



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

**Evaluation of 4G LTE User Side Data QoS
Measurement Tools: The Case of Addis Ababa,
Ethiopia**

By

Adhena Mesele Misgun

Advisor

Dr. Beneyam Berehanu Haile

A Thesis Submitted to the School of Electrical and Computer Engineering, in Partial
Fulfillment of the Requirements for the degree of Master of Science in Telecommunication
Engineering

October 2021

Addis Ababa, Ethiopia



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

Evaluation of 4G LTE User Side data QoS
Measurement Tools: The Case of Addis Ababa,
Ethiopia

By: Adhena Mesele

Approval by Board of Examiners

(Chairman, School Graduate Committee)

Dr. Beneyam Berehanu Haile

Advisor

Internal Examiner

External Examiner

Signature

Signature

Signature

Signature



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.

Adhena Mesele Misgun

Name

Signature

Place: Addis Ababa

Date of Submission: _____

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

Dr. Beneyam Berehanu Haile

Advisor

Signature



Abstract

Measuring Quality of Service (QoS) for a specific mobile network system may be performed by a network operator, Internet service provider, or governmental regulator based on their objectives. Since telecom stakeholders may not use similar measurement methods and tools, their measured values for the selected parameters may not yet be the same. As a result, there is disagreement among stakeholders regarding the benchmark and evaluations. The service coverage and the QoS are critically important for a successful operator business. So, mobile operators should continuously monitor their network performance and undertake optimization and expansion operations on a need basis. The techniques and tools for measuring QoS should be evaluated. The area of study proposed in the thesis is the evaluation of 4G Long Term Evolution (LTE) mobile network system data QoS measurement factors and tools. QoS measurement tools and factors are evaluated based on the selected parameters' reference value.

The thesis evaluates and analyzes the data QoS mobile network measurement tools and factors in the context of Addis Ababa, Ethiopia, using the most relevant throughput and latency performance metrics on Ethio telecom's 4G LTE mobile network. Evaluation of static download speed, upload speed, and latency measurements using Ookla Speedtest, RTR-NetTest, and Nemo Handy-A are conducted using Mean Absolute Percentage Error (MAPE). The achieved evaluated results of download speed measured by the Ookla Speedtest, RTR-NetTest, and Nemo Handy-A to the reference value, 35 Mbps, are 57.35%, 84.22%, and 47.54%, respectively. The evaluated results of upload speed measured by Ookla Speedtest, RTR-NetTest, and Nemo Handy-A to the reference value, 32 Mbps, are 30.42%, 56.47%, and 25.27% correspondingly. The achieved evaluated results of latency measured by the Ookla Speedtest, RTR-NetTest, and Nemo Handy-A to the reference value, 27msec, are 42.06%, 493.56%, and 13.66% respectively. The main reasons for the difference in errors are server location, bottleneck, and parameter estimation methods of the measurement tools.

Keywords-LTE, QoS, Measurement methodology, Measurement factors, QoE, Mobile data services, Measurement tools, Throughput, Latency.



Table of Contents

| | |
|---|------|
| Declaration..... | ii |
| Abstract..... | iii |
| List of Figures..... | vii |
| List of Tables..... | viii |
| Acknowledgment..... | ix |
| List of Acronyms..... | x |
| 1. Introduction..... | 1 |
| 1.1. Statement of the Problem..... | 2 |
| 1.2. Objective of the Thesis..... | 4 |
| 1.2.1. General Objective..... | 4 |
| 1.2.2. Specific Objectives..... | 4 |
| 1.3. Literature Review..... | 5 |
| 1.4. Methodology..... | 6 |
| 1.5. Scope and Limitation..... | 7 |
| 1.5.1. Scope..... | 7 |
| 1.5.2. Limitation..... | 7 |
| 1.6. Contributions..... | 7 |
| 1.7. Thesis Organization..... | 8 |
| 2. Overview of 4G Network..... | 9 |
| 2.1. Background..... | 9 |
| 2.2. Architecture of LTE..... | 11 |
| 2.3. QoS Classes in 4G LTE..... | 13 |
| 2.4. New Enhanced Technologies in 4G LTE..... | 15 |
| 2.5. Modified Shannon Theory on Throughput..... | 17 |
| 2.6. 4G Network Services and Parameters..... | 17 |



| | | |
|--------|---|----|
| 3. | Data QoS Measurement Methodology and Tools | 19 |
| 3.1. | Network Measurement Types | 19 |
| 3.1.1. | Network Performance | 19 |
| 3.1.2. | Overall Network Performance | 19 |
| 3.1.3. | End to End Quality of Service | 20 |
| 3.1.4. | Quality of Experience | 21 |
| 3.2. | Network Measurement Categories | 22 |
| 3.2.1. | Active and Passive Measurement | 22 |
| 3.2.2. | Foreground and Background Measurements | 22 |
| 3.2.3. | Best Effort and the Windowed Approach | 23 |
| 3.2.4. | Network and User Side Measurement | 23 |
| 3.3. | Network Measurement Methodologies | 23 |
| 3.4. | Description of Data Service QoS Parameters | 27 |
| 3.5. | User Side Data Service End-to-end QoS Measurement..... | 28 |
| 3.6. | QoS Measurement Methodologies and Tools | 28 |
| 4. | Methodology and Data Collection..... | 33 |
| 4.1. | Data Collection..... | 34 |
| 4.2. | Data Preprocessing and Analysis | 35 |
| 4.3. | Evaluation Techniques and Reference Values | 36 |
| 4.4. | Data QoS Measurement Factors..... | 37 |
| 4.5. | Map the Factors to the Selected Tools | 39 |
| 4.6. | Identification of Manipulated Factors on the Client Side of the Tools..... | 40 |
| 4.7. | First Scenario: Visualize and Evaluate the Impact of Selected Measurement Factors | 40 |
| 4.8. | Second Scenario: Evaluation of the Three Measurement Tools | 41 |
| 5. | Results and Discussion | 43 |



| | | |
|--------|---|----|
| 5.1. | First Scenario: The Impact of Measurement Factors on the Parameters..... | 43 |
| 5.1.1. | The Impact of Measurement Factors on DL..... | 43 |
| 5.1.2. | The Impact of Measurement Factors on UL..... | 45 |
| 5.1.3. | The Impact of Measurement Factors on Latency | 47 |
| 5.2. | Second Scenario: DL, UL, and Latency Measurements Using the Three Tools | 49 |
| 5.2.1. | DL of the Three Measurement Tools..... | 49 |
| 5.2.2. | UL of the Three Measurement Tools..... | 50 |
| 5.2.3. | Latency of the Three Measurement Tools | 51 |
| 6. | Conclusion and Recommendation | 53 |
| 6.1. | Conclusion..... | 53 |
| 6.2. | Recommendation..... | 54 |
| | Reference | 55 |



List of Figures

Figure 2.1: Shannon capacity of communication system, in bandwidths of 5, 10, and 20 MHz [1]..... 10

Figure 2.2: EPS architecture [12] 12

Figure 2.3: EPS bearer architecture [17]..... 14

Figure 3.1: Classification of NP, QoS, and QoE [3]..... 19

Figure 3.2: Four viewpoints on QoS [3] 21

Figure 3.3: QoS measurement approach..... 28

Figure 3.4: Ookla Speedtest QoS measurement approach..... 29

Figure 3.5: RTR-Nettest QoS measurement approach 30

Figure 3.6: Nemo Handy-A QoS measurement approach 32

Figure 4.1: Proposed approach 33

Figure 4.2: Traffic level Vs Hours from PRS system..... 34

Figure 4.3: Measurement location from RTR server 35

Figure 5.1: Impact of factors on DL using Ookla Speedtest 44

Figure 5.2: Impact of techniques on DL using RTR-NetTest..... 44

Figure 5.3: Impact of no. of threads on DL using Nemo Handy-A..... 45

Figure 5.4: Impact of techniques on UL using Ookla Speedtest 46

Figure 5.5: Impact of techniques on UL using RTR-NetTest..... 46

Figure 5.6: Impact of techniques on latency using Ookla Speedtest 47

Figure 5.7: Impact of server locations on latency using Ookla Speedtest 48

Figure 5.8: Impact of factors on latency using RTR-NetTest..... 48

Figure 5.9: Plot of DL throughput using the three tools 50

Figure 5.10: Plot of UL throughput using the three tools 51

Figure 5.11: Plot of latency throughput using the three tools..... 52



List of Tables

| | |
|---|----|
| Table 2.1: QoS classes in LTE Release 8 | 15 |
| Table 3.1: Different network measurement techniques | 24 |
| Table 4.1: Actual techniques used by the tools..... | 39 |
| Table 4.2: Measurement setup for scenario two | 42 |
| Table 5.1: The average DL and evaluation of the number of threads..... | 45 |
| Table 5.2: Average and evaluation of the techniques using three tools..... | 49 |
| Table 5.3: Average and evaluations of the parameters using the tools..... | 52 |



Acknowledgment

I like to thank God for giving me the wisdom and skill to carry out my thesis work. My special gratitude goes to my advisor Beneyam Berehanu Haile (Dr.), for his excellent guidance and constructive comments up to the submission of my thesis. I would also like to thank my evaluator, Dr. –Ing. Dereje Hailemariam and Yalemzewd Negash (Dr.), for their valuable feedback during the thesis progress presentations.

I am very grateful to thank the employees of the Ethio telecom engineering department who have supported me in collecting data and providing necessary information for the successful delivery of this work. My gratitude goes to Wireless Planning and Engineering Managers, advisors and staff: Melese Fantaye, Yibeltal Addis, Kiflay Berhane, Tewodros Tefera, Dereje Feleke, and Woretaw Chanie for their vital support in the measurement aspect. Next special thanks go to my parents, brothers, and sisters. Without their help, encouragement, and effort in all moments of my life, nothing of this would be possible. I appreciate the support, patience, and encouragement they provided me throughout my studies.

I also wish to appreciate my colleagues and Ethio telecom staff for their support and encouragement throughout the course work. Last but not least, a special acknowledgment extends to Ethio telecom for allowing me to do my master's study and be responsible for my full sponsorship.



List of Acronyms

| | |
|----------------|--|
| 1G | First Generation |
| 2G | Second Generation |
| 3G | Third Generation |
| 3GPP | Third Generation Project Partner |
| 4G | Fourth Generation |
| APN | Access Point Network |
| BER | Bit Error Rate |
| BTC | Bulk Transport Capacity |
| CA | Carrier Aggregation |
| CS | Crowdsourcing |
| CSP | Communication Service Provider |
| DL | Download |
| DT | Drive Test |
| EMS | Element Management System |
| EPC | Evolved Packet Core |
| EPS | Evolved Packet System |
| E-UTRA | Evolved UMTS Terrestrial Radio Access |
| E-UTRAN | Evolved UMTS Terrestrial Radio Access Network |
| FCAPS | Fault Configuration Account Performance Security |
| FCC | Federal Communication Commission |
| FDD | Frequency Division Duplex |
| FTP | File Transfer Protocol |



| | |
|-------------|--|
| GBR | Guaranteed Bit Rate |
| GPS | Global Positioning System |
| GSM | Global System for Mobile |
| GTP | GPRS Tunneling Protocol |
| HSS | Home Subscriber Server |
| HTTP | Hypertext Transfer Protocol |
| ICMP | Internet Control Message Protocol |
| ISP | Internet Service Provider |
| ITU | International Telecommunication Union |
| IXP | Internet Exchange Point |
| KPI | Key Performance Indicator |
| LTE | Long Term Evolution |
| MAPE | Mean Absolute Percentage Error |
| MCS | Modulation and Coding Scheme |
| MDT | Minimization of Drive Test |
| MIMO | Multi Input Multi Output |
| MME | Mobility Management Entity |
| NAS | Non Access Stratum |
| NGN | Next Generation Network |
| NMS | Network Management System |
| NP | Network Performance |
| OFDM | Orthogonal Frequency Division Multiplexing |
| OMC | Operation and Maintenance Centre |
| OPEX | Operating Expenses |



| | |
|----------------|--|
| OSI | Open System Interconnection |
| PCC | Policy and Charging Control |
| PCRF | Policy Control and Charging Rules Function |
| P-GW | Packet Data Network Gateway |
| PMIP | Proxy Mobile IP Protocol |
| PLR | Packet Loss Ratio |
| PRB | Physical Resource Block |
| PRS | Performance Surveillance |
| QCI | QoS Class Identifier |
| QoE | Quality of Experience |
| QoS | Quality of Service |
| RAN | Radio Access Network |
| RE | Resource Element |
| RMSE | Root Mean Square Error |
| RSRP | Reference Signal Received Power |
| RSRQ | Reference Signal Received Quality |
| RSSI | Received Signal Strength Indicator |
| RTT | Round Trip Time |
| SAE | System Architecture Evolution |
| SC-FDMA | Single Carrier- Frequency Division Multiple Access |
| S-GW | Serving-Gateway |
| SIM | Subscriber Identification Module |
| SINR | Signal to Interference plus Noise Ratio |
| TCP | Transmission Control Protocol |



| | |
|--------------|--|
| TDD | Time Division Duplex |
| TEP | Telecom Expansion Project |
| UE | User Equipment |
| UL | Upload |
| UMTS | Universal Mobile Telecommunications System |
| WCDMA | Wideband Code Division Multiple Access |
| WLAN | Wireless Local Area Network |
| WT | Walk Test |



1. Introduction

Mobile networks are used dominantly in telecommunication network systems where the number of customers increased dramatically over time. For several years, voice calls dominated the traffic in mobile telecommunication networks. As stated in [1], the growth of mobile data was initially slow, but within the year 2010, its use began to increase dramatically. The increase in user mobile data demand and the required system capacity are the deriving factors of the 4G mobile network system.

Generally, the driving forces for the technology evolution in a mobile network are services that need mobile Internet with improved bandwidth, spectrum efficiency, lower latency, higher throughput, and capacity. LTE, LTE-Advanced, and 4G+ primarily focus on the provisioning of mobile broadband services with the best spectral efficiency. In indoor environments, the through-wall propagation loss makes difficult to provide mobile services with the required quality.

In Ethiopia, mobile data service usage has shown considerable growth thanks to the various network expansion projects done within the last nine years. Mobile telephone service started with a GSM network deployed and operational within the capital Addis Ababa in 1999 with a capacity of 36,000 subscribers [2]. The GSM-based telephone service attracted more than 6,000 subscribers just after its commercial availability.

Since then, the two major telecom network projects carried out in Ethiopia were the Next Generation Network (NGN) project that started in 2007 and the Telecom Expansion Project (TEP) in 2013. While the NGN project introduced third-generation WCDAM in Addis Ababa, the TEP project expanded it and its evolution like HSPA/HSPA+ to the rest of the regions [2]. Introducing 4G LTE in Addis Ababa was also a part of the TEP project. There were also pilot projects for capacity expansion and coverage improvement for UMTS and 4G LTE-Advanced in the areas that demanded high data traffic in Addis Ababa. Then, 4G LTE mobile network technology was deployed in all areas of Addis Ababa, and mobile subscribers climbed to 54.3 million. Although such different projects were deployed, complaints from the customers still existed, especially from users inside the building.

Data usage increases dramatically from time to time because of the high data rate users demand. The operators deploy the newest cellular network technology to satisfy the higher data rate demand of the users. 4G is the best selected mobile technology in most operators with introducing 5G in some countries. Those



mobile technologies were introduced with better spectral efficiency, higher network capacity, higher network coverage, higher throughput, and lower latency related to the previous technologies.

Operators compete with one another to acquire and retain mobile subscribers in their networks. Then, the operator should utilize various measurement methodologies and tools to assess network performance on the network side and Quality of Service (QoS) on the user side. To optimize their networks, the operator uses the data from Network Performance (NP), QoS, and Quality of Experience (QoE) assessments. Parameter setup, license update, network expansion, and new site deployment are the actions for solving network problems based on the evaluation. Following the appropriate actions, the operator can attract additional users and keep the existing subscribers with their available network.

There are different types of mobile network measurement methodologies and tools. Generally, network measurements can be active or passive [3]. Most passive measurements are performed at the core side of the network and measured using inputs from different network interfaces without injecting traffic. This type of measurement is difficult to use the network interfaces by professionals and third-party bodies. The Network Management System (NMS) and the Minimization Drive Test (MDT) are examples of passive measurement. Active measurements are conducted by injecting traffic on the live network, which may bias selected test area conditions such as network congestion and poor network coverage areas. The traditional Drive Test (DT) or Walk Test (WT) and some crowdsourcing (CS) tools are examples of active measurement. There are measurement platforms that include both active and passive measurement types. Active and passive measurement types depend on the measured parameters.

1.1. Statement of the Problem

In Ethiopia, Ethio telecom is the monopoly operator owned by the Government that controlled the telecom network infrastructure and providing services. The Government began procurement to sell 40% of the Ethio telecom share and invited competent international operators. Ethio telecom delivers telecom services with complaints. There is no regulatory pressure to control customer satisfaction and expected data QoS based on measurement data. But the company reacts using customer complaint feedback from the customer call center or public conferences.

Now Ethiopian Government opens the telecom industry to involve other companies to provide Internet service and telecom network infrastructure. The governmental regulatory body should control and manage the interaction between the stakeholders such as telecom operators, Internet Service Providers (ISPs), Service Layer Agreement (SLA) enterprise customers, and end-users. When we come to the measurement

of QoS in the mobile network, the stakeholders may measure using different measurement techniques and tools then the report may differ. Various QoS measurement reports from several organs may follow disagreement and impermissible benchmark reports.

There were assessments of the mobile network coverage using signal conditions as measurement metrics for optimization and benchmark. However, this is not an end-to-end mobile network service assessment that considers the air interface of the network. To assess mobile network service quality end-to-end, we must use service-based measurement metrics in addition to signal conditions. Telecommunication performance and quality measurements are classified into NP, QoS, and QoE [3]. The QoS of a network is measured from the network side and user side.

Several measurement methods and tools have been applied to measure user-side objective data QoS in mobile networks. The data rate and latency metrics are the most important parameters for the QoS measurement of data service. NMS, DT/WT, MDT, and CS are the methods to measure and collect the required parameters for assessment. Even though the measurements are conducted at the same place and time, the varieties of measurement methods and tools provide different results. Those measurement results for the same measurement condition make controversy between telecom stakeholders such as telecom regulatory body, ISP, SLA enterprise users and end-user, and researchers. Then, this needs to be answered, why the different tools provide different results measured at the same place and time on the selected network technology. So, the various output reports measured by different measurement methodologies and tools in similar locations and times should be studied.

The thesis answers the following questions:

1. What are the main factors on user-side QoS measurement?
2. How do the test tools measure the throughput and latency QoS parameters?
3. How do the tools set the value regarding the identified factors?
4. Is the result of the measurement tools different?
5. How much is the impact of QoS measurement factors that evaluate using Mean Absolute Percentage Error (MAPE)?
6. How much is the accuracy of the CS and Walk test tool related to typical throughput and latency reference value using MAPE?
7. How user side crowdsourcing QoS measurement tools applicable to the telecom stakeholders in Ethiopia?

The gap is the evaluation and investigation of the various mobile network QoS measurement factors and tools that provide different results even though the tests are performed at the same location and time.

1.2. Objective of the Thesis

1.2.1. General Objective

The general objective of the thesis is to evaluate 4G LTE user-side data QoS measurement techniques and tools in the case of Addis Abeba, Ethiopia, in Ethio telecom 4G LTE network.

1.2.2. Specific Objectives

The specific objectives of the thesis are listed as below:

- Read different literature reviews related to the topic.
- Identify and discuss types of network measurements.
- List down the QoS measurement possible factors.
- Discuss the layout of general and specific user side data QoS measurement tools.
- Map the factors to the selected measurement tools and identify the value set by the tools.
- Prepare 4G user side data QoS measurement setup.
- Conduct measurement using crowdsourcing technique: RTR-NetTest and Ookla Speedtest tools, and WT technique: Nemo Handy-A using throughput and latency as parameters according to the measurement framework.
- Collect and analyze measured data parameters using Excel, Actix, and Python
- Evaluate the impact of manipulated factors applied by measurement tools to measure throughput and latency using MAPE.
- Show whether the results of various measurement tools are different or not for selected measurement scenarios.
- Evaluate the accuracy of the QoS measurement tools related to the typical throughput and latency reference value.
- Discuss the reason why different measurement tool results differ and recommend the consideration during QoS measurement.



1.3. Literature Review

Some related studies on mobile network performance and QoS were reviewed in different scenarios. The purpose of the review is to read the background knowledge that is the foundation of current findings. Here are discussed some of the related papers.

The objective of the paper [4] was to investigate the quality of the experience of UMTS enterprise data customers. In this paper, the Network Management System (NMS), Nemo Handy-A, and RTR-NetTest measurement tools were used for QoS measurement. Additionally, the survey technique was used for QoE data collection. The measurement was conducted based on download, upload speed, and latency metrics. The authors analyzed the results based on DL speed, with an average value of 0.57 Mbps using NMS, 0.86 Mbps using Nemo Handy, 2.53 Mbps using RTR NetTest, and 2.67 the MOS using survey. The authors concluded that the QoS of the network contributes to the QoE measurement results. From the paper, we can learn that the average values of different measurement tools for a given network are different.

In the paper [5], user-side QoS measurement techniques were investigated and compared. The paper used Nemo Handy, Speedtest, and OpenSignal measurement tools based on download and latency parameters. The measurement was conducted on an LTE mobile network and the RMSE was used as an evaluation metric. In this paper, they compared the CS measurement with Nemo Handy one by one. The evaluated download results of OpenSignal and Nemo Handy using RMSE are 20.89 and 15.88, respectively, whereas the evaluated download results of Speedtest and Nemo Handy are 8.7 and 15.54 correspondingly. The evaluated latency values of OpenSignal and Nemo Handy using RMSE are 84.64 and 8.55 correspondingly, whereas the evaluated download results of Speedtest and Nemo Handy are 6.85 and 3.75, respectively. We can say that the measurements of Nemo Handy and the CS tools are different. The CS measurement tools were not compared to each other directly.

The objective of the paper [6] was to compare different measurement approaches to find the optimal solution for the quality assessment of Internet access services on mobile networks. The Ookla Speedtest, RTR-NetTest, and Visualware Inc. online measurement tools and Latvian regulatory authority SPRK Internet access QoS measurement system ITEST on three operator networks used data rates and latency as metrics. The average measured latency variations on the mobile network are 180, 46, and 22 in msec for 2G, 3G, and 4G sequentially. The actual measured transmission speed depends on 3G or 4G was provided in a given measurement location. The paper concluded that the proposed measurement methodology applied the same measurement tool, dedicated reference point, and network termination point



to all measurements, allowing for reliable and comparable information about mobile broadband Internet access quality given to consumers. The paper focused on the QoS assessment of the operator's network.

The authors in [7] deal with opportunities and challenges in using CS QoS measurements from smartphones for benchmarking mobile networks. OOKLA SpeedTest, FLOQ, Kyago, Traffic Monitor, and RTR-NetTest using data rate, delay, and data volume as metrics with 100 samples. The analysis was conducted based on the RTR-NetTest tool and DL speed parameter. The RTR-NetTest reported result was 32.391 Mbps, whereas the goodput result was 33.0963 Mbps, which shows that different estimation algorithms yield different results. The paper was concluded that to achieve accurate and fair benchmarking using the CS measurement tools, the post-processing stage must be designed carefully. Although the tests were conducted on the same device using the same operator's technology, there was a significant variance among the measurements. The study only focused on the CS tools, and the test was not taken at the same time as the analysis was done on RTR-NetTest.

The objective of the paper [8] was to explain the broadband speed measurements. ComScore, Speedtest, Akamai, and YouTube test tools are used with data speed as measurement metrics to analyze the factors for different results. Speedtest, ComScore, Akamai, and YouTube have average download throughputs of 7.93 Mbps, 3.4 Mbps, 3.9 Mbps, and 4.21 Mbps, respectively. Although the measurements were conducted in small geographic regions, the results varied significantly. The paper concluded that the main factors for this difference in the results were the proximity of the test server, the number of parallel TCP connections, and the way sample measurement was filtered.

1.4. Methodology

The methodology begins with a review of the literature related to the selected topic area. For the thesis, the data service chosen from several services supported by the 4G LTE mobile network, such as voice calls, video calls, data services, multimedia, online games, etc. To measure 4G LTE mobile network QoS measurement on data service, throughput and latency are selected as the most relevant data QoS metrics from other minimum QoS measurement parameters. The minimum relevant QoS measurement parameters are throughput, latency, jitter, packet loss, and packet error [3]. Then, the selected measurement techniques and specific tools are used. The first and sixth floors of the Ethio telecom Microwave Office Building found in Addis Ababa are selected as measurement locations.



The selected measurement tools from the number of measurement techniques are Nemo Handy-A from WT technique, and RTR-NetTest and Ookla Speedtest from CS measurement technique. The measurement factors for each tool are investigated and evaluated based on the selected reference value of each parameter. To evaluate the measurement tools, the test be performed in the same location and time. The number of samples taken within the selected test locations are 2358. The same smartphone would be selected based on the appropriate 4G LTE terminal category and the SIM cards are locked in a 4G LTE network. Then, the measurement data is collected and analyzed using the throughput and latency parameters on the 4G LTE network using Excel, Actix, and Python. Finally, analyzed and discussed the results of measurement tools comparatively based on measurement factors.

1.5. Scope and Limitation

1.5.1. Scope

The main scope of the thesis is to study the evaluation of QoS measurement techniques and tools using 4G LTE mobile network data services with download speed, upload speed, and latency as the most relevant parameters. The thesis covers the selected test tools and the area of Addis Ababa based on the Ethio telecom 4G LTE mobile network.

1.5.2. Limitation

The limitations of the thesis are:

- The measurement in mobility approach was not considered.
- The category of selected smartphones for WT and CS tools are not the same.

1.6. Contributions

The thesis contributes to the Ethio telecom company as an input why the measurement techniques and tools record different results for the same external conditions. Internet service providers and governmental telecom regulators visualize the QoS measurement factors and evaluation of the measurement tools. Additionally, researchers use it as input to keep knowledge QoS measurement techniques and tools on the mobile network. The thesis helps to use alternative user-side QoS measurement tools confidentially and knowledge-based in telecom stakeholders.



1.7. Thesis Organization

The thesis had been organized into six chapters, including this one. Chapter 2 deals with the overview of the 4G mobile network system. It includes background and architecture of 4G, new enhanced technologies used in 4G LTE, and modified Shannon theorem on maximum throughput. In this chapter discussed the 4G mobile network services and data service QoS measurement parameters. In chapter three, mobile network user side QoS measurement methodologies, categories, and tools are discussed. It deals with factors that affect the data QoS measurement. Chapter 4 contains methodology and data collection that deals with the data sampling, location or area selection, measurement procedures and scenarios, and proposed approach. Chapter 5 covers the results and discussions. Finally, Chapter 6 contains conclusions and recommendations.

2. Overview of 4G Network

2.1. Background

A rapid increase in mobile data usage and the emergence of new applications: Multimedia online gaming, mobile TV, Web 2.0, and streaming content has inspired the Third Generation Partnership Project (3GPP) to work on Long Term Evolution (LTE) on the path to fourth generation mobile system. The evolution was evaluated for interoperability with 3GPP mobile communication systems across a 10-year timeframe and for delivering the high data rates and low latency that future customers would want.

From the years up to 2010, the demand for LTE mobile technology slowed the growth of mobile data substantially. For example, data traffic in Ericsson's global communications networks surged 500 times between 2007 and 2013. As a result, there is a need to expand network capacity. The other cause was that in 2G and 3G mobile technologies, the capacity of a mobile telecommunication system became overloaded after 2010 [1].

$$C = B \log_2(1 + \text{SINR}) \quad (2.1)$$

A communication system can theoretically deliver data without error from the transmitter to the receiver. According to equation (2.1), the data rate is smaller than C, the channel capacity. C is the highest data rate that a cell can support in a mobile communication system, and it equals the total data rate of all the mobiles in the cell. LTE technology is being used to boost data rates by leveraging a bandwidth of up to 20 MHz as shown in Figure 2.1 and SINR-increasing methods.

System capacity can be increase using smaller cells: by adding an extra-base station and reducing the size of each cell, MIMO, by increasing bandwidth. Increased spectrum allocation to 2G and 3G networks has resulted from the increasing use of mobile telecommunications. The radio spectrum, on the other hand, is regulated by the ITU, regional, and national bodies. The third method is to enhance our communication technologies [1].

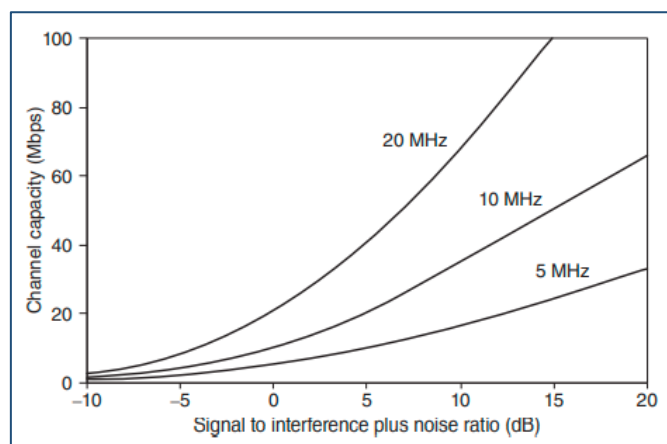


Figure 2.1: Shannon capacity of communication system, in bandwidths of 5, 10, and 20 MHz [1]

Mobile system capacity can be improving by exploiting the higher SINR and greater bandwidth. This continuous advancement in communication technology has been a recurring theme in the evolution of mobile telecommunications, and it is the main reason behind the launch of LTE [1].

LTE stands for long term evolution and was started as a project in 2004 by a telecommunication body known as the 3GPP. LTE evolved from an earlier 3GPP system known as the Universal UMTS, which evolved from the GSM. The Evolved UMTS Terrestrial Radio Access (E-UTRA) and Evolved UMTS Terrestrial Radio Access Network (E-UTRAN) are the two commonly related specifications. Release 8 of the 3GPP specifications contains the first version of LTE.

The ITU launched a set of requirements ITU-R M-2133, ITU-R M-2134, and ITU-R M-2135 for a 4G communication system under IMS-Advanced in 2008. Based on the requirement, the peak data rate is at least 600 Mbps and 270 Mbps on DL and UL using the bandwidth of 40 MHz sequentially. Even though LTE has not met the requirements, LTE-Advanced was met the requirement in 2010 announced [9]. Initially, the ITU intended for the name 4G to be given for systems that met the IMT-Advanced standards. At the end of 2009, the first LTE networks were launched in Norway and Sweden [1]. LTE's primary purpose is to provide high data rates, low latency, and packet-optimized radio access technology that allows for flexible bandwidth deployments.

The main goal of LTE is to provide a high data rate, low latency, and packet-optimized radio access technology supporting flexible bandwidth deployments. Same time its network architecture has been designed to support packet-switched traffic with seamless mobility and better quality of service. LTE was required to support a downlink peak data rate of 100 Mbps and an uplink peak data rate of 50 Mbps. Later, the requirement was exceeded to provide DL and UL peak data rates of 300 Mbps and 75 Mbps

respectively. Latency is another critical issue, particularly for time-critical applications such as voice and interactive games [1] RAN round trip times less than 10ms. The relevant specification suggests a user plane latency of 10 ms for non-roaming mobile, increasing to 50 ms in a typical roaming scenario [1].

LTE downlink modulation is based on OFDM, which allows for multi-user access, radio channel time dispersion resistance, and receiver design simplicity. In the uplink, SC-FDMA is employed because it has a low peak-to-average power ratio [1].

LTE supports both FDD and TDD schemes, resulting in different frame structures in the time domain. In the frequency domain, the different number of Resource Blocks (RBs) address based on the system bandwidth varies from 1.4MHz to 20MHz. A single RB consists of a group of carriers for transferring data. Each RB is 0.5 milliseconds long in the time domain and 180 kHz in the frequency domain. Depending on the cyclic prefix mode and subcarrier spacing, the number of subcarriers and OFDM symbols in downlink (SC-FDMA symbols in uplink direction) is different for each RB [1].

2.2. Architecture of LTE

There are EPC and SAE architectures. The EPC is the direct replacement for the packet-switched domain of UMTS and GSM. The SAE was designed as part of two 3GPP work items that cover the core network and LTE, which include the radio access network, air interface, and mobile [1].

LTE is the next generation of 3GPP network architecture, with E-UTRAN as its air interface. Complementing E-UTRAN in LTE is the System Architecture Evolution (SAE) of the core network that defines EPC (Evolved Packet Core). As shown in Figure 2.2, EPC and E-UTRAN together comprise the Evolved Packet System (EPS). EPC includes the Packet Data Network Gateway (P-GW), Serving Gateway (S-GW), Mobility Management Entity (MME), Home Subscriber Service (HSS), Policy Control and Charging Rules Function (PCRF) [10].

P-GW operates as an IP anchor and connects UE to the external packet data network. P-GW filters the downlink user IP packets and performs policy enforcement. User IP packets are transferred via S-GW and operate as a local mobility anchor when UE moves between eNBs. P-GW can be considered as a mobility anchor between 3GPP and non-3GPP technologies, while S-GW functions as a mobility anchor between other 3GPP technologies. MME is responsible for functionalities like user authentication, bearer management, and idle mode UE tracking UE and core network by Non-Access Stratum (NAS) protocol.

The information on the Packet Data Networks (PDNs) to which the UE can connect is located in HSS. PCRF manages charge functionalities and policy control as well as authorizing QoS [11].

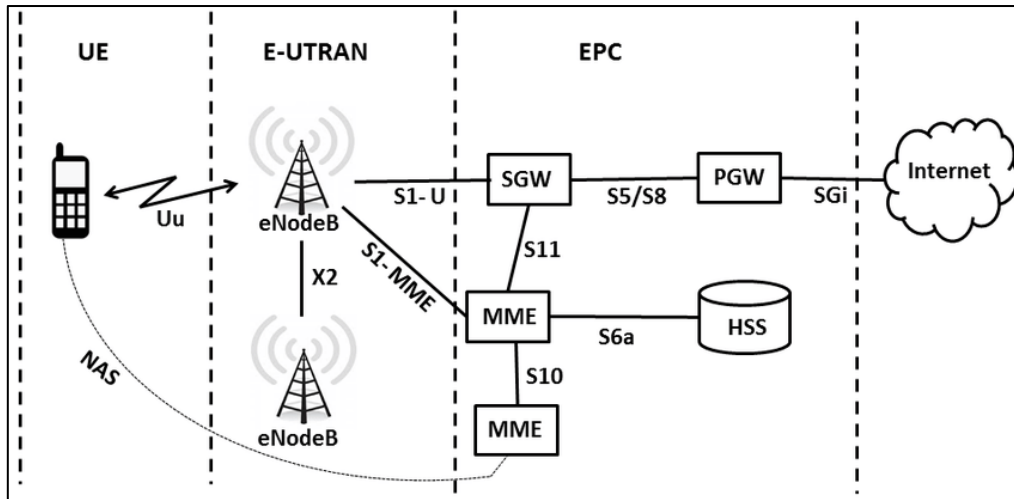


Figure 2.2: EPS architecture [12]

User Equipment (UE) is an access device for the user and provides measurements that indicate channel condition to the network.

E-UTRAN handles the radio communications between the mobile and the evolved packet core and has just one component, called eNB.

Evolved Packet Core (EPC) contains several logical nodes name MME, S-GW, P-GW, PCRF, and HSS. The EPC is responsible for assigning IP addresses to UEs and storing or forwarding packet data between the UE and the external IP network. When a UE turns on, the EPC is in charge of authenticating and establishing the initial connection required for all subsequent communication [13].

Mobility Management Entity (MME) manages the mobile's high-level operation by sending it signaling messages regarding security and data stream management that are unrelated to radio communications. As with the serving gateway, a typical network might contain a handful of MMEs, which looks after a certain geographical region. Each mobile is assigned to a single serving MME, but that can be changed if the mobile moves sufficiently far. The MME also controls the other network elements through signaling messages within the EPC [1].

Serving Gateway (S-GW) acts as a high-level router that transmits data between the base station and the PDN gateway. A typical network might contain a handful of serving gateways, each of which locates after the mobiles in a certain geographical region. Each mobile is assigned to a single serving gateway but can

be changed if the mobile moves far. The S-GW handled overall packet forwarding and routing to and from the UE [13].

PDN Gateway (P-GW) is the EPC's point of contact with the outside world. Through the SGi interface, each PDN gateway exchanges data with one or more external devices or packet data networks, such as the network operator's servers, the Internet, or the IP multimedia subsystem. Each packet data network is identified by an APN [1, 14]. The P-GW is a node that connects a UE to an external PDN such as the Internet and IP Multimedia Subsystem (IMS) [13].

Policy and Charging Rules Function (PCRF) is the most important device which authorizes the treatment that a service data flow will receive by specifying a suitable Policy and Charging Control (PCC) rule [1].

Home Subscriber Server (HSS) is a central database that stores the data of all the network operator's subscribers is a central database that contains information about all the network operator's subscribers [1].

Interfaces: In 4G mobile system architecture new interfaces were introduced. The interfaces are LTE-Uu, S1-U, S1-MME, X2, S5/8, S11, Gx, SGi, and S10. LTE-Uu interface connects the UE, and eNB and S1-U is the interface between S-GW and eNB. S1-MME interface connects MME and eNB. X2 interface connects neighbor eNBs. S5/8 interface connects S-GW and P-GW that S5 and S8 are the same interfaces, the difference being that S8 is used when roaming between different operators while S5 is utilized within the network. S11 interface connects MME and S-GW while Gx connects P-GW to PCRF and provides transfer of QoS policy and charging rules from PCRF to Policy and Charging Enforcement Function (PCEF) in the P-GW. SGi interface connects P-GW to external P-GW and Internet while S10 controls interfaces between two MMEs.

2.3. QoS Classes in 4G LTE

The LTE architecture supports forced QoS with end-to-end quality of service and guaranteed bit rate (GBR) for radio bearers. In 4G LTE, QoS is implemented at the EPS bearer level. It uses EPS bearers with each assigned an EPS bearer ID applied at the level of EPS bearer. EPS bearers are used by LTE to carry data from one section of the system to another [15, 16]. The EPS bearer can be viewed as a bi-directional data pipe that transports data over the network via the best route and with the best QoS. There is a one-to-one mapping, which means that for an EPS Bearer, there is a corresponding EPS Radio Access Bearer (RAB), an S1 bearer, and a radio bearer. If the S5/8 interface depends on GTP, the bearer runs between

the mobile and the P-GW. If the S5/S8 interface depends on PMIP, the bearer runs between the mobile and the S-GW [1].

There are four types of EPS bearers: GBR Bearer, Non-GBR Bearer, Dedicated Bearer, and Default Bearer. Admission control permanently allocates resources to GBR Bearers, whereas non-GBR Bearers do not have admission control. Dedicated bearer associated with specific TFT (GBR or non-GBR) and default bearer Non-GBR, a catch-all for unassigned traffic.

The EPS bearer service architecture is as shown in Figure 2.3.

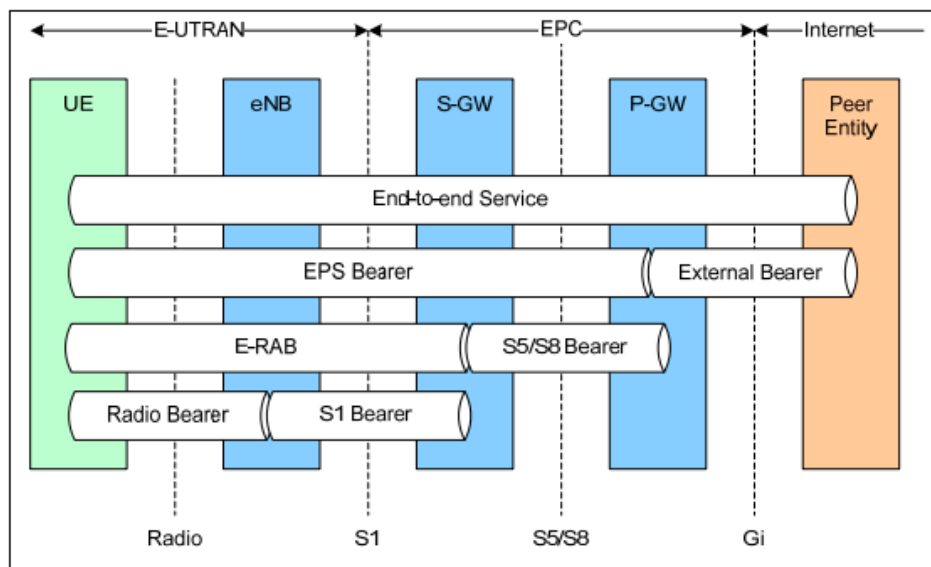


Figure 2.3: EPS bearer architecture [17]

4G Supports differentiated QoS for a large diversity of users and services. LTE networks use a class-based quality of service. The network assigns each bearer type a unique QoS Class Identifier (QCI). In 3GPP LTE networks, QCI is a method that ensures bearer traffic is assigned the appropriate QoS. Various bearer traffic has different QoS requirements and needs different QCI values to be sent. Examples of the QoS parameters include GBR or None-GBR, priority handling, packet delay budget, and packet error loss rate. This overall mechanism is called quality class identifier. In terms of packet forwarding treatment that bearer traffic receives edge-to-edge between the UE and the P-GW, nine QCIs in Release 8 (thirteen QCIs in Release 12, fifteen QCIs in Release 14) are standardized and associated with CQI characteristics according to 3GPP TS 23.203.

Standardized QCIs specified in 3GPP specifications (3GPP 23.203) are shown in Table 2.1

Table 2.1: QoS classes in LTE Release 8

| QCI | Bearer Type | Priority | Packet Delay | PELR | Example | |
|-----|-------------|----------|--------------|-----------|--|--|
| 1 | GBR | 2 | 100 ms | 10^{-2} | Conversational voice (VoIP) | |
| 2 | | 4 | 150 ms | 10^{-3} | Conversational video (live streaming) | |
| 3 | | 3 | 50 ms | 10^{-3} | Real-time games | |
| 4 | | 5 | 300 ms | 10^{-6} | Non-conversational video | |
| 5 | Non-GBR | 1 | 100 ms | 10^{-6} | IMS Signaling | |
| 6 | | 6 | 300 ms | 10^{-6} | Video (buffered streaming), Web, email, ftp. | |
| 7 | | 7 | 100 ms | 10^{-3} | Voice, video (live streaming), interactive games | |
| 8 | | 8 | 8 | 300 ms | 10^{-6} | Video (buffered streaming), Web, email, FTP etc. |
| 9 | | | 9 | | | |

2.4. New Enhanced Technologies in 4G LTE

Multi-Input Multi-Output (MIMO): A MIMO channel provides multiple spatial paths between the transmitter and receiver [18]. The MIMO approach is used in most modern wireless networks to enhance the data rate to the user and support better coverage at the cell edge. To allow the base station to deliver multiple data streams over the same carrier at the same time, all LTE devices must support MIMO transmissions.

Beamforming: It is an RF management and signal processing method in which an access point sends out the same signal via many antennas. In the downlink, LTE uses beamforming to keep away from interference and improve the received signal-to-interference ratio. It can provide significant improvements in coverage range, capacity, reliability, and battery life [13]. Beamforming is similar to the diversity transmission, but has a different geometry and uses various techniques. The base station still has multiple antennas, but their separation is less than before, with typically half a wavelength of the carrier frequency [1].

For radio transmission and reception, LTE was designed for the base station and mobile to use multiple antennas. By employing numerous antennas at the transmitter, receiver, or both, the diversity processing maximizes the received signal power and minimizing the fading level. To enhance the data rate, spatial

multiplexing (MIMO antennas) is used by the transmitter and receiver. Multiple base stations are used in beamforming to increase cell coverage [1].

Carrier Aggregation (CA): LTE-Advanced as of 3GPP Release 10 allows the aggregation of a maximum of five component carriers with up to 20 MHz bandwidth each to attain a total transmission bandwidth of up to 100 MHz. As none of these service providers owns a continuous spectrum of 100 MHz, three different methods of carrier aggregation exist within LTE-Advanced: intra-band contiguous, intra-band non-contiguous, and inter-band carrier aggregation. Intra-band describes the aggregation of component carriers within the same frequency band in a contiguous or non-contiguous way. For inter-band carrier aggregation, the two-component carriers reside in different frequency bands.

Higher Bandwidth: The bandwidth occupies a certain amount of frequency spectrum. If we use wide bandwidth, the data rate will be higher, and LTE can use up to 20 MHz.

Orthogonal Frequency Division Multiplexing (OFDM): is both modulation and multiple access schemes, which optimal for high-rate data transmissions over wideband channels. OFDM is a multi-carrier modulation system, as opposed to single-carrier modulation schemes [13]. This approach is used in the 4G LTE downlink. LTE takes advantage of OFDMA, which is a multi-carrier scheme that allocates radio resources to multiple users based on frequency and time using Orthogonal OFDM. OFDMA allows a network to flexibly assign bandwidth to a user based on bandwidth needs and fees paid on the user's data plan. Unassigned subcarriers are switched off, thus reducing power consumption and interference.

Single Carrier Frequency Domain Multiple Access (SC-FDMA): The scheme provides orthogonal access to multiple users simultaneously accessing the system. It is low signal peakiness due to the single carrier transmission property and avoids intra-cell interference in the uplink. The uplink access is used on SC-FDMA that promises increased uplink coverage due to low peak-to-average power ratio (PAPR) relative to OFDMA. Lower Peak-to-Average Power Ratio for SC-FDMA-based power amplifiers results in extended battery life along with improved uplink performance. In SC-FDMA, data is spreading across multiple subcarriers. SC-FDMA differs from OFDMA, where each subcarrier transports unique data. SC-FDMA is unacceptable for the downlink due to the need for a complex receiver and size and processing power limitations in a wireless device

2.5. Modified Shannon Theory on Throughput

The Modified Shannon Channel capacity theorem defines maximum throughput for a noisy channel as shown in Equation (2.2) [19, 20].

$$TP = \begin{cases} N_{PRB} BW_{PRB} \min (S_{max}, BW_{eff} \log_2 (1 + SINR / SINR_{eff})), & SINR \geq SINR_{min} \\ 0, & SINR < SINR_{min} \end{cases} \quad (2.2)$$

Where; N_{PRB} is number of PRB, TP is throughput in (bps), BW_{PRB} is bandwidth per PRB, BW_{eff} is adjusted BW efficiency, S_{max} is maximum spectral efficiency, $SINR_{min}$ is the minimum required SINR, and $SINR_{eff}$ is adjust for SINR implementation efficiency. The terms included in equation (2.2) have impacts on the throughput QoS measurement result that depends on the location and time during measurement.

2.6. 4G Network Services and Parameters

4G Network Services: LTE is designed to seamlessly support various services, such as interactive multimedia, voice, video, broadband data, and multicast video [21]. End-to-end QoS classes are used to prioritize the various services.

Reference Signal Received Power (RSRP): It's a signal strength-related metric measure for each cell that's employed as a factor in cell reselection and handover decisions. For a particular cell, RSRP is defined as the average power of the Resource Elements (REs) that carry cell-specific Reference Signals (RSs) within the considered bandwidth [22]. It is measured for one subcarrier.

Received Signal Strength Indicator (RSSI): Within the measurement bandwidth over NRBs, it is the linear average of the total received power detected exclusively in OFDM symbols carrying reference symbols by UE from all sources, including co-channel non-serving and serving cells, adjacent channel interference, and thermal noise [22, 23].

Reference Signal Received Quality (RSRQ): It is a cell-specific signal quality metric. RSRQ is using mainly to provide ranking among different candidate cells by their signal quality similar to RSRP. It is employed as an input in making cell reselection and handover decisions in scenarios in which the RSRP measurements are insufficient to decide reliable handover. The RSRQ can be calculated as Equation (2.3) [20]:

$$RSRQ = \frac{N \cdot RSRP}{LTE \text{ Carrier RSSI}} \quad (2.3)$$

Where; N is the number of RBs of the LTE carrier RSSI measurement bandwidth.

Signal to Interference plus Noise Ratio (SINR): It is calculated as the ratio of signal power to the total of average interference power from other cells plus background noise [24]. UE computes SINR on each RB, converts it to Channel Quality Indicator (CQI), and reports the CQI to eNB, where it is used to select the most suitable Modulation and Coding Scheme (MCS) for user data transmission in particular RB. SINR value defines the MCS to be used for an RB. MCS decides the number of bits per modulation symbol to be sent to the user [25].

Channel Quality Indicator (CQI): effectively represents the modulation and coding that UE can support in the current RF condition utilized in DL. It is viable to estimate the DL channel quality and dynamic scheduling on the DL.

To ensure optimized network performance, select the most important metrics for measurement. The data speed and data quality can significantly impact end-user experience and are influenced by several factors [24].

Bandwidth: is a metric for determining the data amount that can be sent in a given amount of time. The smallest unit of radio resource allocated to a user for data transmission is Physical Resource Blocks (PRBs). Among these RRM functions, packet scheduling is the process by which available radio resources are allocated among active users to satisfy their QoS requirements [26].

The main targets of packet scheduling are to maximize the cell capacity and satisfy the minimum QoS needed for connections. Additionally, maintain adequate resources for best-effort users with no strict QoS requirements is among the main objectives of packet scheduling [27]. Additionally, throughput, latency, jitter, packet loss, and packet loss ratio are data service QoS measurement parameters. Those terms will be discussed in chapter three QoS parameters section.

3. Data QoS Measurement Methodology and Tools

3.1. Network Measurement Types

The three main parameters to characterize the actual performance of an IP network are delay, packet loss, and transmission capacity. From the user’s point of view, the throughput is a key performance measure for a dedicated service perceived by the user on the application level. There is a notable variation between network and application throughput because of overheads, which allows statements about network performance (NP) and QoS subsequently.

Quality and performance in the telecommunications field can be divided into several levels, according to Figure 3.1. Those are the classifications of NP, QoS, and QoE according to ITU-T [3].

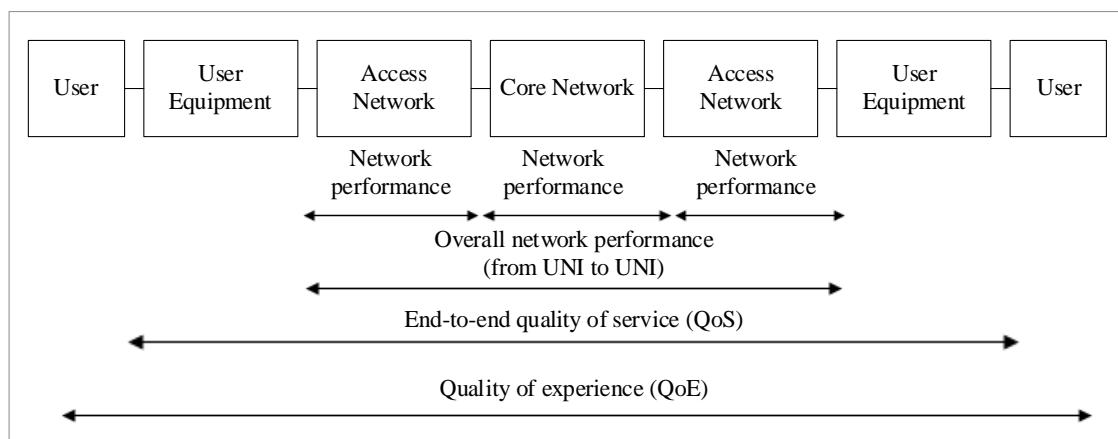


Figure 3.1: Classification of NP, QoS, and QoE [3]

3.1.1. Network Performance

It is the performance of a network element. Typically, the NP is determined by assessing several technical parameters that describe the performance of a network element regarding some particular standard. Examples of network performance parameters are Bit Error Ratio (BER), transmission delay, processing time, and jitter.

3.1.2. Overall Network Performance

The performance of all elements is integrated to form a network and to mention an example, the performance measurements between two user network interfaces.

3.1.3. End to End Quality of Service

Overall performance of the elements and processes are involved in the chain of transmission of a telecommunication service until the service is performed. QoS is measured from the point of view of the user but does not involve the user in the process of measurement and assessment carried out objectively. 4G LTE is a critical technology primarily used to handle higher data rate demand in network traffic. When deploying fixed broadband for multiple data and voice services, QoS has become an essential aspect of network planning and design in LTE. Different service quality may require for different types of services or applications [1].

In mobile networks, QoS is described as a cellular service provider's capacity to deliver a satisfying service that includes voice quality, signal strength, low call blocking and dropping probability, high data speeds for multimedia, data applications, and others. Throughput, latency, packet loss, and jitter are examples of objective QoS measuring parameters. It uses metrics like the guaranteed data rate, maximum error rate, and maximum latency to characterize the performance of a data stream. LTE can provide QoS guarantees and assign different qualities of service to various data streams and users [1].

Throughput, latency, fairness, and bandwidth usage metrics are applied to categorize QoS needs. The traffic characteristics are evaluated to identify the parameters that cause network performance deterioration. The purpose of the QoS provision is to ensure that the user's request is met with quality policies that meet the user's QoS requirements. Network traffic conditions are forecast to ensure, which the requirement-based service is delivering with the fulfillment of required network resources [1].

The main challenge for mobile network operators is to provide high-performance multimedia services. 4G LTE technology was designed to fulfill the needs of users while also ensuring high network performance. The use of KPIs is required to monitor and optimize network performance. KPIs are used to monitor the QoS offered and achieved resource utilization. Accessibility, retainability, mobility, integrity, and availability are among the subcategories of these factors [28].

ITU-T Rec. E.800 defines the quality of service as:

“Totality of characteristics of a telecommunications service that bear on its ability to satisfy stated and implied needs of the user of the service.”

According to ITU Recommendation G.1000, there are four viewpoints of QoS as shown in Figure 3.2.

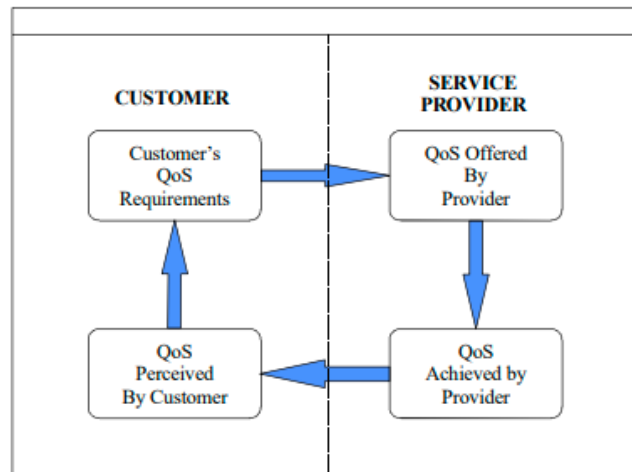


Figure 3.2: Four viewpoints on QoS [3]

The four viewpoints of QoS explain that QoS requirements by the customer state the level of quality required of a particular service, which may be expressed in non-technical language. QoS offered by the service provider is a statement of the level of quality expected to be offered to the customer by the service provider. The quality level is expressed by the values assigned to QoS parameters. The principal use of this form of QoS is for planning and service level agreements. QoS achieved by the service provider is a statement of the level of quality achieved and delivered to the customer. It is expressed by values assigned to parameters, which should be the same as specified for the offered QoS. Then, the two QoS can compare to determine what to achieve to assess the level of performance achieved. QoS perceived by the users is a statement expressing the level of quality they believe they have experienced. Perceived QoS is usually expressed in terms of degrees of satisfaction instead of technical terms. Perceived QoS is assessed by customer surveys and from customer's comments on levels of service.

The QoS of the Internet advertised by ISPs is tested: by ISPs for optimization purposes or by regulators because sometimes the QoS promised is different from the QoS delivered by ISPs. The governmental regulatory body expects to monitor the QoS and QoE of operators and ISPs. Then, the QoS measurement methodologies should be evaluated and select the tool as a benchmark for all stakeholders.

3.1.4. Quality of Experience

It is the overall performance of a telecommunication service by involving users in the process of measurement. QoE assessment is more focused on the subjective part because it is affected by the experience and level of expectations of users of a telecommunications service.

3.2. Network Measurement Categories

3.2.1. Active and Passive Measurement

Active methods are measured using injected traffic that performs analysis based on sending probe packets between the testing client and server. An active measurement generates specific traffic on the network under test and is commonly used on access levels. Probing packets are injected into the network connection to measure the QoS of different services like web brows, file transfer, VoIP, and other over Internet connections. A passive measurement captures real-time user traffic at dedicated points in the network. In other words, Sniffs traffic is routed through a device and performs analysis based on monitoring network traffic between two destinations.

A passive test is a form of background testing that monitors the network usage of other applications on the mobile device itself. It can also measure the network usage of the device overall. The functions required to perform the passive test are not permitted to run on Apple's iOS platform, so the resulting data is limited to Android devices [29].

Passive testing's high benefit comes from its ability to determine the performance of the apps and services a person uses on their device. Passive test results are most indicative of the Internet experience a person receives when using a specific app or service. Commonly, active throughput measurements are performed on the access network side, generating traffic to saturate the available transmission capacity. Passive throughput measurements, on the other hand, are more common in the core network.

Active measurement is mostly adopted methodology by regulators, to measure the QoS of broadband Internet services provided by ISPs, and it allows regulators to produce a benchmarking report regularly. Passive test results are most indicative of the Internet experience a person receives when using a specific app or service, such as a video streaming app. Given that passive testing at its very nature runs in the background and is not a dedicated, standalone test, any data collected through background testing does not and cannot measure the throughput fully on an Internet connection.

3.2.2. Foreground and Background Measurements

Background testing applications, once installed on a device, run background tests periodically whenever the device is on. The tests run at any place and at all times. At the very least, background testing generates significant amounts of data because testing is continuously occurring. In practice, apps or vendors who do not have a large enough user base to rely on consumer-initiated testing results for their analysis alone

typically utilize background testing [29]. Foreground measurements are interactive with the users and measurement is initiated and conducted by the measurement tool users.

3.2.3. Best Effort and the Windowed Approach

The best-effort approach is the best practice method to achieve a maximum of measurement samples for a fixed test file size within a given time. The amount of data per transfer is constant without considering the re-transmission, but the transfer time may or may not be the same for each sample. In the windowed approach, all transfers last for the same duration (the configured window size). As stated in [30], the windowed measurement approach is best suited to achieve a more regular type of sampling for network performance. The best-effort measurement approach is common and best practice for assessing end-to-end quality of service parameters.

3.2.4. Network and User Side Measurement

There are network-side and user-side network measurement types. NMS and MDT measurement methods are network-side measurements, whereas DT and CS measurement methods are user-side measurements.

3.3. Network Measurement Methodologies

Measurement of Internet performance continues to be meaningful for both consumers and service providers. Different measurement methodologies are used in network measurement. However, the various measurement methodologies frequently yield different results. This can confuse consumers and service providers [29].

1. **Consumer-Initiated:** A person intentionally starts consumer-initiated tests. The measurements occur at the time and in the same place recorded for each test sample that is most relevant to the person taking the test. Consumer-initiated testing gives people the flexibility to initiate a test at any time and in any place, while in one place or on the move. Ookla measures more than just throughput; it also measures latency, packet loss, jitter, errors, and other network conditions. The disadvantages are that the needs are many with the geographically dispersed testing server, and users are not randomized. The test volume must be large enough to provide a statistically significant across the desired dimension. It needs skilled and Internet-experienced users.
2. **Drive Testing:** It was initially measured call quality and signal performance for car phone users, later adopted to measure cellular Internet performance. Currently, many mobile network providers are used to measure the strength of their cellular network. The disadvantages of this method are that the test

data is only collected in areas where cars can drive and requires a lot of time and resources to test the whole area. In addition to this, the tested areas may not re-test and cannot measure how connectivity changes throughout the day. It uses similar hardware throughout the tests and has concerns about gaming test results when it comes to driving testing.

3. **Background Testing:** Once the application is installed on the device, run periodic background tests whenever the device is on. The tests run at any place and at all times. The cons of these methodologies are background testing generates a significant amount of data because testing is continuously occurring, the random timing of the test.
4. **Passive Testing:** A form of background testing that monitors the network usage of other applications on the mobile device itself. The results are most indicative of the Internet experience a person receives when using a specific application or service like a video streaming application. The disadvantage of this measurement methodology is that any data collected through background testing cannot measure the maximum throughput of an Internet connection and is not functional on Apple's iOS platform, so the resulting data is limited to android devices.

Testing tools are categorized as hardware and software-based. Hardware-based tools completely replace the end user's equipment in the absence of the users or share the Internet access with normal traffic, while software-based tools can be web-based or dedicated software clients [31].

There are different network measurement methods and tools as listed in Table 3.1.

Table 3.1: Different network measurement techniques

| Measurement methods | Some Measurement Tools |
|---------------------|---|
| NMS | OMC, EMS, PRS |
| MDT | Nastar, Omstar, TEOCO |
| DT/WT | Nemo Outdoor, TEMS & Nemo Handy-A |
| CS | Ookla Speedtest, RTR-NetTest & OpenSignal |

We will discuss each measurement technique as follows:

A. The Network Management System

A network management system is used to control the system and its resources by controlling its usage, access monitoring, and reporting its current and historical status. This helps in judgment, forecasting, decision-making, analyzing data, and making positive efforts to maintain the quality of service.



The network management system covers functions referred to as FCAPS, which stands for fault, configuration, account, performance, and security management, specified by OSI. Fault management provides a fault monitoring function for a mobile network, whereas configuration management includes configuration information management and setting changes of network devices. The account management facilities enable charges to be established for the use of managed objects and costs to be identified for the use of those managed objects. If any threshold is exceeded, the performance management monitors the traffic and wireless resource status of network devices and notifies the user with information about the fault. Security management controls access to resources managed by the network management system. The Element Management System (EMS) is part of the NMS that was developed for specific parts of the network like the core network, base station, and transmission [32, 33].

Performance Surveillance (PRS) is a platform for analyzing mobile network performance, customizing reports, and displaying reports (i.e. packet loss, voice quality indicators, KPIs, counters, and so on). The PRS applies to routine maintenance of mobile networks and it is part of NMS. It can monitor and analyze the overall network performance [34].

B. Drive Test

Drive testing is a method of measuring and assessing the coverage, capacity, and quality of a network. The drive test method records a wide variety of physical and virtual metrics on the mobile network service in a given geographical area [35].

A primary limitation for drive testing is that the method still relies on driving lab test vehicles around vast geographic areas. Even though drive testing only covers areas with roads, it still requires a lot of time and resources. Instead, only a few tests are always taken in a given location and specific roadways which are not re-tested frequently. Additionally, drive testing cannot measure how connectivity changes throughout the day. Test vehicles drive through an area at a specific time, but not all areas at all times of the day.

Telecom operators use drive or walk test methods as basic tools to perform wireless network drive testing, optimization, benchmarking, monitoring, and analysis. This method is used mainly for network benchmarking, optimization, troubleshooting, and service quality monitoring. Some of the tools are Nemo Outdoor, Nemo Handy-A, owned by Keysight Technology, Test Mobile System (TEMS), developed by Ericsson in 1999, and GENEX Probe, owned by Huawei vendor.

C. Minimization Drive Test

MDT is a feature introduced in 3GPP Release 10 that enables operators to utilize users' equipment to collect radio measurements and associated location information, to assess network performance while reducing the OPEX related to traditional drive tests [36]. The collected MDT measurements and geographical location information are utilized to create a map of user throughput throughout the operator's network. This map may be used for the detection of critical QoS problems and the initial analysis of corrective actions.

To collect user throughput measurements with conventional drive tests, typically dummy traffic is generated (e.g., FTP file download, voice, streaming), and the concerned throughput measurements are recorded by the test terminal for later collection and analysis by the operator. In contrast, the throughput measurements collected by MDT functionality are recorded on the RAN side based on real traffic, together with location information reported from the concerned UE. The MDT enables operators to collect measurements indicating the network coverage condition and the users' real-life throughput and connectivity issues. This results in a plethora of applications and use cases that enable network optimization without the need for traditional drive tests. Some examples of minimization drive test platforms are Nastar, Omstar, and TEOCO.

Nastar is one of the MDT platforms, and the solution utilizes the user's equipment to collect field measurements, including radio measurements and location information. The *Nastar* performance analysis system is an intelligent and integrated tool developed by Huawei Technologies. It allows locating and analyzing wireless network quality problems and is applicable for GSM, GPRS, EDGE, UMTS, and LTE networks [37].

TEOCO is the other MDT platforms, and Ethio telecom purchased a new MDT tool that is on the way to implementation in the company. *TEOCO* Mentor solution provides a fully integrated tool allowing communication service providers (CSPs) to face the challenges of radio access networks by maximizing multi-technology network (2G, 3G, 4G) quality, capacity and coverage automatically. Those mobile measurements are collected from the networks [38].

D. Crowdsourcing

Crowdsourcing is a novel paradigm that has applications in a plethora of disciplines. It is an emerging pattern in innovation, problem-solving, and knowledge acquisition, in which tasks are accomplished as a usually large and diverse crowd. In the field of telecommunications, crowdsourcing makes a viable addition to state-of-the-art benchmarking methods for mobile networks, when combined with mobile

sensing approaches to employ smartphones as end nodes [39]. The two issues in conventional mobile testing and measurement are:

- Higher test costs due to the diversity of mobile devices and platforms.
- Difficulty in conducting large-scale user-oriented performance and usability testing.

A new testing approach, known as crowdsourced testing, provides a promising way to address these challenges and issues.

Crowdsourcing test tools are free tools downloaded from Google Play or Apple Store. Some available tools are Ookla Speedtest, RTR-Nettest, OpenSignal, Netradar, and Info Lite. The OpenSignal collects and reports DL and UL speed test, video test (SD or HD), and latency with record whether indoor or outdoor environment and Network availability percentage. Network Cell Info Lite measures the RSSI in dBm with network information, RCSP, LTE RSRP, and LTE RSRQ and sharing this information with their network operator. Netradar is a client-server system that gathers measurements from users around the globe [40].

3.4. Description of Data Service QoS Parameters

The relevance and minimum of technical parameters for the evaluation of typical data QoS measurements are upload throughput, download throughput, latency, jitter, packet loss, and packet error [41]. Their descriptions are:

Throughput: It is the amount of data or number of data packets that can delivered in a predefined period.

Latency: It is simply the amount of time takes for data to travel from one defined location to another. Ideally, the latency of a network is as close to zero as possible [42]. The LTE radio interface and network provide capabilities for less than 10 ms latency from the transmission of a packet from the network to the UE.

Jitter: It is the variation in time delay for the data packets sent over a network. This variable represents an identified disruption in the normal sequencing of data packets. Because jitter causes increased or uneven latency between data packets, it can disrupt network performance, resulting in packet loss and network congestion. Although some level of jitter is to be expected and usually tolerated, quantifying network jitter is an important aspect of comprehensive network performance measurement [42].

Packet Loss: It refers to the number of packets transmitted from one destination to another that fail to transmit. Packet loss can be measured by capturing traffic data on both ends, then identifying missing packets and re-transmission of packets. Packet loss can be caused by network congestion, router performance, and software issues, among other factors [42].

Packet Loss Ratio (PLR): The improvement of PLR estimation is a critical issue, because the packet loss ratio has a high effect on the network performance, especially when dealing with real-time traffic such as VoIP and video flows [18]. The PLR increased during the video flows by users number [43].

3.5. User Side Data Service End-to-end QoS Measurement

A number of the well-known speed testing tools or websites measure speed and some of the issues that arise, and may one prefer one approach to another. The main differentiators include: How close is the test server likely to be to the user? How many TCP connections are used to run the test? And how are sample measurements filtered? The differences in those approaches can have a significant impact on the resulting speed measured [3].

QoS parameters are regularly measured from the network side and user side. Here are user side QoS mobile network measurement tools selected from crowdsourcing and drive test measurement methodologies. Those are drive test and crowdsourcing methods, naming some of the tools as Nemo Handy-A, Ookla Speedtest, and RTR-NetTest.

3.6. QoS Measurement Methodologies and Tools

The general smartphone-based QoS measurement approach is as shown in Figure 3.3.

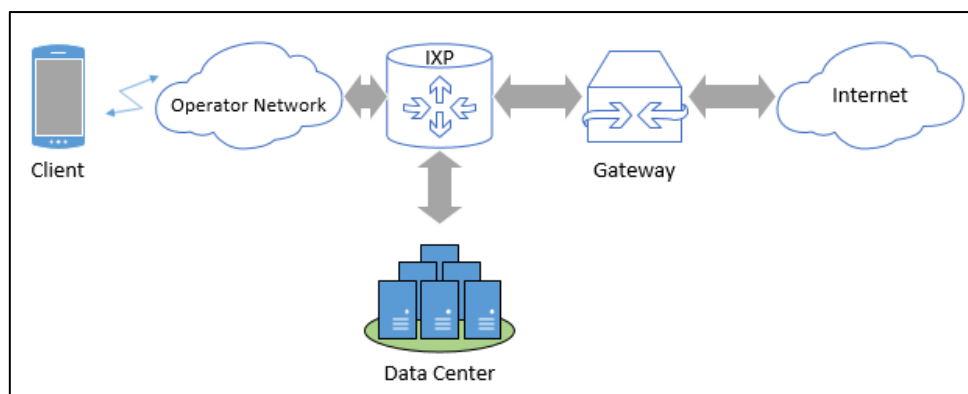


Figure 3.3: QoS measurement approach

A. Ookla Speedtest

ISP and telecom veterans who recognized the pressing need for transparent, high-quality data on Internet access founded Ookla in 2006. The Ookla Speedtest data is based on the test software and measurement infrastructure developed by Ookla Net Metrics. The company behind Speedtest works in fixed broadband and mobile network testing applications, data and analysis, a developer and vendor of networking testing applications [44]. The user interface and measurement approach of the Ookla Speedtest are as shown in Figure 3.4.

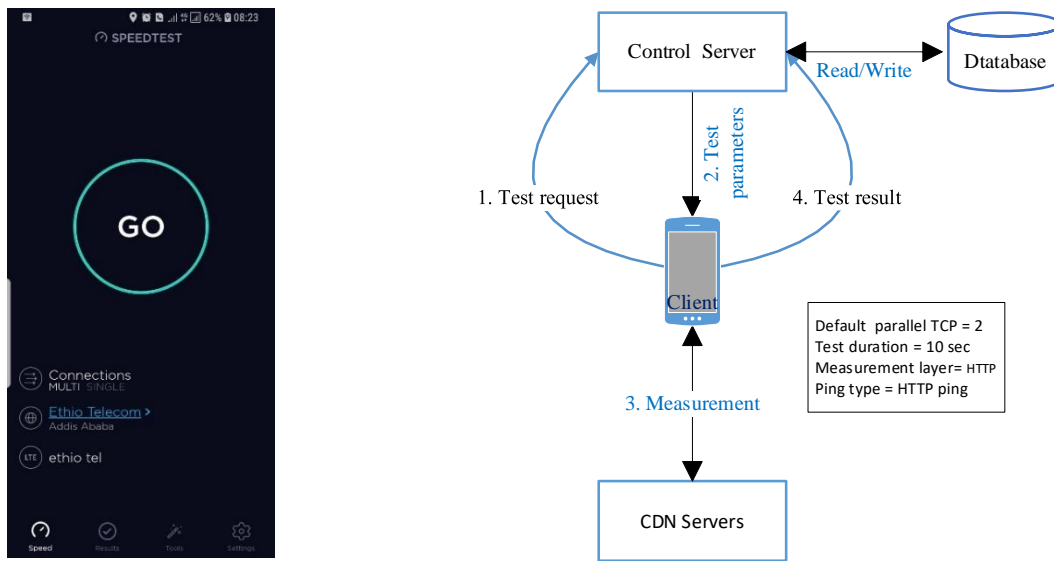


Figure 3.4: Ookla Speedtest QoS measurement approach

To select the measurement server, the control server selects the five nearest servers out of more than 15,000 servers, pings the five servers, and selects the servers with the lowest result in the Ookla Speedtest mobile apps.

Throughput Estimation of Ookla Speedtest: For download, the samples are aggregated into 20 slices that each slices 5% of the samples. By removing 30% of the slowest and 10% of the fastest slices and averaging the remaining slices. For upload, the fastest half samples averaged [8]. For each sample, R is calculated as Equation (3.1):

$$R = \frac{D}{RTT} \tag{3.1}$$

Where D represents the fixed chunk size and RTT represents the round trip time.

If the connection speed is at least 4Mbps, then Ookla uses four threads, unless it will uses the default two threads. The latency test measures the time it takes to set response to an HTTP request sent to the web server. This test is repeated multiple times (six times in Ookla) and the lowest value determines the final result for a specific sample test.

B. RTR-NetTest

RTR-NetTest is a crowdsourcing measurement system for fixed and mobile broadband of Austria’s national regulatory authority for broadcasting and telecommunication [45]. The RTR-NetTest informs users about the current service quality such as upload, download, ping, signal strength for their Internet connection. In addition, a map view and statistics of previous tests can be accessed. It measures Internet network layer performance such as throughput, latency, and signal strength. The user interface and measurement approach of RTR-NetTest are as shown in Figure 3.5.

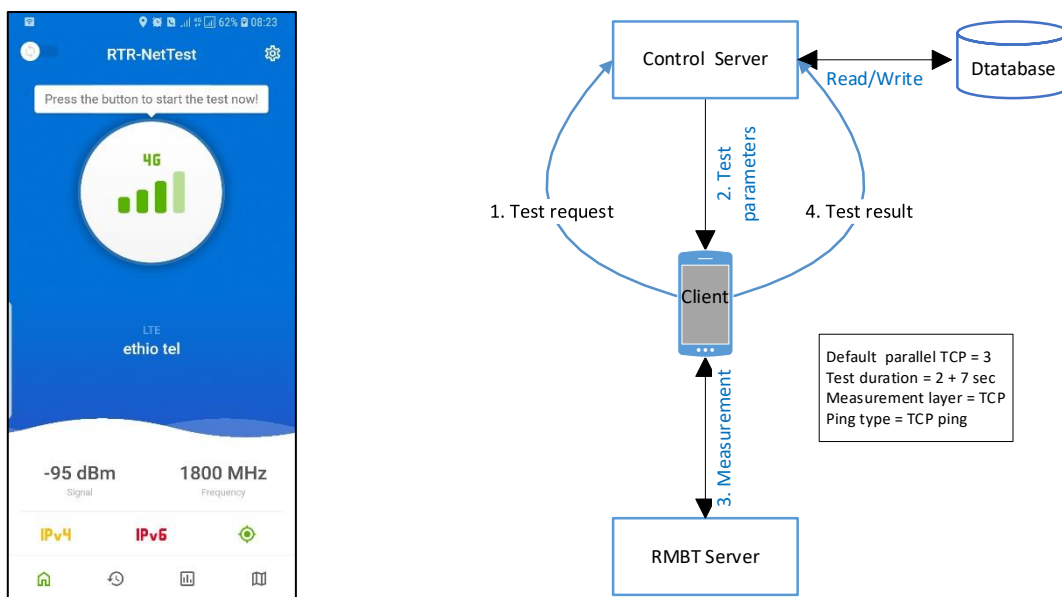


Figure 3.5: RTR-Nettest QoS measurement approach

The RTR-NetTest network measurement platform consists of a test client, a measurement server, a control server, and a database, as shown in Figure 3.5. The client sends a test request to the control server. The control server determines the measurement server used, then generates a token. The control server then transmits the token and all the additional test parameters (DNS name or IP address of the measurement server, number of parallel TCP connections, number of pings, test duration, and others) to the client. A token is used for identifying the test session.

The test consists of seven phases that are carried out one after the other. That means the phases do not overlap. The seven phases are Phase 1: Initialization, Phase 2: Downlink pre-test, Phase 3: Latency test, Phase 4: Downlink test, Phase 5: Uplink pre-test, Phase 6: Uplink test, Phase 7: Finalization. The nominal number of parallel TCP connections used for downlink and uplink measurement starts at one, and the default is three [46].

During the latency test, the client sends N pings in short intervals to the measurement server to test the latency of the connection. One ping consists of the transmission of short strings via one of the TCP connections to the server, which returns short strings as acknowledgment. RTR-NetTest uses ten pings default and takes the median [46].

Throughput estimation of RTR-NetTest: For both download and upload rate, R is calculated as Equation (3.3) [46].

$$b_k := \approx b_k^{(l_{k-1})} + \frac{t^* - t_k^{(l_{k-1})}}{t_k^{(l_k)} - t_k^{(l_{k-1})}} (b_k^{(l_k)} - b_k^{(l_{k-1})}) \quad (3.2)$$

and

$$R := \frac{1}{t^*} \sum_{k=1}^n b_k \quad (3.3)$$

Where; l_k is the index of the chunk received on thread k at t^* or right after t^* , t^* is minimum time for all TCP samples, n is number of TCP connections, K is TCP connection number, and b_k is data size for k TCP connection in t^* calculated as Equation (3.2).

In RTR, the latency is measured using a single TCP connection by transmitting a short string via one TCP connection to the server with a return short string ACT and taking the median out of ten ping tests.

C. Nemo Handy-A

Nemo Handy is the world's most widely used handheld drive test and in-building measurement tool. Nemo Handy is highly suitable for performing measurements both outdoors and indoors. It uses in busy and crowded indoor spaces while simultaneously being used as a regular mobile phone. Nemo Handy provides us with the best real-time measurement visualization on the handheld market [47]. This tool tests throughput in the application measurement layer with options of HTTP and FTP protocols, and the most common used by the operators is FTP. In latency measurement, the tools use ICMP and Iperf protocols

even though most the telecom operators used ICMP ping. The user interface and measurement approach of Nemo Handy-A are as shown in Figure 3.6.

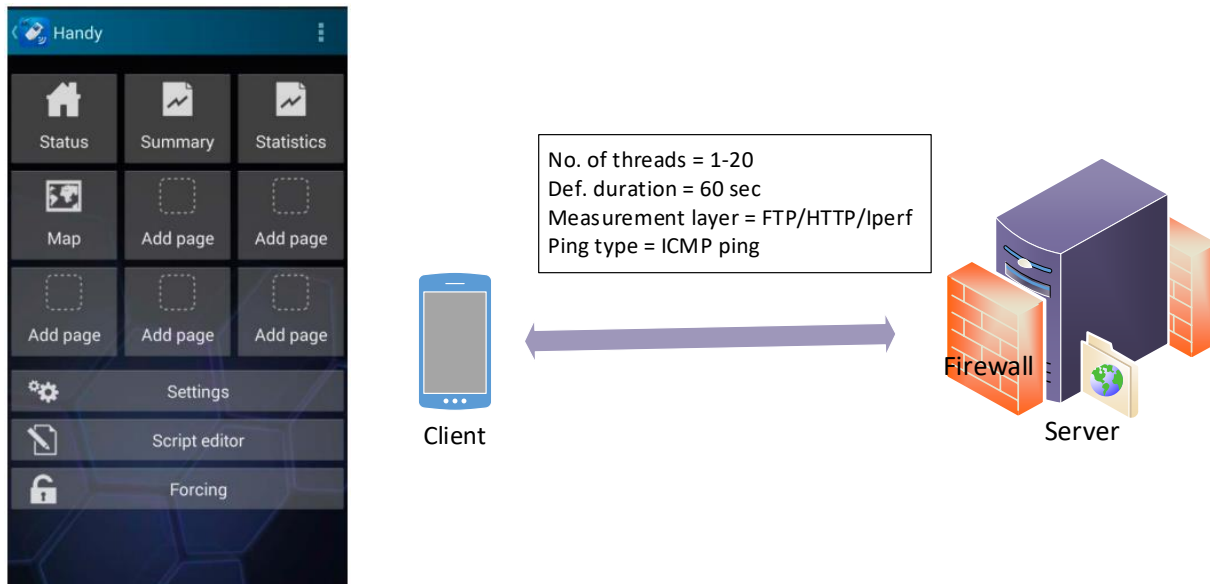


Figure 3.6: Nemo Handy-A QoS measurement approach

This tool has 1 to 20 connection options that can be configured simply during measurement.

Throughput estimation of Nemo Handy-A: As I have observed from the sample log file, set the file size (e.g. 50MB, 100MB) to be transferred in maximum duration. The samples are throughput per second. Throughput, R , is calculated for each sample using Equation (3.5):

$$R = \frac{D}{RTT} \quad (3.5)$$

Total throughput within the configured maximum duration will be the mean of the samples. The latency takes the average of the number of pings.

4. Methodology and Data Collection

The main objective of the thesis work is to evaluate the data QoS measurement tools that uses different measurement techniques and factors based on static measurement on 4G network in Addis Ababa, Ethiopia. The analysis depends on the real-time data collected on the commercial mobile network. Investigate the cause of QoS measurement differences on data service using selected objective parameters to have better understanding by operators, Internet service providers, regulators, and SLA enterprise customers.

Investigate the main possible factors that affect the data QoS measurement on the mobile network. Then, map the identified factors to the three selected testing tools and identify their setup values. To show their impact in different measurement scenarios, choosing the manipulated factors from each test tool is the next step. In scenario one, the measurement is conducted according to the setup and evaluated the impact of the selected manipulated factors using MAPE. In the second scenario, the visualization and evaluation of the selected three tools are done based on the reference value. The measurement performs at the same place and time. The data collection includes using a tool from the DT or WT and two tools from crowdsourcing measurement techniques as shown on the proposed approach in Figure 4.1.

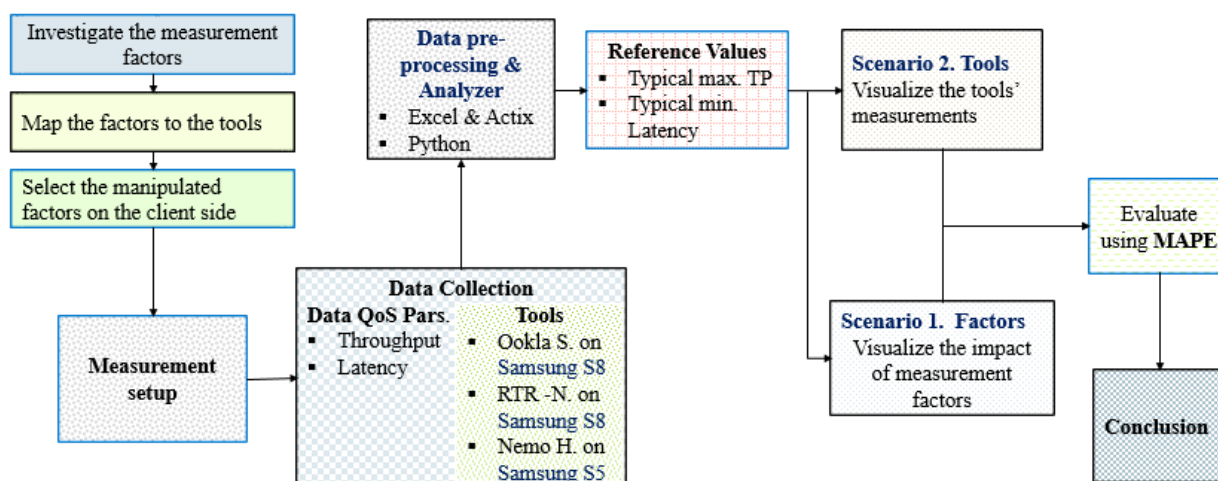


Figure 4.1: Proposed approach

The following sections demonstrate the data collection mechanisms and considerations during the thesis work using Ookla Speedtest, RTR-NetTest, and Nemo Handy-A measurement tools.

4.1. Data Collection

The thesis studies the measurement of the quality of data service in the 4G mobile network. Even though there are several data QoS parameters, the minimum required parameters are throughput, latency, jitter, packet loss, and packet loss ratio. For this work, throughput and latency are taken as QoS measurement parameters. Ookla Speedtest and RTR-NetTest tools are selected from crowdsourcing measurement techniques and Nemo Handy-A from drive test network measurement methods. Those tools are active and user-side measurements, and that use different methodologies to measure the parameters. Ookla Speedtest and RTR-NetTest are chosen because Ookla Speedtest is the most popular network measurement app in Google Play with more than 10 million tests actively initiated per day and more than 15,000 global testing servers available [44]. RTR-NetTest is used by a number of researchers [4, 48-50], and Nemo Handy-A is what most telecom operators are used for optimization, planning, and benchmarking. The other reason for selecting those tools is that the measurement of the three selected tools is conducted in different layers and protocols. The three selected measurement tools can measure the download, upload speed, and latency.

During data measurements and collections, the following conditions are considered. The measurements are conducted during the daytime. For higher traffic, hours from 09:00 to 12:00 and 14:00 to 18:00, whereas for lower traffic, from 12:00 to 14:00, the data is taken from the Ethio telecom 4G mobile network PRS system within 24 hours in a selected measurement location as shown in Figure 4.2. Throughout the thesis, “Mbps” is used for throughput parameter and “msec” for latency as a unit.

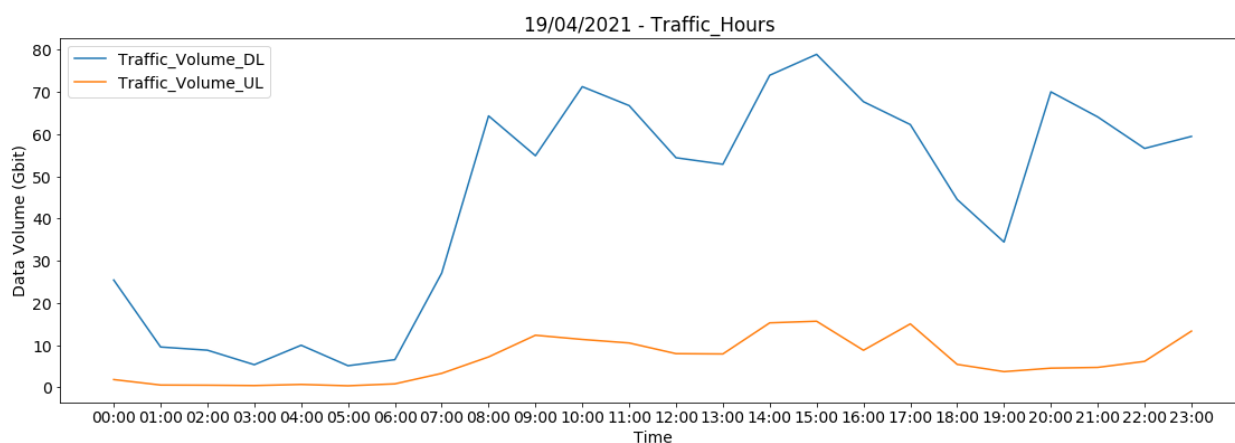


Figure 4.2: Traffic level Vs Hours from PRS system

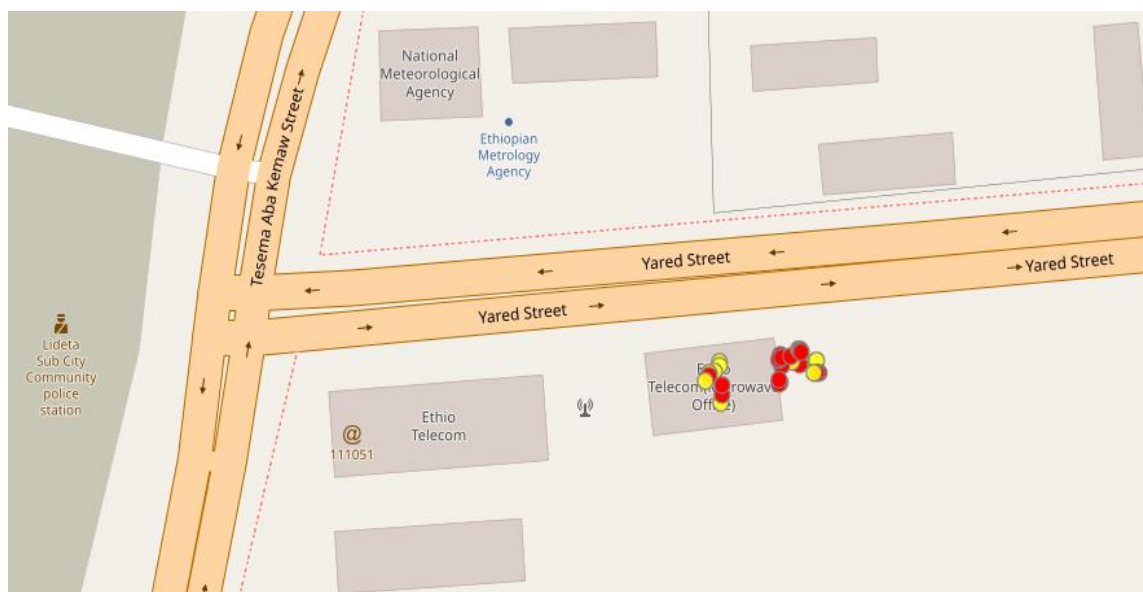


Figure 4.3: Measurement location from RTR server

The static type of measurement is used on the first and sixth floors of the Ethio telecom Microwave Office new building, as shown in Figure 4.3. To show the impact of server location, the number of connections, traffic, and signal level, measurements are taken by Ookla Speedtest. Traffic and signal strength level is measured on RTR-NetTest. The impact of the configured number of threads on the relatively lower and higher signal level is measure by Nemo Handy-A. Those measurements are conducts using the Ookla Speedtest and RTR-NetTest app installed on the Samsung Galaxy S8, whereas the Nemo Handy-A app is installed on the Samsung Galaxy S5 using three different 4G SIM cards. The relatively higher RSRP (-73 dBm) on the sixth floor and the relatively lower RSRP (-94 dBm) on the first floor are used. The nearest serving site for the selected measurement location is Site ID 111051.

Number of Samples: - The thesis includes a total of 2358 measurement samples. Those samples can be grouped based on the tools measured, which are 1130 using Ookla Speedtest, 746 using RTR-NetTest, and 482 using Nemo Handy-A, whereas based on the parameters, 956 samples for DL, 716 samples for UL, and 686 samples for latency measured.

4.2. Data Preprocessing and Analysis

Once the measurement tools and parameters are selected, perform the measurements according to the test setup. Then the data is collected and preprocessed. The data measured using the Nemo Handy-A tool is extracted in Excel format using Actix, and the data measured by the other test tools are sent to the dedicated

Email. Finally, all the data collected from different measurement tools are collected in Excel according to the scenarios and purposes. There are measurements conducted at the same time and location using various tools. To shape the measurement time of the samples, data preprocessing is needed. In scenario one, each factor measures in different network and measurement conditions. As a result, the data is arranged per factor in different measurement conditions. Once the test data is collected and preprocessed, the visualization, analysis, and evaluation are conducted via Python. Python is the most powerful, high-level, open-source, simplest, easiest to read and implement programming language with its visualization.

4.3. Evaluation Techniques and Reference Values

Selection of evaluation techniques from the RMSE and MAPE according to their behaviors and definitions is done. The two evaluation techniques are discussed as follows.

RMSE is the error rate as a square root of MSE. It is the most easily interpreted in terms of measurement units. It is a better measure of fit than a correlation coefficient [51] and calculated as Equation (4.1).

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(\hat{y}_i - y_i)^2}{n}} \quad (4.1)$$

MAPE is the average of the percentage errors and is simple to understand and interpret.

- Mean Absolute Percentage Error (MAPE) calculated as Equation (4.2)

$$MAPE = mean \frac{(|\text{Ref Value} - \text{Measured Value}|)}{\text{Ref Value}} * 100 \quad (4.2)$$

If we are concerned about the impact of large errors rather than the smaller ones, use RMSE that is sensitive to outliers. MAPE is a relative measurement that expresses errors as a percentage of the actual data. The biggest advantage is it provides an easy and intuitive way of judging the extent, or importance of errors. Percentage errors are part of the everyday language, making them easily and intuitively interpretable. Absolute-based error analysis is more robust against outliers in the data sets. Since the measured data contains outliers, MAPE is select as an evaluation technique based on the evaluation reference points of the parameters. The lower the value for MAPE, the better a tool is able to measure the operator's network.

Evaluating the measurement results depends on some standard measurement data. To do this, using the typical maximum value for throughput and the typical minimum value for latency from NMS or Nemo

Handy-A are the two reference value options. Since the measurement scenario is static, selecting values measured by Nemo Handy-A during the data collection of the thesis. Based on this, for DL, the typical maximum and minimum values are 35 Mbps and 10 Mbps sequentially. From those values, 35 Mbps is chosen as the reference value, while 32 Mbps is chosen for UL from the typical maximum and minimum values of 32 Mbps and 15 Mbps, respectively. The evaluation reference value for latency is the relative minimum value of 27 msec.

4.4. Data QoS Measurement Factors

A. Factors beyond the measurement tools: The components in the modified Shannon formula have an impact on QoS measurements. The throughput and latency also depend on air interface parameters: RSRP, SINR, and cell load. The maximum throughput of the 4G mobile network is estimated using the modified Shannon formula as in equation (2.2). RSRP, SINR, the assigned number of physical resource blocks (that depends on the traffic load), the bandwidth, and MCS have an impact on the maximum throughput of the system. The SINR mostly depends on the location of the radio access part, the indoor and outdoor conditions. The traffic level and the signal strength condition are the main factors of the throughput. The throughput is better if both the RSRP and SINR conditions are good and also the traffic level is low, whereas the traffic level is high and the other parameters are good, the expected value decreases. In other words, if both the RSRP and SINR values are lower and the traffic level is lower, we can not say that the maximum throughput is good. It is not considering a mobile device category. Those parameters have an impact on the QoS measurement by operators and ISPs.

B. Notion of the Measurement: The purpose of the measurement tools originally developed have an impact on the measurement. The notions of the measurements are measuring the capacity, available bandwidth, and bulk transfer capacity or throughput. The purpose of the measurement may measure the total capacity of the cell, mostly known as a technology, available bandwidth, or bulk transfer capacity (BTC), or throughput. BTC is different from throughput. Throughput may use several parallel connections, coming from the development of new smartphone technology that supports multiple parallel TCP connections, while BTC supports only one TCP connection. Today, most data QoS mobile network measurements use the throughput one attempts to close the available bandwidth.

C. Measurement Tool Based Factors: The most important factors of data QoS measurement using throughput and latency as parameters are listed. Factors that affect the throughput and latency measurements are [6, 8, 29, 52-55]:

- Factors that affecting throughput measurement are:
 1. Notions of measurement
 2. Measurement server location and measurement area
 3. Layer of measurement
 4. The number of parallel TCP connections used
 5. The way sample measurement is filtered and, as well as statistical aspects
 6. Best effort or Windowed approach
 7. Slow starting mechanism and testing duration
 8. The type and version of the operating system
- Factors that affecting latency measurements are:
 1. Type of ping
 - Native Ping: ICMP based ping
 - Inet Ping: TCP based ping
 - HTTP Ping: HTTP request and response based ping
 2. Server proximity
 - The proximity of the server affects the latency.

When we come to throughput, the notion of the measurement is to clarify which broadband measurement type is to be measured. Measurement of capacity, available bandwidth, bulk transmission capacity, and throughput are the options of broadband speed measurement. The measurement tools fall under one of the notions of measurements. The measurement server location and measurement time are affecting the QoS measurements that the closeness of the server from the measurement client and the traffic condition on different hours and distance from the access network affects the QoS measurement. As stated in [30], there is a significant difference between network throughput and application throughput. So measurement layers have an impact on the QoS measurement.

And the number of parallel connections used during the QoS measurement has an impact on the QoS measurements. It may have several parallel TCP or parallel threads options that depend on the tools. This number differs from one tool to another and has an impact on the result. The way sample measurement is filtered means that tools are used or not filtered during calculating the results and the different tools use different statistical methods such as the median, mean, and minimum or maximum for the same parameters. Active measurement tools use the best effort or windowed types of measurement. The best effort approach fixes the transferred data at various times, whereas the windowed approach fixes the data transfer time with the data may not be the same. The slow start mechanism and test duration have impacted

each other. If the test duration is short, the slow start has an impact on the statistical value of the parameters. As the test duration is longer, the accuracy of the measurement is better but consumes higher data volume during measurement. Different tradeoff setup values of those factors in different tools affect the measurements. The type and version of the operating system of the testing device have an impact on the measurements. Some operating systems support parallel multi-connection.

When we come to latency, the main factor is the server location and estimation approach. The closeness of the measurement server to the test client has an impact on the measurement latency. This affects mainly time-sensitive services. The other one is the ping type used by the tools. There are native ping, Inet ping, and HTTP ping options that have an impact on RTT.

4.5. Map the Factors to the Selected Tools

To map the list of the factors to the selected three tools, investigated the factors from different sources like the wikis of the tools, the papers, product description, and emails. Some of the factors mapped to the selected test tools are listed in Table 4.1.

Table 4.1: Actual techniques used by the tools

| Factors | Ookla Speedtest | RTR-NetTest | Nemo Handy-A |
|--|---|--|---|
| Measurement layer | HTTP [50] | TCP [50] | FTP, HTTP, Iperf [53] |
| Number of parallel TCP or thread connection | Up to 8 parallel HTTP threads & multiple (2-4) TCP connections [52] | Default 3 parallel sessions or TCP connection [46, 52] | For FTP measurement layer, default is 1 and 1-20 threads can be used [53] |
| Filtering of throughput values for each test | For DL, remove fastest 10% and slowest 30%, whereas UL, the fastest half is averaged and the others trimmed [8, 52] | Take the data size within the minimum transfer time t for all TCP samples and normalized it with t | Without filtering, take the average. |
| Nominal test duration for UL and DL speed | 10 Sec [8, 52] | 2 + 7 Sec [46, 52] | HTTP & FTP 60 Sec and Iperf (TCP) 30 sec [53] |
| The number of “Ping” tests | 6 HTTP GET request & response and take lowest value [55] | 10 & take the median [46] | 3 ICMP pings with a 32 KB packet size & take the average [53] |
| Latency test method | By transmitting HTTP packets [52] | Ping-pong using TCP [46] | ICMP ping [53] |

4.6. Identification of Manipulated Factors on the Client Side of the Tools

Once map the measurement factors with the selected user-side data QoS measurement tools in section 4.5, the selection of the manipulated test factors on the client-side of the measurement tool's platform is performed. In the Ookla Speedtest, the number of connections, server location, traffic level, and signal strength condition is select as manipulated factors. In the RTR-Nettest, traffic levels and signal conditions are chosen, whereas, in the Nemo Handy-A, the number of threads is chosen.

In Ookla Speedtest, we can manipulate the number of connections as multi or single. The server location has many options like Hargeisa or Nairobi, among more than 15,000 alternative servers globally. Additionally, at a relatively lower and higher level, signal strength and traffic level are manipulated factors.

In RTR-NetTest, we can manipulate both the signal strength and the traffic level with the option of at a relatively lower or higher level. The other factors are not manipulable from the client-side and set by default according to the measurement tool algorithm developed.

In Nemo Handy-A, the number of threads has options one to twenty is selected even though several factors manipulate on the client-side since other factors are includes in the other tools. To analyze the impact of other test factors beyond server location, traffic level, signal strength condition, and bottleneck, we used Ookla Speedtest and RTR-NetTest on the same server location.

In the thesis, there are two scenarios. The first scenario is the visualization and evaluation of the impact of manipulated measurement factors. The second scenario is the evaluation of the measurement tools based on the reference values and analysis based on the test factors.

4.7. First Scenario: Visualize and Evaluate the Impact of Selected Measurement Factors

To show the impact of the factors in the measurement, we use the manipulated test factors. The following factors can be manipulable on the client-side of the tools. Using Ookla Speedtest, the number of connections in relatively higher and lower traffic are used. Multi and single connections with 204 samples for each combination to all parameters are taken. It shows the impact of the number of parallel connections in different traffic conditions.

In addition to this, server location and signal strength level are controllable factors in Ookla Speedtest. Even though there are more than 15,000 server options, Hargeisa and Nairobi measurement servers are selected for this work. Suddenly, the Addis Ababa, Ethiopia, server option is not found in the server list. The signal strength level selected the relatively higher and lower RSRP levels. The last two factors indicate the impact of server location and signal strength level on the measurement of the parameters. For both factors, 204 samples are taken for all parameter.

Using RTR-NetTest; traffic and signal strength level used during factor evaluation. Both factors are used in relatively higher and lower levels, with 114 samples for each parameter.

Using Nemo Handy-A, the number of threads versus signal strength levels are studied. Even though Nemo Hand-A has from one to twenty thread configuration options, we have used one, two, four, and eight configurable threads in both relatively higher and lower RSRP conditions for this work. It shows the impact of threads at the QoS measurement on the two signal strength levels. In this case, the DL throughput parameter with 240 samples is considered.

Lastly, using Ookla Speedtest and RTR-NetTest on the same measurement server location in Vienna, Austria. This measurement setup enables us to observe the impact of the other factors, like their latency estimation approach, and others.

In the visualization and analysis part, we use labels such as the higher traffic as HT, lower traffic as LT, single connection as S, multi-connection as M, Hargeisa server as Hg, Nairobi server as N, lower RSRP as LRSRP, higher RSRP as HRSRP, thread as TR, and reference as Ref.

4.8. Second Scenario: Evaluation of the Three Measurement Tools

The second scenario is the evaluation and comparison of the three selected measurement tools. The measurement of the second scenario is performed at the same location and time using the three testing tools. For DL, UL, and latency, 100, 100, and 42 samples are measured, respectively. The measurement setup of the three tools is as shown in Table 4.2.

Table 4.2: Measurement setup for scenario two

| | Measurement setup | | |
|----------------------|-------------------|-------------|-----------|
| Factors | Ookla S. | RTR N. | Nemo H. A |
| Duration | 10 | 2 + 7 | 10 |
| Server location | Hargeisa | Vienna | AA |
| Measurement location | MW(G+6) | MW(G+6) | MW(G+6) |
| No. of TCP | 2-3 | 3 | 3 |
| Traffic level | Lower | Lower | Lower |
| Latency Estmation | Lower [55] | Median [46] | Average |

Using the observations from the first scenario, analyze the results in the second scenario. The analysis focuses on the reasons behind the differences in the measurement of the various test tools conducted at the same place and time.

5. Results and Discussion

The result and discussion part is presented in two scenarios, that the impact of measurement factors on selected parameters are visualized and interpreted in the first scenario. The DL, UL, and latency measurements using the three tools are discussing in the second scenario. Subsections 5.1 and 5.2 describe the results and discussions.

5.1. First Scenario: The Impact of Measurement Factors on the Parameters

5.1.1. The Impact of Measurement Factors on DL

To show the impact of the number of connections in relatively higher and lower traffic, we measured data using the Ookla Speedtest tool. The evaluations are conducted based on the DL reference value, 35 Mbps, the yellow line as shown in Figure 5.1. It has two options to manipulate the number of connections that are multi and single connections. The evaluated results for multi and single connections using MAPE are 42.52% and 68.65% in lower-traffic, whereas it is 12.87% and 78.58% in higher traffic sequentially. We can observe from the errors that using multi and single connections during DL QoS measurement has changed in an error of 26.13% in lower traffic and 65.71% in higher traffic.

Hargeisa and Nairobi servers are selected to show the impact of server location on DL throughput using Ookla Speedtest measurement tool. The evaluated results of DL throughput on Nairobi and Hargeisa servers using MAPE are 67.74% and 24.8% sequentially. The data is measured at relatively higher traffic conditions. A change in measurement server location results in a 42.94% change in error. The signal strength level at which the measurement takes place has an impact on the measurement result. According to the evaluated result, the errors are 39.04% and 38.57% on the relatively lower RSRP and higher RSRP sequentially. The change in error is lower in high traffic and it is higher in lower traffic.

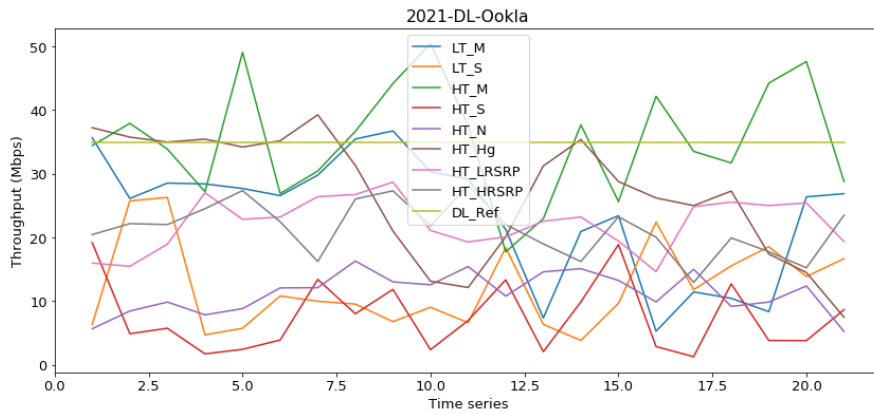


Figure 5.1: Impact of factors on DL using Ookla Speedtest

The client-side of the tool cannot manipulate to demonstrate the impact of measurement techniques using RTR-NetTest. However, we can show the impact of traffic and signal strength levels using a server located in Vienna, Austria. We have used 36 measurement samples for each of relatively lower traffic, higher traffic, lower RSRP, and higher RSRP shown in Figure 5.2. The evaluated results using MAPE are 82.83%, 84.06%, 65.93%, and 48.43%, respectively. The impact of traffic and signal strength level changes in an error of 1.23% and 17.51%, respectively. The change in error in traffic level is smaller when compared with signal strength level because the bottleneck on the link between the measurement client and server is the same.

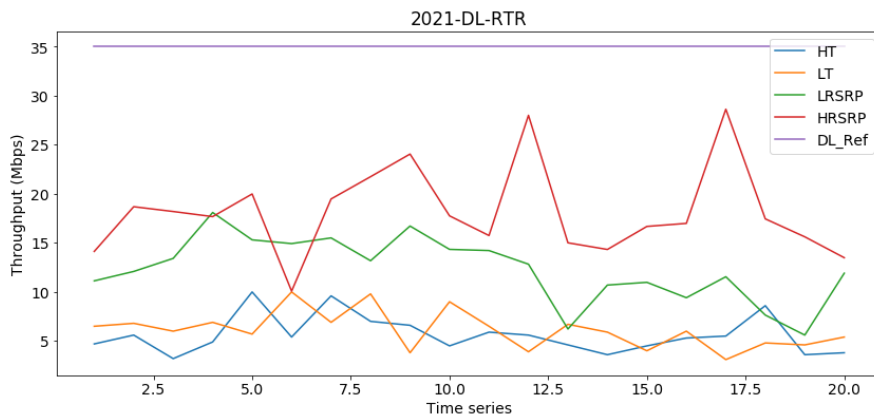


Figure 5.2: Impact of techniques on DL using RTR-NetTest

To show the impact of the number of threads in Nemo Handy-A on the DL parameter, we use the number of threads one, two, four, and eight from 1-20 configurable options with relatively lower and higher RSRP conditions as shown in Figure 5.3 and Table 5.1. In lower RSRP, the evaluated errors are 75.46%, 80.47%, 84.02% and 82.73%. As we have seen in the evaluation, the error trend shows that it is increasing. Therefore, configuring the higher number of threads at a lower signal level leads to an increase in the

errors from the expected value. In higher RSRP, the errors are 55.10%, 44.91%, 30.79%, and 50.25%, respectively. Here we can see that by increasing the number of threads during the measurement at higher signal levels and lower traffic, the error becomes lower, and the result tends to the reference value. During data QoS measurement, configuring the number of parallel TCP or thread connections should be considered based on the traffic and signal strength level that is better if adaptable.

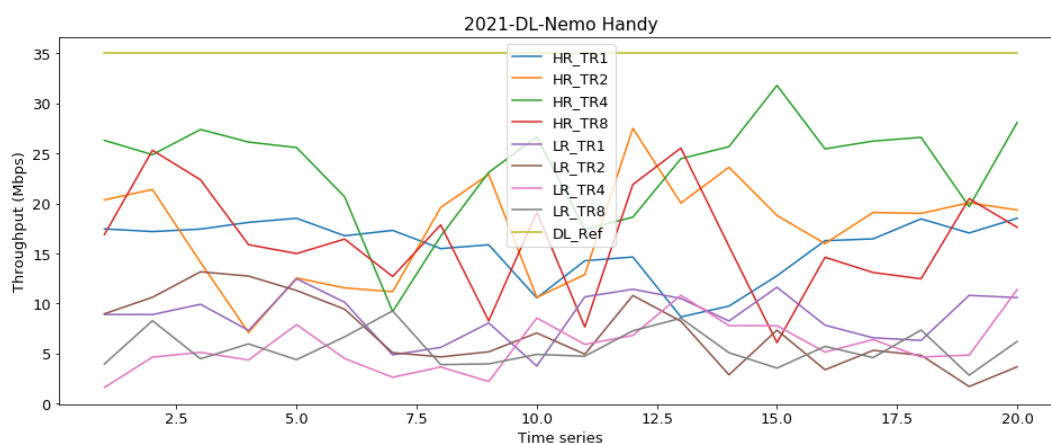


Figure 5.3: Impact of no. of threads on DL using Nemo Handy-A

Table 5.1: The average DL and evaluation of the number of threads

| Nemo H. | Higher RSRP | | | | Lower RSRP | | | |
|----------|-------------|-------|-------|-------|------------|-------|-------|-------|
| No. TR | TR1 | TR2 | TR4 | TR8 | TR1 | TR2 | TR4 | TR8 |
| Avg. DL | 15.72 | 19.28 | 24.23 | 17.41 | 8.6 | 6.84 | 5.6 | 6.04 |
| MAPE (%) | 55.10 | 44.91 | 30.79 | 50.25 | 75.46 | 80.47 | 84.02 | 82.73 |

5.1.2. The Impact of Measurement Factors on UL

When we come to UL, the number of connections in relatively higher and lower traffic, server location, and signal strength level using Ookla Speedtest taken as the measurement techniques. The evaluations are conducted based on the reference maximum DL value of 32 Mbps, measured using Nemo Handy-A as the yellow line shown in Figure 5.4. Based on the data, the errors for multi and single connections in lower-traffic are 40.7% and 31.74%, respectively, whereas, in higher traffic, the errors are 96.78% and 99.07%, correspondingly. In lower traffic, a change in the number of connections follows an 8.96% error difference, and in higher traffic, the change is a 2.29% error. This difference shows that the change in the number of parallel connections in the measurement has a higher impact on lower traffic. Server location is the second technique that can affect the measurements. The errors of the server locations at Hargeisa and Nairobi are 96.9% and 98.91%, respectively, which makes a difference of 2.01% error measured in higher traffic

conditions. The third one is the signal strength level during measurement performed. The evaluation errors in relatively lower and higher RSRP are 59.17% and 15.39%, respectively, and a change in an error of 43.78%. Change in signal strength level incorporated higher measurement errors. Those factors biased the data QoS measurements unless carefully handling the setups and test conditions.

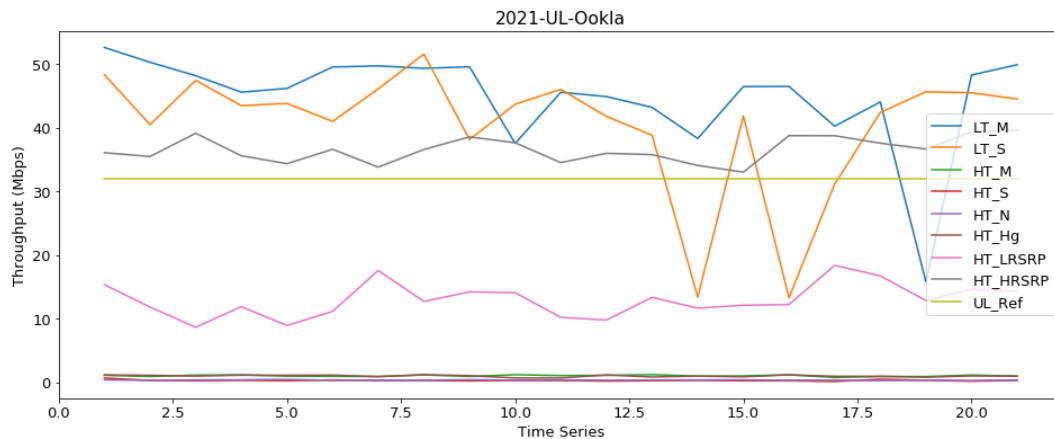


Figure 5.4: Impact of techniques on UL using Ookla Speedtest

The traffic level and signal strength conditions are taken as factors to show their impact on the measurement using the RTR-NetTest shown in Figure 5.5. In relatively lower and higher traffic, the errors to the expected value are 98.78% and 98.85%, respectively, whereas, in relatively lower and higher RSRP, the errors are 99.00% and 98.89%, sequentially. The traffic and signal strength levels change by an error of 0.07% and 0.11%, respectively. In UL, the bottleneck is the cause of error in higher traffic conditions measurements done via RTR-NetTest.

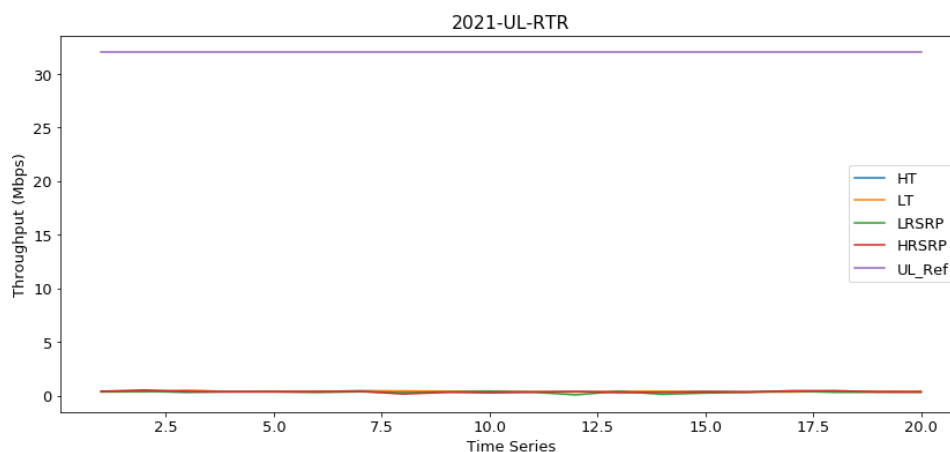


Figure 5.5: Impact of techniques on UL using RTR-NetTest

5.1.3. The Impact of Measurement Factors on Latency

On Ookla Speedtest, the number of parallel connections, server locations, signal strength levels, and traffic levels are selected as measurement factors that illustrated in Figure 5.6. The evaluated errors of latency in multi and single connections are 20.76% and 22.12% at lower traffic, respectively, whereas, in higher traffic, the errors in multi and single connections are 13.11% and 15.88%, respectively. In the higher traffic, the change in the errors is 2.77%, whereas in the lower traffic, the difference in the errors is 1.36%.

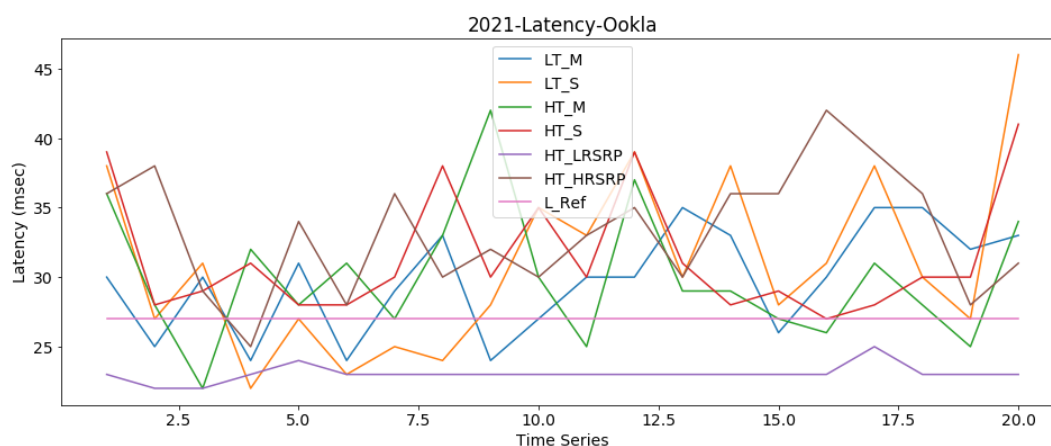


Figure 5.6: Impact of techniques on latency using Ookla Speedtest

The server location affects the latency measurement. Hargeisa and Nairobi measurement servers are selected for this work, as shown in Figure 5.7. In the higher traffic, the evaluated errors for Hargeisa and Nairobi measurement servers are 11.07% and 920.81%, respectively. The other factor is signal strength level, which is the relatively lower and higher RSRP. In the higher traffic level, errors evaluated to the latency reference value are 14.05% and 21.68%, respectively. The change in server location and signal strength level yield a change in errors of 909.74% and 7.63%, respectively. As we have seen in the latency, changes in the server locations have a higher impact on the latency measurements. Secondly, changes in the signal strength level make the difference in the latency measurement values based on the data collected by Ookla Speedtest.

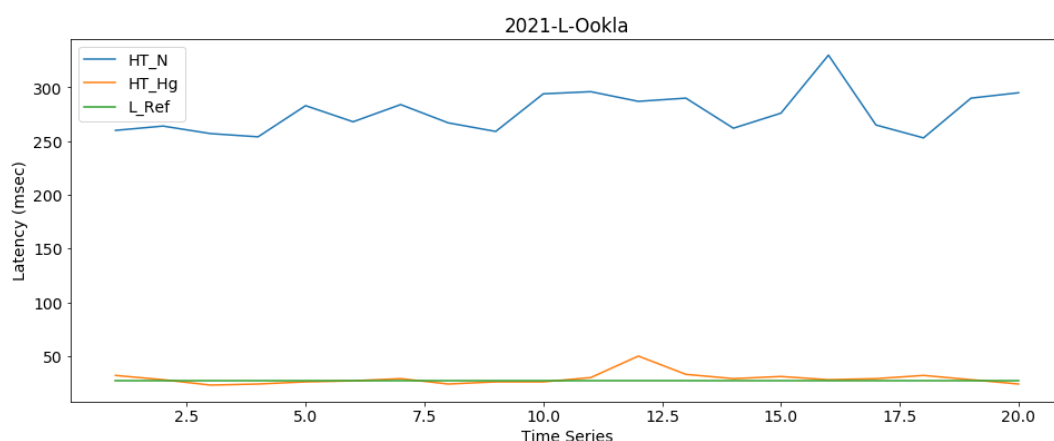


Figure 5.7: Impact of server locations on latency using Ookla Speedtest

We use the RTR-NetTest measurement tool to show the impact of traffic and signal strength levels on latency measurement, as shown in Figure 5.8. The latency errors in relatively lower and higher traffic are 425.25% and 491.41%, respectively. On the other hand, the latency errors in relatively lower and higher RSRP are 541.08% and 535.52% sequentially. The change in traffic and signal strength level yields 66.16% and 5.56% errors, respectively. This result shows us the impact of the traffic level on latency measurements is higher even though the measurements are conducted at different relative signal levels.

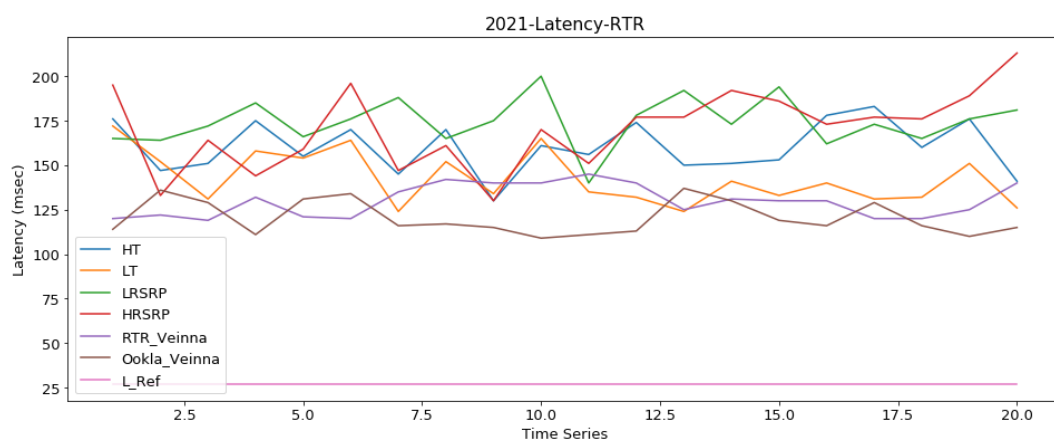


Figure 5.8: Impact of factors on latency using RTR-NetTest

We use data measured by RTR-NetTest and Ookla Speedtest tools at the same traffic level, signal strength level, and measurement server location in Vienna, Austria, as shown in Figure 5.8. We can see the latency measurement errors caused by other internal factors. The latency errors measured using RTR-NetTest and Ookla Speedtest are 380.14% and 347.14%, respectively. Even though the other factors are similar except the measurement tools, there is a difference in an error of 33%. The change in error mainly comes from

the difference in latency estimation methods of the tools that Ookla Speedtest uses the lowest value of six ping samples [55], and the RTR-NetTest uses a median of ten ping samples [46].

Generally, the impact of selected factors on DL, UL, and latency parameters and the errors related to the reference values are summarized in Table 5.2.

Table 5.2: Average and evaluation of the techniques using three tools

| Tools | Factors | Avg. DL | MAPE (%) | Avg. UL | MAPE (%) | Avg. L | MAPE (%) |
|-----------------|-----------------|---------|----------|---------|----------|--------|----------|
| Ookla Speedtest | LT_M | 20.29 | 42.52 | 42.63 | 40.7 | 31.78 | 20.76 |
| | LT_S | 10.97 | 68.65 | 38.92 | 31.74 | 32.03 | 22.12 |
| | HT_M | 35.28 | 12.87 | 1.03 | 96.78 | 29.95 | 13.11 |
| | HT_S | 7.50 | 78.58 | 0.3 | 99.07 | 31.29 | 15.88 |
| | HT_N | 11.29 | 67.74 | 0.35 | 98.91 | 275.62 | 920.81 |
| | HT_Hg | 26.81 | 24.80 | 0.99 | 96.9 | 29.05 | 11.07 |
| | LRSRP | 21.34 | 39.04 | 13.07 | 59.17 | 23.21 | 14.05 |
| | HRSRP | 21.50 | 38.57 | 36.97 | 15.39 | 32.68 | 21.68 |
| RTR-NetTest | LT | 6.01 | 82.83 | 0.39 | 98.78 | 141.82 | 425.25 |
| | HT | 5.58 | 84.06 | 0.37 | 98.85 | 159.68 | 491.41 |
| | LRSRP | 11.92 | 65.93 | 0.32 | 99.00 | 173.1 | 541.08 |
| | HRSRP | 18.05 | 48.43 | 0.36 | 98.89 | 171.59 | 535.52 |
| Ookla & RTR | RTR-NetTest | | | | | 129.64 | 380.14 |
| | Ookla Speedtest | | | | | 120.73 | 347.14 |

5.2. Second Scenario: DL, UL, and Latency Measurements Using the Three Tools

The measurement setup of this scenario is according to Table 4.2 that the measurement takes place at the same location and time statically. Here we discuss the results and evaluation for download, upload, and latency QoS measurement parameters.

5.2.1. DL of the Three Measurement Tools

The DL throughput is measured statically rather than mobility. The average DL throughputs measured using Ookla SpeedTest, RTR-NetTest, and Nemo Handy-A are 14.93, 5.52, and 18.36 Mbps, respectively, based on the measurement setup in Table 4.2. Their evaluated results to the DL reference value, the red color shown in Figure 5.9, 35 Mbps, are 57.35%, 84.22%, and 47.54%, respectively. Even though the measurement place and time are the same for the similar 4G mobile network, the DL measurement values

are different. These can see in each plot trend in Figure 5.9. This difference is also stated in [4, 5] even though their measurement scenarios vary from this paper. According to the discussions in scenario one, we can say that the bottleneck at the link to the measurement server due to the difference in server location is the main reason for the difference. The other contributing factors are the measurement layer and the DL throughput estimation methods. From this, we can say that to evaluate the QoS of the network, the client and measurement server should be available within the operator's network area. The difference in estimation techniques causes various measurement results of different tools shared in [6-8]. The other significant causes are server location and the number of parallel connections used during the measurement and shared by [7, 8].

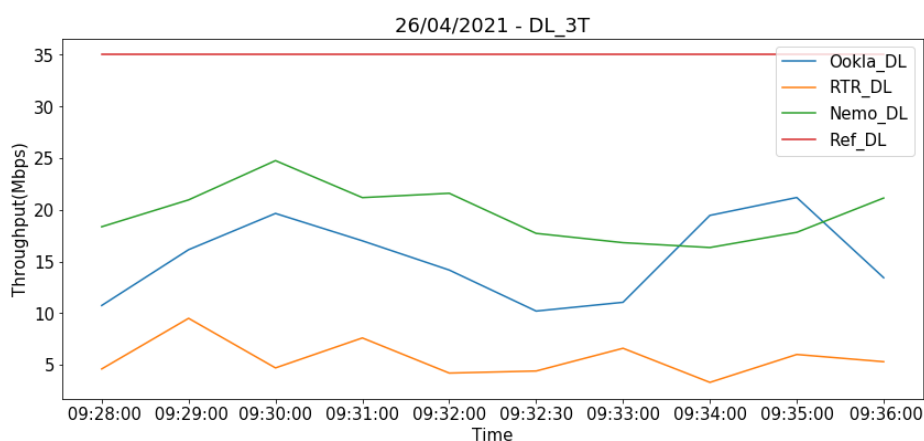


Figure 5.9: Plot of DL throughput using the three tools

5.2.2. UL of the Three Measurement Tools

In UL, the measurement setup is the same as in DL, as shown in Table 4.2. The values of UL throughput measured by Ookla Speedtest, RTR-NetTest, and Nemo Handy-A are 22.27, 13.93, and 23.91 Mbps, respectively, while their evaluations are 30.42%, 56.47%, and 25.27%, respectively. As we observe from Figure 5.10, the UL average value of measured samples of RTR-NetTest is lower than others, and the error is higher than the other tools. This difference is due to a bottleneck between the measurement client and the measurement server located in Vienna, Austria. The measurement server location of the Nemo Handy-A is used server in Addis Ababa, Ethiopia. That is why the measurement errors to reference value are lower. The other one is Ookla Speedtest selects the CDN measurement server located in Hargeisa, Somaliland. As the measurement server is located remotely from the evaluated network, the measurement is vulnerable to a bottleneck problem.

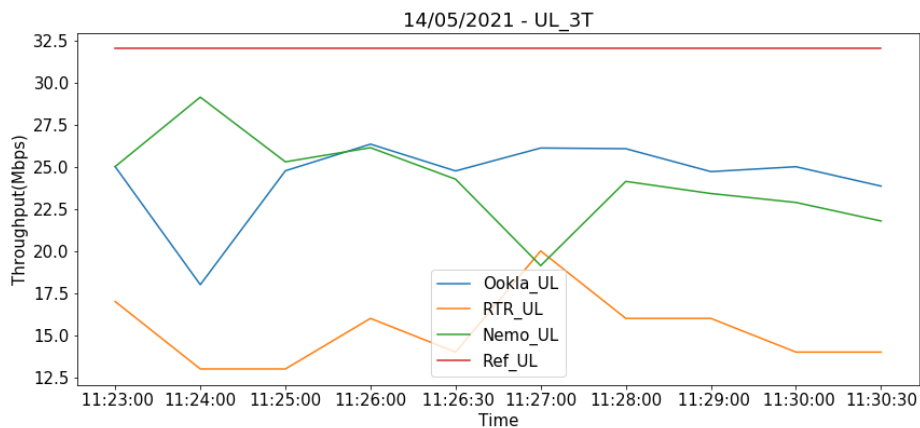


Figure 5.10: Plot of UL throughput using the three tools

5.2.3. Latency of the Three Measurement Tools

In the case of latencies, the reference value is the minimum value for standard tools, Nemo Handy-A, and the value is 27 msec that the red color line shown in Figure 5.11. According to Table 5.3, the average latencies of the Ookla Speedtest, RTR-NetTest, and Nemo Handy-A measurement tools are 38.21, 160.26, and 30.69 msec, respectively, while the evaluated errors are 42.06%, 493.56%, and 13.66%, respectively. Not only the throughput but also the latency of the measurement tools yield significant differences. When we see the errors, the results measured by RTR-NetTest have a higher error on the Ethio telecom 4G network system since the server location is in Vienna, Austria. The measurement server of Nemo Handy-A is located in Addis Ababa, Ethiopia, whereas the Ookla Speedtest measurement server location is in Hargeisa, Somaliland. We observe that the server location highly affects the latency than the throughput. The impact of server location on throughput measurement depends on the bottleneck of the links, network nodes, and measurement servers. The other reason for the latency difference is the estimation methods of the tools. The tool's latency estimation methods are in Ookla Speedtest uses the lowest value of six ping samples [55], RTR-NetTest uses a median of ten ping samples [46], and Nemo Handy-A uses the average of the samples.

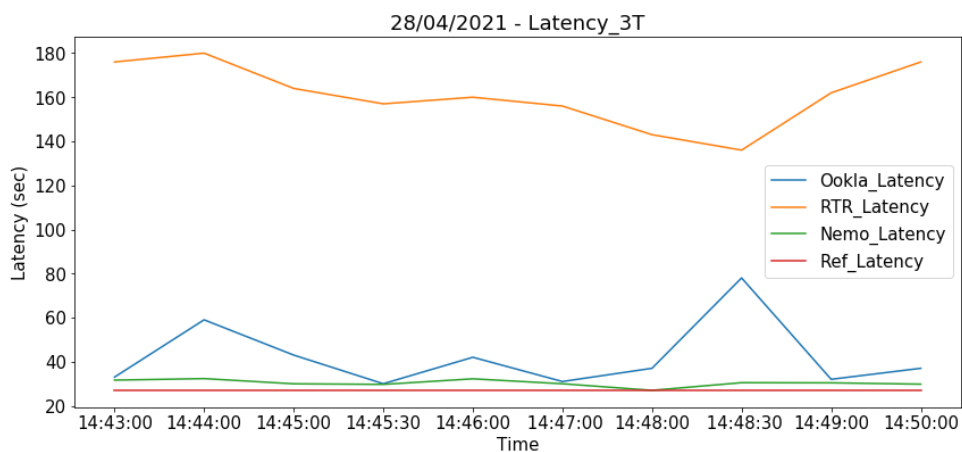


Figure 5.11: Plot of latency throughput using the three tools

Table 5.3: Average and evaluations of the parameters using the tools

| Tools | Evaluation | | | | | |
|-----------------|------------|----------|---------|----------|--------------|----------|
| | Avg. DL | MAPE (%) | Avg. UL | MAPE (%) | Avg. Latency | MAPE (%) |
| Ookla Speedtest | 14.93 | 57.35 | 22.27 | 30.42 | 38.21 | 42.06 |
| RTR-NetTest | 5.52 | 84.22 | 13.93 | 56.47 | 160.26 | 493.56 |
| Nemo Handy-A | 18.36 | 47.54 | 23.91 | 25.27 | 30.69 | 13.66 |

6. Conclusion and Recommendation

6.1. Conclusion

The objective of the thesis was to evaluate data service QoS measurement factors and tools based on the DL, UL, and latency measurement parameters. The measurement factors, which are the notion of measurement tool, measurement techniques of the tools, and network status, have a significant impact on QoS measurements. We concluded that data QoS measurement using various techniques and tools for the same specific operator's network is different. The reasons for the measurement difference, which include factors beyond the measurement tools, the notion of measurement, and the tool's measurement techniques, are discussed in this thesis. According to the results, the throughput measurement is highly affected by the number of parallel connections or threads and traffic level that indicate the load status of the network. The main factors for latency measurement are server location and the estimation methods used by the measurement tools. When the QoS assessment tool requires configuration for the number of parallel connections or threads, the traffic and signal strength level could be taken into consideration. The measurement error of the Nemo Handy-A is less than the other tools in all parameters based on the Ethio telecom 4G LTE network. The main factors for those measurement differences are the location of the measurement server, the bottleneck, and the parameter estimation technique of the tools.

Even though Nemo Handy-A is a standard measurement tool, it costs time, money, analysis time, unfair representation measurement area, and hours relative to the crowdsourcing methods. If the challenges of CS measurement tools are handled carefully, then using the CS tool as a QoS measurement is advantageous. When we select QoS measurement tools for the dedicated network, we should have considered the setup of their measurement factors.

Using the crowdsourcing measurement method can overcome the limitations of the drive test method, even though limited in QoS parameters, and inject additional traffic into the active network.



6.2. Recommendation

Some recommendations from the process and work of the thesis are as follows. During evaluating the QoS of the Operator's network, the measurement server should be within the network service area. When telecom operators, Internet service providers, and governmental regulators use QoS measurement tools, they should evaluate the measurement techniques and tools based on test factors.

Even though the Nemo Handy-A measurement tool showed better accuracy for a given scenario, by applying careful post-processing using crowdsourcing measurement tools during data QoS measurement has an advantageous. We recommended Ethio telecom to use the crowdsourcing measurement tools for fair distribution of the measurements, real-time operation, reduced measurement cost, and access to measurement data at any time and location that used for network troubleshooting, which is a fast input for optimization and maintenance.

Here are the recommendations for further thesis on this topic:

- Compare the QoS measurement platforms on local measurement servers installed on the network coverage of the operator network.
- Evaluate the QoS measurement tools in the case of mobility through the selected area of network coverage.

Reference

- [1] C. Cox, “An Introduction to LTE, LTE-Advanced, SAE, VoLTE and 4G Mobile Communications.,” *Published by John Wiley & Sons Ltd, UK, 2014.*
- [2] E. Telecomi, “Ethio telecom Annual Business Performance Report (01 July 2019 – 30 June 2020),” *2019.*
- [3] S. Haryadi, “Telecommunication Quality of Service Concept: Network Performance and Quality of Service,” *Lecture Note, 2016.*
- [4] R. Bisrat et al, "Quality Evaluation for Indoor Mobile Data Customers in Addis Ababa Business Area Using Data from Network Management System, Walk Test, Crowdsourcing and Subjective Survey; In E. Nigussie, T. Tegegne, & F. Mekuria (Eds.), *Information and Communication Technology for Development for Africa - 2nd International Conference, ICT4DA 2019, Revised Selected Papers* (pp. 164-175), 2019.
- [5] A. Bethelhem, “Evaluation of User Side Quality of Service Measurement Techniques for Addis Ababa LTE Data Service,” *Thesis paper, Addis Ababa University, 2018.*
- [6] E. Lipenbergs, V. Bobrovs, and G. Ivanovs, “Investigation of Service Quality of Measurement Reference Points for the Internet Services on Mobile Networks,” *Journal Paper, Latv. J. Phys. Tech. Sci., 2016.*
- [7] C. Midoglu and P. Svoboda, “Opportunities and Challenges of Using Crowdsourced Measurements for Mobile Network Benchmarking a Case Study on RTR Open Data,” in *Proceedings of 2016 SAI Computing Conference, SAI 2016, 2016.*
- [8] S. Bauer, D. Clark, and W. Lehr, “Understanding broadband speed measurements,” *Telecommun. Policy Res. Conf.*, pp. 1–39, 2010.
- [9] M. P. Wylie-Green et al, “LTE Measurements,” *LTE UMTS Evol. to LTE-Advanced Second Ed.*, pp. 303–323, 2011.
- [10] F. Rezaei, M. Hempel, and H. Sharif, “LTE PHY performance analysis under 3GPP standards parameters,” *2011 IEEE 16th Int. Work. Comput. Aided Model. Des. Commun. Links Networks, CAMAD 2011, no. to, pp. 102–106, 2011.*
- [11] 3GPP TS 23.402, “Technical Specification Group Services and System Aspects; Architecture enhancements for non-3GPP accesses, (Release 13),” *2016.*

-
- [12] A. Kitana, I. Traore, and I. Woungang, “Impact Study of a Mobile Botnet over LTE Networks,” *J. Internet Serv. Inf. Secur.*, vol. 6, no. 2, pp. 1–22, 2016. Available: <http://isyu.info/jisis/vol6/no2/jisis-2016-vol6-no2-01.pdf>.
- [13] J. Acharya, L. Gao and S. Gaur, “Heterogeneous Networks in LTE-Advanced; Fundamental of LTE,” *Published by John Wiley & Sons, Ltd., First Edition, 2014*.
- [14] 3GPP TS 23.003, “Numbering, Addressing and Identification, Release 13, Section 13,” *March 2016*.
- [15] 3GPP TS 36.300, “Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall Description; Stage 2, Release 16, Section 13,” *September 2021*.
- [16] 3GPP TS 23.401, “General Packet Radio Service (GPRS) Enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) Access, Release 17,” *2021*.
- [17] Ericsson “LTE/SAE System Overview,” *Vendor Book, 2009*.
- [18] E. Dahlman *et al.*, “The 3G long-term evolution - Radio interface concepts and performance evaluation,” *IEEE Veh. Technol. Conf.*, vol. 1, no. c, pp. 137–141, 2006.
- [19] B. B. Haile, E. Mutafungwa, and J. Hamalainen, “LTE-Advanced enhancements for self-backhauled LTE-U small cells: An Addis Ababa case study,” *IEEE AFRICON Conf.*, vol. 2015-Novem, 2015.
- [20] P. Mogensen *et al.*, “LTE capacity compared to the shannon bound,” *IEEE Veh. Technol. Conf.*, no. 1, pp. 1234–1238, 2007.
- [21] A. Ghosh *et al.*, “Fundamentals of LTE,” *Prentice Hall Press, Upper Saddle River, NJ, USA, 2011*.
- [22] S. Sesia, I. Toufik, and M. Baker, “LTE- The UMTS Long Term Evolution: From Theory to Practice,” *Book, Published in John Wiley & Sons, 2009*.
- [23] 3GPP TS 36.214 version 16.1.0 Release 16, “LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer; Measurements,” *2021*.
- [24] F. Afroz *et al.*, “SINR, RSRP, RSSI and RSRQ Measurements in Long Term Evolution Networks,” *Int. J. Wirel. Mob. Networks*, vol. 7, no. 4, pp. 113–123, 2015.



-
- [25] H. Holma and A. Toskala, “LTE for UMTS: Evolution to LTE-Advanced,” *Book, Chichester, West Sussex: Published by John Wiley & Sons, 2011.*
- [26] G. Piro et al, “Simulating LTE cellular systems: An open-source framework,” *IEEE Trans. Veh. Technol., vol. 60, no. 2, pp. 498–513, 2011.*
- [27] H. Holma and A. Toskala, “LTE for UMTS: OFDMA and SC-FDMA based radio access,” *John Wiley & Sons Ltd, 2009.*
- [28] F. Krasniqi, L. Gavrilovska, A. Maraj, “The Analysis of Key Performance Indicators (KPI) in 4G LTE Networks. In: Poulkov V. (eds) Future Access Enablers for Ubiquitous and Intelligent Infrastructures. FABULOUS 2019.” *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 283. Springer, Cham, 2019.*
- [29] B. Connelly, “How Ookla Ensures Accurate, Reliable Data: A Guide to Our Metrics and Methodology.” Available: <https://www.speedtest.net/about/knowledge/test-methods>. [Accessed May 23, 2021].
- [30] ETSI EG 203 165, “Speech and multimedia Transmission Quality (STQ); Throughput Measurement Guidelines,” 2016.
- [31] Y. Umutoni, “Tools and Methods for Testing the QoS of Internet services: in ITU Regional Standardization Forum for Africa,” 2015.
- [32] K. Kawamura et al, “Management system for mobile networks,” *Fujitsu Sci. Tech. J., vol. 48, no. 1, pp. 47–53, 2012.*
- [33] B. S. Elenbogen, “Computer network management: Theory and practice,” *SIGCSE 1999 - Proc. 13th SIGCSE Tech. Symp. Comput. Sci. Educ., no. June, pp. 119–121, 1999.*
- [34] C.Yuhan & S.Ruhao, “PRS makes network O&M less art than science,” *Mag., Available: https://www.huawei.com/mediafiles/CORPORATE/PDF/Magazine/communicate/69/HW_259763.pdf. [Accessed Mar. 16, 2021]*
- [35] A. Tadesse, “Evaluating Quality of Services of 4G LTE Cellular Data Network: The Case of Addis Ababa,” *MSc Thesis, 2019.*
- [36] J. Johansson et al, “Minimization of drive tests in 3GPP release 11,” *IEEE Commun. Mag., vol. 50, no. 11, pp. 36–43, 2012.*



- [37] D. López-Pérez et al, “OFDMA Femtocells: A Roadmap on Interference Avoidance,” *IEEE Commun. Mag.*, vol. 47, no. September, pp. 41–48, 2009.
- [38] RTR-NetTest. Available: <https://www.netztest.at/en/>. [Accessed Apr. 21, 2021]
- [39] P.-H. Lehne, “3rd nordic workshop of system & network optimization (snow): In Proceedings of the Mobile Network Operator Challenges to Meet Future Demand for Mobile Broadband”, 2012.
- [40] S. Sonntag, L. Schulte, and J. Manner, “Mobile network measurements -It’s not all about signal strength,” *IEEE Wirel. Commun. Netw. Conf. WCNC*, pp. 4624–4629, 2013.
- [41] ECC Report 195, “Minimum Set of Quality of Service Parameters and Measurement Methods for Retail Internet Access Services”, 2013.
- [42] A. Rusan and R. Vasiiu, “Emulation of backhaul packet loss on the LTE S1-U interface and impact on end user throughput,” *Proc. - 2015 IEEE 11th Int. Conf. Intell. Comput. Commun. Process. ICCP 2015*, pp. 529–536, 2015.
- [43] M. Mahfoudi et al, “LTE network capacity analysis to avoid congestion for real time traffic,” *Mediterr. Microw. Symp.*, vol. 2015-April, pp. 3–7, 2015.
- [44] The Speedtest Server Network. May 27, 2021, [online]. Available: <https://www.ookla.com/speedtest-servers>. [Accessed May 27, 2021]
- [45] W. Li et al, “Toward Accurate Network Delay Measurement on Android Phones,” *IEEE Trans. Mob. Comput.*, vol. 17, no. 3, pp. 717–732, 2018.
- [46] A. Gaber, “alladin-IT Nettek Measurement Algorithm”, *Description*, 2017.
- [47] Keysight Technologies, “Nemo Handy Handheld Measurement Solution,” [keysight.com](https://www.keysight.com/zz/en/product/NTH00000B/nemo-handy-handheld-measurement-solution.html). Available: <https://www.keysight.com/zz/en/product/NTH00000B/nemo-handy-handheld-measurement-solution.html>. [Accessed Jun. 05, 2021].
- [48] V. Raida et al, “Constant Rate Ultra Short Probing (CRUSP): Measurements in Live LTE Networks,” *IEEE Int. Conf. Commun.*, vol. 2019-May, 2019.
- [49] K. Kousias et al, “Estimating downlink throughput from end-user measurements in mobile broadband networks,” *20th IEEE Int. Symp. A World Wireless, Mob. Multimed. Networks, WoWMoM 2019*, no. M1, 2019.



-
- [50] R. KLUS, “Big data analytics in the context of mobile network performance optimization.”, *Thesis, Brno University of Technology, The Faculty of Electrical Engineering and Communication, Department of Radio Electronics, 2019.*
- [51] U. Khair et al, “Forecasting Error Calculation with Mean Absolute Deviation and Mean Absolute Percentage Error,” *J. Phys. Conf. Ser.*, vol. 930, no. 1, 2017.
- [52] I. Koukoutsidis, “Public QoS and Net Neutrality Measurements,” *J. Inf. Policy*, vol. 5, pp.245–286, 2015. Available: <http://www.jstor.org/stable/10.5325/jinfopoli.5.2015.0245>.
- [53] User Manual, “Nemo Handy-A,” pp. 1–242, 2015.
- [54] S. Hodayouni et al, “The Impact of Duration and Settings of TCP Measurements on Bandwidth Estimation in Mobile Networks,” *IEEE Int. Symp. Pers. Indoor Mob. Radio Commun. PIMRC*, vol. 2017-October, pp. 1–6, 2018.
- [55] W. Li et al, “On the Accuracy of Smartphone-based Mobile Network Measurement,” *Proc. - IEEE INFOCOM*, vol. 26, pp. 370–378, 2015.

Evaluation of 4G LTE User Side Data QoS Measurement Tools: The Case of Addis Ababa, Ethiopia

Adhena Mesele Misgun
Ethio Telecom
Addis Ababa, Ethiopia
adhana.mesele@ethiotelcom.et

Dr. Beneyam Berehanu Haile
Addis Ababa Institute of Technology
Addis Ababa University, Ethiopia
beneyamb.haile@aait.edu.et

Abstract: Measuring Quality of Service (QoS) for a specific mobile network system may be performed by a network operator, Internet service provider, or governmental regulator based on their objectives. Since telecom stakeholders may not use similar measurement methods and tools, their measured values for the selected parameters may not yet be the same. As a result, there is disagreement among stakeholders regarding the benchmark and evaluations. The service coverage and the QoS are critically important for a successful operator business. So, mobile operators should continuously monitor their network performance and undertake optimization and expansion operations on a need basis. The techniques and tools for measuring QoS should be evaluated. The area of study proposed in the thesis is the evaluation of 4G Long Term Evolution (LTE) mobile network system data QoS measurement factors and tools. QoS measurement tools and factors are evaluated based on the selected parameters' reference value.

The thesis evaluates and analyzes the data QoS mobile network measurement tools and factors in the context of Addis Ababa, Ethiopia, using the most relevant throughput and latency performance metrics on Ethio telecom's 4G LTE mobile network. Evaluation of static download speed, upload speed, and latency measurements using Ookla Speedtest, RTR-NetTest, and Nemo Handy-A are conducted using Mean Absolute Percentage Error (MAPE). The achieved evaluated results of download speed measured by the Ookla Speedtest, RTR-NetTest, and Nemo Handy-A to the reference value, 35 Mbps, are 57.35%, 84.22%, and 47.54%, respectively. The evaluated results of upload speed measured by Ookla Speedtest, RTR-NetTest, and Nemo Handy-A to the reference value, 32 Mbps, are 30.42%, 56.47%, and 25.27% correspondingly. The achieved evaluated results of latency measured by the Ookla Speedtest, RTR-NetTest, and Nemo Handy-A to the reference value, 27msec, are 42.06%, 493.56%, and 13.66% respectively. The main reasons for the difference in errors are server location, bottleneck, and parameter estimation methods of the measurement tools.

Keywords-LTE, QoS, Measurement methodology, Measurement factors, QoE, Mobile data services, Measurement tools, Throughput, Latency.

I. INTRODUCTION

Data usage increases dramatically from time to time because of the high data rate users demand. The operators deployed the newest cellular network technology to satisfy the higher data rate user demand. 4G is the best selected

mobile technology in most operators through introducing 5G in some countries. Those mobile technologies were introduced with better spectral efficiency, higher network capacity, higher network coverage, higher throughput, and lower latency related to the previous technologies.

A rapid increase in mobile data usage and emergence of new applications: Multimedia Online Gaming, mobile TV, Web 2.0, streaming content have motivated the Third Generation Partnership Project (3GPP) to work on the Long-Term Evolution (LTE) on the way towards fourth-generation mobile. The evolution was considered compatibility with 3GPP's mobile communication systems competitive over timescales of 10 years and by delivering the high data rates and low latencies that future users would require [1].

The main goal of LTE is to provide a high data rate, low latency, and packet-optimized radio access technology supporting flexible bandwidth deployments. Same time its network architecture has been designed to support packet-switched traffic with seamless mobility and better quality of service.

LTE was required to deliver a peak data rate of 100 Mbps in the downlink and 50 Mbps in the uplink direction. Later, the requirement was exceeded to provide DL and UL peak data rates of 300 Mbps and 75 Mbps respectively. The relevant specification suggests a user plan latency of 10 ms for non-roaming mobile, increasing to 50 ms in a typical roaming scenario [1].

LTE architecture supports forced QoS, with end-to-end quality of service and guaranteed bit rate (GBR) for radio bearers. In 4G LTE, QoS implement at the EPS bearer level. It uses EPS bearers with each assigned an EPS bearer ID applied at the level of EPS bearer. There is a one-to-one mapping, which means that for an EPS Bearer, there is a corresponding EPS Radio Access Bearer (RAB), an S1 bearer, and a Radio Bearer (RB). There are four types of EPS bearers: GBR bearer, Non-GBR bearer, dedicated bearer, and default bearer. GBR bearer that resources are permanently allocated by admission control whereas Non-GBR bearer no admission control. Dedicated bearer associated with specific TFT (GBR or non-GBR) and Default Bearer Non-GBR, a catch-all for unassigned traffic.

The Operators compete with one another to acquire and retain mobile subscribers on their networks. Then, the operator should utilize various measurement methodologies and tools to assess network performance on the network side and QoS on the user side. To optimize their networks, the operator uses the data from NP, QoS, and QoE assessments. Parameter setup, license update, network expansion, and new site deployment are part of the optimization process. Following the appropriate optimization, the operator can attract additional users and keep the existing subscribers with their available network.

There are different types of mobile network measurement methodologies and tools. Generally, network measurements can be active or passive [2]. Mostly passive measurements are performed at the core side of the network and measured using inputs from different network interfaces without injecting traffic. This type of measurement is difficult to use the network interfaces by professionals simply and third-party bodies. Examples of passive measurement techniques are Network Management System (NMS) and Minimization Drive test (MDT). Active measurements are conducted by injecting traffic on the live network but may bias selected test network area conditions such as congestion, poor network coverage. Examples of active measurement techniques are traditional drive tests or walk tests, some crowdsourcing tools. There are measurement platforms that include both active and passive measurement types. Active and passive measurement types depend on the measured parameters. Some related studies on mobile network performance and QoS were reviewed in different scenarios. The purpose of the review is to read the background knowledge that is the foundation of current findings. Here are some of the related papers:

The objective of the paper [3] was to investigate the quality of the experience of UMTS enterprise data customers. In this paper, the Network Management System (NMS), Nemo Handy-A, and RTR-NetTest measurement tools were used for QoS measurement. Additionally, the survey technique was used for QoE data collection. The measurement was conducted based on download, upload speed, and latency metrics. The authors analyzed the results based on DL speed, with an average value of 0.57 Mbps using NMS, 0.86 Mbps using Nemo Handy, 2.53 Mbps using RTR NetTest, and 2.67 the MOS using survey. The authors concluded that the QoS of the network contributes to the QoE measurement results. From the paper, we can learn that the average values of different measurement tools for a given network are different. Whereas paper [4] evaluate the QoS of data service in 4G mobile network system in Addis Ababa using Nemo Handy –A, Speedtest, and OpenSignal. The download speed and latency were used as measurement metrics for the evaluation. The author concluded that the measurement result of Nemo Handy –A, and selected crowdsourcing measurement tools were varied significantly. In paper [5], investigated the practical implementation for developing a harmonizing approach to quality monitoring to obtain objective information on quality of Internet access services on mobile networks. The authors concluded that the

proposed measurement methodology applied the same measurement tools and dedicated reference point to all measurements allows providing reliable and comparable information about mobile broadband Internet access quality given to customers.

The authors in [6] deal with opportunities and challenges in using CS QoS measurements from smartphones for benchmarking mobile networks. Ookla SpeedTest, FLOQ, Kyago, Traffic Monitor, and RTR-NetTest using data rate, delay, and data volume as metrics with 100 samples. The analysis was conducted based on the RTR-NetTest tool and DL speed parameter. The RTR-NetTest reported result was 32.391 Mbps, whereas the goodput result was 33.0963 Mbps, which shows that different estimation algorithms yield different results. The paper was concluded that to achieve accurate and fair benchmarking using the CS measurement tools, the post-processing stage must be designed carefully. Although the tests were conducted on the same device using the same operator's technology, there was a significant variance among the measurements. The study only focused on the CS tools, and the test was not taken at the same time as the analysis was done on RTR-NetTest. Whereas the objective of the paper [7] was to explain the broadband speed measurements. ComScore, Speedtest, Akamai, and YouTube test tools are used with data speed as measurement metrics to analyze the factors for different results. Speedtest, ComScore, Akamai, and YouTube have average download throughputs of 7.93 Mbps, 3.4 Mbps, 3.9 Mbps, and 4.21 Mbps, respectively. Although the measurements were conducted in small geographic regions, the results varied significantly. The paper concluded that the main factors for this difference in the results were the proximity of the test server, the number of parallel TCP connections, and the way sample measurement was filtered.

Most papers had evaluated the QoS of the given mobile network system and the measurement scenario and setups are different from this thesis. The averaged measured value of selected metrics data QoS measurement for different measurement tools are different that stated in paper [3-7]. The differences are some of the papers used crowdsourcing tools only or the measurement locations and time not same. The gaps are why the various QoS measurement tools provide different results even though the measurement takes place at a similar location and time for the same specific operator's or ISP's network technology. Therefore, the evaluation of QoS measurement factors in mobile network system done on this thesis.

Now Ethiopian government opens the telecom industry to involve other companies to provide internet service and telecom network infrastructure. The governmental regulatory body should control and manage the interaction between the stockholders such as telecom operators, ISPs, SLA enterprise customers, and end-users. When we come to the QoS measurement in the mobile network, the stockholders may measure using different measurement techniques and tools that the report may not be the same. Various QoS measurement reports from several organs may follow disagreement and impermissible of benchmark reports.

There were assessments of the mobile network coverage using signal conditions as measurement metrics for optimization and benchmark. However, this is not an end-to-end mobile network service assessment that considers the access network node to the user equipment. To assess mobile network service coverage end to end, we must use service-based mobile network metrics in addition to signal conditions.

Several measurement methods and tools had applied to measure user-side objective data service QoS mobile network data rate and latency used as metrics. NMS, DT/WT, MDT, and CS are the methods to measure and collect the required assessment parameters. Even though the measurements conduct at the same place and time, the varieties of measurement methods and tools provide different results. Those measurement results for the same measurement condition make controversy between telecom stockholders such as telecom regulatory bodies, ISP, SLA between the service provider and end-user or enterprise users, and researchers. Then, this needs to answer why the different tools provide different results measured at the same place and time on the selected network technology. So, the various output reports measured by different measurement methodologies and tools in similar locations and times should study. The thesis answers the questions: What are the main factors on user side QoS measurement? How do the test tools measure the throughput and latency QoS parameters? How do the tools set the value regarding the identified factors? Is the result of measurement tools conducted on the same location and time differ? How much the impact of QoS measurement factors? In addition, how much the accuracy of CS and Walk test tools related to reference value?

The remaining parts of the papers are organized as follows. Section II and III provides description of quality measurement methods, factors and data collection method for measurement setup. Next Section IV presents obtained results and interpretations. Finally, Section V presents concluding remarks of the work.

II. DATA QOS MEASUREMENT TOOLS AND FACTORS

The relevance and minimum of technical parameters for the evaluation of typical data QoS measurements are throughput, latency, Jitter, Packet Loss, and packet error [8].

Throughput: - It is the amount of data or number of data packets that can be delivered in a pre-defined period.

Latency: - It is simply the amount of time takes for data to travel from one defined location to another. Ideally, the latency of a network is as close to zero as possible [9]. The LTE radio interface and network provide capabilities for less than 10 ms latency from the transmission of a packet from the network to the UE.

Jitter: - It is the variation in time delay for the data packets sent over a network. This variable represents an identified disruption in the normal sequencing of data packets. Since the jitter manifests itself in increased or uneven latency between data packets, which can disrupt network performance, lead to packet loss, and network congestion. Although some level of

jitter is to be expected and usually tolerated, quantifying network jitter is an important aspect of comprehensive network performance measurement [9].

Packet Loss: - It refers to the number of packets transmitted from one destination to another that fails to transmit. Packet loss can be measured by capturing traffic data on both ends, then identifying missing packets and retransmission of packets. Packet loss can be caused by network congestion, router performance, and software issues, among other factors [9].

Packet Loss Ratio (PLR): - The improvement of PLR estimation is a critical issue, because the Packet Loss Ratio has a high effect on the network performances, especially when dealing with real-time traffic such as VoIP and Video Flows [10]. The PLR increased during the video flows by the increase of the user's number [11].

The general smartphone based QoS measurement approach is as shown in Figure 2.1.

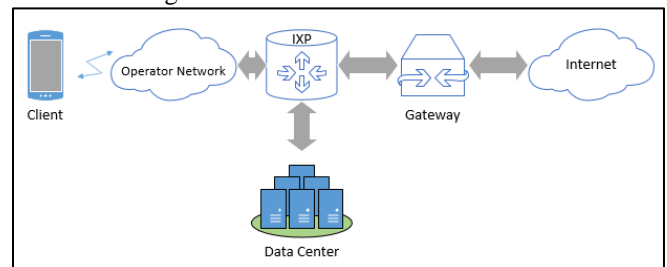


Figure 2.1: QoS Measurement Approach

1. Ookla Speedtest

ISP and telecom experts who recognized the pressing need for transparent, high-quality data on Internet access founded Ookla in 2006. The Ookla Speedtest data is based on the test software and measurement infrastructure developed by Ookla Net Metrics, the company behind Speedtest, works in fixed broadband and mobile network testing applications, data and analysis, a developer and vendor of networking testing applications [12]. The measurement approach of Ookla Speedtest is as shown in Figure 2.2.

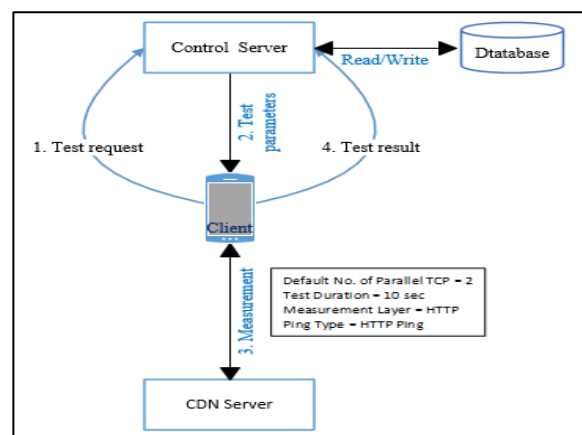


Figure 2.2: Ookla Speedtest QoS Measurement Approach

To select the measurement server, the control server selects the five nearest servers out of more than 15,000 servers, pings

the five servers, and selects the servers with the lowest result in the Ookla Speedtest mobile apps.

Throughput Estimation of Ookla Speedtest (HTTP): For Download, the samples were aggregated into 20 slices that each slices 5% of the samples. By removing 30% of the slowest & 10% of the fastest slice and averaged the remaining slices. For UL, the fastest half samples averaged [7]. For each sample, R is calculated as Equation (1):

$$R = \frac{D}{RTT} \quad (1)$$

Where D is fixed chunk size and RTT is round trip time. If the connection speed is at least 4Mbps, then, Ookla uses four threads unless it will use the default two threads. The latency test performs by measuring the time it takes to set a response for an HTTP request sent to the webserver. This test is repeated multiple times (Six times in Ookla) that the lowest value determines the final result for a specific sample test.

2. RTR NetTest

RTR-NetTest is a crowdsourcing measurement system for fixed and mobile broadband of Austria’s national regulatory authority for broadcasting and telecommunication [13]. The RTR-NetTest informs users about the current service quality such as upload, download, ping, signal strength for their Internet connection. In addition, a map view and statistics of previous tests can be accessed. It measures Internet network layer performance such as Throughput (download and upload), Latency, and Signal Strength.

The Measurement Approach of RTR NetTest is as shown in Figure 2.3.

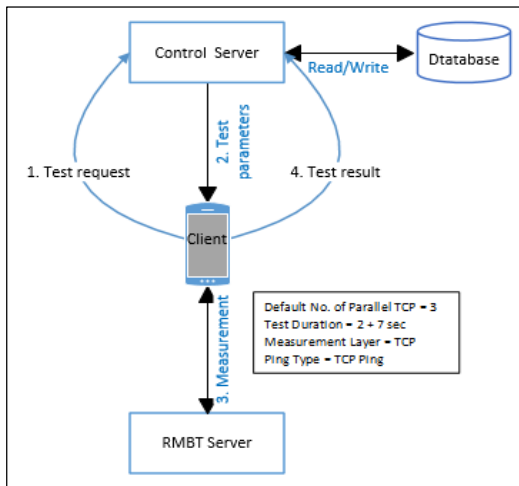


Figure 2.3: RTR Nettest QoS Measurement Approach

RTR NetTest network measurement platform consists of test client, measurement server, control server, and database as shown in Figure 2.3. The client sends a test request to the Control Server. The control server determines the measurement server used, then generates a token. The Control Server then transmits the token and all the additional test parameters (DNS name or IP address of the Measurement Server, number of parallel TCP connections, number of

pings, test duration, and others) to the client. The token is used for identifying the test session.

The test consists of seven phases that are carried out one after the other. That means the phases do not overlap. The seven phases are Phase 1: Initialization, Phase 2: Downlink pre-test, Phase 3: Latency test, Phase 4: Downlink test, Phase 5: Uplink pre-test, Phase 6: Uplink test, Phase 7: Finalization. The nominal number of parallel TCP connections used for downlink and uplink measurement starts from one, and the default is three [14].

During the Latency test, the client sends N pings in short intervals to the Measurement Server to test the latency of the connection. One ping consists of the transmission of short strings via one of the TCP connections to the server, which returns short strings as acknowledgment. RTR NetTest uses ten pings default and takes the median [14].

Throughput Estimation of RTR NetTest (TCP): For both Download and Upload Rate, R, calculated as Equation (3) [14].

$$b_k := b_k^{(l_k-1)} + \frac{t^* - t_k^{(l_k-1)}}{t_k^{(l_k)} - t_k^{(l_k-1)}} (b_k^{(l_k)} - b_k^{(l_k-1)}) \quad (2)$$

$$R := \frac{1}{t^*} \sum_{k=1}^n b_k \quad (3)$$

Whereas, l_k = The index of the chunk received on thread k at t^* or right after t^*

t^* = Minimum time for all TCP samples

n = Number of TCP connections

K = TCP connection number

b_k = Data size for k TCP connection in t^* calculated as Equation (2)

In RTR, the latency measured use single TCP by transmitting short string via one TCP connection to the server with return short string ACK and take the median out of ten Ping test.

3. Nemo Handy –A

Nemo Handy is the world’s most widely used handheld drive test and in-building measurement tool. The handheld drive test tool, Nemo Handy, is highly suitable for performing measurements both outdoors and in busy and crowded indoor spaces while being simultaneously used as a regular mobile phone. Nemo Handy provides you with the best real-time measurement visualization on the handheld market [15]. This tool tests throughput in the application measurement layer with options of HTTP and FTP protocols, and the most common used by the operators is FTP. In latency measurement, the tool uses ICMP ping and Iperf protocols even though mostly the telecom operators used ICMP ping.

The Measurement Approach of Nemo Handy –A is as shown in Figure 2.4.

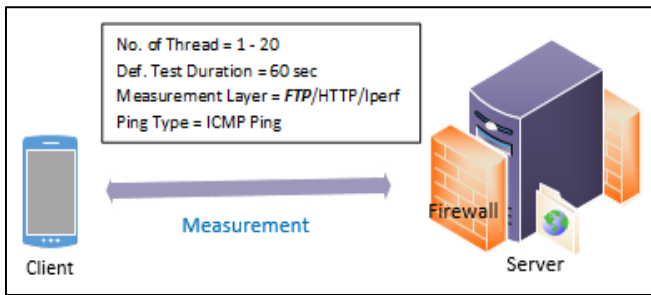


Figure 2.4: Nemo Handy – A QoS Measurement Approach

The tool has 1 to 20 number of connection options that can be configured simply during measurement.

Throughput Estimation of Nemo Handy –A (FTP): As I have observed from the sample log file, set the file size (e.g. 50MB, 100MB) to be transfer in maximum duration. The samples are throughput per second. Throughput, R, for each sample calculated as Equation (4):

$$R = \frac{D}{RTT} \quad (4)$$

Total throughput within configured maximum duration will be the mean of the samples. The latency take the average of number of pings.

A. Factors beyond the measurement tools: - The components in the modified Shannon formula have an impact on QoS measurements. The throughput and latency also depend on air interface parameters: RSRP, SINR, and cell load. The maximum throughput of the 4G mobile network is estimated using the modified Shannon formula. RSRP, SINR, the assigned number of physical resource blocks (that depends on the traffic load), the bandwidth, and MCS have an impact on the maximum throughput of the system. The SINR mostly depends on the location of the radio access part, the indoor and outdoor conditions. The traffic level and the signal strength condition are the main factors of the throughput. The throughput is better if both the RSRP and SINR conditions are good and also the traffic level is low, whereas the traffic level is high and the other parameters are good, the expected value decreases. In other words, if both the RSRP and SINR values are lower and the traffic level is lower, we can not say that the maximum throughput is good. It is not considering a mobile device category. Those parameters have an impact on the QoS measurement by operators and ISPs.

B. Notion of the Measurement: - The purpose of the measurement tools originally developed have an impact on the measurement. The notions of the measurements are measuring the capacity, available bandwidth, and bulk transfer capacity or throughput. The purpose of the measurement may measure the total capacity of the cell, mostly known as a technology, available bandwidth, or bulk transfer capacity (BTC), or throughput. BTC is different from throughput. Throughput may use several parallel connections, coming from the development of new smartphone technology that supports multiple parallel TCP connections, while BTC supports only one TCP connection. Today, most data QoS mobile network measurements use the throughput one attempts to close the available bandwidth.

C. QoS Measurement Tool Based Factors: - The most important factors of data QoS measurement using throughput and latency as parameters are listed. Factors that affect the throughput and latency measurements are [5] [7] [16-20]:

- Factors on throughput measurement:
 - A. Notions of the measurement
 - B. Measurement server location & measurement area
 - C. Measurement layer
 - D. Number of parallel TCP connection used
 - E. The way sample measurement filtered & statistical aspect
 - F. Best effort or Windowed approach
 - G. Slow start mechanism & testing duration
 - H. The type and version of the operating System
- Factors on latency measurement:
 - A. Ping type
 - Native Ping: ICMP based ping
 - Inet Ping: TCP based ping
 - HTTP Ping: HTTP request and response based ping
 - B. Server proximity
 - The proximity of the server affects the latency

When we come to throughput, the notion of the measurement is to clarify which broadband measurement type is to be measured. Measurement of capacity, available bandwidth, bulk transmission capacity, and throughput are the options of broadband speed measurement. The measurement tools fall under one of the notions of measurements. The measurement server location and measurement time are affecting the QoS measurements that the closeness of the server from the measurement client and the traffic condition on different hours and distance from the access network affects the QoS measurement. As stated in [21], there is a significant difference between network throughput and application throughput. Then, measurement layers have an impact on the QoS measurement.

Additionally, the number of parallel connections used during the QoS measurement has an impact on the QoS measurements. It may have several parallel TCP or parallel threads options that depend on the tools. This number differs from one tool to another and has an impact on the result. The way sample measurement is filtered means that tools are used or not filtered during calculating the results and the different tools use different statistical methods such as the median, mean, and minimum or maximum for the same parameters. Active measurement tools use the best effort or windowed types of measurement. The best effort approach fixes the transferred data at various times, whereas the windowed approach fixes the data transfer time with the data may not be the same. The slow start mechanism and test duration have impacted each other. If the test duration is short, the slow start has an impact on the statistical value of the parameters. As the test duration is longer, the accuracy of the measurement is better but consumes higher data volume during measurement. Different tradeoff setup values of those factors in different tools affect the measurements. The type and

version of the operating system of the testing device have an impact on the measurements. Some operating systems support parallel multi-connection, whereas others may not. When we come to latency, the main factor is the server location and estimation approach. The closeness of the measurement server to the test client has an impact on the measurement latency. This affects mainly time-sensitive services. The other one is the ping type used by the tools. There are native ping, Inet ping, and HTTP ping options that have an impact on RTT.

III. DATA COLLECTION

The thesis studies the measurement of the quality of data service in the 4G mobile network. Even though there are several data QoS parameters, the minimum required parameters are throughput, latency, jitter, packet loss, and packet loss ratio. For this work, throughput and latency are taken as QoS measurement parameters. Ookla Speedtest and RTR-NetTest tools are selected from crowdsourcing measurement techniques and Nemo Handy-A from drive test network measurement methods. Those tools are active and user-side measurements, and that use different methodologies to measure the parameters. Ookla Speedtest and RTR-NetTest are chosen because Ookla Speedtest is the most popular network measurement app in Google Play with more than 10 million tests actively initiated per day and more than 15,000 global testing servers available [12]. RTR-NetTest is used by a number of researchers [3, 22-24], and Nemo Handy-A is what most telecom operators are used for optimization, planning, and benchmarking. The other reason for selecting those tools is that the measurement of the three selected tools is conducted in different layers and protocols. The three selected measurement tools can measure the download, upload speed, and latency.

During data measurements and collections, the following conditions are considered. The measurements are conducted during the daytime. For higher traffic, hours from 09:00 to 12:00 and 14:00 to 18:00, whereas for lower traffic, from 12:00 to 14:00, the data is taken from the Ethio telecom 4G mobile network PRS system within 24 hours in a selected measurement location as shown in Figure 4.2. Throughout the thesis, “Mbps” is used for throughput parameter and “msec” for latency as a unit.

The static type of measurement is used on the first and sixth floors of the Ethio telecom Microwave Office new building. To show the impact of server location, the number of connections, traffic, and signal level, measurements are taken by Ookla Speedtest. Traffic and signal strength level is measured on RTR-NetTest. The impact of the configured number of threads on the relatively lower and higher signal level is measure by Nemo Handy-A. Those measurements are conducts using the Ookla Speedtest and RTR-NetTest app installed on the Samsung Galaxy S8, whereas the Nemo Handy-A app is installed on the Samsung Galaxy S5 using different three 4G SIM cards. The relatively higher RSRP (-73 dBm) on the sixth floor and the relatively lower RSRP (-

94 dBm) on the first floor are used. The nearest serving site for the selected measurement location is Site ID 111051.

Number of Samples: - The thesis includes a total of 2358 measurement samples. Those samples can be grouped based on the tools measured, which are 1130 using Ookla Speedtest, 746 using RTR-NetTest, and 482 using Nemo Handy-A, whereas based on the parameters, 956 samples for DL, 716 samples for UL, and 686 samples for latency measured.

Data preprocessing steps: - Once the measurement tools and parameters are selected, perform the measurements according to the test setup. Then the data is collected and preprocessed. The data measured using the Nemo Handy-A tool is extracted in Excel format using Actix, and the data measured by the other test tools are sent to the dedicated Email. Finally, all the data collected from different measurement tools are collected in Excel according to the scenarios and purposes. There are measurements conducted at the same time and location using various tools. To shape the measurement time of the samples, data preprocessing is needed. In scenario one, each factor measures in different network and measurement conditions. As a result, the data is arranged per factor in different measurement conditions. Once the test data is collected and preprocessed, the visualization, analysis, and evaluation are conducted via Python. Python is the most powerful, high-level, open-source, simplest, easiest to read and implement programing language with it’s visualization.

Evaluation techniques and reference value: - Since the measured data contains outliers, MAPE selected as evaluation techniques based on the evaluation reference points of the parameters. MAPE is the average of the percentage errors and is simple to understand and interpret. Mean Absolute Percentage Error (MAPE) calculated as Equation (5):

$$MAPE = mean \frac{(|Ref Value - Measured Value|)}{Ref Value} * 100 \quad (5)$$

To evaluate the measurement results depend on some standard measured data. To do this, using the typical maximum value for throughput and the typical minimum value for latency from NMS or Nemo Handy-A are the two reference value options. Since the measurement scenario is static, selecting values measured by Nemo Handy-A during the data collection of the thesis. Based on this, for DL, the typical maximum and minimum values are 35 Mbps and 10 Mbps sequentially. From those values, 35 Mbps chosen as the reference value, while 32 Mbps is chosen for UL from the typical maximum and minimum values of 32 Mbps and 15 Mbps, respectively. The evaluation reference value for latency is the relative minimum value of 27 msec.

Data QoS measurement factors can be grouped as factors beyond the measurement tools, notion of the measurement, QoS measurement tool based factors. Once the factors are investigated, map the factors to the selected tools and identify the manipulated factors on client side of the tools.

1. **FIRST SCENARIO: VISUALIZE AND EVALUATE THE IMPACT OF SELECTED MEASUREMENT FACTORS**

To show the impact of the factors in the measurement, we use the manipulated test factors. The following factors can be manipulable on the client-side of the tools. Using Ookla Speedtest, the number of connections in relatively higher and lower traffic are used. Multi and single connections with 204 samples for each combination to all parameters are taken. It shows the impact of the number of parallel connections in different traffic conditions.

In addition to this, server location and signal strength level are controllable factors in Ookla Speedtest. Even though there are more than 15,000 server options, Hargeisa and Nairobi measurement servers are selected for this work. Suddenly, the Addis Ababa, Ethiopia, server option is not found in the server list. The signal strength level selected the relatively higher and lower RSRP levels. The last two factors indicate the impact of server location and signal strength level on the measurement of the parameters. For both factors, 204 samples are taken for all parameter.

Using RTR-NetTest; traffic and signal strength level used during factor evaluation. Both factors are used in relatively higher and lower levels, with 114 samples for each parameter. Using Nemo Handy-A, the number of threads versus signal strength levels are studied. Even though Nemo Hand-A has from one to twenty thread configuration options, we have used one, two, four, and eight configurable threads in both relatively higher and lower RSRP conditions for this work. It shows the impact of threads at the QoS measurement on the two signal strength levels. In this case, the DL throughput parameter with 240 samples is considered.

Lastly, using Ookla Speedtest and RTR-NetTest on the same measurement server location in Vienna, Austria. This measurement setup enables us to observe the impact of the other factors, like their latency estimation approach.

In the visualization and analysis part, we use labels such as the higher traffic as HT, lower traffic as LT, single connection as S, multi-connection as M, Hargeisa server as Hg, Nairobi server as N, lower RSRP as LRSRP, higher RSRP as HRSRP, thread as TR, and reference as Ref.

2. **SECOND SCENARIO: EVALUATION OF THE THREE MEASUREMENT TOOLS**

The second scenario is the evaluation and comparison of the three selected measurement tools. The measurement of the second scenario is performed at the same location and time using the three testing tools. For DL, UL, and latency, 100, 100, and 42 samples are measured, respectively. The measurement setup of the three tools is as shown in Table 3.1.

Table 3.1: Measurement setup for scenario two

| | Measurement setup | | |
|-------------|-------------------|----------|-----------|
| Factors | Ookla S. | RTR N. | Nemo H. A |
| Duration | 10 | 2 + 7 | 10 |
| Server Loc. | Vienna | Hargeisa | AA |
| Mnt Loc. | MW (G+6) | MW (G+6) | MW (G+6) |

| | | | |
|---------------|------------|-------------|---------|
| No. of TCP | 2 - 3 | 3 | 3 |
| Traffic Level | Lower | Lower | Lower |
| Latency Est. | Lower [20] | Median [14] | Average |

Using the observations from the first scenario, analyze the results in the second scenario. The analysis focuses on the reasons behind the differences in the measurement of the various test tools conducted at the same place and time.

IV. **RESULT AND DISCUSSION**

The result and discussion part is presented in two scenarios, that the impact of measurement factors on selected parameters are visualized and interpreted in the first scenario. The DL, UL, and latency measurements using the three tools are discussing in the second scenario. Subsections 2.1 and 2.2 describe the results and discussions.

1. **FIRST SCENARIO: THE IMPACT OF MEASUREMENT FACTORS ON THE PARAMETERS**

a) **THE IMPACT OF MEASUREMENT FACTORS ON DL**

To show the impact of the number of connections in relatively higher and lower traffic, we measured data using the Ookla Speedtest tool. The evaluations are conducted based on the DL reference value, 35 Mbps, the yellow line as shown in Figure 4.1. It has two options to manipulate the number of connections that are multi and single connections. The evaluated results for multi and single connections using MAPE are 42.52% and 68.65% in lower-traffic, whereas it is 12.87% and 78.58% in higher traffic sequentially. We can observe from the errors that using multi and single connections during DL QoS measurement has changed in an error of 26.13% in lower traffic and 65.71% in higher traffic. Hargeisa and Nairobi servers are selected to show the impact of server location on DL throughput using Ookla Speedtest measurement tool. The evaluated results of DL throughput on Nairobi and Hargeisa servers using MAPE are 67.74% and 24.8% sequentially. The data is measured at relatively higher traffic conditions. A change in measurement server location results in a 42.94% change in error. The signal strength level at which the measurement takes place has an impact on the measurement result. According to the evaluated result, the errors are 39.04% and 38.57% on the relatively lower RSRP and higher RSRP sequentially. The change in error is lower in high traffic and it is higher in lower traffic.

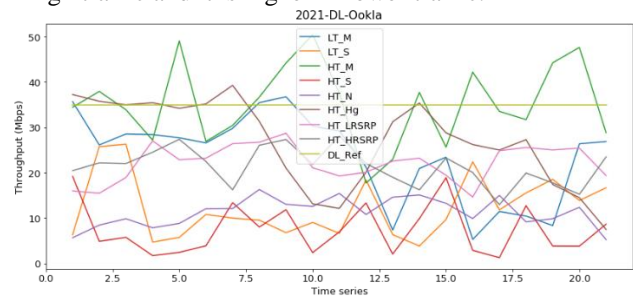


Figure 4.1: Impact of factors on DL using Ookla Speedtest

The client-side of the tool cannot manipulate to demonstrate the impact of measurement techniques using RTR-NetTest. However, we can show the impact of traffic and signal strength levels using a server located in Vienna, Austria. We have used 36 measurement samples for each of relatively lower traffic, higher traffic, lower RSRP, and higher RSRP shown in Figure 4.2. The evaluated results using MAPE are 82.83%, 84.06%, 65.93%, and 48.43%, respectively. The impact of traffic and signal strength level changes in an error of 1.23% and 17.51%, respectively. The change in error in traffic level is smaller when compared with signal strength level because the bottleneck on the link between the measurement client and server is the same.

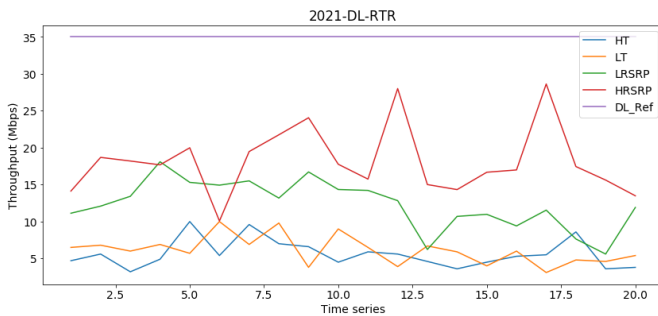


Figure 4.2: Impact of factors on DL using RTR NetTest

To show the impact of the number of threads in Nemo Handy-A on the DL parameter, we use the number of threads one, two, four, and eight from 1-20 configurable options with relatively lower and higher RSRP conditions as shown in Figure 4.3 and Table 4.1. In lower RSRP, the evaluated errors are 75.46%, 80.47%, 84.02% and 82.73%. As we have seen in the evaluation, the error trend shows that it is increasing. Therefore, configuring the higher number of threads at a lower signal level leads to an increase in the errors from the expected value. In higher RSRP, the errors are 55.10%, 44.91%, 30.79%, and 50.25%, respectively. Here we can see that by increasing the number of threads during the measurement at higher signal levels and lower traffic, the error becomes lower, and the result tends to the reference value. During data QoS measurement, configuring the number of parallel TCP or thread connections should be considered based on the traffic and signal strength level that is better if adaptable.

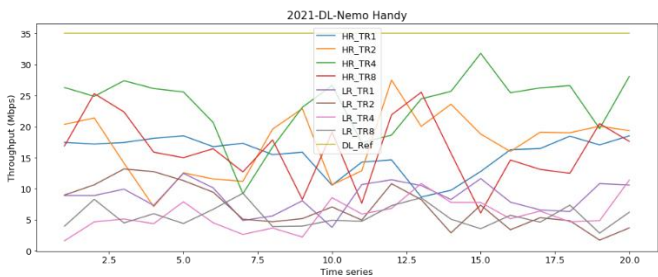


Figure 4.3: Impact of threads on DL using Nemo Handy -A

Table 4.1: The average DL and evaluation of number of threads

| Nemo H. | Higher RSRP | | | | Lower RSRP | | | |
|----------|-------------|------|------|------|------------|------|------|------|
| | TR1 | TR2 | TR4 | TR8 | TR1 | TR2 | TR4 | TR8 |
| No. TR | TR1 | TR2 | TR4 | TR8 | TR1 | TR2 | TR4 | TR8 |
| Avg. DL | 15.7 | 19.2 | 24.2 | 17.4 | 8.6 | 6.84 | 5.6 | 6.04 |
| MAPE (%) | 55.1 | 44.9 | 30.7 | 50.2 | 75.4 | 80.4 | 84.0 | 82.7 |
| | 0 | 1 | 9 | 5 | 6 | 7 | 2 | 3 |

b) THE IMPACT OF MEASUREMENT FACTORS ON UL

When we come to UL, the number of connections in relatively higher and lower traffic, server location, and signal strength level using Ookla Speedtest taken as the measurement techniques. The evaluations are conducted based on the reference maximum DL value of 32 Mbps, measured using Nemo Handy-A as the yellow line shown in Figure 4.4. Based on the data, the errors for multi and single connections in lower-traffic are 40.7% and 31.74%, respectively, whereas, in higher traffic, the errors are 96.78% and 99.07%, correspondingly. In lower traffic, a change in the number of connections follows an 8.96% error difference, and in higher traffic, the change is a 2.29% error. This difference shows that the change in the number of parallel connections in the measurement has a higher impact on lower traffic. Server location is the second technique that can affect the measurements. The errors of the server locations at Hargeisa and Nairobi are 96.9% and 98.91%, respectively, which makes a difference of 2.01% error measured in higher traffic conditions. The third one is the signal strength level during measurement performed. The evaluation errors in relatively lower and higher RSRP are 59.17% and 15.39%, respectively, and a change in an error of 43.78%. Change in signal strength level incorporated higher measurement errors. Those factors biased the data QoS measurements unless carefully handling the setups and test conditions.

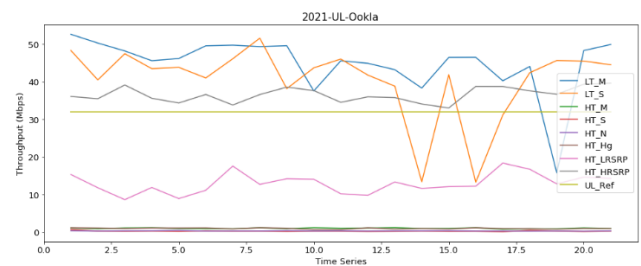


Figure 4.4: Impact of factors on UL using Ookla Speedtest

The traffic level and signal strength conditions are taken as factors to show their impact on the measurement using the RTR-NetTest shown in Figure 4.5. In relatively lower and higher traffic, the errors to the expected value are 98.78% and 98.85%, respectively, whereas, in relatively lower and higher RSRP, the errors are 99.00% and 98.89%, sequentially. The traffic and signal strength levels change by an error of 0.07% and 0.11%, respectively. In UL, the bottleneck is the cause of error in higher traffic conditions measurements done via RTR-NetTest.

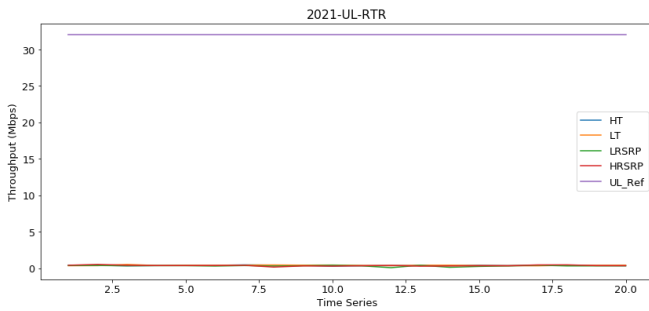


Figure 4.5: Impact of factors on UL using RTR Nettest

c) **THE IMPACT OF MEASUREMENT FACTORS ON LATENCY**

On Ookla Speedtest, the number of parallel connections, server locations, signal strength levels, and traffic levels are selected as measurement factors that illustrated in Figure 4.6. The evaluated errors of latency in multi and single connections are 20.76% and 22.12% at lower traffic, respectively, whereas, in higher traffic, the errors in multi and single connections are 13.11% and 15.88%, respectively. In the higher traffic, the change in the errors is 2.77%, whereas in the lower traffic, the difference in the errors is 1.36%.

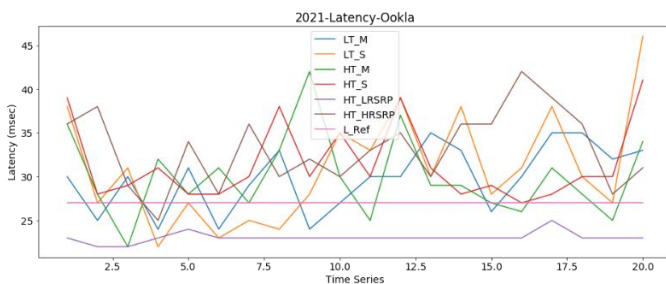


Figure 4.6: Impact of factors on latency using Ookla Speedtest

The server location affects the latency measurement. Hargeisa and Nairobi measurement servers are selected for this work, as shown in Figure 4.7. In the higher traffic, the evaluated errors for Hargeisa and Nairobi measurement servers are 11.07% and 920.81%, respectively. The other factor is signal strength level, which is the relatively lower and higher RSRP. In the higher traffic level, errors evaluated to the latency reference value are 14.05% and 21.68%, respectively. The change in server location and signal strength level yield a change in errors of 909.74% and 7.63%, respectively. As we have seen in the latency, changes in the server locations have a higher impact on the latency measurements. Secondly, changes in the signal strength level make the difference in the latency measurement values based on the data collected by Ookla Speedtest.

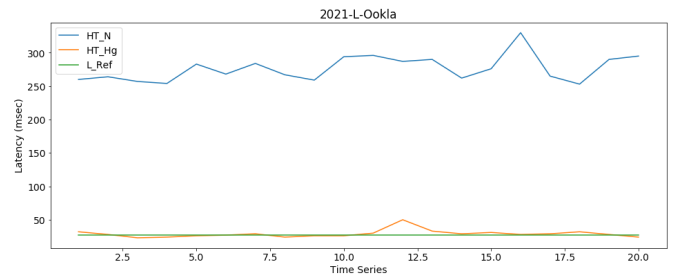


Figure 4.7: Impact of server locations on latency using Ookla Speedtest

We use the RTR-NetTest measurement tool to show the impact of traffic and signal strength levels on latency measurement, as shown in Figure 4.8. The latency errors in relatively lower and higher traffic are 425.25% and 491.41%, respectively. On the other hand, the latency errors in relatively lower and higher RSRP are 541.08% and 535.52% sequentially. The change in traffic and signal strength level yields 66.16% and 5.56% errors, respectively. This result shows us the impact of the traffic level on latency measurements is higher even though the measurements are conducted at different relative signal levels.

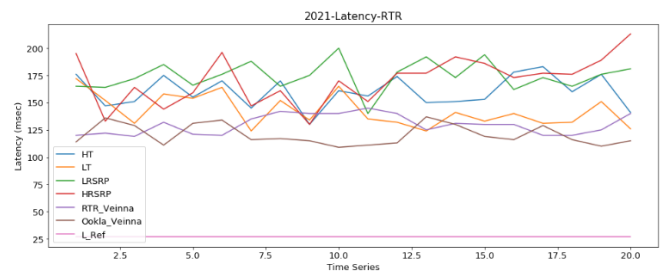


Figure 4.8: Impact of factors on latency using RTR Nettest

We use data measured by RTR-NetTest and Ookla Speedtest tools at the same traffic level, signal strength level, and measurement server location in Vienna, Austria, as shown in Figure 4.8. We can see the latency measurement errors caused by other internal factors. The latency errors measured using RTR-NetTest and Ookla Speedtest are 380.14% and 347.14%, respectively. Even though the other factors are similar except the measurement tools, there is a difference in an error of 33%. The change in error mainly comes from the difference in latency estimation methods of the tools that Ookla Speedtest uses the lowest value of six ping samples [20], and the RTR-NetTest uses a median of ten ping samples [14].

2. **SECOND SCENARIO: DL, UL AND LATENCY MEASUREMENT RESULT USING THE THREE TOOLS**

The measurement setup of this scenario is according to Table 3.1 that the measurement takes place at the same location and time statically. Here we discuss the results and evaluation for download, upload, and latency QoS measurement parameters.

a) **DL OF THE THREE MEASUREMENT TOOLS**

The DL throughput is measured statically rather than mobility. The average DL throughputs measured using Ookla

SpeedTest, RTR-NetTest, and Nemo Handy-A are 14.93, 5.52, and 18.36 Mbps, respectively, based on the measurement setup in Table 3.1. Their evaluated results to the DL reference value, the red color shown in Figure 4.9, 35 Mbps, are 57.35%, 84.22%, and 47.54%, respectively. Even though the measurement place and time are the same for the similar 4G mobile network, the DL measurement values are different. These can see in each plot trend in Figure 4.9. This difference is also stated in [3, 4] even though their measurement scenarios vary from this paper. According to the discussions in scenario one, we can say that the bottleneck at the link to the measurement server due to the difference in server location is the main reason for the difference. The other contributing factors are the measurement layer and the DL throughput estimation methods. From this, we can say that to evaluate the QoS of the network, the client and measurement server should be available within the operator's network area. The difference in estimation techniques causes various measurement results of different tools shared in [5-7]. The other significant causes are server location and the number of parallel connections used during the measurement and shared by [6, 7].

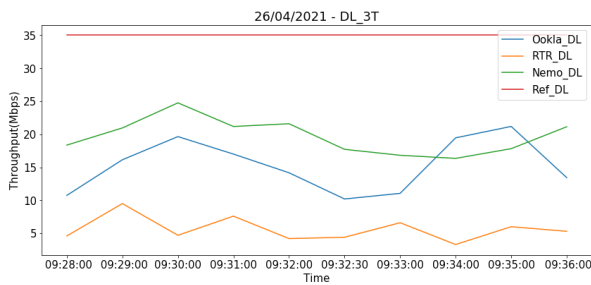


Figure 4.9: Plot of DL throughput using the three tools

b) UL OF THE THREE MEASUREMENT TOOLS

In UL, the measurement setup is the same as in DL, as shown in Table 3.1. The values of UL throughput measured by Ookla Speedtest, RTR-NetTest, and Nemo Handy-A are 22.27, 13.93, and 23.91 Mbps, respectively, while their evaluations are 30.42%, 56.47%, and 25.27%, respectively. As we observe from Figure 4.10, the UL average value of measured samples of RTR-NetTest is lower than other, and the error is higher than the other tools. This difference is due to a bottleneck between the measurement client and the measurement server located in Vienna, Austria. The measurement server location of the Nemo Hand-A is used server in Addis Ababa, Ethiopia. That is why the measurement errors to reference value are lower. The other one is Ookla Speedtest selects the CDN measurement server located in Hargeisa, Somaliland. As the measurement server is located remotely from the evaluated network, the measurement is vulnerable to a bottleneck problem.

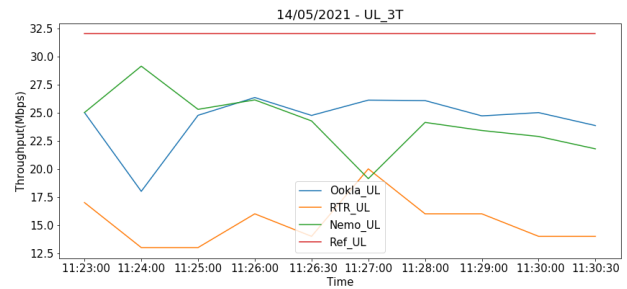


Figure 4.10: Plot of UL throughput using the three tools

c) LATENCY OF THE THREE MEASUREMENT TOOLS

In the case of latencies, the reference value is the minimum value for standard tools, Nemo Handy-A, and the value is 27 msec that the red color line shown in Figure 4.11. According to Table 4.2, the average latencies of the Ookla Speedtest, RTR-NetTest, and Nemo Handy-A measurement tools are 38.21, 160.26, and 30.69 msec, respectively, while the evaluated errors are 42.06%, 493.56%, and 13.66%, respectively. Not only the throughput but also the latency of the measurement tools yield significant differences. When we see the errors, the results measured by RTR-NetTest have a higher error on the Ethio telecom 4G network system since the server location is in Vienna, Austria. The measurement server of Nemo Handy-A is located in Addis Ababa, Ethiopia, whereas the Ookla Speedtest measurement server location is in Hargeisa, Somaliland. We observe that the server location highly affects the latency than the throughput. The impact of server location on throughput measurement depends on the bottleneck of the links, network nodes, and measurement servers. The other reason for the latency difference is the estimation methods of the tools. The tool's latency estimation methods are in Ookla Speedtest uses the lowest value of six ping samples [20], RTR-NetTest uses a median of ten ping samples [14], and Nemo Handy-A uses the average of the samples.

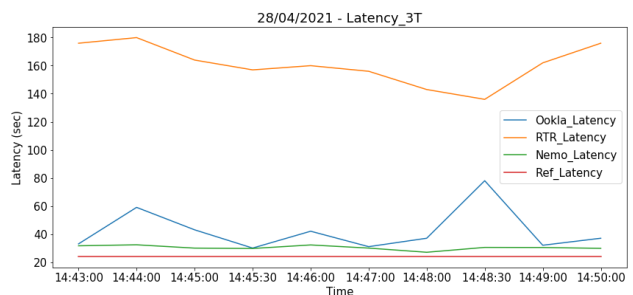


Figure 4.11: Plot of latency throughput using the three tools

Table 4.2: Average and evaluations of the parameters using the tools

| Tools | Evaluation | | | | | |
|-----------------|------------|----------|---------|----------|--------------|----------|
| | Avg. DL | MAPE (%) | Avg. UL | MAPE (%) | Avg. Latency | MAPE (%) |
| Ookla Speedtest | 14.93 | 57.35 | 22.27 | 30.42 | 38.21 | 42.06 |
| RTR NetTest | 5.52 | 84.22 | 13.93 | 56.47 | 160.26 | 493.56 |
| Nemo Handy -A | 18.36 | 47.54 | 23.91 | 25.27 | 30.69 | 13.66 |

V. CONCLUSION AND RECOMMENDATION

1. CONCLUSION

The objective of the thesis was to evaluate data service QoS measurement factors and tools based on the DL, UL, and latency measurement parameters. The measurement factors, which are the notion of measurement tool, measurement techniques of the tools, and network status, have a significant impact on QoS measurements. We concluded that data QoS measurement using various techniques and tools for the same specific operator’s network is different. The reasons for the measurement difference, which include factors beyond the measurement tools, the notion of measurement, and the tool’s measurement techniques, are discussed in this thesis. According to the results, the throughput measurement is highly affected by the number of parallel connections or threads and traffic level that indicate the load status of the network. The main factors for latency measurement are server location and the estimation methods used by the measurement tools. When the QoS assessment tool requires configuration for the number of parallel connections or threads, the traffic and signal strength level should be taken into consideration. The measurement error of the Nemo Handy-A is less than the other tools in all parameters based on the Ethio telecom 4G LTE network. The main factors for those measurement differences are the location of the measurement server, the bottleneck, and the parameter estimation technique of the tools.

Even though Nemo Handy-A is a standard measurement tool, it costs time, money, analysis time, unfair representation measurement area, and hours relative to the crowdsourcing methods. If the challenges of CS measurement tools are handled carefully, then using the CS tool as a QoS measurement is advantageous. When we select QoS measurement tools for the dedicated network, we should have considered the setup of their measurement factors. Using the crowdsourcing measurement method can overcome the limitations of the drive test method, even though limited in QoS parameters, and inject additional traffic into the active network.

2. RECOMMENDATION

Some recommendations from the process and work of the thesis are as follows. During evaluating the QoS of the Operator's network, the measurement server should be within the network service area. When telecom operators, Internet service providers, and governmental regulators use QoS

measurement tools, they should evaluate the measurement techniques and tools based on test factors.

Even though the Nemo Handy-A measurement tool showed better accuracy for a given scenario, by applying careful post-processing using crowdsourcing measurement tools during data QoS measurement has an advantage. We recommended Ethio telecom to use the crowdsourcing measurement tools for fair distribution of the measurements, real-time operation, reduced measurement cost, and access to measurement data at any time and location that used for network troubleshooting, which is a fast input for optimization and maintenance.

Here are the recommendations for further thesis on this topic:

- Compare the QoS measurement platforms on local measurement servers that installed on the network coverage of the operator’s network.
- Evaluate the QoS measurement tools in case of mobility through the selected area network coverage.

REFERENCE

- [1] C. Cox, “An Introduction to LTE, LTE-Advanced, SAE, VoLTE and 4G Mobile Communications,.” Published by John Wiley & Sons Ltd, UK, 2014.
- [2] S. Haryadi, “Telecommunication Quality of Service Concept: Network Performance and Quality of Service,” 2016.
- [3] Reesom Bisrat A., Haile B. B., Mutafungwa E., & Hämäläinen J., “Quality Evaluation for Indoor Mobile Data Customers in Addis Ababa Business Area Using Data from Network Management System, Walk Test, Crowdsourcing and Subjective Survey; In E. Nigussie, T. Tegegne, & F. Mekuria (Eds.), *Information and Communication Technology for Development for Africa - 2nd International Conference, ICT4DA 2019, Revised Selected Papers* (pp. 164-175), (2019).
- [4] B. Alemayehu, “Evaluation of User Side Quality of Service Measurement Techniques for Addis Ababa LTE Data Service,” *MSc Thesis, AAU, 2018*.
- [5] E. Lipenbergs, V. Bobrovs, and G. Ivanovs, “Investigation of Service Quality of Measurement Reference Points for the Internet Services on Mobile Networks,” *Journal Paper, Latv. J. Phys. Tech. Sci.*, 2016, doi: 10.1515/lpts-2016-0037.
- [6] C. Midoglu and P. Svoboda, “Opportunities and challenges of using crowdsourced measurements for mobile network benchmarking a case study on RTR open data,” in *Proceedings of 2016 SAI Computing Conference, SAI 2016*, 2016, doi: 10.1109/SAI.16.7556101.
- [7] S. Bauer, D. Clark, and W. Lehr, “Understanding broadband speed measurements,” *Telecommun. Policy Res. Conf.*, pp. 1–39, 2010.
- [8] ECC Report 195, “Minimum Set of Quality of Service Parameters and Measurement Methods for Retail Internet Access Services”, 2013.
- [9] A. Rusan and R. Vasiu, “Emulation of backhaul packet loss on the LTE S1-U interface and impact on end user throughput,” *Proc. - 2015 IEEE 11th Int. Conf. Intell. Comput. Commun. Process. ICCP 2015*, pp. 529–536, 2015, doi: 10.1109/ICCP.2015.7312715.
- [10] E. Dahlman *et al.*, “The 3G long-term evolution - Radio interface concepts and performance evaluation,” *IEEE Veh. Technol. Conf.*, vol. 1, no. c, pp. 137–141, 2006, doi: 10.1109/VETECS.2004.1387928.
- [11] M. Mahfoudi *et al.*, “LTE network capacity analysis to avoid congestion for real time traffic,” *Mediterr. Microw. Symp.*, vol. 2015-April, pp. 3–7, 2015, doi: 10.1109/MMS.2014.7088980.
- [12] The Speedtest Server Network. May 27, 2021, [online]. Available: <https://www.ookla.com/speedtest-servers>.
- [13] W. Li *et al.*, “Toward Accurate Network Delay Measurement on Android Phones,” *IEEE Trans. Mob. Comput.*, vol. 17, no. 3, pp. 717–732, 2018, doi: 10.1109/TMC.2017.2737990.



- [14] A. Gaber, “alladin-IT Nettetst Measurement Algorithm”, Description, 2017.
- [15] Keysight Technologies, “Nemo Handy Handheld Measurement Solution,”keysight.com
<https://www.keysight.com/zz/en/product/NTH00000B/nemo-handy-handheld-measurement-solution.html>. (accessed June 05, 2021).
- [16] B. Connelly, “How Ookla Ensures Accurate, Reliable Data: A Guide to Our Metrics and Methodology.” Available:
<https://www.speedtest.net/about/knowledge/test-methods>. [Accessed May 23, 2021].
- [17] I. Koukoutsidis, “Public QoS and Net Neutrality Measurements,” *J. Inf. Policy*, vol. 5, no. 2015, pp. 245–286, 2015, [Online]. Available: <http://www.jstor.org/stable/10.5325/jinfopoli.5.2015.0245>.
- [18] U. Manual, “Nemo Handy-a,” pp. 1–242, 2015, [Online]. Available: www.Anite.com/nemo.
- [19] S. Homayouni et al, “The impact of duration and settings of TCP measurements on bandwidth estimation in mobile networks,” *IEEE Int. Symp. Pers. Indoor Mob. Radio Commun. PIMRC*, vol. 2017-Octob, pp. 1–6, 2018, doi: 10.1109/PIMRC.2017.8292522.
- [20] W. Li et al, “On the accuracy of smartphone-based mobile network measurement,” *Proc. - IEEE INFOCOM*, vol. 26, pp. 370–378, 2015, doi: 10.1109/INFOCOM.2015.7218402.
- [21] ETSI EG 203 165, “Speech and multimedia Transmission Quality (STQ); Throughput Measurement Guidelines,” 2016.
- [22] V. Raida et al, “Constant Rate Ultra Short Probing (CRUSP): Measurements in Live LTE Networks,” *IEEE Int. Conf. Commun.*, vol. 2019-May, 2019, doi: 10.1109/ICC.2019.8761179.
- [23] K. Kousias et al, “Estimating downlink throughput from end-user measurements in mobile broadband networks,” *20th IEEE Int. Symp. A World Wireless, Mob. Multimed. Networks, WoWMoM 2019*, no. M1, 2019, doi: 10.1109/WoWMoM.2019.8792968.
- [24] R. KLUS, “Big data analytics in the context of mobile network performance optimization.”, Thesis, Brno University of Technology, The Faculty of Electrical Engineering and Communication, Department of Radio Electronics, 2019.

