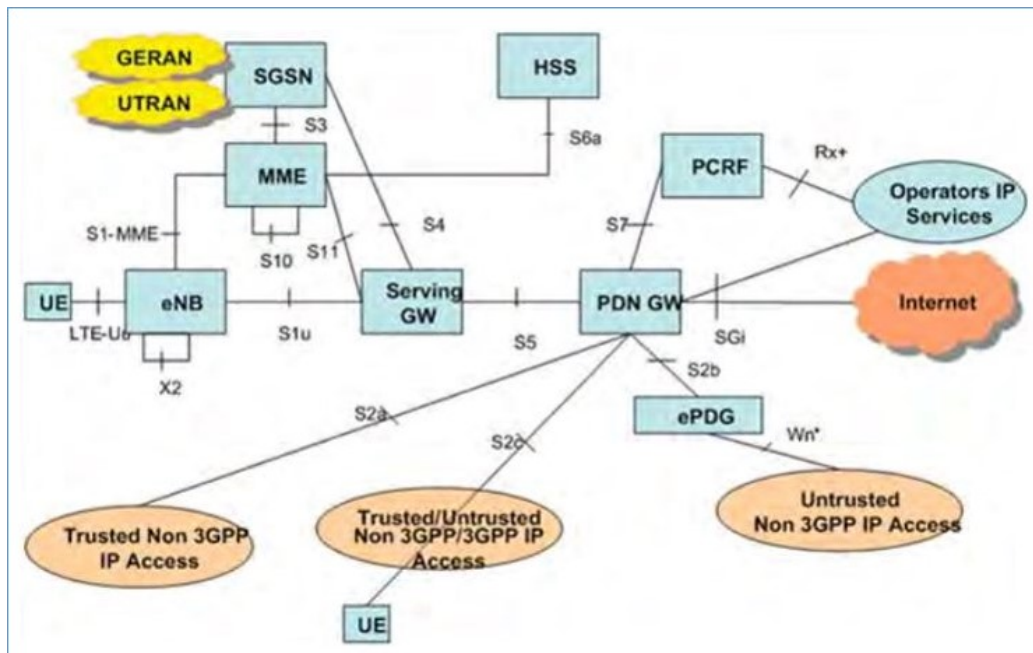


Figure 2.1: LTE Network Architecture [26].

The EPC is composed of several functional entities such as:

- MME
- Home Subscriber Server(HSS)
- Serving Gateway (S-GW).
- Packet Data Network Gateway (PDN-GW)
- Policy and Charging Rules Function (PCRF)

**MME:** is in charge of all subscriber and session management control plane operations. In summary, MME is responsible for mobility, UE identification, security, and idle state mobility control.

**HSS:** It is the permanent user subscription data store. The HSS keeps the master copy of the subscriber profile, which includes information about the services that the user is authorized. It also saves the user's location in the level of visited network control node, such as MME.

**S-GW:** The S-GW is the point at which the packet data interface connects to the E-UTRAN. Its functionalities are acting as a local mobility anchor, exchanging packets with eNodeBs, allowing packet routing and forwarding, and providing QCI granularity.

**PDN-GW:** Serves as a bridge between the core network and the PDN. IP address allocation, policy enforcement, packet classification, routing, uplink and downlink service level charge are the primary functions of P-GW. It also serves as a mobility anchor for access networks that are not 3GPP.

**PCRF:** PCRF's main responsibility is to start QoS if a user requests better QoS, as well as to identify service flow and enable charging policy and IP Multimedia Subsystem (IMS) configuration for each user.

[26] provides additional information about the LTE architecture.

### 2.1.3 LTE Quality of Service

QoS is implemented between the UE and the PDN Gateway in LTE-based networks [28] and applied through a set of bearers. These bearers groups comprise radio bearers, S1 bearers, and S5/S8 carriers referred to as EPS. A bearer is a traffic separation element that allows traffic with different QoS requirements to be treated differently. A bearer establishes a virtual path between a UE and the PDN Gateway. As shown in Figure 2.2 below, Bearers are classified into two categories: Default Bearer and Dedicated Bearer.

**Default Bearer:** When a mobile device initiates the connection to the LTE network for the first time, the default bearer will be assigned to the mobile device according to its service requirements and will remain connected until the UE disconnects from the network. The default bearer does not provide any guaranteed bit

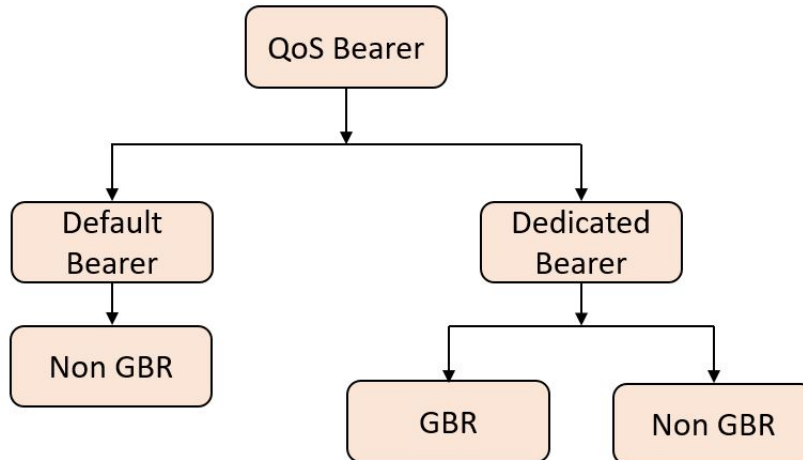


Figure 2.2: QoS Bearer [27].

rate service; instead, it provides best-effort service. Each default bearer is assigned a unique IP address. The default bearer can be assigned a QoS Class Identifier (QCI) of 5 to 9 (Non-Guaranteed Bit Rate).

**Dedicated bearer:** Serves as a dedicated tunnel for specific services to receive appropriate treatment like VoIP, video. It does not require a new IP address, although it does share the default bearer’s address. Both Guaranteed Bit Rate (GBR) and Non-Guaranteed Bit Rate (Non-GBR) Service are available from Dedicated Bearer. Dedicated bearers to give special treatment to specific services use Traffic Flow Templates (TFT).

**GBR:** are commonly utilized in applications such as VoIP and other real-time voice calling apps. GBR QoS parameter values are assigned to each bearer. There is no probability of congestion-related packet loss for a service that uses the GBR bearer if the traffic carried by the GBR bearer conforms to its value. Because it reserves all transmission resources during an admission control function, a Guaranteed Bit Rate (GBR) bearer is frequently created ”on demand.

**Non-GBR:** It does not guarantee a certain bit rate service, however. This bearer is primarily used for web browsing and FTP transfers. Non-GBR bearer services are prone to packet loss due to congestion. It does not block any specific transmission resources. the non-GBR bearer is established as the default or dedicated bearer and stays for a longer time.

Different QoS parameters are used in an LTE network to evaluate the overall performance of a service experienced by the users, depending on whether it provides real-time or best-effort service. A bearer has two or four QoS parameters, such as:

- QoS Class Indicator (QCI)
- Allocation and Retention Priority (ARP)
- GBR – real-time services only
- MBR – real-time services only

[26] Provides more information on the QoS factors described above, as well as other parameters included in the 3GPP standard.

## 2.2 Video Streaming Service

### 2.2.1 Background of Video Streaming Service

Recent advancements in computing technology, compression technology, high bandwidth storage devices, and high-speed networks have enabled the delivery of real-time multimedia services over the Internet. There are two modes of media transmission over the Internet: download mode and streaming mode. In the download mode, the user can only play a downloaded file after the entire file has been downloaded from a server. Transmitting the entire video file takes a long and possibly unacceptable amount of time. On the other hand, in the streaming mode, the entire video file does not have to be downloaded. That is, the video loads a little bit at a time and the browser plays the video rather than copying and saving it. Details of historical background and video streaming protocols are described in [19].

### 2.2.2 Video Streaming Architecture

Figure 2.3 depicts a video streaming architecture that shows how raw video and audio data are pre-compressed using video and audio compression techniques before being saved in storage devices. A streaming server receives compressed video/audio data from storage devices at the request of the client, and then the application-layer QoS control module adapts the video/audio bit-streams to the network state and QoS requirements. The transport protocols then packetize the compressed bit-streams and send the video/audio packets to the Internet after adaption. Due to congestion, packets may be dropped or delayed on the Internet. Packets that are successfully transmitted to the receiver pass through the transport layers and are then processed by the application layer and finally decoded at the video/audio decoder.

Other video streaming-related factors are briefly discussed in [19], including media synchronization mechanisms, QoS control, continuous media distribution services, streaming servers, and video compression and streaming media protocols.

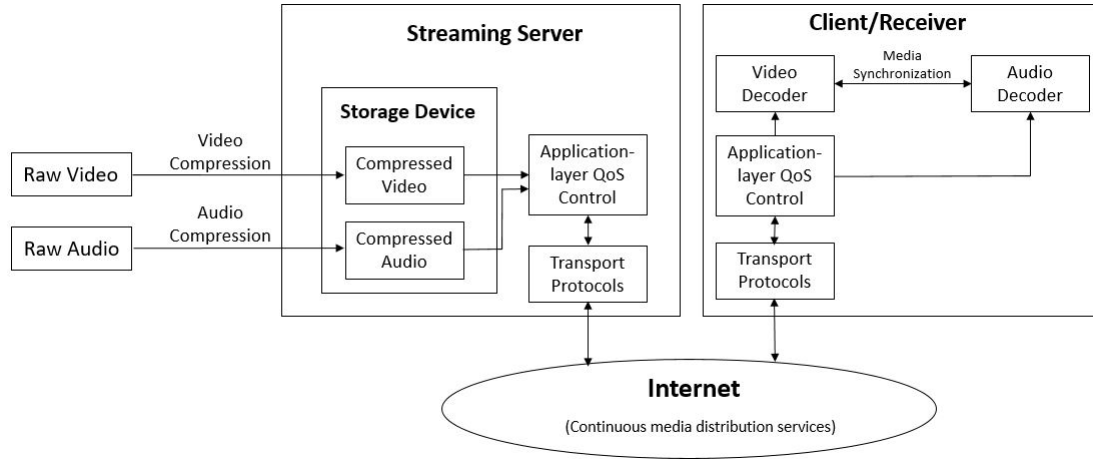


Figure 2.3: Video streaming Architecture [19].

### 2.2.3 Video Streaming Quality Metrix

The network operators must be able to react on quality issues, before the user perceives them. To address such issues 3GPP has defined the LTE technology, to provide wireless broadband with the respect of QoS support. There are two types of quality concepts: QoS and QoE. For Video/Audio quality [29] identified the below KQIs which are the most important for the video streaming service and it is defined as the ability and immediacy of the end user to understand the video content. The measurement is taken during the active session service.

- **Delay:** The total time takes from a packet to travel from source to destination.
- **Jitter:** The delay variation.
- **Packet Loss Rate:** It is defined as the number of packets that are successful received, or not, at the destination.
- **Mobility:** The mobility success rate in terms of handover performance of the service.
- **Utilization:** The EPC side like the EPC bearer utilization. As mentioned in Chapter 1, QoS is used to determine network performance from a technical perspective by measuring a variety of characteristics such as throughput,

available bandwidth, delay, error probability, jitter, packet loss, and so on. On the other hand, QoE is an evaluation of individuals' experience when using the service, in [20] the fundamental QoE metrics for adaptive video streaming is presented, which are:

- **Startup delay:** It is the time in seconds between when a user initiates a video session and when the media player starts playing video frames.
- **Stalling/Buffering events:** When the player buffer is empty, the video stops playing since there are no more video frames to display on the screen. It is also known as buffer starvation.
- **Stalling/buffering time:** It is the time duration of buffering events occurs during the video session. It is computed as the time duration within each video player interruptions.
- **Stalled/Buffered time rate:** This measure represents the percentage of total stall time out of the entire length of the played audio or video multimedia file.

# Chapter 3

## QoS Measurement and Analysis

### 3.1 QoS Measurement

QoS can be measured using network-side quality metrics to determine whether a service is providing the requested service at the expected quality level or not. However in the context of real-time video streaming apps, QoS is no longer sufficient to evaluate the service quality. Therefore, in addition to QoS, recent studies have turned their focus to QoE. [8, 20], and it helps to determine what improvements in service performance are needed to assure customer satisfaction. QoS can be measured objectively by monitoring the network side quality parameters or subjectively by using opinion scores.

#### 3.1.1 Objective measurements

Depending on the accessibility of the source signal, they are classified into three approaches [8]:

- **Full reference (FR):** A reference video is compared frame-by-frame (e.g., color processing, spatial and temporal features, contrast features) with a distorted video sequence to obtain the quality commonly used in lab testing environments.
- **Reduced reference (RR):** Only some features of the reference signal are extracted and employed to evaluate the quality of the distorted signal.
- **No reference (NR):** The reference video is inessential while evaluating the distorted video sequences Quality. (commonly used for real-time quality as-

assessment of videos.

Some of the most well-known objective quality assessment approaches are: Peak Signal to Noise Ratio (PSNR), Structural Similarity Metric (SSIM), Multi-Scale Structural similarity, Video Quality Model (VQM), Natural Image Quality Evaluator (NIQE), etc.

### 3.1.2 Subjective measurements

Subjective assessment is considered as the most accurate approach to measure the QoE perceived by the end user. This method uses volunteer video streaming service users to evaluate video sequences and then score them based on their point of view and perception; the MOS is the average of the values obtained for each test sequence. However, this method is very resource intensive in terms of human resources, cost, and time. Different subjective measurement approaches can be used to assess video quality; some of the widely known subjective measuring methods are introduced in [8], which are:

- **Double Stimulus Continuous Quality Scale (DSCQS):** A reference and a processed video sequence are shown to the evaluator twice, after which the evaluator is asked to rate the video's quality on a scale of zero to one hundred, and the difference between the video evaluation values is computed to determine the video's quality. This quality assessment take a long time to implement.
- **Single Stimulus Continuous Quality Evaluation (SSCQE):** The user rates the quality of a 20 to 30-minute video and it is possible to observe the variation in quality over time by calculating an average quality evaluation of the subjects' responses over time. It requires well-trained observers to achieve stable assessment outcomes.
- **Absolute Category Rating (ACR):** It is considered as a single stimulus approaches. It is shown for 10 seconds and then rated on a five-grade quality scale expressed as MOS.
- **Absolute Category Rating-Hidden Reference (ACR-HR):** This is a similar technique to the ACR. Except that, the participants are also shown the reference version of each distorted test sequence. A differential quality score is then used to calculate a final quality evaluation.

- **Absolute Category Rating-Hidden Reference (ACR-HR):** Pair Comparison: The subjects are shown a pair of videos to compare and subsequently evaluate (i.e., which pair has the best quality). There are differences in the findings according on which one was shown first. These and other details about subjective and objective evaluation approaches is presented in [8].

## 3.2 QoS Measurement Techniques

QoS can be evaluated in a variety of ways, including network and user-side approaches. To measure the QoS on the network side, different network management tools can be utilized [20]. Drive Test and Crowd-sourcing techniques, on the other hand, can be used as part of mobile side measuring techniques [29].

### 3.2.1 Opensignal

Opensignal is a free crowdsourcing app that measures the real-world mobile network experience as it is perceived by users, such as video experience, data speed, network availability status, and other metrics illustrated in Figure 3.1. It is available for free download from the official Google Play store and can be used on 3G, 4G, 5G, and wifi networks [30].

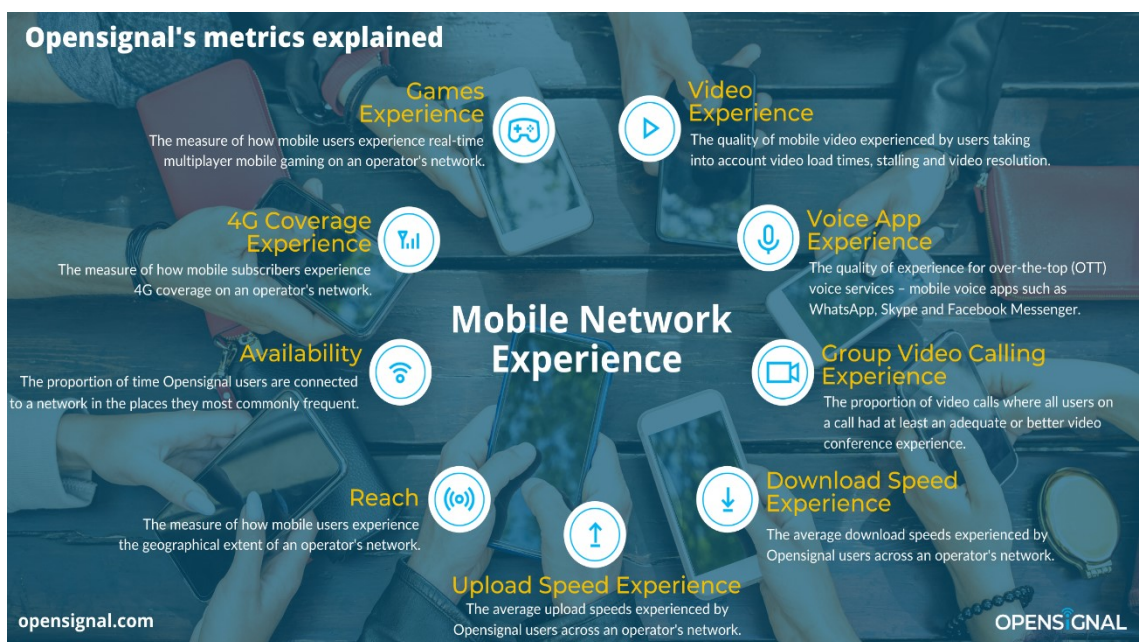


Figure 3.1: Opensignal's Metrics [31].

As presented in [31], Opensignal has several interesting key features, some of these are:

- Test tab with speed test and video streaming quality to analyze mobile video performance with results for time spent buffering, loading and playback.
- Signal dashboard and ping test - for testing latency response.
- Connectivity map to show locations of 3G and 4G/LTE tests.
- Network stats with average speeds for download, upload and latency and network strength on major providers.
- Historical log of your WiFi and 3G, 4G/LTE and 5G speed tests to check connectivity over time etc.

### 3.2.2 Nemo Handy

Nemo Handy is the handheld drive test tool used to measure and monitor the air interface of wireless networks, and mobile applications QoS and QoE. It is an extensive application capable of testing features that offer full application-level metrics on voice calls and quality, File Transfer Protocol (FTP)/Hypertext Transfer Protocol (HTTP) data transfers, YouTube video streaming, video quality testing, Facebook testing, Hypertext Markup Language (HTML) browsing, Netflix testing, Instagram testing, WhatsApp testing, Viber testing, email testing, ping, and also it can measure the indoor and outdoor wireless network with a smart and inconspicuous application installed on an android-based device, etc. [32, 33].

The Nemo Handy tool supports the latest wireless cellular technologies, for example, Fifth Generation (5G) New Radio (NR), LTE/LTE-Advanced (LTE-A), Evolved High Speed Packet Access (HSPA+), High Speed Packet Access (HSPA), Wideband Code Division Multiple Access (WCDMA), Code Division Multiple Access (CDMA), Evolution-Data Optimized (EVDO), Global System for Mobile Communications (GSM) and Wireless Fidelity (WiFi) or Heterogeneous Networks (HetNet). In the LTE network, different QoS attributes can be measure [32, 34], some of these are:

- **Received Signal Strength Indicator (RSSI):** It comprises the linear average of the total received power (in [W]) observed only in certain OFDM symbols of measurement sub frames, in the measurement bandwidth, over N

number of resource blocks by the UE from all sources, including co-channel serving and non-serving cells, adjacent channel interference, thermal noise, etc.

- **Reference Signal Received Power (RSRP):** It is defined as the linear average over the power contributions (in [W]) of the resource elements that carry cell-specific reference signals within the considered measurement frequency bandwidth.
- **Reference Signal Received Quality (RSRQ):** It is defined as the linear average over the power contribution (in [W]) of the resource elements carrying cell-specific reference signals divided by the linear average of the noise and interference power contribution (in [W]) over the resource elements carrying cell-specific reference signals within the same frequency bandwidth.
- **Signal-to-Interference-plus-Noise Ratio (SINR):** It is defined as the ratio  $N \times \text{E-UTRA RSRP} / (\text{E-UTRA carrier RSSI})$ , where N is the number of RBs of the E-UTRA carrier RSSI measurement bandwidth.
- **Throughput:** It is the time-averaged number of correctly received transport blocks in a communication system running an application, where a Transport Block is defined in the reference measurement channel According to 3GPP technical specification,

The amount of data sent or received in one unit of time is referred to as throughput. It is the measure of how much data packets successfully travel through the network [15].

Throughput is calculated with the following formula:

$$\text{Throughput} = \frac{\text{Transmittedblockssize} * \text{Numberofblocksacknowledged}}{\text{Measuringtime}} \quad (3.1)$$

- Verify Location of serving cell
- Single Cell functionality check
- Search good dominant RF environment and etc.

### 3.2.3 Cloudflare Speed Test

The Cloudflare speed test is a tool that allows you to measure the speed and consistency of your connection to the Internet. You can use it to confirm that the speed provided by your internet service provider (ISP) is the speed you are receiving. Compare different ISPs or test network connectivity in different parts of your house. The measurements run on the Cloudflare network, which spans data centers in over 200+ cities worldwide. This ensures you are testing against a server that is close to you, which means you are measuring only the speed of your ISP, with minimal networks in between that may affect your score [35].

## 3.3 Correlation Coefficients

In many studies, correlation analysis is performed to determine the degree of association between study variables. To analyze correlation coefficients, several software programs are employed, and among them, SPSS is one of the best. According to [36], SPSS can provide statistics ranging from simple descriptive numbers to complicated studies of multivariate matrices. The Spearman's Rank Correlation Coefficient and Pearson's Product Moment Correlation Coefficients are the most widely utilized correlation coefficients in researches.

### 3.3.1 Pearson's correlation coefficient

The bivariate Pearson Correlation provides a sample correlation coefficient  $r$ , which indicates the strength and direction of linear correlations between two continuous variables. By extension, the Pearson Correlation examines whether there is statistical evidence for a linear relationship among the same pairs of variables in the population, as indicated by a population correlation coefficient,  $(\rho)$ . Pearson correlation is a type of parametric measure [37].

Pearson's mathematical formulation to quantify the degree of relationship ( $r$ ) between variables, namely,  $X$  and  $Y$ , can be given as [38, 39]:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (3.2)$$

Where  $x_i$  and  $y_i$  are the values of  $x$  and  $y$  for the  $i$ th individual.

Pearson Correlation coefficient is affected by extreme values, which may exaggerate

or dampen the strength of relationship, and is therefore inappropriate when either or both variables are not normally distributed. [39].

Fundamentally, the correlation coefficient R is between -1 and +1 and it can be interpreted based on its value, as illustrated in Table 3.1.

Table 3.1: A Conventional Approach to Interpreting a Correlation Coefficient [40].

Absolute Magnitude of the Observed Correlation Coefficient	Interpretation
0.00-0.10	Negligible
0.10-0.39	Weak correlation
0.40-0.69	Moderate correlation
0.70-0.89	Strong correlation
0.90-1.00	Very strong correlation

### 3.3.2 Spearman's rank-order correlation coefficient

Spearman's rank correlation coefficient is a nonparametric (distribution-free) rank statistic proposed as a measure of the strength of the association between two variables. It is a measure of a monotone association that is used when the distribution of data makes Pearson's correlation coefficient undesirable or misleading [40]. It is appropriate when one or both variables are skewed or ordinal and is robust when extreme values are present.

The formula for computing the Spearman's correlation coefficient for a correlation between variables x and y is given by:

$$r_s = \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} \quad (3.3)$$

Where  $d_i$  is the difference in ranks for x and y. Correlation coefficients are used for various sorts of data, thus there are several factors to consider while selecting one. Using a correlation statistic, without first checking or when the normality assumption is violated, inferences and interpretations may become unreliable or invalid. Three common normality test methods are discussed in [41]: histograms, boxplots, and Q-Q-plots, which are graphical methods, as well as numerical methods (skewness and kurtosis indexes) and formal normality tests.

Other details, including the correlation types discussed above and normality check methodologies can be found in [40, 41, 42].

# Chapter 4

## Data Collection and Analysis Methodologies

In this study, we used subjective and objective data collection methods; the subjective data (video streaming user's perception) has been collected using a survey questionnaire. It includes the respondent demographic status like age, gender, job status, educational background, etc. On the other hand, the objective data has been collected using a standard handheld network measurement device called Nemo Handy, Opensignal and Cloudflare speed test. It comprises three layers of quality metrics, which are application, network and physical layers. The data collected from six geographically distributed locations based on their monthly traffic distribution and geographic location from April 20, 2021 to May 28, 2021 for six weeks and each performance measurement was taken at four different times of the day: morning, afternoon, evening and night.

### 4.1 Objective Data

**Opensignal:** As part of this research, Opensignal was employed as a data collection strategy and its free version is available for the mobile platform and for personal computer (PC) as well, in this research we use only the mobile platform. The installation and measurement procedures are quite simple and straightforward. Simply download and install the App like other standard mobile applications, after successful installation when a measurement is need we can run the application by a single click on the corresponding button as shown below in figure 4.1a, and the sample measured outputs for video streaming and data speed experience are depicted in Figure 4.1b and Figure 4.1c.

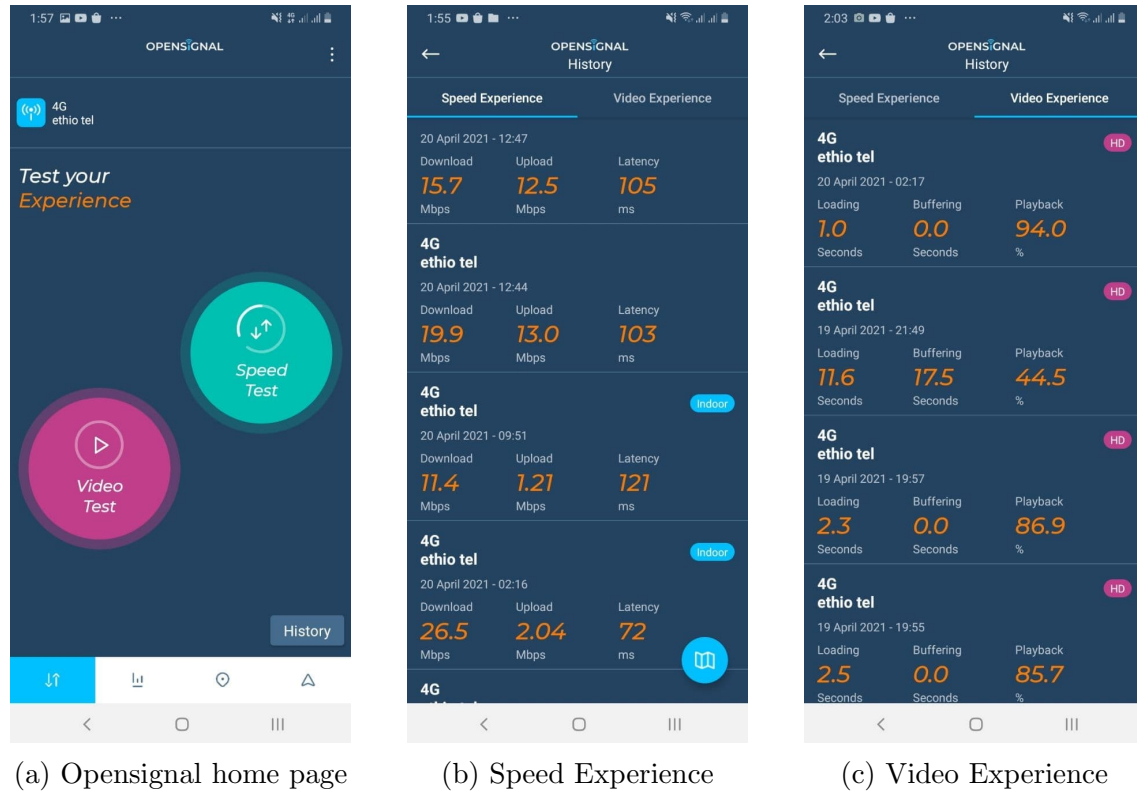


Figure 4.1: Opensignal Home Page and Sample Measured Outputs

As described in chapter 3 above, Opensignal allows you to measure various real-world mobile network experience metrics, however, only the video streaming experience and data speed experience metrics are included in this study:

- Video Experience:** It is derived based on the ITU standard and quantify the quality of videos by measuring real-world video streams from end-user devices over different networks. In this study startup delay, stalling duration and playback video streaming experiences have been selected as an application layer performance and the startup delay and stalling duration measured in second (Sec) and the playback in percentage (%).
- Speed Experience:** Opensignal speed test measures the average upload, download speed experience, and latency for each operator across their mobile data networks. The upload and download speed experience metrics are measured in megabits per second (Mbps), is calculated on the basis of the middle range of samples and the latency is measured in milliseconds (ms). In order to test the download and upload speed the Opensignal speed test runs five seconds for each measurement and for consistent and accurate latency mea-

surement a ping test is used. Opensignal speed test doesn't have a dedicated test server to measure the end-to-end consumer network experience, it runs on a common internet Content Delivery Network (CDN) servers.

**Nemo Handy:** In this research RSRP, SINR and Download throughput network performance metrics were collected using the Nemo Handy tool. The collected data is logged to the phone's internal storage and the log files are made available in wide range of file formats that allows easy post-processing with a Nemo Analyze or third-party post-processing Actix Analyzer tools. Actix Analyzer is a network performance analysis desktop solution, providing advanced drive test survey analytics, and supporting network optimization, network acceptance and validation used in Second-Generation (2G), Third-Generation (3G), LTE, Voice over Long-Term Evolution (VLoTE) and Fifth-Generation (5G) rollouts and it is capable of plotting parameters on maps and display them on charts, tables, and in workbooks [42]. Figure 4.2 and shows a sample Actix Analyzer route map for a specific log file.

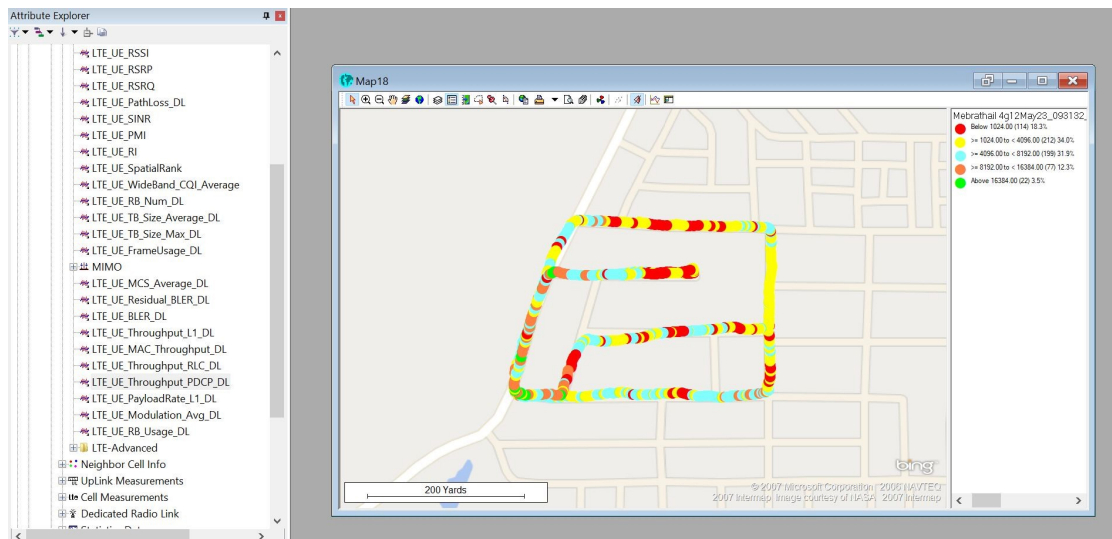


Figure 4.2: Sample Actix Analyzer route map

**Cloudflare speed test:** In this research Cloudflare speed test is used to measure the Jitter, it's calculated as the average latency delta between any two consecutive measurements and the speed is measured by downloading/uploading progressively larger files. Figure 4.3 shows sample jitter test result measured by a Cloudflare speed test.

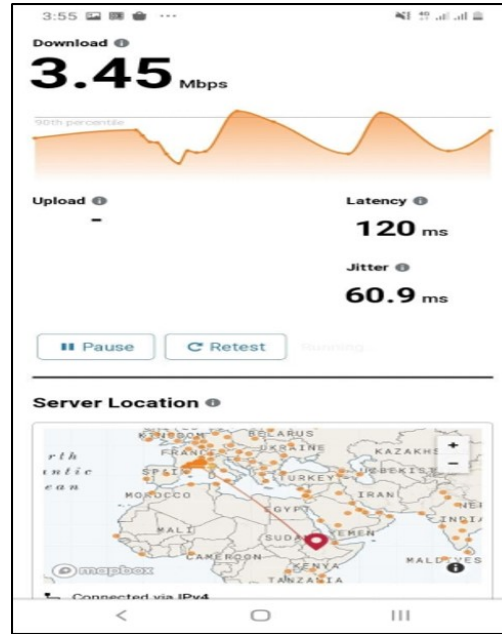


Figure 4.3: Cloudflare speed test sample measured output

## 4.2 Subjective Data

The subjective data collection took place in Ethiopia's capital, Addis Ababa. The data was collected at six distinct locations in four different zones, which were chosen based on their monthly traffic distribution. Central Addis Ababa Zone (CAAZ), South Addis Ababa Zone (SWAAZ), East Addis Ababa Zone (EAAZ), and South West Addis Ababa Zone (SWAAZ) are the four study zones with the six individual study locations Gofa Mebrathail, Gelan, Jemo, Torhailoch, Piassa, and Atlas. The survey questionnaire was contextually formulated to capture the LTE video streaming users' perceptions of stalling duration, stalling frequency and the overall satisfaction level of video streaming users with 300 LTE video streaming service users participating in the subjective data collection.

### Determining Sample Size:

The LTE user base of ET as of November 2020 was 1,492,650, and we used the formula in [43] to calculate the sample size.

$$n = \frac{N}{(1 + N(e)^2)} \quad (4.1)$$

Where  $n$  - sample size,  $N$  - Total customer,  $e$  - detection error expressed in to percentage (5% - 10%) and thus Total customer ( $N$ ) is 1,492,650, Detection error

(e) = 6%, any value for e from 5% - 10% is acceptable.

$$n = \frac{1492650}{1 + 1492650(0.05)^2} = 278 \text{ Respondents}$$

This means, the respondents should not to be less than 278. In order to make the research feasible the sample size of 300 subscribers from four Addis Ababa zones were participated in the survey and random sampling was used to select the 300 respondents. The survey’s 300 participants (204 men and 96 women) were chosen at random from the places where the research was done, and their distribution was based on demographic information such as gender, age, educational background, and so on. Table 4.1 provides an overview of the personal profiles of the survey participants.

Table 4.1: Overview of the subscribers’ personal profiles.

Gender	Male	204
	Female	96
Age	Under 18 years	41
	18 – 25 years	89
	26 – 35 years	72
	36 – 45 years	77
	46 – 55 years	15
	Above 55 years	6
Educational Backgrounds	Master’s degree and above	10
	Bachelor’s degree	50
	University or college student	70
	Diploma or Certificate	120
	Grade 11 – 12	30
	Grade 10 and below	20

The survey data was collected from the six selected specific locations at different times of the day (morning, afternoon, evening, and night). The time from 06:00 AM to 11:59 AM is considered a morning session, the time from 12:00 PM to 04:59 PM is considered an afternoon, the time from 05:00 PM to 07:59 PM is considered an evening, and the time from 08:00 PM to 12:00 AM is considered a night session. The experiment periods of the day are depicted in Table 4.2 with a respective number of participants and percentage of correspondents who participated in the experiment. The majority of the video experiments (72%) were conducted in the morning and afternoon periods.

Table 4.2: Experiment Parts of the Day with Respective No. of Subjects Percentage.

Time Period		No. of Respondents	Percentage(%)
Morning	06:00AM - 11:59AM	108	36
Afternoon	12:00PM - 04:59PM	108	36
Evening	05:00PM - 07:59PM	48	16
Night	08:00PM - 12:00AM	36	12
Total		300	100

**Video Selection:** To collect the user feedback, four YouTube videos have been selected considering various types of video contents, which are news, sport, movie and music with a time duration of 0:03:27, 0:02:28, 0:02:22 and 0:05:09 respectively. The average time duration of selected videos is 0:3:21. All selected videos were MPEG-4 files encoded using the Advanced Video Codec (AVC) at 30 frames per second (FPS), with a video resolution of 480 pixels. Table 4.3 shows characteristics of videos selected for the Experiment.

Table 4.3: Characteristics of videos used in subjective tests.

Category	Video Resolution	Frame Rate	Video Codec	Video Length
News	854 x 480	30 fps	H.264	0:03:27
Sport	854 x 480	30 fps	H.264	0:02:28
Movie	854 x 480	30 fps	H.264	0:02:22
Music	854 x 480	30 fps	H.264	0:05:09

**Questionnaire:** The ultimate goal of this research is to understand how network QoS affects the QoE of LTE video streaming users. To investigate the relationships between different levels of QoS and QoE, both subjective and objective data were collected. For subjective data collection, a systematically designed survey questionnaire consisting of demographic information and users' perceptions was used, and for the evaluation of multimedia technologies such as video streaming, it is necessary to convert user perception into quantitative data.

There are several subjective test methods for video quality assessment, most of them described in the ITU recommendations. In this study, we adopt one of the most widely used scales: the 5-point MOS scale described in the Absolute Category Rating. (ACR) method standardized by ITU-T, as indicated in Table 4.4.

Table 4.4: ITU-T 5-Point Scale – Absolute Category Rating.

Rating scale	Estimated Quality	Perceived Impairment
5	Excellent	Imperceptible
4	Good	Not Annoying
3	Fair	Slightly annoying
2	Poor	Noticeable, Annoying
1	Bad	Perceptible or Very Annoying

**Demographic survey questions:** The first part of the questionnaire (Demographic information) includes five questions age, gender, Job status and educational background, which helped us to consider all the target groups and makes certain generalizations about them, it also included the types of terminal used to watch the video, the exact date, time and place where the video watched. In total, 300 participants responded to the first 10 questions. The demographic questions depicted in the hard copy are shown in Table 4.5.

Table 4.5: Contents of Demographic Questions.

Questions No.	Questionnaire Contents
Q1	Age group
Q2	Gender
Q3	Job Status
Q4	Educational background
Q5	Study zones
Q6	The name of the experiment's location (site).
Q7	Specific location of the experiment (Indoor/Outdoor)
Q8	Terminal type
Q9	Time period
Q10	Date of experiment

**Perception Survey Questions:** When end users watch a streaming video, from the service provider side it is difficult to know the QoE perceived by the end-users or whether they are pleased or annoyed with the service. In this research, to conduct a subjective quality test, A Closed ended Scalar/Likert questions were used [44] and we use MOS, the widely known perceptual video quality measurement technique. This method was used to collect the level of annoyance of selected participants based on their point of view and perception, which was then quantified using a five-point discrete scale (1:bad, 2:poor; 3: fair; 4:good; 5:excellent), with the average of the values obtained from each participant being considered a MOS.

As it is depicted in Table 4.6 under the user perception survey, a total of four ques-

tions related with video streaming quality metrics are included and each participant was asked these questions after watching their favorite videos from the available video contents.

Table 4.6: Contents of User Perception Survey Questions.

Questions No.	Questionnaire Contents
Q1	Perception of annoyance level of initial delay
Q2	Perception of annoyance level of frequency of interruptions(stalling)
Q3	Perception of annoyance level of duration time of interruptions (stalling)
Q4	The overall satisfaction level of the LTE video streaming service

**Evaluation Process:** First, we created awareness of the survey’s objective and provided clarity on each survey question to all participants, which ultimately assisted us in completing the survey without trouble, and we installed the Open Signal apps on the user’s smartphone. Following that, every participant filled out the first part of the questionnaire (demographic questions) before watching the video, and then they selected only one of the four video contents provided and watched the video using their own mobile devices, at the same time all the application, network and physical layer quality metrics were measured together. Once the users have finished viewing the video, they filled the second part of the questionnaire based on the video they just saw. Finally, the data collected from all the survey locations was converted to an integer value according to the MOS standard.

In this survey, around 309 users were participated but outliers were removed from the total number of participants for a variety of reasons. Outliers were identified based on user’s feedback, such as:

- If the selected user perception score by the subscriber is more than one options
- If there is an incomplete information in both demographic and user perception surveys

Based on the aforementioned criteria we identified 9 outliers from a total of 309 respondents; 8 of them were missing one or more user perception scores, while the other was due to incomplete demographic information.

### 4.3 System Model

Different layers QoS determining factors have various effects on video streaming services, to understand their impact on the QoE of video streaming users, we studied the correlation among QoE and QoS parameters. To undertake the study, we collected different quality metrics from the Application, Network, and Physical layers, while the user perception was gathered through a survey questionnaire. From the application layer, startup delay, stalling length and playback performance metrics were collected using Opensignal crowdsourcing tool, latency, jitter, and download throughput, as well as RSRP and SINR performance metrics, were collected from the network and physical layers, respectively, using Cloudflare speed tests and Nemo-handly tools.

The LTE network QoS and QoE were evaluated in each of the study locations and the collected data was used to create a correlation among QoE and different layers of QoS parameters. First, we have done the correlation between MOS and application-layer metrics; then, we studied the relationship of MOS with network and physical-layer metrics. On the other hand, the correlation among application, network and physical-layer metrics (the correlation between one parameter with another layer parameter) have been studied. In general, the collected quality metrics were used as input for the spearman rho statistical algorithm, and its result was interpreted using the standard approach to interpreting a correlation coefficient. Figure 4.4 illustrates system model for video streaming correlation analysis.

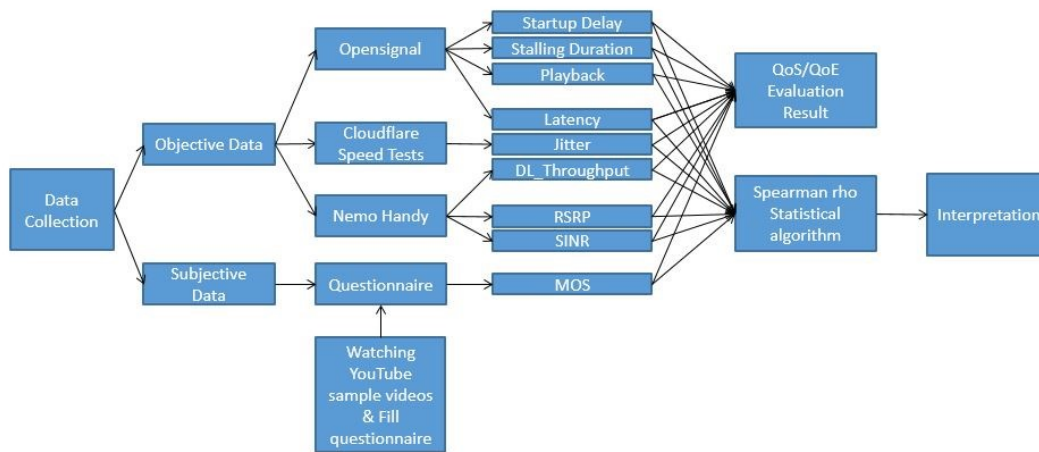


Figure 4.4: System Model for LTE Video Streaming Correlation Analysis

## 4.4 Correlation Coefficient Selection

Various correlation coefficients can be used to measure correlations between different variables. Pearson's coefficient, Spearman's rho coefficient, and Kendall's tau coefficient are the most widely used. We selected the appropriate coefficient using the Shapiro-Wilk test presented in Study [45], which was selected as a most powerful test for all types of distribution and sample sizes. According to the results only RSRP, SINR and Latency quality metrics were normally distributed with the significance values of .362, .038 and .553.

Therefore, as illustrated in [38] and discussed above in chapter 3, because Pearson Correlation Coefficient is affected by extreme values, which can exaggerate or dampen the strength of the relationship, if we must use it, our variables should be normally distributed; otherwise, interpretation and inferences may be unreliable or invalid.

On the other hand, Spearman's rho coefficient is suitable when one or both variables are skewed or ordinal, and it is robust when extreme values are present. As a result, it was decided that Spearman's rho coefficient would be used in this research.

# Chapter 5

## Analysis, Results and Interpretations

This chapter includes the LTE network's spatial and temporal evaluation results, as well as the correlation results and their discussion of the impact of each application, network, and physical layer QoS metrics on video streaming QoE. Furthermore, each result is thoroughly discussed.

### 5.1 Subjective Evaluation Results

Subjective metrics are the results of participant responses to the sample videos, which is ranked based on the video streaming metrics discussed in chapter four of this thesis work. The subjective evaluation conducted at six separate study locations in Addis Ababa using four perceptual survey questions. The ratings used to calculate the overall MOS value were collected from the paper-based ballot questionnaire using a 5-point likert scale presented in chapter four.

Figure 5.1 shows that, the MOS scores varied from 2.0 to 5.0, 70.83% of participants rated the MOS value in between 3.5 to 4.8 and the remaining 29.17% of participants rated a score of less than 3.5. On average, a MOS result of 3.9 is obtained therefore according to ITU-T standard [46], the average MOS result indicated that some users are unhappy with the current video streaming quality.

The spatial MOS result shows where the video streaming service customers perceived the worst and best quality in the LTE cellular mobile communication network. According to the findings of the experimental evaluation, 66.67% of the sample locations (4 locations out of 6) the survey participants were not satisfied on the video streaming quality provided by ethio telecom. In the rest, two study locations at

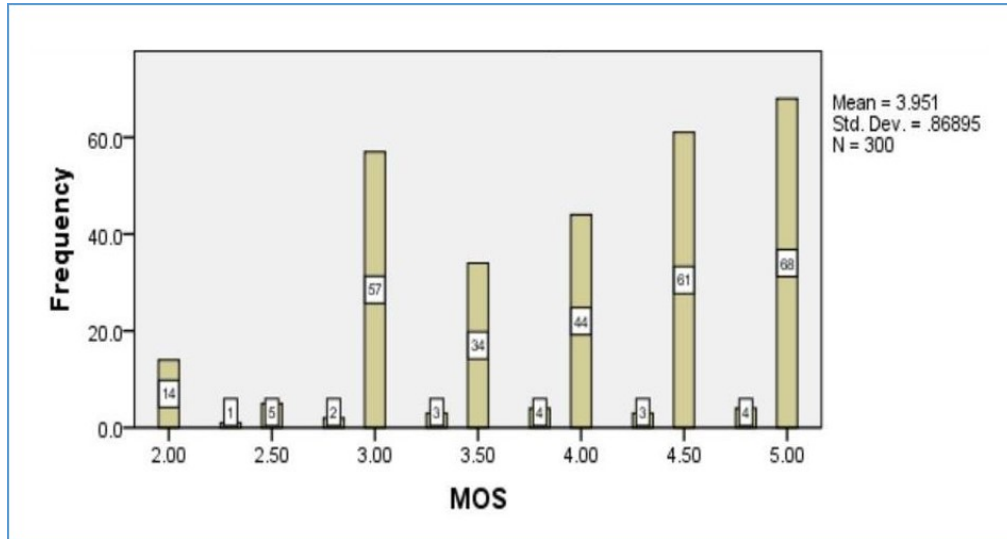


Figure 5.1: Average MOS per frequency of participants

Atlas and Gelan areas the average MOS values were 4.1 and 4.3 respectively, which means they were relatively satisfied with the service, in accordance with the above-mentioned ITU-T standard.

The maximum, average, and minimum MOS values achieved in all study locations are shown in Figure 5.2 below. In all research areas, the maximum MOS value was 5. Except Gelan and Atlas in the remaining 4 study locations the lowest MOS value was 2.0. Atlas and Mebrat had the highest and lowest average MOS, with the values of 4.38 and 3.72, respectively.

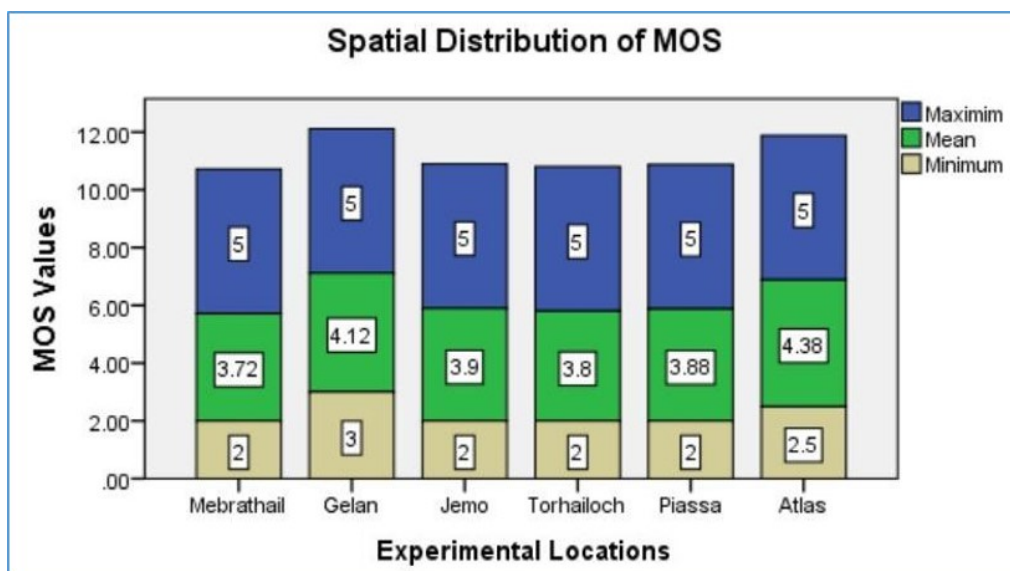


Figure 5.2: Result of Average MOS per Experiment Locations

As per the survey findings, the indoor and outdoor MOS scores were reasonably comparable, however indoor measurements indicating somewhat better results with the average indoor and outdoor MOS values of 3.69 and 4.12 respectively. Despite the fact that the measurements taken indoor showed better results on average, there were several locations where the MOS was quite low. MOS can be affected by a variety of different factors; in this case, the environment context (indoor/outdoor) is one of the possible reasons. As shown in Figure 5.3, the LTE video streaming customers who used the service indoor were less satisfied than those who used it outside.

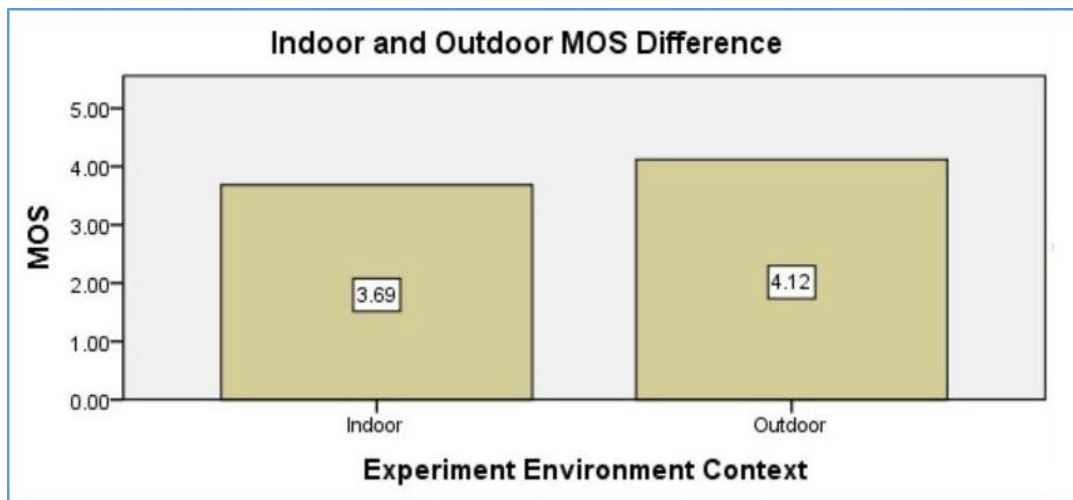


Figure 5.3: Indoor and Outdoor MOS Difference

Morning, Afternoon, Evening, and Night are the four time periods in which the responses of the participants were collected. As it is depicted in Figure 5.4, the smallest MOS was measured in the majority of the research locations during the evening and night hours. Due to the fact that the evening and night time periods are not normal working hours, several users were expected to access the network at the same time, therefore the MOS values at that time appears to be very small. Of the six places surveyed, the lowest MOS value was 2.5 achieved at Mebrathail area. Furthermore, the survey result shows the highest MOS value of 4.8 achieved during the period of the morning at Atlas area.

The MOS results taken during the afternoon time periods were better than from the results of evening and night time periods but they were lower than those in the morning because users have a tendency of accessing the network during the break time and at the end of the day it impacts the QoE of end users.

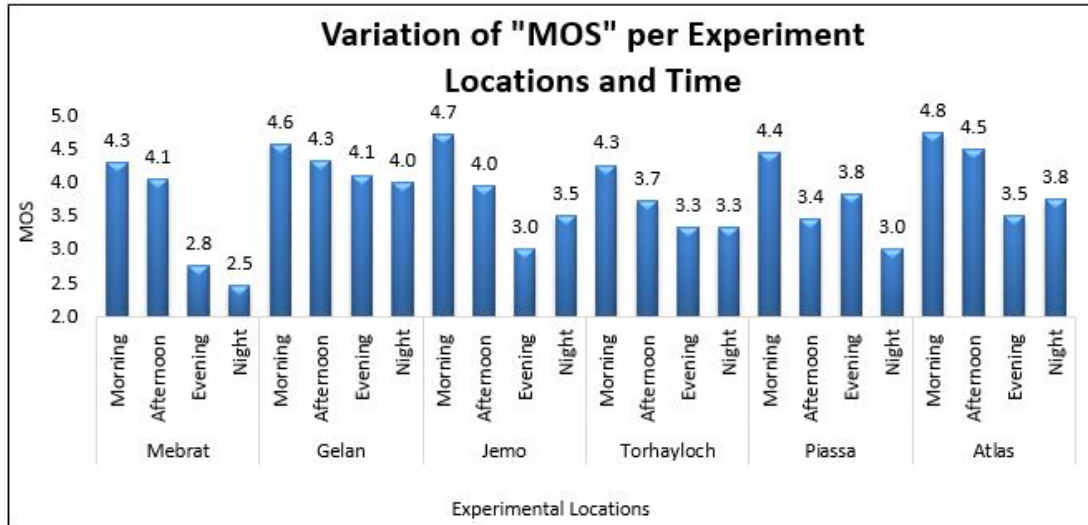


Figure 5.4: Variation of "MOS" per Experiment Locations and Time

Figure 5.5 shows how the application layer KIPs influence the video streaming QoE. Users with average startup delays between 1.1 - 2.0 and 2.1 - 3.0 seconds had the highest MOS values of 4.3 and 3.6 respectively, while users with average startup delays above 3.1 relatively had the lowest MOS of 3.1. The result shows that, participants with longer initial delay length had the lowest MOS value, whereas participants with shorter initial delay period had the highest MOS value. The result also shows that the startup delay had a negative impact on video streaming perceived quality.

Users those who watched the sample videos without stalling duration had the highest average MOS of 4.7, whereas the MOS score dropped to 3.7 for those who have experienced average stalling duration ranged between 0.1-1.8 seconds. Users with a MOS value 3.0 were impacted by the average stalling duration above 1.9 seconds. From this result, we observed that video streaming perceived quality was negatively impacted by an increase in the average duration of stalls.

The lowest MOS values of 3.5 were captured in the range of playback values between 64.9 and 80.5 percent, on the other hand the highest MOS value were achieved when the playback was greater than 90.0 percent. Furthermore, users who experienced playback between 81.0 and 89.9 percent, their QoE have increased significantly with a MOS value of 4.0. The result shows that the playback impact on video streaming perceived quality was positive, in contrast to startup delay and stalling duration.

In addition to technical reasons, QoE can be affected by other influential factors.

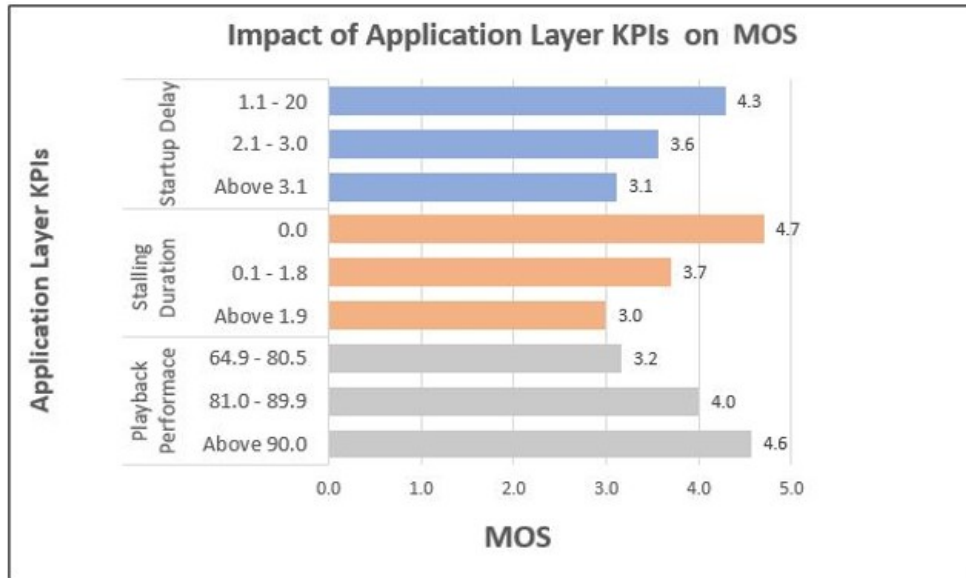


Figure 5.5: Impact of Application KPIs on MOS

As part of this research, age and gender factors have been considered and analyzed for their impact on MOS value.

96 female and 204 male totally 300 volunteers were participated in the survey and as it is depicted in Figure 5.6 females respondents achieved an average of MOS 3.91, whereas males average MOS value was 4.05, it is almost the same results. However, considering that the number of male participants is more than double that of female participants, it is recommended that both genders be proportional to achieve the best results of the study.

On the other hand, participants under the age of 35 achieved a better MOS values than those over 35, with a MOS score of 4.01 and 3.97 respectively. 201 participants were under 35 years of age, while 99 individuals were over the age of 35. Therefore, the age distribution of participants should be balanced.

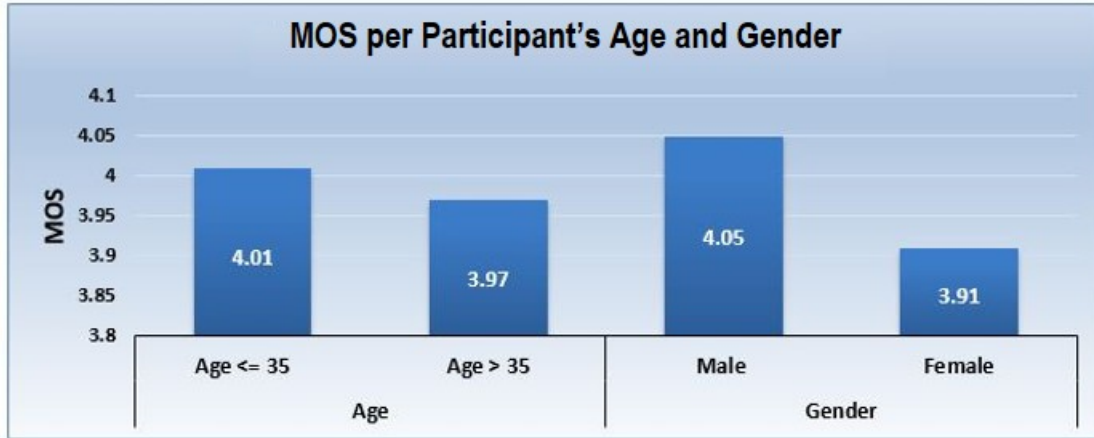


Figure 5.6: MOS per Participant's Age and Gender

## 5.2 Objective Evaluation Results

In order to measure the application, network and physical layer video streaming quality metrics the Nemohandy, Opensignal and Cloudflare speed tests apps were used. Startup delay and stalling duration in seconds and playback performance in percentage were evaluated from the application layer, while download throughput, latency and Jitter, and RSRP and SINR were considered from the network and physical layers respectively. For each of the six research locations, all measurements were taken on the same day and at the same time in order to study their relationships from a user's perspective.

The Nemohandy recorded the measurements per second from the moment we started measuring. For each measurement, it took an average of 8 minutes to complete. The measured data were extracted with an excel file format using a post-processing Actix Analyzer tool. The application and network layer quality metrics, measured by Opensignal and Cloudflare speed tests apps, were manually collected from the customer's phone and analyzed using Microsoft Excel and SPSS tool.

### 5.2.1 Application Layer Evaluation Results

The KPIs included in the application layer are startup delay, stalling duration and playback performance, all the three quality metrics were measured using the Opensignal app taking into account all four time periods and six study locations. About 17% of the users in each of the four locations, (Mebrathail, Gelan, Jemo and Torhailoch) participated in the study, while the remaining 32% participants were from Piassa and Atlas area, which is 16% for each.

In all four-time periods we didn't consider the same numbers of participants, because in the evening and night time periods, it was very challenging to find LTE service users who were interested to participate in the field survey. Because of that, only 29.4% of participants took part in the evening and night time periods, while 70.6% participated in the morning and afternoon time periods. All participants were used their own smartphone and the LTE cellular access network to stream the sample YouTube videos and only a resolution of 854 x 480-pixel (480p) standard definition sample videos with H264-MPEG-4 AVC video codecs used in all study areas.

Table 5.1. depicts Opensignal evaluation results for 300 YouTube mobile streaming sessions. Based on the experiment, the average startup delay was 2.1 seconds, with maximum and minimum startup delays of 6.5 seconds and 0.7 seconds, respectively. On the other hand, there was an average stalling duration of 0.8 seconds, while the maximum stalling duration was 5.1 seconds and the average playback was 85.7 percent.

Table 5.1: Results of Application Layer Quality Metrics.

KPIS	Maximum	Mean	Minimum	Std. Deviation
Startup Delay	6.5	2.1	0.7	1.3
Stalling Duration	5.1	0.8	0	1.3
Playback	95.7	85.7	60	8.8

Although the startup delay of up to 6.5 seconds was recorded, the maximum average startup delay was 2.7 seconds, captured at Jemo area. The minimum average initial delay of 2.2 seconds were registered in both Torhailoch and Piassa areas and on the rest three study locations an average startup delay of 2.4 seconds was captured during the period of evaluation and all participants at each study location experienced an initial startup delay while watching the sample videos.

The next observation was how the video streaming users experience the stalling duration. The lowest average stalling duration was 0.4 seconds at Jemo and Gelan, while the maximum stalling duration of 1.8 seconds was captured at Mebrathail and on the other three study locations, the average stalling duration was 1.2 seconds. Of the 300 participants, 62% did not experience stalling duration.

On the other hand, Atlas and Mebrathail had the highest and lowest playback, with percentages of 86.6 and 80.2 respectively. In the rest of the four study locations, on average, 83.9 percent playback was registered.

Overall, 32.3% of the participants were experienced a playback of less than 85%.Fig-

Figure 5.7 shows the application layer quality metrics measured results per experimental location.

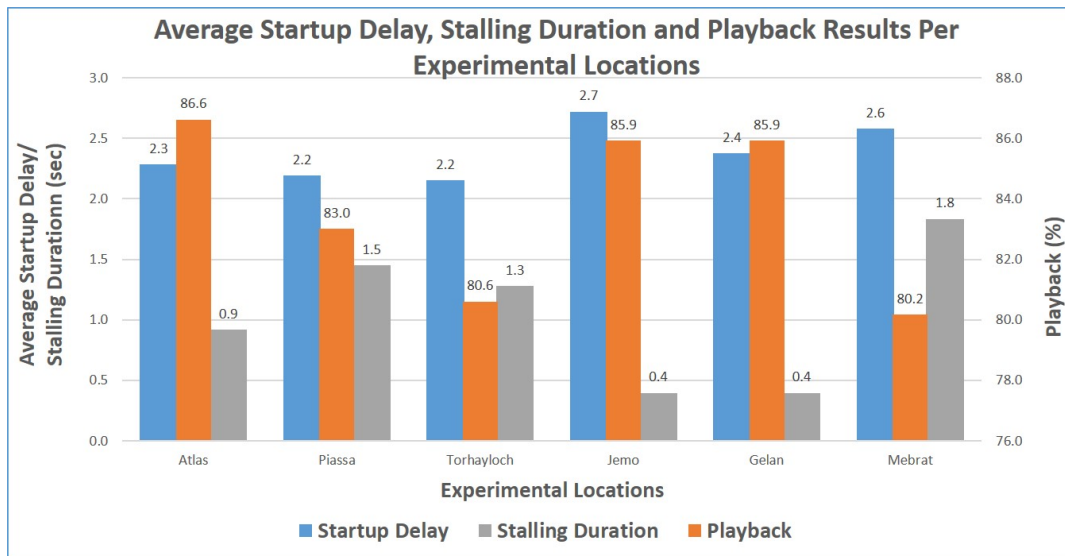


Figure 5.7: Average Startup Delay, Stalling Duration and Playback per Experimental Locations.

In addition, we have seen what the LTE video streaming service looks like in terms of application layer quality metrics over different periods. Figure 5.8 below shows the results of average startup delay, stalling duration and playback over different time periods. The average startup delay and stalling duration for the night session were 3.3 and 1.7 seconds, respectively. Conversely, the minimum startup delay and stalling duration were captured in the morning session and the results were 1.4 and 0.3 seconds. Similarly, the morning and afternoon playback results were 90.8 and 85.8, while the evening and night playback results were 81.2 and 77.1. In general, the LTE video users perceived better experience in the morning and afternoon sessions, while the evening and night time periods quality of service remained bad.

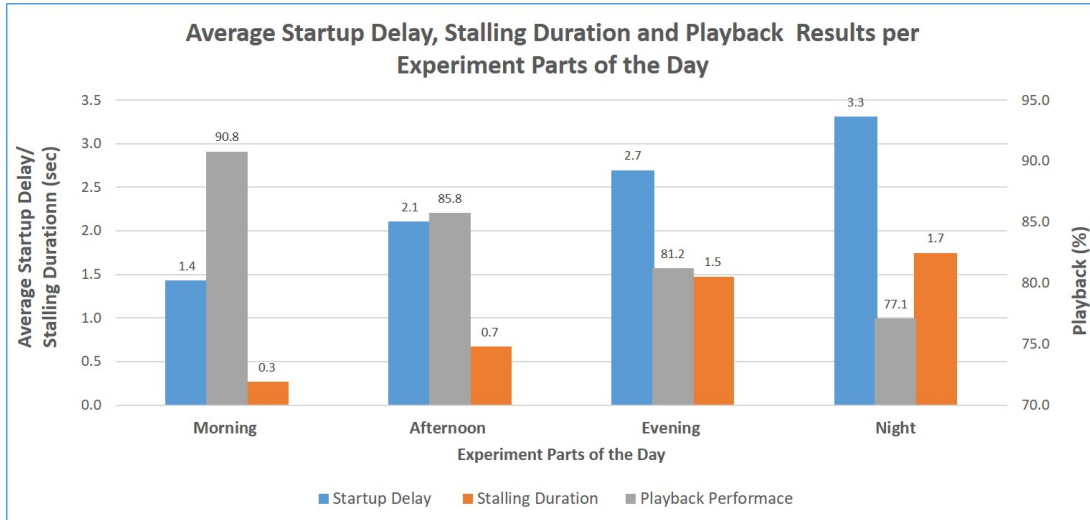


Figure 5.8: Average Startup Delay, Stalling Duration and Playback Results per Experiment Parts of the Day.

Figure 5.9 illustrate the application layer KPIs results per experimental environment context. Therefore, in an outdoor environment, the best average startup delay and stalling duration were 2.0 and 0.7 seconds, while in an indoor context, they were 2.5 and 1.2 seconds. Maximum playback was obtained in an outdoor environment on average with a score of 87.0 percent and in the indoor environment, a value of 81.9 percent was measured. The findings of this study shows that the measurements taken indoor were smaller than those measured outdoor. According to [47], indoor measurements of wireless networks were smaller than outdoor measures due to walls and doors, metallic furniture, and other considerations.

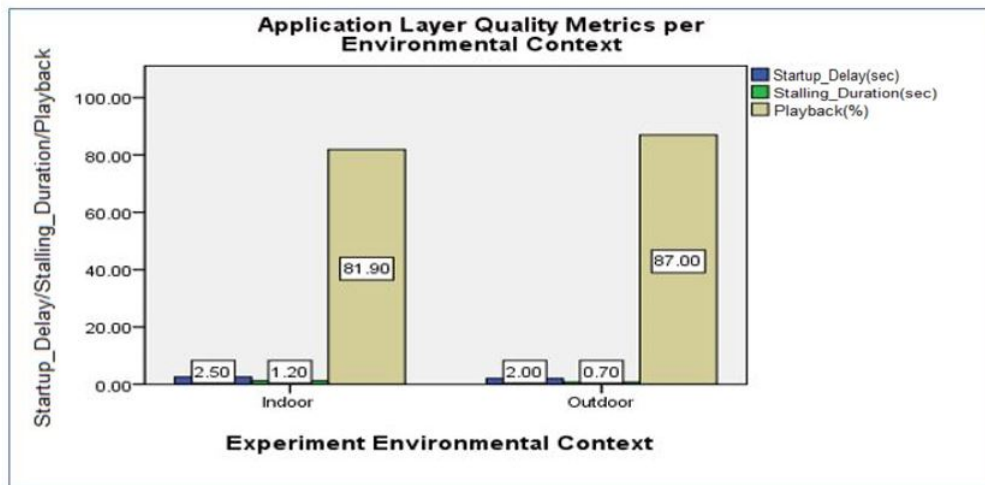


Figure 5.9: Application Layer KPIs per the Experimental Environment Context.

## 5.2.2 Network Layer Evaluation Results

Opensignal, Cloudflare speed test tool and Nemohandy were used to measure the network layer quality metrics (Download throughput, Latency and Jitter) while survey participants watching YouTube video streaming sessions. The measurements were performed in all six study locations, as shown in Figure 5.10 below which was taken from the Huawei KML file, generated by google earth, and the locations were selected based on their geographical distribution and monthly traffic as described in chapter four.

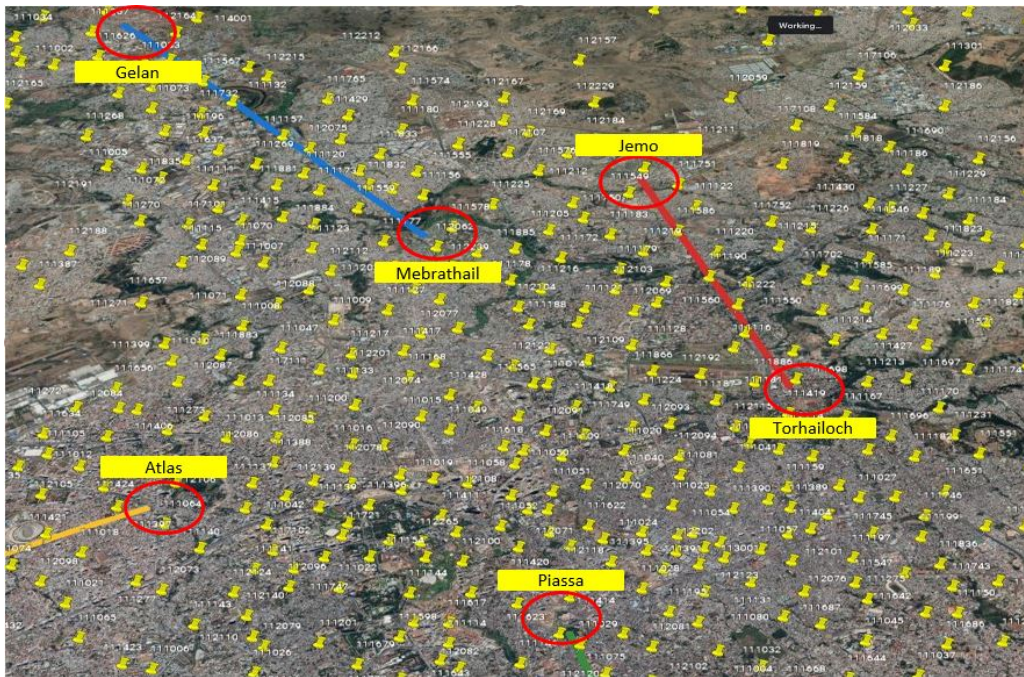


Figure 5.10: Geographical Locations of the Experiments.

Figure 5.11 shows the bar chart that illustrates network layer QoS metrics per experimental locations. The Maximum and Minimum average download throughput of 7.82 Mbps and 5.40 Mbps were recorded at Gelan and Torhailoch area respectively, these are an average results but there were also some locations where the download throughput was less than 1 Mbps, thus this can have substantial effects on the video streaming service. According to a study [23], latency and jitter have a negative impact on the quality of experience of video streaming viewers when the delay variation is greater than 10 milliseconds. As part of this investigation, On average, we measured the lowest and highest latency of 112.8 and 120.9ms at Jemo and Mebrathail. Similarly, the highest and lowest Jitters scores were 67.6 and 29.0ms at Piassa and Jemo; accordingly, their impact on video streaming services will be determined.

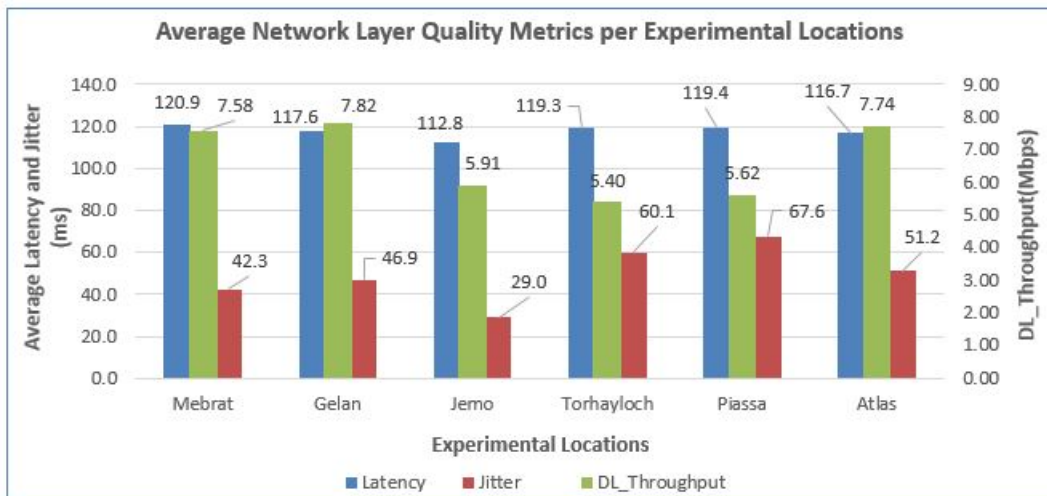


Figure 5.11: Average Network Layer Quality Metrics per Experimental Locations.

Figure 5.12 shows the results of the network layer metrics measured in the four different time periods when participants were watching the sample video streaming sessions. Out of the 4 time periods, 9.5 Mbps and 7.2 Mbps download throughput were achieved in the morning and afternoon measurements and also with these time periods, the latency and jitter were (106.1 and 112.8 ms) and (41.4 and 47.4 ms) respectively. On the other hand, in the night hours of the day, the download throughput was lower with the value of 4.8 Mbps and the latency and jitter increased to 133.0 ms and 48.7 ms. The decreasing of the download throughput in the evening and night hours and at the same time the increase in latency and jitter affect the overall customer satisfaction, and as we have shown above in this study, the MOS score was also decreased in the evening and nighttime periods.

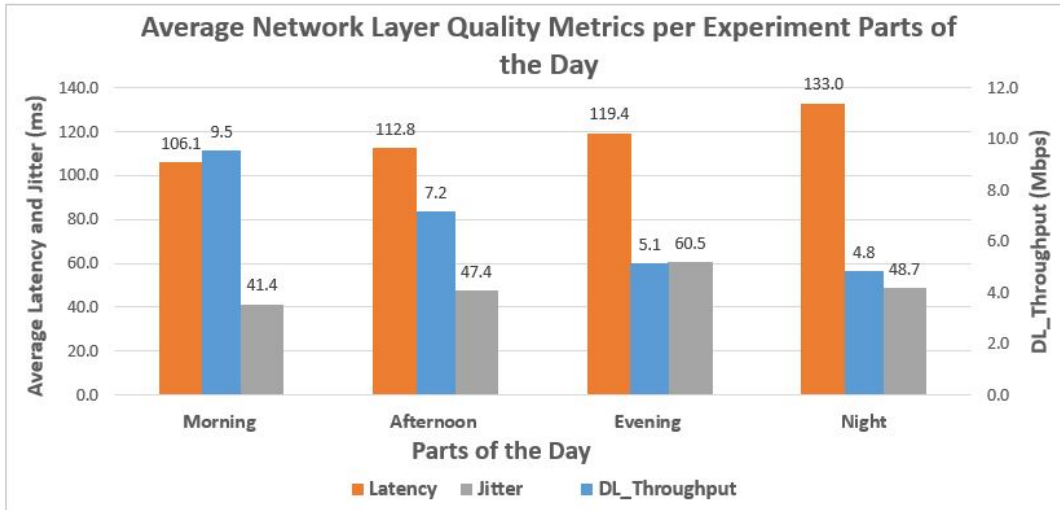


Figure 5.12: Average Network Layer Quality Metrics per Experiment Parts of the Day.

Looking at Figure 5.13, it can be observed that there is a difference between the indoor and outdoor measurements. As previously mentioned, because of the problems developed due to strong multi-path phenomena the MOS values were decreased in an indoor scenario. In this research, the maximum download throughput was achieved in the outdoor environment which is 8.04 Mbps, whereas the latency and jitter were 110.40 ms and 47.20 ms with in this environment.

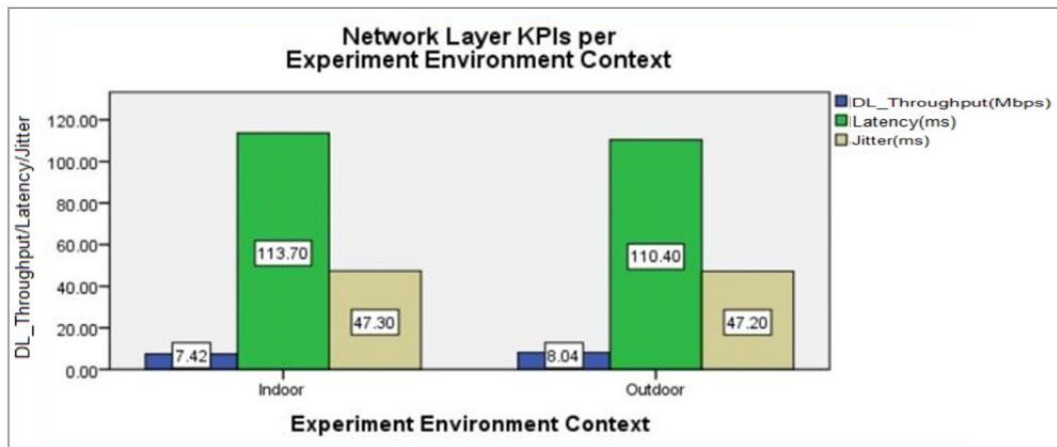


Figure 5.13: Average Network Layer Quality Metrics per Experiment Environmental Context.

Figure 5.14 shows the Pia chart of the actual download throughput measured by Nemohandy under different environmental conditions. Based on our data analysis,

47% of users experienced a download throughput in between 2.08 Mbps and 5.98 Mbps, 31% of users were experienced a download throughput between 6.04 Mbps and 9.94 Mbps and the rest 22% of the users experienced between 10.08 Mbps and 14.6 Mbps. Nearly half of the participants experienced a download throughput of less than 5.98 Mbps, and some of them had even lower results, which are insufficient for watching high-quality video streaming in particular, and this has a direct impact on the customers' satisfaction with the service.

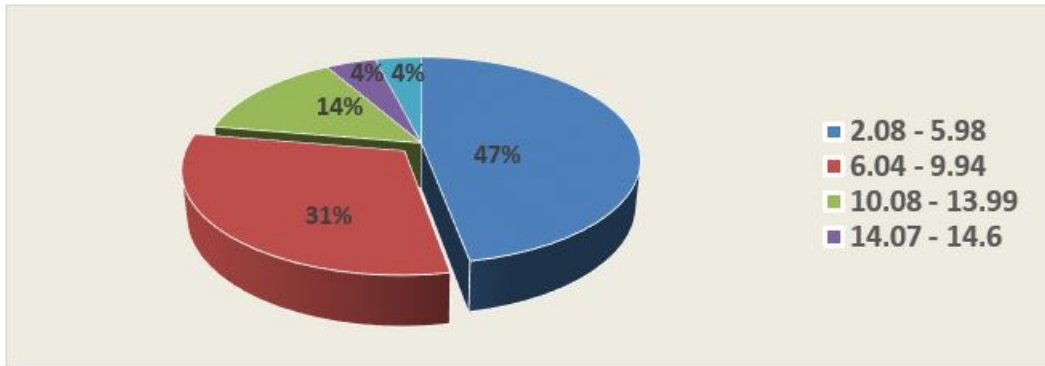


Figure 5.14: Evaluation Results of Average Download Throughput in Mbps.

According to [48], the latency requirement should not exceed 100 milliseconds for SD videos and 50 milliseconds for HD videos. As we have seen in Figure 5.15, in this study 39% of participants were experienced a latency between 117 ms and 142 ms, which was the largest percentage and a latency between 91ms and 116ms, was experienced by 38 percent of participant, which was the second highest percentage. Generally, 70% of participants experienced more than 100 ms of latency therefore, since this result did not satisfy the requirements of latency and jitter, we were unable to watch videos at the expected quality level.

Similarly, the jitter requirement for SD and HD videos were described in study [47], the jitter requirement for SD videos should not exceed 50 milliseconds, whereas the jitter requirement for HD videos should not exceed 30 milliseconds. Figure 5.16 shows that, 41% participants had experienced the jitter between 9 and 34 ms. The rest 59% of the LTE video streaming users had experienced jitter between 35 ms and 108 ms. Most customers experienced jitter that was greater than 35 milliseconds, which suggests that many people did not obtain the service they wanted because of the jitter.

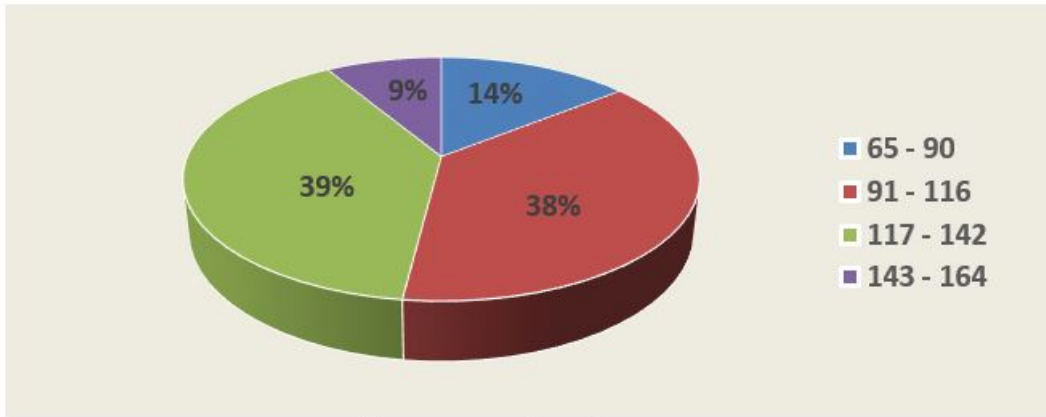


Figure 5.15: Evaluation Results of Average Latency in ms.

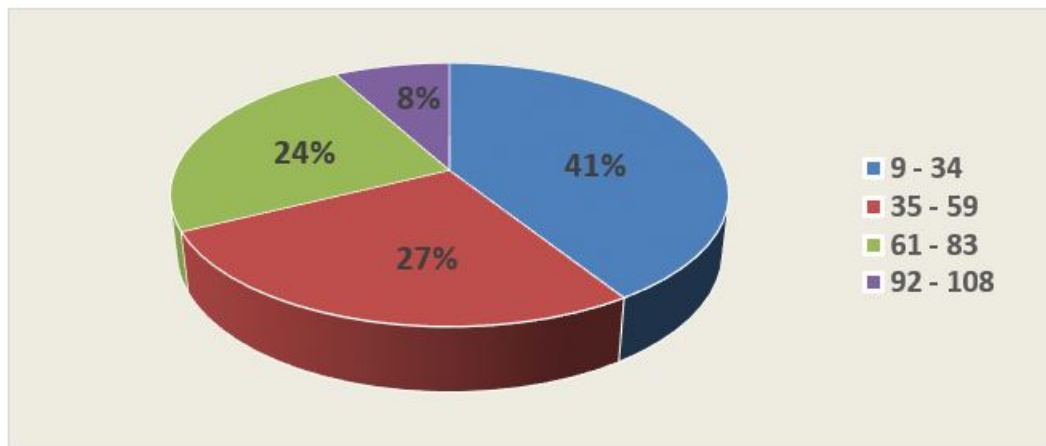


Figure 5.16: Evaluation Results of Average Jitter in ms.

### 5.2.3 Physical Layer Evaluation Results

Figure 5.17 below shows the results of the RSRP and SINR measured values for each of the study locations. At Gelan, the maximum average RSRP and SINR were -83.64dBm and 13.21dB, respectively. On the other hand, the minimum average RSRP and SINR values were -92.04dBm and 8.66dB at Piassa, respectively. As we have seen above, the results obtained here have also been implicated in the MOS results obtained in the same locations.

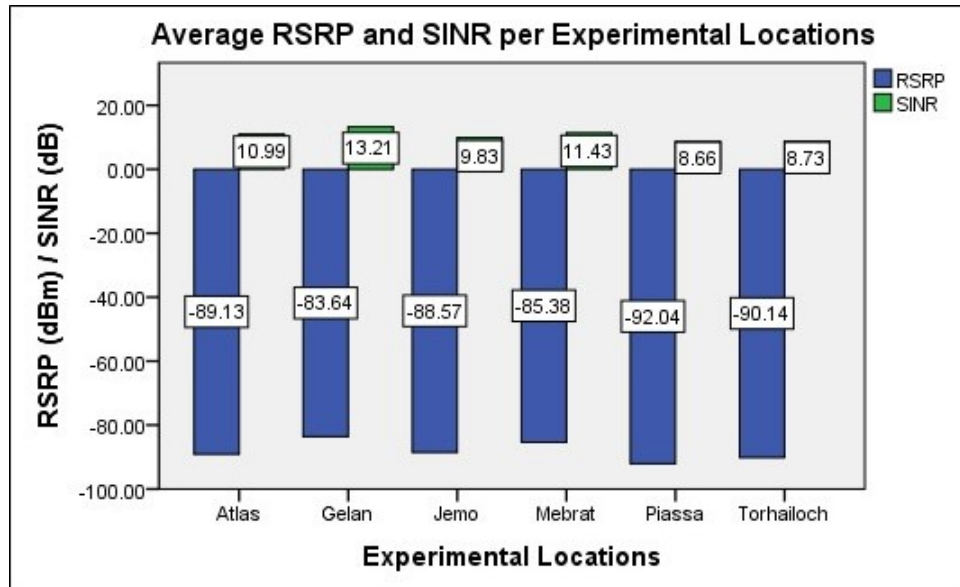


Figure 5.17: Average RSRP/SINR achieved per Experimental Locations.

As illustrated in Figure 5.18 The best average RSRP and SINR were found to be -87.59dBm and 10.94dB in an outdoor situation, respectively and for the indoor environment the average RSRP and SINR were -89.57dBm and 9.20dB.

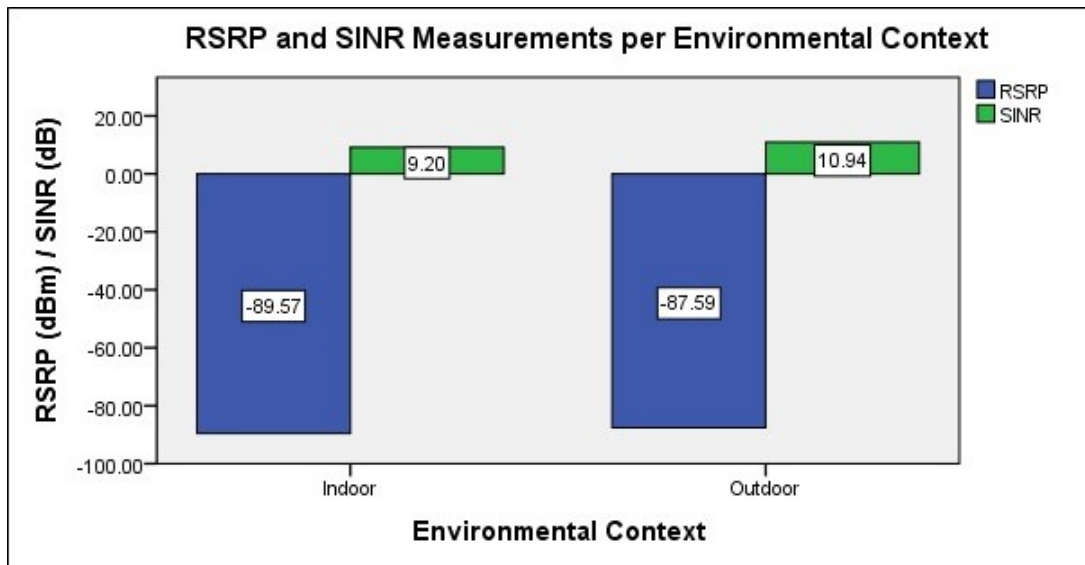


Figure 5.18: Average RSRP/SINR Achieved per Environmental Context.

Figure 5.19 depicts the average RSRP and SINR results achieved per parts of the day. Participants who watched the videos had the best RSRP and SINR measurement results in the morning and afternoon, with average results of (-85.71 and -88.17dBm) and ( 12.09 and 10.03dB), respectively. Participants who watched sam-

ple videos in the evening and at nighttime periods had relatively poor RSRP and SINR measurement results of (-89.88 and -92.69dBm) and (9.80 and 7.97dB), respectively. So that users who access mobile data can easily linked to a nearest cell tower, which means the connection (RSRP) you get will be better. However, the average RSRP and SINR is reduced in the evening and nighttime periods. This result shows that video streaming users are not getting good service in the evening and night.

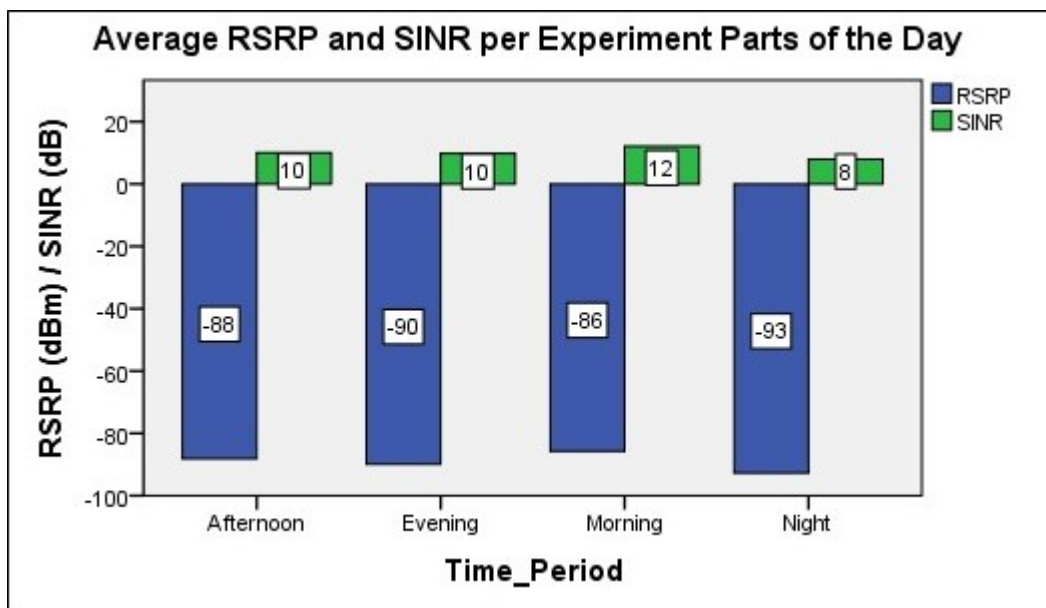


Figure 5.19: Average RSRP/SINR Achieved per Parts of the Day.

As shown in Table 5.2 below, 78% of the RSRP values were between -94.95 to -80.13, whereas 12 percent and 11 percent, respectively, fell between -102.85 and -95.10 and -79.88 and -71.78. In [49], it is revealed that for RSRP values less than or equal to -100 dBm, the performance drop rapidly. In our study, even though the average measured RSRP values were larger than -102.9dBm, in some study locations we captured less results than -102.9dBm values. For SINR values, around 33% of users captured less than 13dB. The recommendation in [49] reveals that, with these values, reliable data speeds may be attained, but marginal data with dropouts is possible. When this value gets close to 0, performance will drop drastically.

Table 5.2: Results of Physical Layer Quality Metrics.

RSRP	Percentage (%)	SINR	Percentage (%)
-102.85 to -95.10	12%	-5 to 7.8	33%
-94.95 to -80.13	78%	8 to 13.4	43%
-79.88 to -71.78	11%	14.1 to 20.5	23%

## 5.3 Correlation Analysis Results

### 5.3.1 Correlation among MOS and Application, Network and Physical Layer Quality Metrics.

We collected the users' perception using a survey Questionnaire, and at the same time, the application, network and physical layers quality metrics were collected to conduct a correlation study among them. Accordingly, on the basis of the correlation coefficient and P-value results, we identified which quality metrics are more influencing the MOS. This study considered the correlation between MOS and each of the layer metrics included in this study, to identify how much each quality parameter affects the MOS.

Tables 5.3 shows the results of spearman correlation among MOS and Application, Network and Physical Layer Quality Metrics. When we look at the correlation results among MOS and the application layer quality metrics, on average, the spearman correlation of MOS and Playback was found to be a positive strong correlation and there is enough evidence to show the correlation is statistically significant with the correlation coefficient and level of significance values of ( $r = .746$ ,  $p < 0.001$ ). On the contrary the correlation of MOS with startup delay and stalling duration was found to be a negative moderate correlation with a value of correlation coefficient ( $r = -.667$  and  $-.593$ ) respectively and the significance values for both correlation results were  $P < 0.001$ .

Furthermore, this research examined the relationship among MOS and Network Layer Quality Metrics. On average the Download Throughput and Latency have a moderate correlation with MOS and their correlation coefficient magnitudes were 0.684 and -0.419 respectively with  $p < .001$  significance value for both. The negative sign (-) of the correlation coefficient value indicates the converse relationship between MOS and Latency. The next one the correlation result between jitter and MOS, their relationship was negative, weaker in strength and statistically significant ( $r = -.335$ ,  $p < .001$ ).

Table 5.3: Correlation results among MOS and Application, Network and Physical Layer Quality Metrics.

	MOS						
	Mebrathail	Gelan	Jemo	Torhailoch	Piassa	Atlas	Average
Startup Delay	-.818	-.916	-.618	-.377	-.582	-.694	-.667
Stalling Duration	-.776	-.786	-.588	-.379	-.603	-.428	-.593
Playback	.810	.902	.654	.730	.664	.716	.746
DL Throughput	.763	.763	.611	.626	.624	.718	.684
Latency	-.329	-.735	-.381	-.413	-.338	-.319	-.419
Jitter	-.394	-.570	-.415	-.162	-.066	-.402	-.335
RSRP	.433	.702	.597	.371	.600	.674	.563
SINR	.353	.506	.561	.457	.571	.637	.514

In [50] the author investigated how physical layer quality metric (RSRP) affect the QoE in video streaming for different video content types over LTE networks by analyzing the relationship between RSRP and QoE and they find out, for RSRP values less than -104 dBm, the MOS values dropped rapidly. Similarly in this study, regarding the physical layer metrics and MOS relationship, on average the highest degree of correlation achieved between RSRP and MOS, they have a strong positive correlation with values of ( $r = .563$ ,  $p < .001$ ). Similarly, MOS and SINR have a positive relationship, moderate in strength and statistically significant ( $r = .514$ ,  $p < .001$ ). For the RSRP value between -100 dBm and -90 dBm the average MOS value was around 3.6. It was also shown that MOS values declined rapidly when the RSRP was less than -100dBm.

With respect to study locations, the highest degree of correlation was achieved at Mebrathail between start up delay and MOS, with correlation coefficient magnitude of ( $r = -.916$ ,  $p < .001$ ) and apart from Mebrathail and Gelan, for the remaining four study locations we observed that the next highest degree of association between playback and MOS.

In accordance with the results of this study, we identified that the playback quality metric to be the main factor responsible for the MOS variation, that is, Playback is the most influential factor on QoE of video streaming users. Figure 5.20 shows the scatter plot representing the correlation between playback and MOS, that indicates a strong positive association between playback and MOS scores, that is, when the playback increases, the level of the video streaming user's QoE increases

or vice versa. It should be emphasized that playback refers to the percentage of the test time that the video is playing without loading and buffering therefore, it is the results of the combined effect of startup delay and stalling duration. In comparison, the startup delay has a greater impact on the MOS than the stalling duration. After the playback, download throughput is the second most influential factor on MOS.

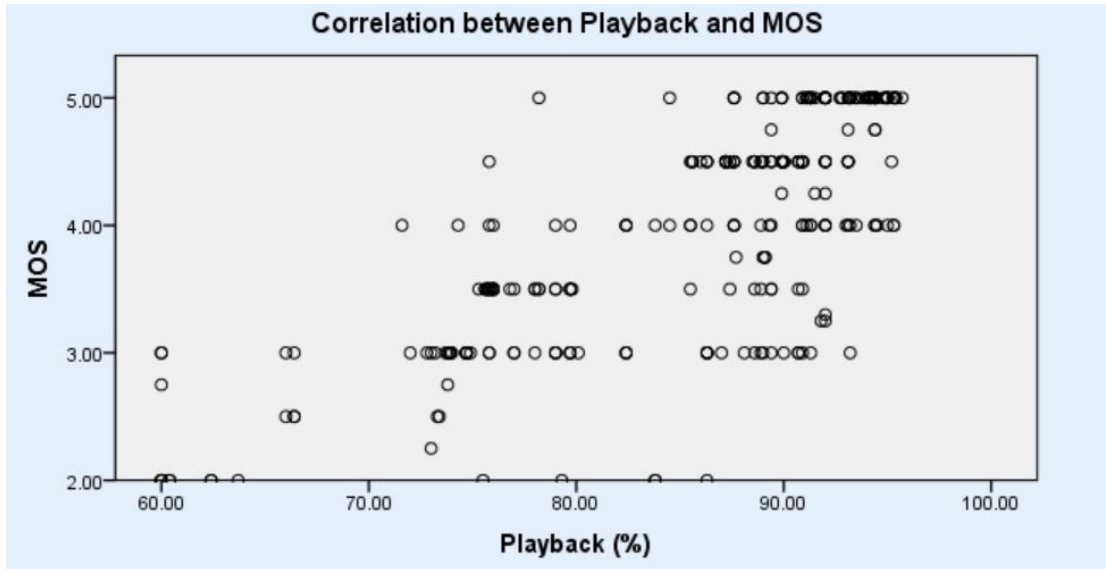


Figure 5.20: Scatter Plot Representing the Correlation between Playback and MOS.

We have also performed a regression analysis to acquire the mathematical relationship between MOS and playback quality metric. As shown in Equation 5.1, the coefficient of playback is positive, thus a higher level of Playback giving a higher MOS. 0.077 is the marginal effect of one unit of playback on MOS.

$$Y = \beta_0 + \beta_1 X + \epsilon \quad (5.1)$$

$$\begin{aligned} \text{MOS} &= \beta_0 + \beta_1 \text{Playback} + \epsilon \\ \text{MOS} &= -2.55 + .077 \text{Playback} + \epsilon \end{aligned}$$

Where, X and Y are independent and dependent variables respectively.  $\beta_0$  is the constant,  $\beta_1$  is X's coefficient and  $\epsilon$  is the error term.

### 5.3.2 Correlation between Application and Network Layer Quality Metrics.

Users QoE could be influenced by different objective (network quality, device capacity) and subjective parameters (expectations, emotion and location) [51]. In terms of video streaming services, there are also a number of problems when the video packets travel through the wireless network, from the lower layer all the way up to the application layer. In this study, based on the results of a correlational analysis, we determined which network metrics have the most influence on application layer performance. The correlation result depicted in Table 5.4 shows, download throughput has a higher degree of correlation with the playback quality metric with a values of ( $r = .711$ ,  $p < .001$ ), which is strong positive correlation. The second highest degree of correlation achieved in between the download throughput and stalling duration ( $r = -.647$ ,  $p < .001$ ), their relationship was negative, moderate in strength and statistically significant.

Table 5.4: Correlation results among Application and Network Layer Quality Metrics.

	Startup Delay	Stalling Duration	Playback	DL-Throughput	Latency	Jitter
Startup Delay	1.00	.681	-.719	-.576	.444	.360
Stalling Duration	.681	1.00	-.715	-.647	.374	.424
Playback	-.719	-.715	1.00	.711	-.540	-.427
DL-Throughput	-.576	-.647	.711	1.00	-.499	-.411
Latency	.444	.374	-.540	-.499	1.00	.419
Jitter	.360	.424	-.427	-.411	.419	1.00

According to the correlation analysis results, it is observed that the download throughput plays a major role for the variation of application layer quality metrics. In other word, if there is change in download throughput then there is corresponding significance change in the application quality metrics because they have the highest degree of association between them. In specific, Figure 5.21 shows the scatter plot representation of the highest degree of correlation result, which is the correlation between download throughput and playback quality metrics.

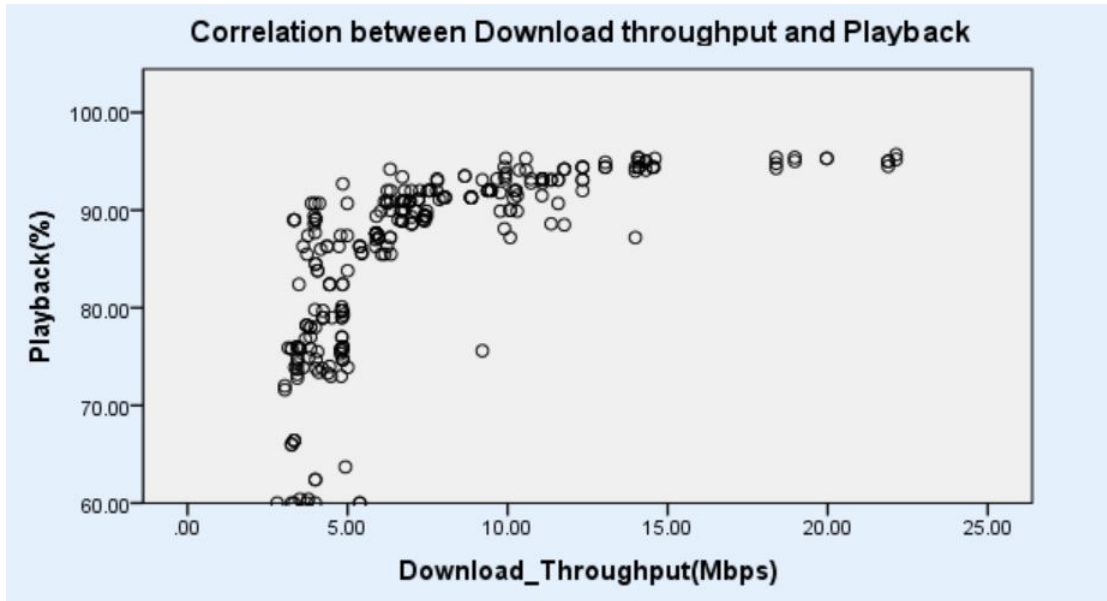


Figure 5.21: Scatter Plot Representing the Correlation between Download throughput and Playback.

### 5.3.3 Correlation between Network and Physical Layer Quality Metrics.

By measuring the LTE physical layer quality metrics RSRP and SINR, article [52] investigated their relationship with a download throughput for different network loads. Accordingly, the study shows that the download throughput results within the RSRP ranges of -130 to -125 dBm were 5286.42 Kbps and also 101295.2 Kbps. Similarly, as depicted in Table 5.5, we studied the correlation among physical layer quality metrics and network layer performance metrics in order to determine which quality metric is the more influential on network layer quality metrics. As a result, RSRP and SINR had positive strong correlation with download throughput having a correlation coefficient values of ( $r = .778$  and  $r = .773$ ) respectively, while they

have negative moderate correlation with latency and jitter ( $r = -.421$  and  $r = -.407$ ), the significance level for both conditions was  $p < 0.001$ ).

Table 5.5: Correlation results among Network and Physical Layer Quality Metrics.

	DL Throughput	Latency	Jitter	RSRP	SINR
Download Throughput	1.00	-0.499	-0.411	0.778	0.773
Latency	-0.499	1.00	0.419	-0.421	-0.407
Jitter	-0.411	0.419	1.00	-0.487	-0.432
RSRP	0.778	-0.421	-0.487	1.00	0.798
SINR	0.773	-0.407	-0.432	0.798	1.00

Looking at the correlation results obtained separately from all study locations, except Mebrathail and Jemo on the other study areas the download throughput has a higher degree of correlation with RSRP and SINR. In this case, at Mebrathail and Jemo, the download throughput mostly affected by other wireless network impairments, mainly the increase in the number of users and the like, but the influence of RSRP and SINR is not negligible. For the rest four locations, RSRP was the most influential factor. The scatter plot in Figure 5.22 below shows that RSRP and SINR have a positive relationship with Download throughput, thus when RSRP and SINR increase, the Download throughput increases as well.

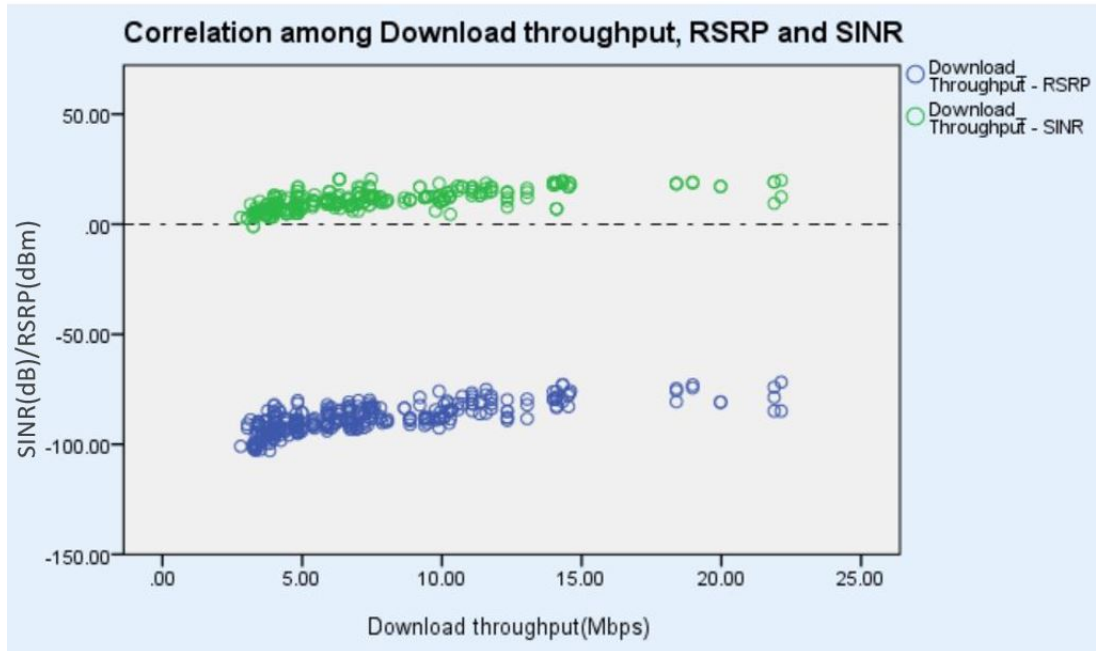


Figure 5.22: Scatter Plot Representing the Correlation among Download Throughput with RSRP and SINR.

SINR is one of the impairments that might lower the efficiency of an LTE network's download throughput, and we determined that it has a significant relationship with download throughput in this study. On the other hand, their mathematical relationship can be represented using the below equation, and the download throughput efficiency was calculated as Shannon channel capacity formula[53, 54], In the case of the single-input single-output (SISO) a Shannon capacity formula used:

$$S_{max}(bits/s/Hz) = \log_2(1 + SNR) \quad (5.2)$$

The Shannon capacity for general MIMO with perfect transmitted knowledge is:

$$S_{max}(bits/s/Hz) = \sum_{k=1}^{\min(n_T, n_R)} \log_2(1 + SNR_k) \quad (5.3)$$

Where  $n_T$  and  $n_R$  are the number of transmit and receive antennas respectively and  $SNR_k$  denotes the resulting SNR of the  $k$ th spatial sub-channel which is in-

fluenced by the eigenvalue, the noise/interference, as well as the allocated transmit power on that sub-channel. However, for LTE MIMO antenna system and to address different implementation issues a modified Shannon capacity formula were used:

$$S_{max}(bits/s/Hz) = BW\_eff * \eta * \log_2(1 + SNR/SNR\_eff) \quad (5.4)$$

Here BW\_eff adjusts for the system bandwidth (BW) efficiency of LTE and SNR\_eff adjusts for the SNR implementation efficiency of LTE. The factor  $\eta$  is a correction factor which nominally should be equal to one [53].

# Chapter 6

## Conclusion and Future Work

### 6.1 Conclusion

Even though the video streaming service is influenced by human, system, context and content related influencing factors, studying the correlation among QoE and QoS different factors is the main purpose of this thesis. Additionally, we investigate the spatio-temporal network measurement variations in terms of video streaming service. In order to see how one layer influences the other layers, we divided the correlation study into three sections: The correlation of MOS with the three layers of quality metrics (application, network, and physical layer), application and network layer quality metrics correlation, and network and physical layers quality metrics correlation.

The study was conducted in six geographically distributed locations in Addis Ababa, under four zones. Totally eight basic quality metrics have been identified: startup delay, stalling duration and playback performance as well as latency, jitter and download throughput from application and network layers respectively and RSRP and SINR from physical layer. Along with quality metrics, subjective data have been collected that show us how users perceive the quality of service. Data has been collected using Nemohandy, Opensignal and Cloudflare speed test together with crowdsourced approach. For subjective quality metrics three hundred LTE subscribers of ethio telecom in Addis Ababa were asked to complete a questionnaire after watching the sample video.

In this study, we determined that there is a MOS problem because its value was found to be 3.9, therefore to determine the cause of this problem; we studied the correlation among MOS and other QoS parameters. The results of the quality metrics evaluation shows that, the average, startup delay and stalling duration are

2.1 and 0.8 seconds, and the average playback is 85.7%. On the network side, the average download throughput, latency, and jitter results are 7.42Mbps, 113.7ms, and 47.29ms, respectively, and the average RSRP and SINR are -88.10dBm and 10.49dB from physical layer.

In order to increase customer satisfaction, HD video streaming services are recommended, but to watch HD and greater video quality, a latency of less than 50ms and jitter values of less than 30ms are required. However, the findings obtained here are higher than the recommended values. As a result, this consequence may decrease the satisfaction level of video streaming users and the MOS result reflected this trend as well.

In the first subsection, a higher degree of correlation achieved between MOS and playback with a strongly positive and statistically significant values of ( $r = .746$ ,  $p < 0.001$ ). Here our main finding is that MOS is mainly lowered by playback, implying that, playback is the main factor responsible for the MOS variation. In the second subsection, it is observed that there is a higher correlation between download throughput and playback ( $r = .711$ ,  $p < 0.001$ ). As a result, download throughput is the primary cause of variation in application layer quality metrics. The RSRP and SINR results for Mebrathail and Jemo were good, but the network quality measurements were poor, and furthermore, network-layer quality metrics have a low correlation with RSRP and SINR. This implies that while RSRP and SINR are good, network layer metrics may not be good due to other factors like increase in the number of users and etc. and spatial structure is an important factor affecting QoS as well as QoE.

In the morning time period, we were able to get the highest MOS values and the MOS values in the afternoon were higher than those in the evening and nighttime measurements. Because users have a tendency to connect to the network during downtime, which has a detrimental influence on end-user QoE. The study also shows that at Mebrathail and Jemo in particular during the evening and nighttime periods, it is not suitable for video streaming services, possibly due to congestion or higher distance from the serving cell. This implies that the QoS varies at different locations and times of the day on the cellular network.

This study helps ET in understanding the current network quality challenges in various locations and times in terms of video streaming services and currently, when video streaming related problems occur in ET quality and optimization teams, they take immediate measurements and corrective actions, but there is a gap in investigating the root cause of the problems. Therefore, from this thesis work,

Ethio telecom can understand the gap of recent root cause analysis approaches. This approach is a customer centric approach; it allows to analyze the correlation among QoS parameters and QoE from different perspectives. Moreover, this approach shows how spatial and temporal aspects are important factors influencing the QoE and it is easily adaptable to other application environments. Eventually, it helps to improve customer satisfaction and, in the future competitive environment; it helps to mitigate customer migration concerns.

## 6.2 Future Work

The following areas are recommended for future work.

- This research was limited to Addis Abeba, the capital of Ethiopia. Therefore, in order to have a complete picture, future works may include other cities where the service is available.
- As the video streaming service is available at different quality levels, future research could benefit from including multiple resolution videos, such as 720P, 1080p, etc.
- A variety of technologies, such as tablets and comparable video-enabled devices, as well as smart TVs, could be used in future studies.
- This thesis work is limited to YouTube-based video streaming services over LTE; however, future work may incorporate other Internet services and technologies, such as LTE Advanced.
- Consider the joint effect of quality metrics on QoE.

# Reference

- [1] F. M. Adam. “Information and Communication Technology (ICT), Social Changes and Transformation,” in: *Ugandan Technology and Management University Kampala Uganda*, (2016).
- [2] Cisco. *Cisco Annual Internet Report (2018–2023)*. URL: <https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html>. [Accessed on April 10, 2021.]
- [3] A. R. Bisrat. “Quality of Experience Evaluation for Addis Ababa UMTS Enterprise Data Customers”. In: *Addis Ababa University, Addis Ababa Institute of Technology School of Electrical and Computer Engineering* (October, 2018).
- [4] Ethio Telecom. *Statistics*. URL: <http://www.ethiotelecom.et..> [Accessed on April 12, 2021.]
- [5] B. B. Haile, B. M. Zerihun, and D. A. Bulti. “On the relevance of capacity enhancing 5G technologies for Ethiopia”. In: 3 (2017).
- [6] P. Uppu. “QoE of video streaming over LTE network”. In: *School of Computing Blekinge Institute of Technology, Karlskrona, Sweden*, (2013).
- [7] 3gpp. *The Mobile Broadband Standard*, URL: <http://www.3gpp.org/LTE>. [Accessed on April 10, 2021.]
- [8] B. Khadija, S. Essaid, and S. Mohamed.. “Quality of experience for streaming services: measurements, challenges and insights”. In: *IEEE Access* vol. 8 (2020).
- [9] ITU. “Terms and Definitions Related to Quality of Service and Network Performance Including Dependability”. In: *ITU-T Recommendation E.800* (1994).

- [10] S. Michael, W. Sarah, and C. Pedro. “Considering user behavior in the quality of experience cycle: Towards proactive QoE-aware traffic management”. In: *IEEE Communications Letters*, vol. 23, .no. 7, (1994), pp. 1145–1148. (August, 1994).
- [11] S.Rene, C. Eduardo, and C. Marilia.. “An overview of quality of experience measurement challenges for video applications in IP networks”. In: , (2010), pp. 252–263. (June, 2010).
- [12] K. Brunnstrom, S. A. Beker, and M. K. De... “Qualinet white paper on definitions of quality of experience,” in: (2013).
- [13] J. Parikshit, T. Venkatesh, and M. Deep. “Measurement of Quality of Experience of Video-on-Demand Services: A Survey”. In: *IEEE Communications Surveys Tutorials* 18.1 (2016), pp. 401–418. DOI: 10.1109/COMST.2015.2401424.
- [14] D. S. Tibebu. “Measurement of Quality of Experience of Video-on-Demand Services: A SurveyQoE Model for Addis Ababa LTE Web Browsing Service Using Neural Network Approach”. In: *Addis Ababa University, Addis Ababa Institute of Technology School of Electrical and Computer Engineering* (December 2019).
- [15] D.Weldu. “Machine Learning Based QoE Estimation Model for Video Streaming over UMTS Network”. In: *Addis Ababa University, Addis Ababa Institute of Technology School of Electrical and Computer Engineering* (February, 2020).
- [16] Sandvine. “The Global Internet Phenomena Report”. In: (October 2018).
- [17] Cisco Visual Networking Index. “Global Mobile Data Traffic Forecast Update, 2017–2022”. In: (February, 2019).
- [18] M.Sajid Mushtaq and A. Mellouk. “Methodologies for Subjective Video Streaming QoE Assessment”. In: Jan. 2017, pp. 27–57. ISBN: 9781785481093. DOI: 10.1016/B978-1-78548-109-3.50002-3.
- [19] W. Dapeng et al. “Streaming video over the Internet: approaches and directions”. In: *IEEE Transactions on Circuits and Systems for Video Technology* 11.3 (2001), pp. 282–300. DOI: 10.1109/76.911156.
- [20] K. P. Ricky et al. “Measuring the quality of experience of HTTP video streaming”. In: *12th IFIP/IEEE International Symposium on Integrated Network Management (IM 2011) and Workshops*. 2011, pp. 485–492. DOI: 10.1109/INM.2011.5990550.

- [21] V. D. B. Swetha Kilari. “Effect of Delay/ Delay Variation on QoE in Video Streaming”. In: *School of Computing Blekinge Institute of Technology, Karlskrona Sweden* (May, 2010).
- [22] P. Orosz et al. “A case study on correlating video QoS and QoE”. In: *2014 IEEE Network Operations and Management Symposium (NOMS)*. 2014, pp. 1–5. DOI: 10.1109/NOMS.2014.6838399.
- [23] Z. F. G/Mariam. “Quality of Experience Evaluation for Smartphone Video Streaming over 3G Network: In the case of Addis Ababa, Ethiopia”. In: *Addis Ababa University, Addis Ababa Institute of Technology School of Electrical and Computer Engineering* (December 2019).
- [24] V. Eerola. “LTE Network Architecture Evolution”. In: *Helsinki University of Technology* ().
- [25] D. Shah. “A Tutorial On LTE Evolved UTRAN (EUTRAN) and LTE Self Organizing Networks (SON)”. In: 2011.
- [26] V. Srinivasa Rao. *Protocol Signaling Procedures in LTE*. September 2011.
- [27] A. Navita. “A Survey on Quality of Service in LTE Networks”. In: *International Journal of Science and Research (IJSR)* 4.5 (May 2015), pp. 370–375. URL: [https://www.ijsr.net/search\\_index\\_results\\_paperid.php?id=SUB154155](https://www.ijsr.net/search_index_results_paperid.php?id=SUB154155),.
- [28] Alcatel-Lucent. “The LTE Network Architecture A comprehensive tutorial”. In: *STRATEGIC WHITE PAPER*) (May 2015). URL: <https://www.nokia.com/networks/portfolio/>,.
- [29] M. Vaser and S. Forconi. “QoS KPI and QoE KQI Relationship for LTE Video Streaming and VoLTE Services”. In: *2015 9th International Conference on Next Generation Mobile Applications, Services and Technologies*. 2015, pp. 318–323. DOI: 10.1109/NGMAST.2015.34.
- [30] Opensignal. *Understanding mobile network experience: What do Opensignal’s metrics mean?*, May 26, 2021. URL: <https://www.opensignal.com/>. [Accessed on June 10, 2021.]
- [31] Opensignal. *Opensignal Internet Speed Test*, URL: <https://apps.apple.com/us/app/opensignal-internet-speed-test/id598298030>. [Accessed on June 10, 2021.]

- [32] B. Alemayehu. “Evaluation of User Side Quality of Service Measurement Techniques for Addis Ababa LTE Data Service”. In: *Addis Ababa University, Addis Ababa Institute of Technology School of Electrical and Computer Engineering* (July, 2018).
- [33] KEYSIGHT TECHNOLOGIES. *Nemohandy Technical Overview*, January 14, 2020. URL: [www.keysight.com](http://www.keysight.com). [Accessed on April 10, 2021.]
- [34] F. Afroz et al. “SINR, RSRP, RSSI and RSRQ Measurements in Long Term Evolution Networks”. In: *International Journal of Wireless Mobile Networks* 7 (Aug. 2015), pp. 113–123. DOI: 10.5121/ijwmn.2015.7409.
- [35] Cloudflare. *Speed Test*, URL: <https://speed.cloudflare.com/about/>. [Accessed on July 15, 2021.]
- [36] D. Arkkelin. “Using SPSS to Understand Research and Data Analysis”. In: *Valparaiso University, sychology Curricular Materials. 1.* (2014).
- [37] Kent State University. *SPSS TUTORIALS: PEARSON CORRELATION*, URL: <https://libguides.library.kent.edu/SPSS/PearsonCorr>. [Accessed on May 04, 2021.]
- [38] S. Senthilnathan. “Usefulness of Correlation Analysis”. In: *SSRN Electronic Journal* (July 2019). DOI: 10.2139/ssrn.3416918.
- [39] M. Mukaka. “Statistics Corner: A guide to appropriate use of Correlation coefficient in medical research”. In: *Malawi medical journal : the journal of Medical Association of Malawi* 24 (Sept. 2012), pp. 69–71.
- [40] P. Schober, C. Boer, and L. Schwarte. “Correlation Coefficients: Appropriate Use and Interpretation”. In: *Anesthesia& Analgesia* 126 (2018), pp. 1763–1768.
- [41] J. Hauke and T. Kossowski. “Comparison of Values of Pearson’s and Spearman’s Correlation Coefficients on the Same Sets of Data”. In: *Quaestiones Geographicae* 30 (June 2011), pp. 87–93. DOI: 10.2478/v10117-011-0021-1.
- [42] Actix. *ACTIX WIRELESS NETWORK SOLUTIONS:ACTIX ANALYZER Over View*, URL: <https://libguides.library.kent.edu/SPSS/PearsonCorr>. [Accessed on May 04, 2021.]
- [43] B. Adam Mtaho and F. Ishengoma. “Factors Affecting QoS in Tanzania Cellular Networks”. In: *ArXiv abs/1410.0533* (2014).

- [44] Jessica L. O’Leary and Glenn D. Israel. “Constructing Closed-Ended Items for a Questionnaire,” in: *UF/IFAS: University of Florida* ().
- [45] N. Mohd Razali and B. Yap. “Power Comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Tests”. In: *J. Stat. Model. Analytics* 2 (Jan. 2011).
- [46] ITU. “TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS”. In: *ITU-T Recommendation G.107* (1994).
- [47] G. Ferrari et al. “Wireless Sensor Networks: Performance Analysis in Indoor Scenarios”. In: *EURASIP Journal on Wireless Communications and Networking* 2007 (Dec. 2007), pp. 41–41. DOI: 10.1155/2007/81864.
- [48] Actix. *State of Mobile Networks Report: Chapter 2 Consistent quality* , URL: <https://www.tutela.com/state-of-mobile-networks-southeast-asia/consistent-quality>. [Accessed on May 12 ,2021.]
- [49] Teltonika. *State of Mobile Networks Report: Chapter 2 Consistent quality* , URL: [https://wiki.teltonika-networks.com/view/Mobile\\_Signal\\_Strength\\_Recommendations](https://wiki.teltonika-networks.com/view/Mobile_Signal_Strength_Recommendations). [Accessed on September 17 ,2021.]
- [50] N. Awad and I. Mkwawa. “The impact of the reference signal received power to quality of experience for video streaming over LTE network”. In: *2017 Annual Conference on New Trends in Information Communications Technology Applications (NTICT)*. 2017, pp. 192–196. DOI: 10.1109/NTICT.2017.7976145.
- [51] R. Dhungana and D. Shahi. “A Quantitative Study of Quality of Experience (QoE) for Decreasing and Increasing Network Performance.” In: 2011.
- [52] M. Davidovic, S. Bjekovic, and I. Tomic. “On the impact of network load on LTE network downlink performance”. In: June 2015.
- [53] F. Afroz et al. “SINR, RSRP, RSSI and RSRQ Measurements in Long Term Evolution Networks”. In: *International Journal of Wireless Mobile Networks* 7 (Aug. 2015), pp. 113–123. DOI: 10.5121/ijwmn.2015.7409.
- [54] P. Mogensen et al. “LTE Capacity Compared to the Shannon Bound”. In: *2007 IEEE 65th Vehicular Technology Conference - VTC2007-Spring*. 2007, pp. 1234–1238. DOI: 10.1109/VETECS.2007.260.

# Appendix A

## Sample SPSS Correlation results

		Startup Delay	Stalling Duration	Play-back	Throughput	Latency	Jitter	RSRP	SINR
	Corr. Coefficient	1	0.736	-0.985	-0.839	0.355	0.194	-0.736	-0.653
	Sig. (2-tailed)		0	0	0	0.011	0.173	0	0
Startup Delay	N	51	51	51	51	51	51	51	51
	Corr. Coefficient	0.736	1	-0.705	-0.605	0.278	-0.074	-0.578	-0.468
	Sig. (2-tailed)	0		0	0	0.049	0.604	0	0.001
Stalling Duration	N	51	51	51	51	51	51	51	51
	Corr. Coefficient	-0.985	-0.705	1	0.826	-0.388	-0.206	0.7	0.647
	Sig. (2-tailed)	0	0		0	0.005	0.146	0	0
Playback	N	51	51	51	51	51	51	51	51

	Corr. Coefficient	-0.736	-0.578	0.7	0.871	-0.328	-0.433	1	0.836
	Sig. (2-tailed)	0	0	0	0	0.019	0.002		0
RSRP	N	51	51	51	51	51	51	51	51
	Corr. Coefficient	-0.653	-0.468	0.647	0.819	-0.379	-0.436	0.836	1
	Sig. (2-tailed)	0	0.001	0	0	0.006	0.001	0	
SINR	N	51	51	51	51	51	51	51	51
	Corr. Coefficient	0.355	0.278	-0.388	-0.384	1	0.046	-0.328	-0.379
	Sig. (2-tailed)	0.011	0.049	0.005	0.005		0.747	0.019	0.006
Latency	N	51	51	51	51	51	51	51	51
	Corr. Coefficient	0.194	-0.074	-0.206	-0.292	0.046	1	-0.433	-0.436
	Sig. (2-tailed)	0.173	0.604	0.146	0.037	0.747		0.002	0.001
Jitter	N	51	51	51	51	51	51	51	51
	Corr. Coefficient	-0.694	-0.428	0.716	0.718	-0.319	-0.402	0.674	0.637
	Sig. (2-tailed)	0	0.002	0	0	0.027	0.005	0	0
MOS	N	48	48	48	48	48	48	48	48

# Appendix B

## Subjective Metrics Questionnaire

**Part I: Please select one choice from the given options for your demographic status.**

Age?

- Under 18 years
- 18 - 25 years
- 26 - 35 years
- 36 - 45 years
- 46 - 55 years
- Above 55 years

Gender?

- Male
- Female

Current job status?

- Investor
- Employed
- Not employed
- Business Owner

Educational background?

- Master's degree and above
- Bachelor's degree
- University or college student
- Diploma or Certificate
- Grade 11 - 12
- Grade 10 and below

Zones?

- SAAZ
- NAAZ
- SWAAZ
- EAAZ

Specific location where you watched the video?

- Mebrat Haile
- Gellan
- Jemo
- Torhailoch
- Piassa
- Atlas

Where do you watched the video?

- Indoor
- Outdoor

Terminal type (The type of device you used to watch the video)?

- Smartphones
- Tablets
- Computers
- Smart TV
- Others

Select the time period you watched the video?

- Morning (06:00AM - 11:59AM)
- Afternoon (12:00PM - 04:59PM)
- Evening (05:00PM - 07:59PM)
- Night (08:00PM - 12:00AM)

Date\_\_\_\_\_

## **Part II: User Perception.**

Please select one choice from the given options corresponding to the stated questions.

1. How would you evaluate the perception of annoyance level of initial delay (the time spent from the moment you request the video until the playback begins?)

- Imperceptible
- Perceptible, but not annoying
- Slightly annoying
- Annoying
- Very annoying

2. How do you perceive the annoyance level of frequency of pausing (stalling) during the video session?

- Imperceptible
- Perceptible, but not annoying
- Slightly annoying
- Annoying
- Very annoying

3. How do you perceive the annoyance level of duration time of each pausing (stalling) during the video session?

- Imperceptible
- Perceptible, but not annoying
- Slightly annoying
- Annoying
- Very annoying

4. Overall, how satisfied or dissatisfied are you with the LTE network video streaming service.

- Very Satisfied
- Somewhat satisfied
- Neither satisfied nor dissatisfied
- Somewhat dissatisfied
- Very dissatisfied

5. Any additional experience or comments. \_\_\_\_\_

# Appendix C

## IEEE Manuscript

### Correlation among Quality of Experience and Quality of Service Parameters for LTE Video Streaming Service: In case of Addis Ababa, Ethiopia

Alemayehu Tiruneh Mamo  
Addis Ababa Institute of Technology  
School of Electrical and Computer  
Engineering  
Addis Ababa University  
Ethiopia, Addis Ababa  
[alemayehu.nathan@gmail.com](mailto:alemayehu.nathan@gmail.com)

**Abstract**— The growth in prevalence of smartphones and the increasing popularity of multimedia video streaming services resulted in impressive demand growth and a significant need for Quality of Service (QoS); following that making the video streaming services available to users at a cost effective way without compromising quality is a big challenge. In this paper, we analyze the correlation among Quality of Experience (QoE) and QoS parameters for video streaming service over LTE network in Addis Ababa, to identify the level of impact of QoS parameters on QoE of video streaming users and indicate measurement variations depending on video streaming service requirements. The survey included 300 Long Term Evolution (LTE) video streaming users from six different locations, and we collected data using the Nemo Handy, Opensignal, and the Cloudflare speed test. On average, a MOS result of 3.9 is obtained, and the analysis result reveals that the playback is the main factor responsible for the variation of Mean Opinion Score (MOS). The main reason for the variance in application layer quality metrics is download throughput, and network layer quality matrices are influenced by an increase in the number of users and other factors, rather than Reference Signal Received Power (RSRP) and Signal-to-Interference-plus-Noise Ratio (SINR), in some of the study locations. On the other hand, measured latency and jitter results, in particular during the evening and night periods are insufficient to watch standard and higher definition videos at the desired quality level. Within these time periods, the MOS values are less than 3.3.

**Keywords**— *Key words: LTE, QoE, QoS, Correlation, Subjective Measurement and Objective Measurement.*

#### I. INTRODUCTION

Most of the connections to video platforms and social networking applications originate from wireless cellular mobile communication networks. This massive access from wireless networks introduces several issues. Following that, telecom service providers are increasing the access speed that allows a huge number of users and changing the way of managing the mobile communication network to guarantee QoS. The 3GPP has defined the 4th mobile generation radio technology, named LTE or the Evolved Universal Terrestrial

Access Network (E-UTRAN), to provide wireless broadband with respect to QoS support [1,2].

The quality concept takes two forms, which are QoS and QoE. QoS is the network quality evaluated from a technical perspective to determine network performance, through measuring several factors (i.e., throughput, available bandwidth, delay, error probability, jitter, packet loss, etc.) [3]. According to the International Telecommunication Union (ITU) recommendation, the quality of service was defined as a collective effect of service performance, which determines the degree of satisfaction of a user of the service. On the other hand the QoE is the Degree of delight or annoyance of the user of an application or service as perceived subjectively includes the complete end-to-end system effects ... maybe influenced by user state, content and context [4].

Video streaming services are one of the dominant services around the world [5]. Similarly, a recent study in the Ethiopian capital, Addis Ababa, shows that video streaming services have the largest share of LTE services [6]. Although the service has a large number of users and generates better revenue, there are many complaints about the service [7, 8]. Hence, it is very important to investigate the current network performance to demonstrate the extent to which the areas surveyed are sufficient for different video quality and study correlation among QoE and QoS parameters to identify the most influential factors on the QoE of end users. In this paper, we study the correlation among application, network and physical layer quality metrics as well as their relationship with QoE. Figure 1 shows the methodology used for this research.

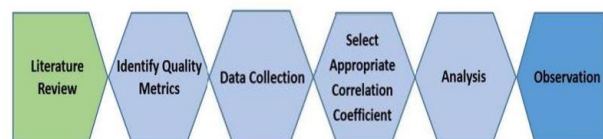


Fig. 1. Methodology

The rest of this paper is structured as follows. Section II highlights the related works. Section III discusses QoS/QoE measurement and correlation analysis approaches. Section IV addresses analysis results and Interpretations and finally conclude this paper in section V.

## II. RELATED WORK

In [9], the author investigated the relationship among three levels of QoS (network, application, and user QoS) for Hypertext Transfer Protocol (HTTP) video streaming services to understand how the network QoS affects the QoE of video streaming users. As a network QoS Round-trip Time (RTT), packet loss rate, and network bandwidth have been used, Initial buffering time, mean rebuffering duration, and rebuffering frequency are considered as application QoS and from the user side MOS is used. The author finally proposed three-level performance metrics for HTTP video streaming and identified the rebuffering frequency to be the main factor responsible for the MOS variance.

V. D. Bhamidipati investigated the effect of delay and delay variation on the QoE of video streaming end-users in [10]. The perceived QoE for different streaming sessions has been collected using MOS and finally the author figures out the threshold level of delay that is acceptable by the users, the tolerance of users for different delay levels, and the effect of packet delay. The acceptable tolerance was found to be  $\pm 8$ ms and  $\pm 10$ ms for different video types. and the author concluded that when delay and delay variation increase above  $\pm 10$ ms the QoE decreases.

The author of [11] evaluated user perception of videos (MOS), objective video quality metrics (Video Quality Metric (VQM), Structural Similarity (SSIM), and Peak Signal Noise Ratio (PSNR)), and network QoS (Packet loss, jitter, and reordering) to find the correlation between QoS and QoE. The video resolutions of 480p, 720p, and 1080p are considered, and a three-phase evaluation method is introduced. Finally, the author concluded that MOS is affected by the previous experiences of the watchers and that linear combinations of some QoS metrics can correlate well with MOS values.

[12] presents a video streaming QoE evaluation over 3G networks for Android-based smartphone users. This paper considered the application and network layer quality metrics (average video resolution watched, average initial latency, the average number of stalling events, average total stalling length, and average buffer level) as well as network-level parameters such as average download throughput to show the relationship between QoE and QoS. As a result, the correlation between the average acceptability range and the MOS has been found to be strong with a correlation coefficient and significant level values of ( $r=0.852$ ,  $p=0.01$ ) and the download throughput was lower than the recommended values, with an overall MOS score of 2.79.

## III. DATA COLLECTION AND CORRELATION ANALYSIS APPROACHES

### A. Subjective and Objective Data Collection

The data collection process involves spatial and temporal aspects. From video experience, we measured startup delay and stalling duration in seconds (Sec) and playback in percentages (%). From the speed experience, latency in milliseconds (ms) was measured. RSRP, SINR, and Download throughput measurements were collected using

Nemohandy hand held measurement tool, and Actix Analyzer is used to retrieve data from Nemohandy in a variety of ways and also it has a capability of plotting parameters on maps and display them on charts, tables, etc. The jitter measurement is collected by a Cloudflare speed test, calculated as the average latency delta between any two consecutive measurements and it is measured by downloading/uploading progressively larger files.

On the other hand, the subjective survey questionnaire was carefully developed to capture LTE video streaming users' perceptions of stalling duration, stalling frequency, and overall satisfaction level. The data was collected from six distinct places in four different zones which were chosen based on their monthly traffic distribution and at different times of a day (morning, afternoon, evening and night).

There are several subjective tests methods for video quality assessment, most of them described in the ITU recommendations. In this study, we adopt one of the most widely used scales: the 5-point MOS scale described in the Absolute Category Rating (ACR) method (1:bad, 2:poor; 3: fair; 4:good; 5:excellent), which are standardized by ITU-T. Table I provides an overview of the personal profiles of the survey participants.

TABLE I. A SUMMARY OF THE PARTICIPANTS' PERSONAL PROFILES.

Gender	Male	204
	Female	96
Age	Under 18 years	41
	18 – 25 years	89
	26 – 35 years	72
	36 – 45 years	77
	46 – 55 years	15
	Above 55 years	6
Educational Backgrounds	Master's degree and above	10
	Bachelor's degree	50
	University or college student	70
	Diploma or Certificate	120
	Grade 11 – 12	30
	Grade 10 and below	20

### B. Determining Sample Size:

The sample size was calculated using the formula found in [13]

$$n = \frac{N}{(1+N(e)^2)} \quad (1)$$

Where n - sample size, N – Total customer, e – detection error expressed in to percentage (5% -10%) and thus the detection error (e) = 6%, any value for e from 5% - 10% is acceptable.

$$\approx 278 \text{ Respondents}$$

The respondents should not to be less than 278. In order to make the research feasible the sample size of 300 subscribers from four Addis Ababa zones were participated in the survey.

### C. Correlation Coefficient Selection

Various correlation coefficients can be used to measure correlations between different variables. Pearson's coefficient, Spearman's rho coefficient, and Kendall's tau coefficient are the most widely used. We selected the appropriate coefficient using the Shapiro-Wilk test presented in study [4], which was selected as a most powerful test for all types of distribution

and sample sizes. According to the results only RSRP, SINR and Latency quality metrics were normally distributed with the significance values of .362, .038 and .553. Therefore, because Pearson Correlation Coefficient is affected by extreme values, which can exaggerate or dampen the strength of the relationship, if we must use it, our variables should be normally distributed; otherwise, interpretation and inferences may be unreliable or invalid.

On the other hand, Spearman's rho coefficient is suitable when one or both variables are skewed or ordinal, and it is robust when extreme values are present. Therefore, As a result, it was decided that Spearman's rho coefficient would be used in this research. The formula for computing the Spearman's correlation coefficient for a correlation between variables x and y is given by:

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2-1)} \quad (2)$$

Where  $d_i$  is the difference in ranks for x and y.

#### IV. ANALYSIS, RESULTS AND INTERPRETATIONS

##### A. Subjective Evaluation Results

Subjective metrics are the results of participant responses to the sample watched videos. On average, a MOS result of 3.9 is obtained and the MOS scores varied from 2.0 to 5.0. 70.83% of participants rated the MOS value in between 3.5 to 4.8 and the remaining 29.17% of participants rated a score of less than 3.5. According to ITU-T standard [14], the average MOS result indicated that some users are unhappy with the current video streaming quality. MOS can be affected by a variety of different factors; in this case, the environment context (indoor/outdoor) is one of the possible reasons. As shown in Figure 2, the LTE video streaming customers who used the service indoor were less satisfied than those who used it outside.

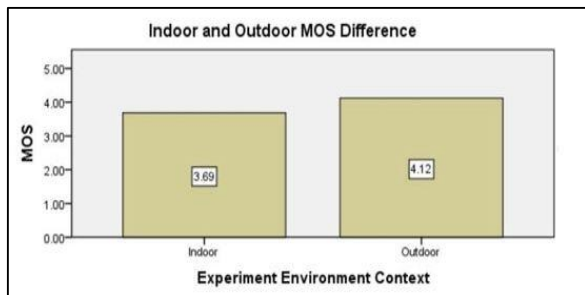


Fig. 2. Indoor and Outdoor MOS Difference

Morning, Afternoon, Evening, and Night are the four periods in which the responses of the participants were collected. The smallest MOS was measured in the majority of the research locations during the evening and night hours. Due to the fact that the evening and night time periods are not normal working hours, several users were expected to access the network at the same time, therefore, of the six places surveyed, the lowest MOS value was 2.5 achieved at Mebrathail area.

In addition to technical reasons, QoE can be affected by other influential factors. As part of this research, age and gender factors have been considered and analyzed for their impact on MOS values.

Female respondents achieved an average of MOS 3.91, whereas males average MOS value was 4.05. On the other hand, participants under the age of 35 achieved a better MOS values than those over 35, with a MOS score of 4.01 and 3.97 respectively. The number of male participants is more than double that of females and 201 participants were under 35 years of age, while 99 individuals were over the age of 35. Therefore, it is recommended that both age and gender distribution of participants should be proportional to achieve the best results of the study.

##### B. Objective Evaluation Results

###### 1) Application Layer Evaluation Results

Table II depicts Opensignal evaluation results for 300 YouTube mobile streaming sessions. Based on the experiment, the average length of the startup delay was 2.1 seconds, with maximum and minimum startup delays of 6.5 seconds and 0.7 seconds, respectively. On the other hand, there was an average stalling duration of 0.8 seconds, while the maximum and the minimum stalling duration were 5.1 and 0 seconds and the average playback was 85.7 percent.

TABLE II. OPENSIGNAL EVALUATION RESULTS OF VIDEO STREAMING METRICS

KPIS	Maximum	Mean	Minimum	Std. Deviation
Startup Delay	6.5	2.1	0.7	1.3
Stalling Duration	5.1	0.8	0	1.3
Playback	95.7	85.7	60	8.8

When we observe the results of average startup delay, stalling duration and playback over different time periods: The average startup delay and stalling duration were recorded during the night session with a score of 3.3 and 1.7 seconds, their minimum scores were 1.4 and 0.3 seconds captured in the morning session. Similarly, the morning and afternoon playback results were 90.8 and 85.8, while the evening and night playback results were 81.2 and 77.1. In general, the LTE video users perceived better experience in the morning and afternoon sessions, while the evening and night time periods quality of service remained bad.

On the other hand, the measurements taken indoor were smaller than those measured in outdoor. The average startup delay and stalling duration achieved in an outdoor environment were 2.0 and 0.7 seconds, respectively, while 2.5 and 1.2 seconds were got from the indoor measurement. According to [15], indoor measurements of wireless networks were smaller than outdoor measures due to walls, doors, metallic furniture, and other considerations.

###### 2) Network Layer Evaluation Results

According to a study [10], latency and jitter have a negative impact on the quality of experience of video streaming viewers when the delay variation is greater than 10 milliseconds. As part of this investigation, morning and afternoon measurements showed 9.5 Mbps and 7.2 Mbps download throughput, respectively, and the latency and jitter during this time periods were (106.1 and 41.4 ms) and (112.8 and 47.4 ms), respectively.

On the other hand, download throughput was low at 4.8 Mbps during the night hours of the day, while latency and jitter increased to 133.0 ms and 48.7 ms, respectively.

The decreasing of the download throughput in the evening and night hours and at the same time the increase in latency and jitter affect the overall customer satisfaction, and within these time periods, the MOS score was also decreased. From a geographical standpoint, the maximum and minimum average download throughputs of 7.82 Mbps and 5.40 Mbps, recorded in the Gelan and Torhailoch areas respectively. These are average results, but there were some locations where the download throughput was less than 1 Mbps, which could have a significant impact on the video streaming service.

As we have seen in Figure 3, the maximum download throughput was achieved in the outdoor environment that is 8.04 Mbps, whereas the latency and jitter were 110.4 and 47.2 ms with in this environment.

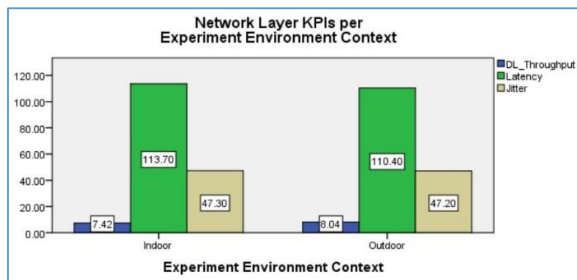


Fig. 3. Average DL\_Throughput/Latency/Jitter results per Experiment Environmental Context.

According to [16], the latency requirement should not exceed 100 milliseconds for SD videos and 50 milliseconds for HD videos. In this study, 39 percent of participants experienced a latency between 117 and 142 milliseconds, which was the greatest percentage, and 38 percent of participants experienced a latency between 91 and 116 milliseconds, which was the second highest percentage. Generally, 70% of participants experienced more than 100 ms of latency. As a result, users were unable to watch movies at the desired quality level because this result did not meet the latency and jitter requirements.

### 3) Physical Layer Evaluation Results

When we look at RSRP and SINR results in terms of location, Gelan has the highest average RSRP and SINR with 83.64dBm and 13.21dB, respectively. At Piassa, we achieved the minimum average RSRP and SINR values of -92.04dBm and 8.66dBm, respectively. As we can see from the MOS results above, the results obtained here have been implicated in the MOS results obtained in the same locations.

Participants who watched the videos had the best RSRP and SINR measurement results in the morning and afternoon, with average results of (-85.71 and -88.17dBm) and ( 12.09 and 10.03dB), respectively, therefore, customers who access mobile data with in these periods can simply connect to a nearby cell tower, resulting in a better connection. The RSRP and SINR measurement values for participants who watched sample videos in the evening and at night were (-89.88 and -92.69dBm) and (9.80 and 7.97dB), respectively, which were lower than those in the morning and afternoon. In the evenings and at night, the average RSRP and SINR were lower. This

finding indicates that video streaming users do not have reliable service in the evenings and at night.

Even though the average measured RSRP values in this study were higher than -102.9dBm, in some study locations we captured poor results. For SINR values, around 33% of users captured less than 13dB. The recommendation in [17] reveals that, with these values, reliable data speeds may be attained, but marginal data with dropouts is possible. When this value gets close to zero, performance will drop drastically.

### C. Correlation Analysis Results

#### 1) Correlation among MOS and Application, Network and Physical Layer Quality Metrics.

When we look at the correlation results among MOS and the application layer quality metrics, on average, the spearman correlation of MOS and Playback was found to be a positive strong correlation and there is enough evidence to show the correlation is statistically significant with the correlation coefficient and level of significance values of ( $r = .746$ ,  $p < 0.001$ ). On the contrary the correlation of MOS with startup delay and stalling duration was found to be a negative moderate correlation with a value of correlation coefficient ( $r = -.667$  and  $-.593$ ) respectively and the significance values for both correlation results were  $P < 0.001$ .

With respect to study locations, the highest degree of correlation was achieved at Mebrathail between startup delay and MOS, with correlation coefficient magnitude of ( $r = -.916$ ,  $p < .001$ ) and apart from Mebrathail and Gelan, for the remaining four study locations we observed that the next highest degree of association between playback and MOS.

In accordance with the results of this study, we identified that the playback quality metric to be the main factor responsible for the MOS variation, that is, Playback is the most influential factor on QoE of video streaming users. Figure 4 shows the scatter plot representing the correlation between playback and MOS, that indicates a strong positive association between playback and MOS scores, when the payback increases, the level of the video streaming user's QoE increases or vice versa.

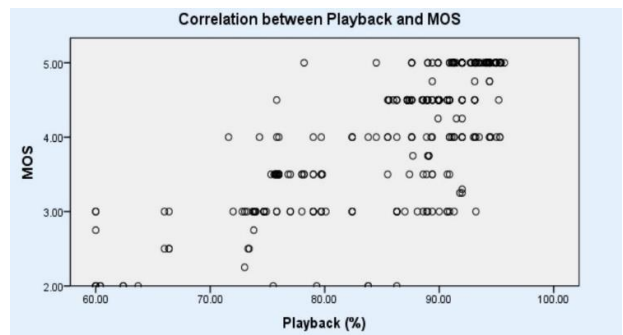


Fig. 4. Scatter Plot Representing the Correlation between Playback and MOS.

#### 2) Correlation between Application and Network Layer Quality Metrics.

Users QoE could be influenced by different objective (network quality, device capacity) and subjective parameters (expectations, emotion and location) [18]. In terms of video streaming services, there are also a number of problems when the video packets travel through the wireless network, from the lower layer all the way up to the application layer. In this

study, based on the results of a correlational analysis, we determined which network metrics have the most influence on application layer performance.

Download throughput has a higher degree of correlation with the playback quality metric with a values of ( $r = .711$ ,  $p < .001$ ), which is strong and positive correlation. The second highest degree of correlation achieved in between the download throughput and stalling duration ( $r = -.647$ ,  $p < .001$ ), their relationship was negative, moderate in strength and statistically significant.

According to the correlation analysis results, it is observed that the download throughput plays a major role for the variation of application layer quality metrics than latency and jitter . In other word, if there is change in download throughput then there is corresponding significance change in the application quality metrics because they have the highest degree of association between them. Figure 5 shows the scatter plot representation of correlation between download throughput and playback quality metrics.

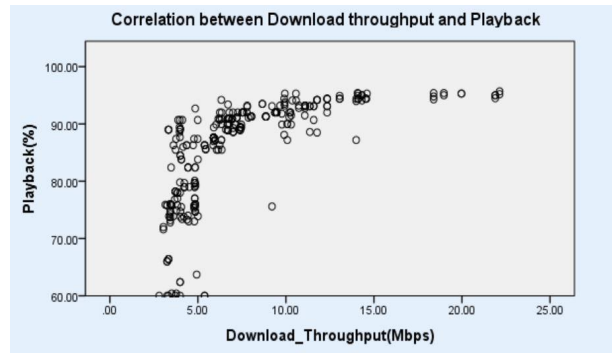


Fig. 5. Scatter Plot Representing the Correlation between Download throughput and Playback

### 3) Correlation between Network and Physical Layer Quality Metrics.

RSRP and SINR had positive strong correlation with download throughput having a correlation coefficient values of ( $r = .778$  and  $r = .773$ ) respectively, while they have negative moderate correlation with latency and jitter ( $r = -.421$  and  $r = -.407$ ), the significance level for both conditions was  $p < 0.001$ ).

Looking at the correlation results obtained separately from all study locations, except Mebrathail and Jemo on the other study areas the download throughput has a higher degree of correlation with RSRP and SINR. In this case, at Mebrathail and Jemo, the download throughput mostly affected by other wireless network impairments, mainly the increase in the number of users and the like, but the influence of RSRP and SINR is not negligible. For the rest four locations, RSRP was the most influential factor. The scatter plot in Figure 6 below shows that RSRP and SINR have a positive relationship with Download throughput, thus when RSRP and SINR increase, the Download throughput increases as well.

SINR is one of the impairments that might lower the efficiency of an LTE network's download throughput, and we determined that it has a significant relationship with download throughput in this study.

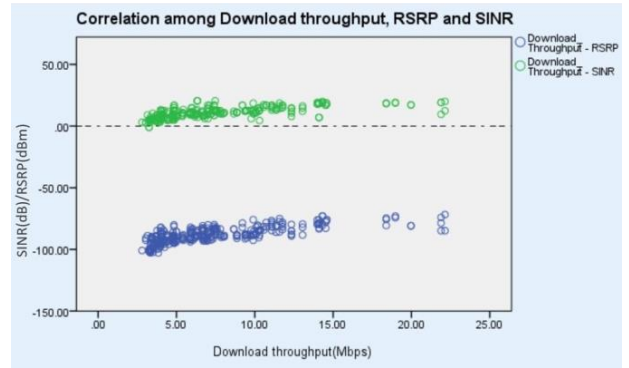


Fig. 6. Scatter Plot Representing the Correlation among Download Throughput with RSRP and SINR.

On the other hand, their mathematical relationship can be represented using the below equation 3. For LTE Multiple-input-multiple-output (MIMO) antenna systems, the download throughput efficiency was calculated using the Shannon channel capacity formula [19, 20], and to address different implementation issues, a modified Shannon capacity formula was used:

$$S(\text{bits/s/Hz}) = BW\_eff * \eta * \log_2(1 + SNR / SNR\_eff) \quad (3)$$

Here  $BW\_eff$  adjusts for the system bandwidth (BW) efficiency of LTE and  $SNR\_eff$  adjusts for the SNR implementation efficiency of LTE. The factor  $\eta$  is a correction factor which nominally should be equal to one.

## V. CONCLUSION AND FUTURE WORK

### A. Conclusion

Even though the video streaming service is influenced by human, system, context and content related influencing factors, studying the correlation among QoE and QoS different factors is the main purpose of this thesis. Additionally, we investigate the spatio-temporal network measurement variations in terms of video streaming service.

As we can see from the results of the measurement, the latency and jitter results are not sufficient for video streaming services in most of the study locations, especially in the evening and night hours. For example, in order to watch HD and greater video quality, a latency of less than 50ms and jitter values of less than 30ms are required. However, the findings obtained here are higher than the recommended values. As a result, this consequence may decrease the satisfaction level of video streaming users and the MOS result reflected this trend as well, which is 3.9 on average.

The correlation results show that MOS is mainly lowered by playback, implying that playback is the main factor responsible for the MOS variation. On the other hand, it is observed that there is a higher correlation between download throughput and playback ( $r = .711$ ,  $p.001$ ) than latency and jitter. As a result, download throughput is the primary cause of variation in application layer quality metrics. While the RSRP and SINR results were good in the two study locations, Mebrathail and Jemo, the network quality measures were poor. In addition, RSRP and SINR have a weak relationship with network-layer quality indicators. This implies that while RSRP and SINR are good, network layer metrics may not be good due to other factors like increase in the number of users

and etc. and spatial structure is an important factor affecting QoS as well as QoE.

The study also shows that at Mebrathail and Jemo in particular during the evening and nighttime periods, it is not suitable for video streaming services, possibly due to congestion or higher distance from the serving cell. This implies that the QoS varies at different locations and times of the day on the cellular network. In addition, from this thesis work, Ethio telecom can understand the gap of recent root cause analysis approaches.

### B. Future Work

The following areas are recommended for future work.

- This research was limited to Addis Abeba, the capital of Ethiopia. Therefore, in order to have a complete picture, future works may include other cities where the service is available.
- As the video streaming service is available at different quality levels, future research could benefit from including multiple resolution videos, such as 720P, 1080p, etc.
- A variety of technologies, such as tablets and comparable video-enabled devices, as well as smart TVs, could be used in future studies.
- This thesis work is limited to YouTube-based video streaming services over LTE; however, future work may incorporate other Internet services and technologies, such as LTE Advanced.
- Consider the joint effect of quality metrics on QoE.

### REFERENCES

- [1] P. Uppu. "QoE of video streaming over LTE network". In: School of Computing, Blekinge Institute of Technology, Karlskrona, Sweden, (2013).
- [2] 3gpp. The Mobile Broadband Standard, url: <http://www.3gpp.org/LTE>, [Accessed on April 10, 2021.], pp. 271–350. G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving".
- [3] B. Khadija, S. Essaid, and S. Mohamed. "Quality of experience for streaming services: measurements, challenges and insights". In: IEEE Access vol. 8 (2020).
- [4] ITU. "Terms and Definitions Related to Quality of Service and Network Performance Including Dependability". In: ITU-T Recommendation E.800 (1994).
- [5] J. Parikshit, T. Venkatesh, and M. Deep. "Measurement of Quality of Experience of Video-on-Demand Services: A Survey". In: IEEE Communications Surveys Tutorials 18.1 (2016), pp. 401{418. doi: 10.1109/COMST.2015.2401424.
- [6] D. S. Tibebu. "Measurement of Quality of Experience of Video-on-Demand Services: A SurveyQoE Model for Addis Ababa LTE Web Browsing Service Using Neural Network Approach". In: Addis Ababa University, Addis Ababa Institute of Technology School of Electrical and Computer Engineering, (December 2019).  
A. R. Bisrat. "Quality of Experience Evaluation for Addis Ababa UMTS Enterprise Data Customers". In: Addis Ababa University, Addis Ababa Institute of Technology School of Electrical and Computer Engineering, (October, 2018).
- [7] D.Weldu. "Machine Learning Based QoE Estimation Model for Video Streaming over UMTS Network". In: Addis Ababa University, Addis Ababa Institute of Technology School of Electrical and Computer Engineering, (February, 2020).
- [8] Ricky K. P. Mok, Edmond W. W. Chan, and Rocky K. C. Chang. "Measuring the quality of experience of HTTP video streaming". In: 12th IFIP/IEEE International Symposium on Integrated Network Management (IM 2011) and Workshops. 2011, pp. 485{492. doi: 10.1109/INM.2011.5990550.
- [9] V. D. B. Swetha Kilari. "Effect of Delay/ Delay Variation on QoE in Video Streaming". In: School of Computing Blekinge Institute of Technology, Karlskrona Sweden (May, 2010).
- [10] P. Orosz et al. "A case study on correlating video QoS and QoE". In: 2014 IEEE Network Operations and Management Symposium (NOMS). 2014, pp. 1{5. doi: 10.1109/NOMS.2014.6838399.
- [11] Z. F. G/Mariam. "Quality of Experience Evaluation for Smartphone Video Streaming over 3G Network: In the case of Addis Ababa, Ethiopia". In: Addis Ababa University, Addis Ababa Institute of Technology School of Electrical and Computer Engineering, (December 2019).
- [12] Adam B. Mtaho and F. Ishengoma. "Factors Affecting QoS in Tanzania Cellular Networks". In: ArXiv abs/1410.0533 (2014).
- [13] ITU. "TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS". In: ITU-T Recommendation G.107 (1994).
- [14] J. Parikshit, T. Venkatesh, and M. Deep. "Measurement of Quality of Experience of Video-on-Demand Services: A Survey". In: IEEE Communications Surveys Tutorials 18.1 (2016), pp. 401{418. doi: 10.1109/COMST.2015.2401424.
- [15] Actix. State of Mobile Networks Report: Chapter 2 Consistent quality , url: <https://www.tutela.com/state-of-mobile-networks-southeast-asia/consistent-quality>. [Accessed on May 12, 2021.]
- [16] Teltonika. "State of Mobile Networks Report: Chapter 2 Consistent quality," url:[https://wiki.teltonikanetworks.com/view/Mobile\\_Signal\\_Strength\\_Recommendations](https://wiki.teltonikanetworks.com/view/Mobile_Signal_Strength_Recommendations). [Accessed on September 17, 2021.]
- [17] R. Dhungana and D. Shahi. "A Quantitative Study of Quality of Experience (QoE) for Decreasing and Increasing Network Performance." In: 2011.
- [18] F. Afroz et al. "SINR, RSRP, RSSI and RSRQ Measurements in Long Term Evolution Networks". In: International Journal of Wireless Mobile Networks 7 (Aug. 2015), pp. 113{123. doi: 10.5121/ijwmn.2015.7409.
- [19] P. Mogensen et al. "LTE Capacity Compared to the Shannon Bound". In: 2007 IEEE 65th Vehicular Technology Conference - VTC2007-Spring. 2007, pp. 1234{1238. doi: 10.1109/VETECS.2007.260.