Addis Ababa University

Addis Ababa Institute of Technology

School of Electrical and Computer Engineering

Telecommunication Engineering Graduate Program

QoE Model for Addis Ababa LTE Web Browsing Service Using Neural Network Approach

By

Dereje Solomon Tibebu

Advisor

Dr. Eng. Yihenew Wondie

A Thesis Submitted to School of Electrical and Computer Engineering, in Partial Fulfillment of the Requirements for the Degree of Masters of Science in Telecommunications Engineering

December 2019

Addis Ababa, Ethiopia

Addis Ababa University

Addis Ababa Institute of Technology

School of Electrical and Computer Engineering

QoE Model for Addis Ababa LTE Web Browsing Service Using Neural Network Approach

Signed by the Examining Committee:

____________________	__________	________
Chair or School Dean	Signature	Date
Dr. Eng. Yihenew Wondie	__________	________
Advisor	Signature	Date
____________________	__________	________
Examiner	Signature	Date
____________________	__________	________
Examiner	Signature	Date

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.

Dereje Solomon Tibebu ________________

Name Signature

Place: Addis Ababa

Date of Submission: ________________

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

Dr. Eng. Yihenew Wondie ________________

Advisor Signature

Abstract

In order to address the customer's satisfaction, mobile operators try to find out what the customer needs and what quality makes the customer satisfied. The customer satisfaction can be measured or estimated by Quality of Experience (QoE) measurement. Its estimation and measurement is important to identify the network problems, to understand causes and contributing factors.

Web browsing is one of the widely used application on Long Term Evolution (LTE) networks. Therefore, it is essential for service providers to ensure a better QoE on web browsing service. Web QoE can measure the user satisfaction by subjective or objective measurement. Subjective test suffers from some drawbacks, such as it has high cost in terms of time, money, and manual effort and also cannot be used for real-time QoE evaluation. In Ethiopia only subjective measurement is used, to know the level of customer satisfaction. Due to that, the company is exposed for high expenses and also can not perform the real time measurement of QoE.

To overcome the problem on subjective test, this thesis developed a web browsing QoE model, using Neural Network algorithm that is implemented in matlab software. The model takes the following QoS metrics as input parameters: page response delay, page content browsing delay and page download throughput. The model map these metrics to QoE interms of Mean Opinion Score (MOS).

The model performed an estimation of QoE with a Mean Square Error (MSE) of 0.002 and correlation of 97.2%, relatively to the target QoE. As the result indicates, the estimated and measured QoE values are highly correlated. And the error between them is very low. So, this model can be used for estimating the web browsing QoE for the mobile operators, to get objective measurement advantages. Also, it can be used for operators to identify the network factors that most influence the web browsing QoE.

Keywords: *QoE, QoS, LTE, MOS, Estimation, Web Browsing, Model, Neural Network, Subjective Measurment and Objective Measurment.*

Acknowledgment

I praise God, the almighty for providing me this chance and granting me the capability to continue successfully. This research appears in its current form due to the support and guidance of several people. Therefore, I would like to offer my heartfelt thanks to all of them.

I am grateful to my advisor Dr. Eng. Yihenew Wondie, for his continuous support, patience and guidance throughout our thesis work to finish successfully. He used to teach me a lot about the research work and guide me to work in the right approach by correcting my mistakes to improve the quality of my thesis. He gave me his valuable feedback at each and every moment which helped me to do my work more.

I am also very grateful to thank employees of ethio telecom (Departments of Service Management Center, Engineering and Customer Service) who have supported me in collecting data and for their technical guidance. Also, my deepest thank goes to my wife and daughters for their emotional support and encouragement throughout my postgraduate program.

Finally, a special acknowledgment to ethio telecom for giving me the opportunity to study my postgraduate program at Addis Ababa University, Addis Ababa Institute of Technology (AAiT).

Table of Contents

List of Figures

List of Tables

Abbreviations and Acronyms

2G	Second Generation
3G	Third Generation
3GPP	Third Generation Partnership Project
4G	Fourth Generation
ACK	Acknowledgement
AMOS	Application Mean Opinion Score
ANN	Artificial Neural Network
AuC	Authentication Center
CAAZ	Central Addis Ababa Zone
CDN	Content Delivery Network
CS	Core Network
CSP	Communication Service Provider
DNS	Domain Name System
DOM	Document Object Model
DT	Derive Test
EAAZ	East Addis Ababa Zone
EPC	Evolved Packet Core
EPS	Evolved Packet System
E-RAB	Evolved Radio Access Bearer
E-SMLC	Evolved Serving Mobile Location Centre
E-UTRA	Evolved UMTS Terrestrial Radio Access
FTP	File Transfer Protocol
GBR	Guaranteed Bit Rate
GMLC	Gateway Mobile Location Centre
GSM	Global System for Mobile

GTP	GPRS Tuning Protocol
HSDPA	High Speed Download Packet Access
HSS	Home Subscriber Server
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
IF	Influence Factors
IP	Internet Protocol
Kbps	Kilo bit per second
KQI	Key Quality Indicator
MME	Mobility Management Entity
MOS	Mean Opinion Score
MSE	Mean Square Error
NAAZ	North Addis Ababa Zone
NAS	Non Access Stratum
NN	Neural Network
OS	Operating System
PBSR	Page Browsing Success Rate
PCBD	Page Content Browsing Delay
PCRF	Policy Control and Charging Rules
PDN	Packet Data Network
PDP	Packet Data Protocol
PDTH	Page Download Throughput
P-GW	PDN Gateway
PRD	Page Response Delay
PRSR	Page Response Success Rate
PS	Packet Switch

QCI	QoS Class Identifier
QoE	Quality of Experience
QoS	Quality of Service
R	Correlation Coefficient
R Value	Correlation Value
R99	Release-99
RAB	Radio Access Bearer
RAN	Radio Access Network
RMSE	Root Mean Squire Error
RTT	Round Trip Time
SAAZ	South Addis Ababa Zone
SAE	System Architecture Evolution
S-GW	Serving Gate Way
SSR	Sum of Squares of Regression
SST	Sum of Squares of Total
SVR	Support Vector Regression
SWAAZ	South West Addis Ababa Zone
SYN	Synchronization
TCP	Transport Control Protocol
UE	User Equipment
UMTS	Universal Mobile Telecommunications System
URL	Uniform Resource Locator
UTRA	UMTS Terrestrial Radio Access
VoIP	Voice over Internet Protocol
WAP	Wireless Application Protocol

Chapter 1 Introduction

This chapter presents background of the thesis and explains its motivation, objectives, scope and contribution of this thesis. Also, it discusses applied methodology and explains selected reviewed related literatures. Finally, thesis organization is outlined.

1.1. Background

Penetration of mobile data service is significantly increasing due to the expansion of mobile network and the advancement of the mobile device technologies. In the case of Addis Ababa, Ethiopia, the mobile data service demand is exponentially increasing [1]. Figure 1-1 shows, the mobile data demand forecast in Addis Ababa, Ethiopia from 2017 to 2021.

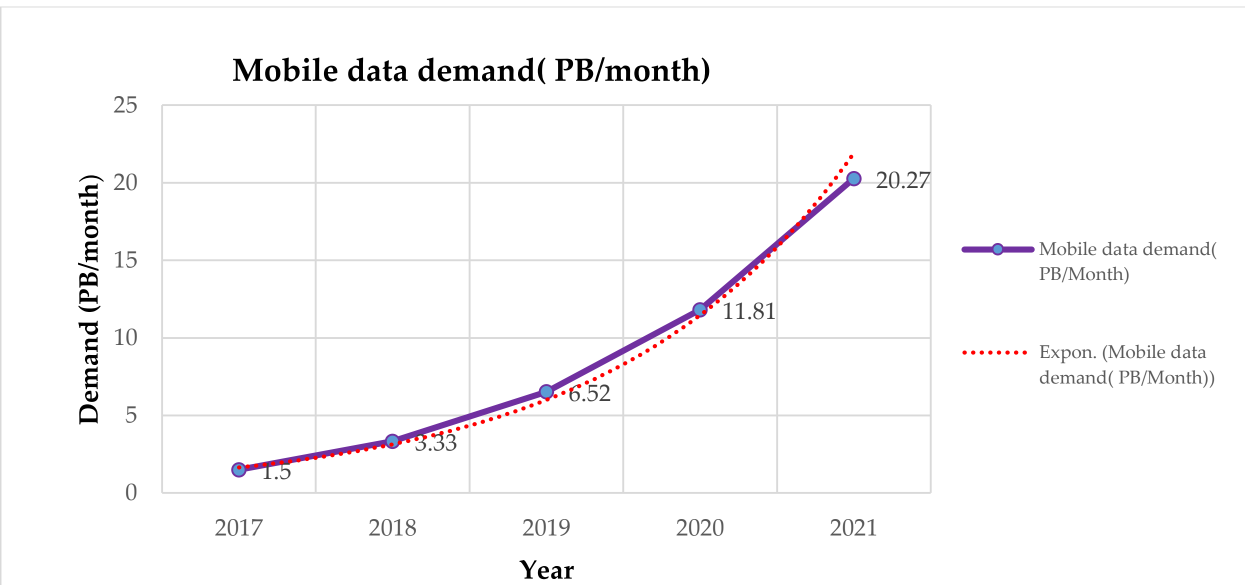


Figure 1-1 Mobile Data Demand Forecast in Addis Ababa (2017-2021) [1].

Ethio telecom deployed Universal Mobile Telecommunication System (UMTS) technology to cover large area of Ethiopia but Long Term Evolution (LTE) technology deployed only in Addis Ababa city to deliver high speed data service [2]. LTE standard was introduced in Release 8 of Third Generation Partnership Project (3GPP) [3]. The above LTE standard was developed to completely dedicated to packet-switched services. LTE permits higher throughput and spectral efficiency in addition to lower latency and a more flexible channel bandwidth, compared to UMTS [4]. As we know, there are different mobile data service, some of them are streaming, web browsing, email etc. Figure 1-2 depicts the mobile data traffic in Addis Ababa for different mobile data services. Among all the data services, web browsing is one of the dominant data service on LTE network.

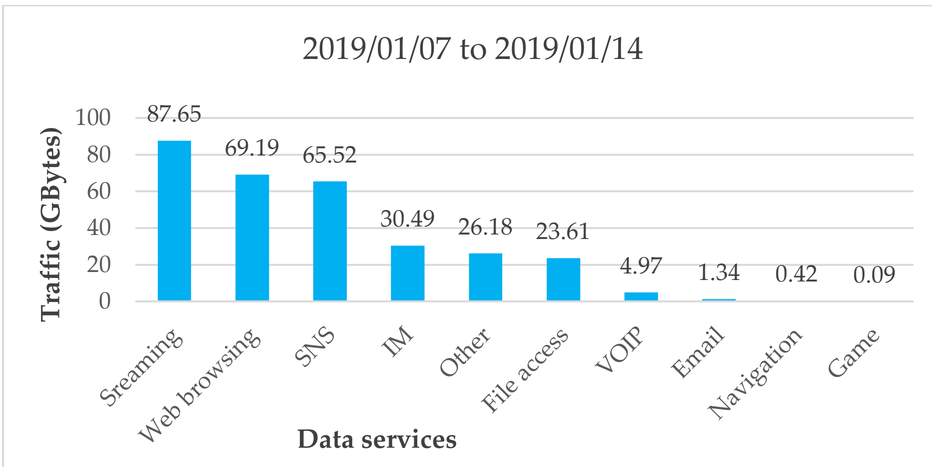


Figure 1-2 Addis Ababa Traffic Distribution for Mobile Data Service.

Insuring a better Quality of Experience (QoE) is essential for service providers, to successfully serve mobile data customers, when they use web browsing service. Survice providers must seek new approaches to find out what quality makes the customers satisfied [5].

The Quality of Service (QoS) has been studied for a long time, but attention to QoE is more recent. Among service providers, network operators and equipment manufacturers the term QoS has been in use for a long time and has reached a high level of common understanding. QoS work is based on technical performance (i.e., it is mainly technology-centered) whereas QoE is based on end-user behavior (i.e., user-centered). It is argued that task on QoS is essential, but not satisfactory, for measuring user experience [6].

Studying and measuring QoE is characteristically difficult task, due to its subjective nature. Now, there are two methods to determine the QoE or the satisfaction level of a customer with a service. These methods are subjective and objective methods. Subjective method is the most informative approach, where researchers using questionnaires and interviews to directly obtain user opinion. Objective measurment is performed by understanding the relationship between QoS parameters with QoE.

There are factors that influence web browsing user perceptions. Before studying objective methods of QoE for web browsing, these influencing factors must be identified. Characteristically, web browsing QoE influence factors (IFs) are classified into three main categories. The influence factors are user influence factors, context influence factors and system influence factors. User influence factors includes usage history, demographic background and expectation. Context influence factors are a serious influence on the customer behavior and web browsing QoE. Some of the context influence factors are location (e.g., cafeteria, office and home), interactivity (e.g., high level interactivity vs. low level interactivity), task type (e.g., business and entertainment) and task urgency (e.g., urgent and casual) [6].

System influence factors are the type of influence factors that includes server related, delivery network, content related and client influence factors. The ability to know the customer satisfaction or QoE for the web browsing service is needed to identifying network conditions that are contributing for degraded the QoE [6]. Considering the correlation between web browsing QoE and network influences helps the operators to detect problems, to know causes and contributing factors.

1.2. Statement of the Problem

Even though the network performance in LTE is better than GSM and UMTS, knowing the user experience is another issue to be considered. In Ethiopia where the case study is conducted, QoE measurment is done based on customer survey or subjective measurment to know the customer satisfaction on web browsing service. The data from Table 1-1 illustrates, the perception of the customers' in different technologies and time, the ratings are made on a 1 to 10-point scale. As one can see the data from Table 1-1, the user perception on LTE is lower than 3G on some rounds. So measuring the user perception on LTE web browsing service can be an issue like UMTS web browsing services.

Table 1-1 User Perception on Loading Time of Web Page on 3G and 4G [7].

Network Performance Indicators	User Perception (Scored from 1 to 10) in Different Survey Round			
	5th Round	6th Round	7th Round	9th Round
Loading time of Web Page on 4G	8.2	5.69	7.3	7.4
Loading time of Web Page on 3G	7.4	7.14	7.5	6.8

QoE in LTE web browsing services can measured by subjective or objective measurement. But Subjective test has high cost in terms of time, money, and manual effort. It is conducted in laboratory environment with limited test and the test cannot be used for real time.

In the case of ethio telecom, QoE measurement is done based on customer survey or subjective measurements [7]. Due to that the company is exposed for high expenses and also cannot perform the real time measurement of QoE. In order to avoid the high cost of subjective test, developing objective QoE measurement or model is required, by understanding the relationship between web browsing QoE and network factors. The develop model is required for telecom operators get the benefit of objective measurement such as less cost in terms of money, time and manual effort and also the company can do real time measurements by using the developed model. In addition to that the model helps the company to identify problems, to understand causes and contributing factors.

1.3. Objective of the Thesis

1.3.1. General Objective

The purpose of this research is modeling QoE for mobile web browsing based on QoS measurements on LTE networks in Addis Ababa using Neural Network.

1.3.2. Specific Objectives

To accomplish the general objective, the following specific objectives were set.

- To grasp the global trend on web browsing QoE, estimation mode and Neural Network;
- To perform web browsing QoS data collection from LTE network in Addis Ab-aba;
- To identify input parameters for the model;
- To obtain the target value to train the Neural Network;
- To develop the web browsing QoE estimation model by using Neural Network algorithm on MATLAB;
- To do performance analysis and evaluation of the estimated and measured QoE.

1.4. Methodology

Once the problem is known and objectives are set, this research presents the following methodology to address the identified problem. The methodology is presented by using work flow diagram in Figure 3-1 with some explanations on this thesis work.

The research was conducting by modelling the QoE for LTE web browsing services in Addis Ababa. Literature review, data collection and data analysis or data pre-processing activities were performed before starting to develop modeling QoE. The model was de-veloping using Neural Network approach, more specifically the algorithm used is Artificial Neural Networks (ANN) in Matlab software.

The web browsing QoS parameters and the MOS values were inputs to learn the ANN for developing the web QoE model. The testing step is the test to verify whether the proposed model is valid or not. Take the measured QoS parameters and then put these parameters to QoE model to get the estimated QoE value. Before, we use the proposed model for estimating the QoE, the model test will be performing to verify whether the proposed model is valid or not. The performance of the proposed model will identify the level of accuracy to estimate the web QoE value. The metrics to calculate the performance of the developed model are MSE and R-Squared.

The MSE tells how close the model output is to the target values, the smaller the MSE, the model value is nearer to the target and vice verse. And R-Squared indicates the correlation between the model output and the target value; high value of R- Squared represent a relationship between model output and target value while low values mean that the proposed estimated model is not appropriate for required purpose. Finally, based on the obtained results the model will be ready to estimate the web QoE values.

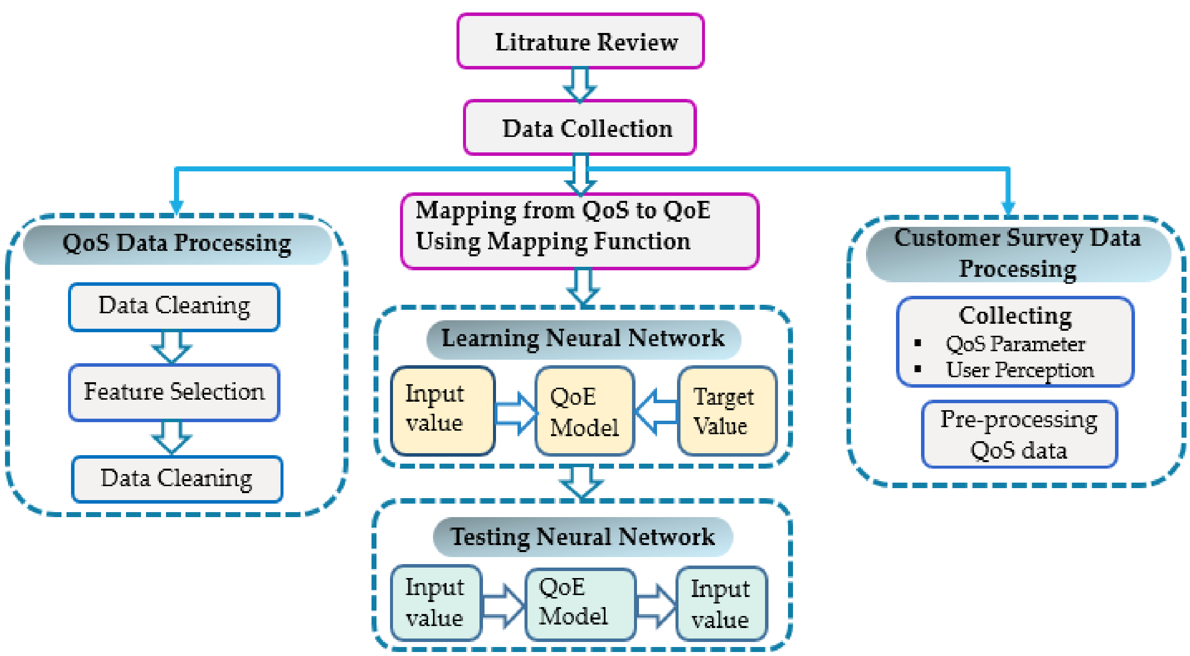


Figure 1-3 The Methodology Used.

1.5. Selected Related Literature Review

To understand state-of-the-art on mobile QoE, authors in [8] presents literatures survey focused on QoE modeling and measurement. They show the measurement and the modeling by quantifying the QoE, mapping of network-level QoS to QoE and taking QoE models for specific applications such as web browsing, video application, audio application, online multiplayer gaming. They have also taken a more in depth look in to mobile QoE, motivated by the fact that QoE measurement in mobile networks is not only rising in importance but also more challenging compared to fixed networks. Inaddition to presenting the state-of-the-art on mobile QoE, they defined some of the key exceptional challenges.

To develop the QoE model from the QoS metrics the authors in [9] presents a novel QoE prediction model, for web browsing based on real LTE data. The model estimates the MOS value by evaluating the network measurements and QoS parameter to know user perceived quality. The authors used the real Drive Test (DT) data and MOS measurements were used as reference data, in order to produce a new QoE prediction model, using machine learning techniques. More specifically, they used Support Vector Regression (SVR) algorithm to map the QoS metrics in to MOS. The developed model showed a correlation of 92% and RMSE of 10%. The new developed model enables the application of QoE as a more realistic network optimization criterion.

To demonstrate the use of Neural Network approach to create QoE model in the assessment of user satisfaction resulted from the QoS parameters instead of directly measuring the MOS from users, authors in [10] proposes the estimation model of user satisfaction in terms of QoE by using neural network approach. The author analyzes the data by ANN tool box in matlab software and the input of this model obtained from QoS parameter and Application Mean Opinion Score (AMOS) in each user of different networks were used to learn inneural networks to get the right weights in QoE correlation model. The final step to develop this model was testing the model by comparing the value from the proposed model and MOS value to validate the model. The result shows that the good agreement between QoE value from the proposed model and QoE value from the AMOS applications.

Web QoE model is used to know the relationship between web QoE and and network characteristics to identify when and where degreded on the web browsing QoE. Authors in [11] presented a large-scale study that studied web QoE, such as session length, abandonment rate and partial download ratio, from a “cellular network operator" point of view. The first step they performed is developing a machine learning method to estimate web QoE metrics from network trace appropriately. The machine learning mechanism used to precisely model the influence of radio network characteristics on user experience metrics. This model can be used by cellular network operators to detect the improvement of network elements that most influence web QoE.

In [12], authors assessing the web browsing QoE by studing the existing objective quality assessement model. Also, they compare the studied existing model performances with simulations to observe their individual advantages and limitations. The result from the Simulation depicts that the recommended QoE model can be useful for evaluating the quality of different web sources in the LTE networks by considering some property of mobile networks. A fitting model is presented to show the relationship between network QoS and user QoE gained in subjective lab test. To avoid the limitations of the existing models, this paper also presents an enhanced QoE model which incorporate the effects of parameters such as page download size, the packet losses and connection throughput in QoE assessment.

1.6. Scope of the Thesis

The scope of this thesis is to develop a model that estimate LTE web browsing QoE by using Neural Network approach based on web browsing QoS measurement. Even though QoE modeling concept is the same for all mobile technologies, this thesis work is restricted to only LTE network.

1.7. Contribution

Estimating model for QoE of LTE web browsing service is essential for any telecom operators to get the benefit of objective measurement. The benefit of objective measurements is avoiding high cost in terms of time and money, also it allows real time QoE measurement. Currently, ethio telecom uses only subjective measurement to know the customer satisfaction, due to that the proposed objective web browsing QoE model is required to obtain the above objective model advantage. The other contribution of the proposed work is identifying individual impact of QoS parameters on LTE web browsing QoE.

1.8. Thesis Organization

This thesis is organized in six chapters. The first chapter states a brief introduction to the proposed work. Chapter 2 presents technical background on LTE networks, web browsing service in LTE and concepts on QoE. Chapter 3 introduces the neural network approach that includes the data preprocessing and validation metrics. Chapter 4 presents all the steps that conducted to the development of the proposed model and identifies the QoE when using LTE. Chapter 5 presents the result obtained with this model and discussion on the obtained results. Finally, Chapter 6 presents conclusion of the thesis and the future work.

Chapter 2 Technical Background

This chapter delivers the technical background that are essential to understand this research. Since, this research is performing to model QoE for LTE web browsing service, having theoretical background on LTE network and the service delived by LTE network is essential. So, to understand the theoretical concept on LTE, Web browsing service and QoE for LTE web browsing service are discussing in this chapter. Also this chapter presents, the technical definitions related to the LTE network, web browsing service and QoE on the web browsing service, to understand their signigicance.

2.1. Long Term Evolution

The LTE standard was introduced in Release 8 of 3GPP [13]. It has been proposed to support only Packet Switched (PS) services, in contrast to the Circuit-Switched (CS) model of previous cellular systems. LTE technology provides seamless Internet Protocol (IP) connectivity between User Equipment (UE) and the Packet Data Network (PDN), without any disturbance to the end users' applications during mobility. Also it permit higher throughput and spectral efficiency in addition to lower latency and a more flexible channel bandwidth, relatively to UMTS. While the term 'LTE' consist of the evolution of the radio access through the Evolved-UTRAN (E-UTRAN), it is accompanied by an evolution of the non-radio aspects under the term 'System Architecture Evolution' (SAE) which contains the Evolved Packet Core (EPC) network. Together LTE and SAE comprise the Evolved Packet System (EPS) [14].

2.1.1. LTE Architectural Overview

EPS provides the user with IP connectivity to a PDN for accessing the Internet, in addition to for running services. An EPS bearer is typically related with a QoS. Multiple bearers can be set for a user in order to deliver different QoS streams or connectivity to different

PDNs. Figure 2-1 shows the overall network architecture including the network elements and the standardized interfaces. At a high level, the network is consist of of the CN (i.e. EPC) and the access network (i.e. E-UTRAN). While the CN contains many logical nodes, the access network is made up of essentially just one node, the evolved NodeB (eNodeB), which connects to the UEs. Each of these network elements is inter-connected by means of interfaces which are standardized to permit multivendor interoperability [14].

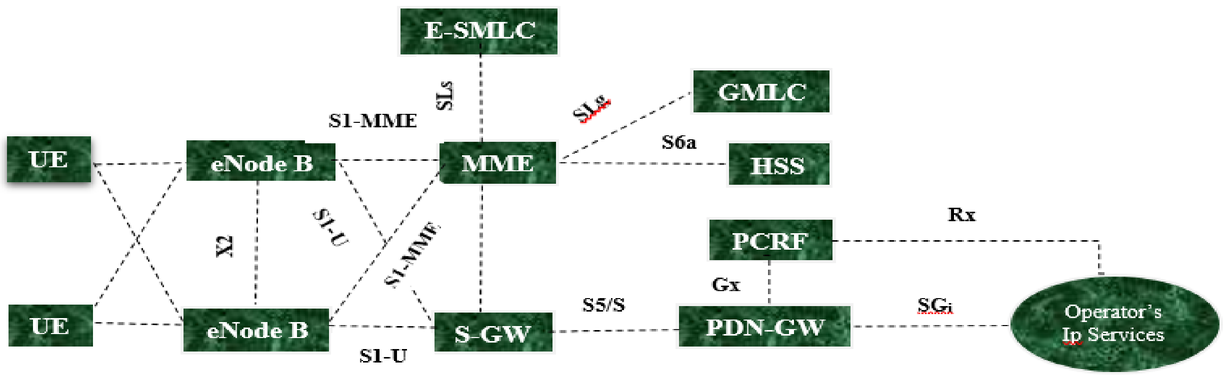


Figure 2-1 LTE Network Architecture [14].

The Core Network

The core network or EPS is responsible for the overall control of the UE and the establishment of the bearers. The EPC is composed by the following network elements[14]:

PDN Gateway (PDN-GW) - links the EPS with the PDN. It also assigns the Internet Protocol (IP) addresses designated for the UE.

Serving Gateway (SGW) - responsible for keeping information concerning the bearers during the UE's iddle mode and serves as the local mobility anchor for the data bearers when the UE moves between eNodeBs.

Mobility Management Entity (MME) - responsible for the transition between the access radio network and the EPC. This network element processes the signaling between the UE and the core network.

Home Subscriber Server (HSS) - The HSS has users' SAE subscription data such as the EPS-subscribed QoS profile and any access restraints for roaming.

Evolved Serving Mobile Location Centre (E-SMLC) - Controls the whole coordination and scheduling of resources necessary to get the location of a UE that is attached to E-UTRAN. It also computes the final location based on the estimates it receives, and it estimates the UE speed and the achieved accuracy.

Gateway Mobile Location Centre (GMLC) - comprises functionalities essential to support Location Services (LCS). After accomplishing authorization, it sends positioning requests to the MME and obtains the final location estimates.

Policy Control and Charging Rules Function (PCRF) - responsible for policy control decision making and controlling the flow-based charging functionalities.

The Access Network

The access network of LTE, E-UTRAN, simply has only eNodeBs. For ordinary user traffic (as opposed to broadcast), there is no integrated controller in E-UTRAN; hence the E-UTRAN architecture is said to be flat [14].

The eNodeBs are inter-connected with each other by means of an interface called X2, and to the EPC by means of the S1 interface specifically, to the MME by means of the S1-MME interface and to the S-GW by using S1-U interface. The E-UTRAN is accountable for all radio-related activities [14].

2.1.2. LTE QoS Architecture

The EPS bearer service layered architecture is depicted in Figure 2-2. It does not look very different from the bearer service architecture defined in Release 99. However, there is a major difference that is not obvious at first vision. In 3G UMTS a subscribe request for a defined QoS of an end-to- end service starts the QoS negotiation procedure. This depends on the subscriber's subscribed QoS stored in the Home Location Registor (HLR) and the available network resources which QoS is granted to a particular connection at the end. In LTE different to 3G PS connections, a default bearer with a default QoS is already established when the UE attaches to the network [15].

The QoS attributes of this default bearer are determined by the subscribed QoS parameters stored in the HSS [15]. It is essential to know that one UE in LTE can have multiple end-to-end services active and each of these services will have its own individual bearer.

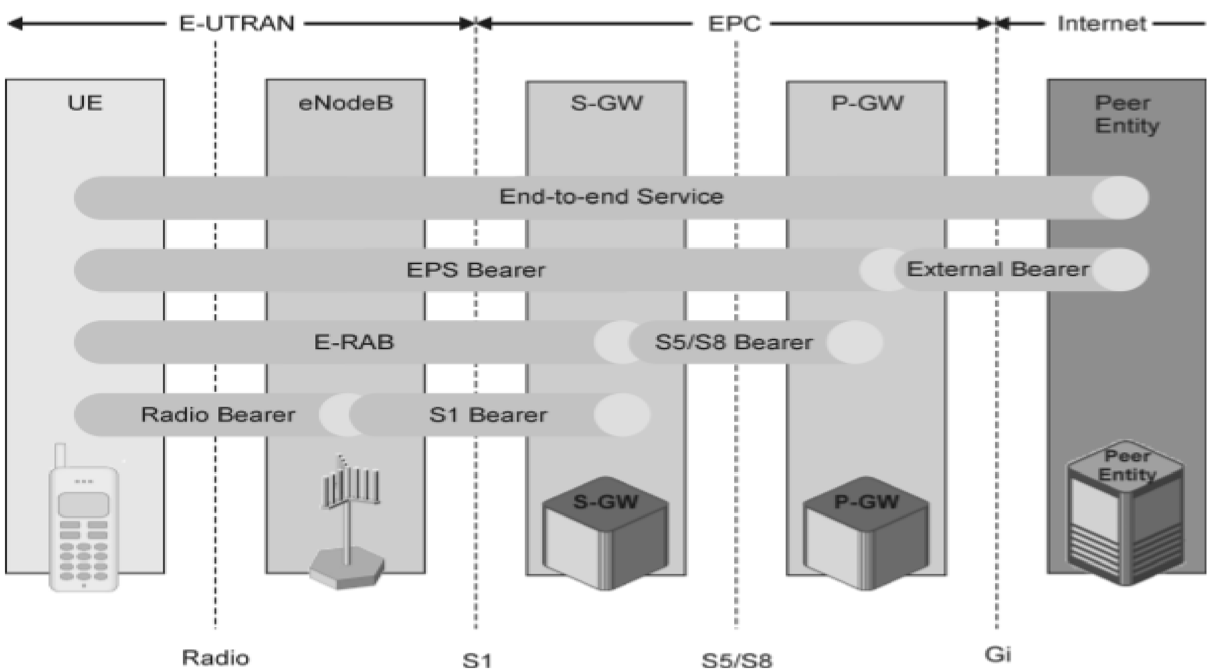


Figure 2-2 LTE QoS Architecture [15]

Therefore, 256 individual E-RABs for a single UE can be addressed by E-UTRAN protocols whereas in UMTS only 15 different RAB-IDs had been defined by the standard organizations. To standardize the QoS handling, nine QCIs have been defined by 3GPP. Among the nine classes, four classes with a Guaranteed Bit Rate (GBR) and five classes with a Non-Guaranteed Bit Rate (Non-GBR). Besides the bit rate, the parameter priority, packet delay budget, and packet error loss rate are critical influences [15].

2.2. Web Browsing Service

This section describes some basic concepts on the web browsing service. Defining and explaining the web browsing QoS parameters that are related with the web browsing services also discussed, to take the parameters as inpts for the developing model.

2.2.1. Web browsing Service Overview

The browsing service permits a mobile user to browse the web by using a browser installed on a mobile phone. The user may either type in a Uniform Resource Locator (URL) or click a link in order to access a web page. A mobile phone may use Hyper Text Transfer Protocol (HTTP) to communicate directly with a web server [16].

As shown in Figure 2-3, a web page is requested by sending a GET request message to a web server or a Wireless Application Protocol (WAP) gateway that in turn sends the request to the server. If the operation is successful, the web server will reply with a GET response message that contains the requested web page. A web page may have multiple embedded objects such as images, each object is fetched with a separate GET request message. A GET response message from a web server comprises a status code to indicate the result of the operation [16].

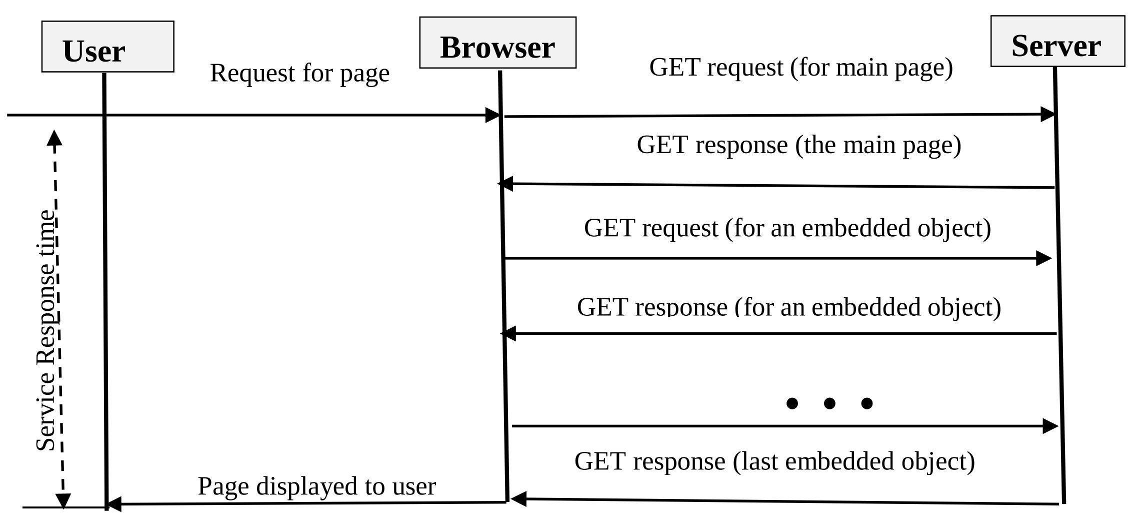


Figure 2-3 Browsing Service Message Flow [16].

Web Page Transfer Time (Total Page Load Time)

This indicator quantifies the average time required for the transfer of web page. This indicator only takes account of successfully completed sessions (session that allowed the full transfer of web pages) [17].

$$Total\ Page\ Load\ Time\ [s] = (t_{\text{end reception}} - t_{\text{webpage request}})\ [s] \quad (2.1)$$

Where:

$t_{\text{webpage request}}$ - Moment when the user equipment makes the webpage perception request.

$t_{\text{end reception}}$ - Moment when the full webpage received by the user equipment.

2.2.2. Key Phases of Total Load Time

In web browsing, users transfer a number of HTTP objects from Server to their web browser. It is take place within a few particular phases. These phases are Domain Name Sytem (DNS) Lookup, Transport Control Protocol (TCP) Handshake, request for Base Hyper Text Markup Language (HTML) and request for Embedded Objects. According to ITU-T G.1030 [18], Total page load time is divided into two parts: the time to search and the time to download a web page (contents). The searching time or the page response delay includes DNS Lookup, TCP Handshake and the time to request for Base HTML. The remaining time is that are not included in page response delay is the time to download the content or content browsing delay. Figure 2-4 illustrates the above mentioned phases of HTTP Object transfer [18]. Below are brief explanation for each of the phase of total page load time.

DNS Lookup

To access a website from any browser, a user generally refers to the web address by name, not by IP address. This named web address is then queried to a DNS server to get the server IP for the particular web site. DNS queries are resolved in a few different ways.

A browser can resolve a query locally from the cached information in the browser or Operating System (OS) cache from a previous lookup. When a local query is unsuccessful then it is passed to a DNS server to resolve the name and then send a response back [19].

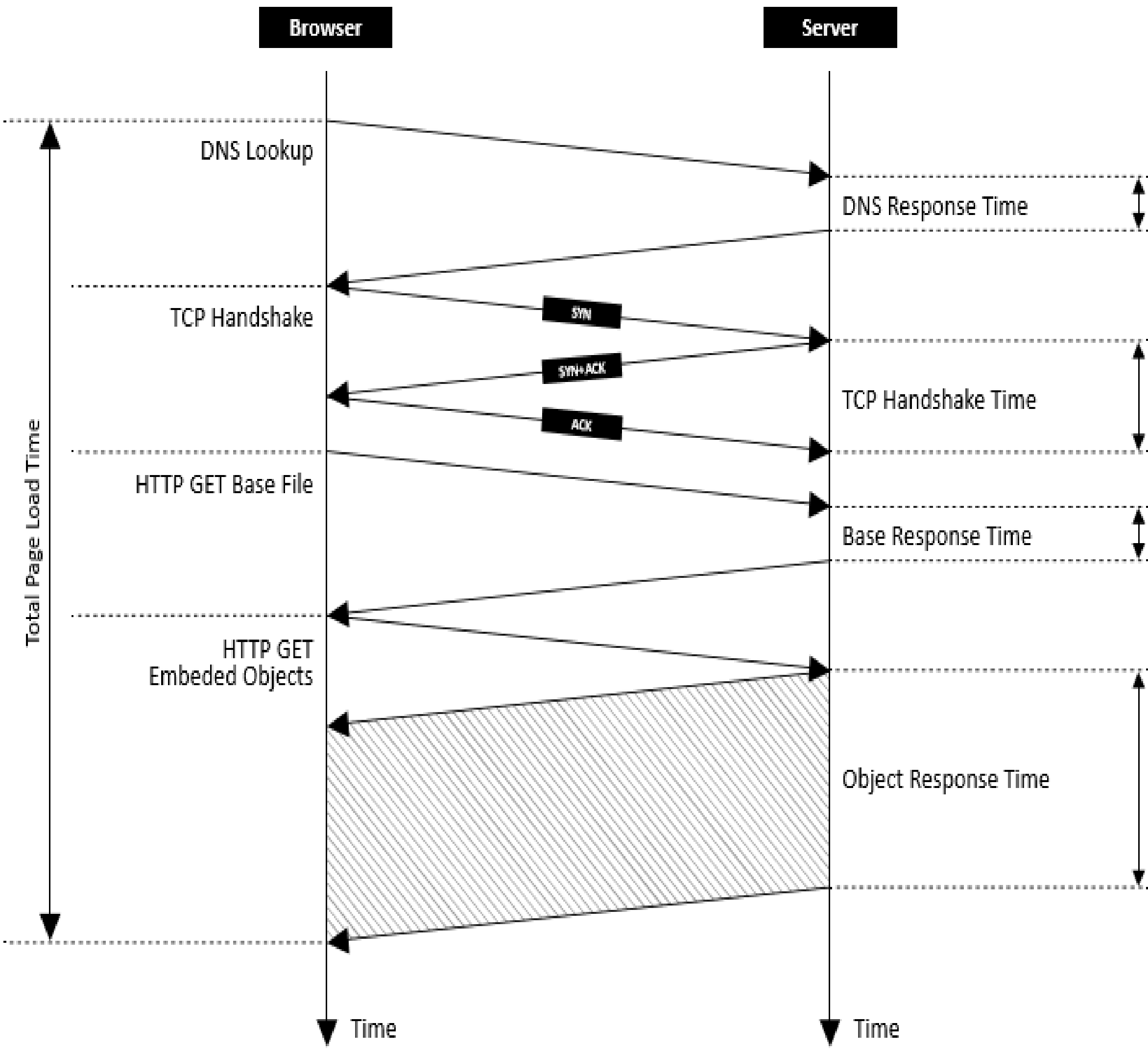


Figure 2-4 Key Phases of Total Load Time [18].

TCP Handshake

TCP Handshake is a fundamental phase of a web transfer. This is a 3-Way Handshake which is performed by a server and a client to establish a TCP connection. If any packet gets lost during this phase, the initial connection establishment time increased significantly. Because, when the Round Trip Time (RTT) between the server and client is higher than average, the time TCP takes for Handshake become crucial. Due to this the waiting time of the user will be longer [19].

Request for Base HTML File

Base HTML file is the first reply from any web server for the first HTTP GET request. This first response is usually a HTML file which is represented as Base HTML. Base HTML file is needed by the web browser render engine to create the Document Object Model (DOM) for building the skeleton of the requested page. The browser requires some time to create the DOM, process all the other requests and display the content on the screen. So, this phase is essential from the user perception as this phase is responsible for the first sign of service delivery [19].

Request for Embedded Objects

Embedded Objects are the Nodes in the DOM tree which want to load from a web server. When the browser comes across with any such Node, it straight away sends a request down to the network-level to get that object from the specified address. As soon as the object gets transferred, browser displays the object on the screen. This phase has a specific significance if the users are browsing on a task-driven mode. Network disturbance can hold any Embedded Object, but if that specific object is the desired object for the user, the QoE may degrade significantly. The Embedded Objects are the contents of the currently requested page. Which is an important factor for determining the user perception of that particular web site [19].

2.2.3. Web Browsing QoS Metrics

This sub section presents the web browsing QoS metrics that are used for this thesis to develop the web QoE estimation model.

Page Response Delay [s]

In interactive applications, it is important to look separately at the average response time. The average response time is the time it takes the first packets of information to arrive after a request has been sent (provided the PDP context has already been established). It does not apply to certain applications such as voice and video conferencing [16].

Specifically, for web browsing services, this parameter capture how long it takes your browser to receive the first byte of a response from a web server when you request a particular URL. In other words, this metric indicates, the amount of time a subscriber waits before the web page information starts to display in the title bar of the browser after the subscriber types a Uniform Resource Locator (URL) in the address bar [20]. The page response time encompasses the network latency of sending your request to the web server, the amount of time the web server spent processing and generating a response, and amount of time it took to send the first byte of that response back from the server to your browser [21]. OR, It is the IP-service setup time. It is the time period needed to establish a TCP/IP connection to the server of a service, from sending the initial query to a server to the point of time when the content is sent or received [22].

$$HTTP\ IP - Service\ Set\ up\ Time\ [s] = (t_{\text{IP – Service success successful}} - t_{\text{IP – Service access start}})\ [s] \quad (2.2)$$

Where:

$t_{\text{IP – Service success successful}}$ is time of IP – service access attempt, or user enters the URL and hits return, or first [SYN] sent.

$t_{\text{IP – Service access start}}$ is time of successful IP service access, or web page download starts, or reception of the first data packet, or sending the first GET command.

Page Content Browsing Delay [s] For the case of web browsing service, after a data link has been successfully established, this parameter describes the average data transfer rate measured throughout the entire connect time to the service. The data transfer shall be successfully terminated. The requirement for this parameter is network and service access. Throughput is measured from opening the data connection to the final successful transfer of the content (web page) [22].

$$Content\ Browsing\ Delay\ [s] = (t_{\text{Data transferd complete}} - t_{\text{Data transfer start}})\ [s] \tag{2.3}$$

Where:

$t_{data\ transfer\ start}$ is time of successfully started data transfer, or web page download starts, or start reception of the first data packet containing content.

$t_{data\ transferred\ complete}$ is time when data transfer complete, or web page download successfully completed, or reception of last data packet containing content

Page Download Throughput [kb/s]

It is normally supposed that an end-user mobile throughput [kb/s] who has a high throughput towards his/her mobile will always have a high QoE. This assumption is applicape for most interactive and background class applications and hence it is important to include this measure in QoE calculations [16].

$$Throughput\ [kbps] = \frac{User\ data\ transferred\ [Kbit]}{(t_{data\ transferred\ complete} - t_{data\ transfer\ start})\ [s]} \tag{2.4}$$

Whare:

$t_{data\ transfer\ start}$ is time of successfully started data transfer, or web page download starts, or start reception of the first data packet containing content.

$t_{data\ transferred\ complete}$ is time when data transfer complete, or web page download successfully completed, or reception of last data packet containing content

$User\ data\ transferred\ [Kbit]$ is the total webpage size in kilo bit.

Different literatures gave different meanings for the same web browsing QoS parameters. So to avoid this confusion, operator's definition for each QoS parameters are listed below [20].

Page Response Delay (PRD) [ms]: Indicates the amount of time a subscriber waits before the web page information starts to display in the title bar of the browser after the user types a Uniform Resource Locator (URL) in the address bar.

Page Response Success Rate (PRSR) [%]: Indicates the rate at which website access requests are successfully responded to after subscribers type a Uniform Resource Locator (URL) in the address bar of a web browser.

Page Content Browsing Delay (PCBD) [S]: The amount of time a subscriber waits before a web page is displayed after the server starts to respond.

Page Browsing Success Rate (PBSR) [%]: Indicates the rate at which a browser successfully displays a web page after a subscriber types a Uniform Resource Locator (URL) in the address bar or refreshes a page for multiple times.

Page Download Throughput (PDTH) [kbps]: Indicates the speed at which a web page is collected to a subscriber's mobile equipment after the subscriber types a Uniform Resource Locator (URL) or refreshes a web page.

2.3. Quality of Experiance

In this section, we see the difference between QoE and QoS, then QoE assessment methodologies and its assessment scales. Finally, we discuss on some QoE metrics.

2.3.1. Qualiy of Service and Quality of Experiance

QoS is defined as the ability of the network to provide a service at guaranteed service level. QoS encompasses all functions, mechanisms and procedures in the cellular network and terminal that ensure the provision of the negotiated service quality between the user equipment (UE) and the core network (CN) [16].

QoE is how a user perceives the usability of a service when they use the service. Or, how satisfied a user is with a service in terms of, for example, usability, accessibility, retainability and integrity of the service. Service integrity concerns throughput, delay, delay variation (or jitter) and data loss during user data transmission; service accessibility relates to unavailability, security (authentication, authorization and accounting), activation, access, coverage, blocking, and setup time of the related bearer service; service retain ability, in general, characterizes connection losses [16].

The term 'QoE' refers to the perception of the user about the quality of a particular service or network. It is expressed in human feelings like 'excellent', 'good', 'poor', etc. On the other hand, QoS is intrinsically a technical concept. It is measured, expressed and understood in terms of networks and network elements, which usually has little meaning to a user [16].

Although better network QoS in many cases will result in better QoE, fulfilling all the traffic QoS parameters will not guarantee a satisfied user. A good throughput in one part of a network might not help if there is no coverage a short distance away [16].

Delivering high QoE depends on the understanding of the factors contributing to the user's perception of the target services, and applying that knowledge to define the operating requirements. This top-down method decreases development costs and the risks of user rejection and complaint, by ensuring that the device or system will meet user requirements [16].

2.3.2. QoE Assessment Methodologies

Studying and measuring QoE is characteristically difficult task, due to its subjective nature. Now, there are two methods to determine the QoE or the satisfaction level of a customer with a service. These methods are subjective and objective methods [23].

Subjective quality assessment. In this methodology, researchers evaluate the quality of a system, application or service through the subjective evaluation of users during controlled experiments. This is often the most informative method, with researchers using questionnaires and interviews to directly obtain user perception [24].

Objective quality assessment. In this type assessement, researchers evaluate the quality of a system by monitoring application QoS metrics that are known to highly correlate with user QoE [24]. We use objective quality assessment to study the relationship between LTE Web browsing service quality and QoE. Then, we propose QoE model that estimate user QoE based on QoS metrics.

By its nature, QoE is subjective, at the same time it is essential to have it measured. Waiting for end-users to vote with their money might turn out to be very expensive for interested party. As such, a approach has to be devised to measure QoE as realistically as possible. The top-down methodology could be useful in this regard [16].

- The key is to know the factors (metrics) contributing to user perception.
- Apply that knowledge to define the operating requirements (values).
- Create a methodology to measure these factors constantly (tools, location, statistical sampling) and improve them as and when needed.

2.3.3. Assessment Scale for Quality of Experiance

The scale recommended by ITU-T P.800 [25] [26], is used to collect the user satisfaction about how they perceive any service. In the survey, a user gave MOS using this five points scale. The scale is defined as, Excellent (5), Good (4), Fair (3), Poor (2) and Bad (1) and detailed in Table 2- 1.

Table 2-1 ITU- T Scale in for Quality Impairment [25].

Scale (MOS)	QoE	Impairment
5	Excellent	Imperceptible
4	Good	Good perceptible but not annoying
3	Fair	Fair slightly, Annoying
2	Poor	Poor annoying
1	Bad	Very annoying

Chapter 3 Neural Network Approach

The Neural Network approach, each of the data pre-processing steps, and the validation metrics with their detail discussion are presented. This helps to understand the basic concepts for developing the web browsing QoE estimation model using neural network approach. Also, to give basic understanding on related issues before and after modeling.

3.1. Neural Network

A neural network is a machine that is intended to model the way in which the brain performs a certain task or function of interest. The network is usually applied by using electronic components or is simulated in software on a digital computer. It is an enormously parallel distributed processor consists of simple processing units that has a natural tendency for storing experiential knowledge and making it available for use [27]. Knowledge is learned by the network from its environment through a learning process, interneuron connection strengths, called synaptic weights, are used to store the learned knowledge.

3.1.1. Model of Neuron

A neuron is an information-processing unit that is essential to the task of a neural network. The three basic elements of the neural model are set of synapses, or connecting links, An adder, activation function [27].

A set of connecting links, each of which is considered by a weight of its own. Specifically, a signal x_j at the input of connecting link j connected to neuron k is multiplied by the connecting link weight w_{kj}. The second basic element is adder for summing the input signals, weighted by the respective link strengths of the neuron, the operations described here constitute a linear combiner. And third is activation function used for limiting the amplitude of the output of a neuron. The activation function is also known as a squashing function, in that it squashes (limits) the permissible amplitude range of the output signal to some finite value.

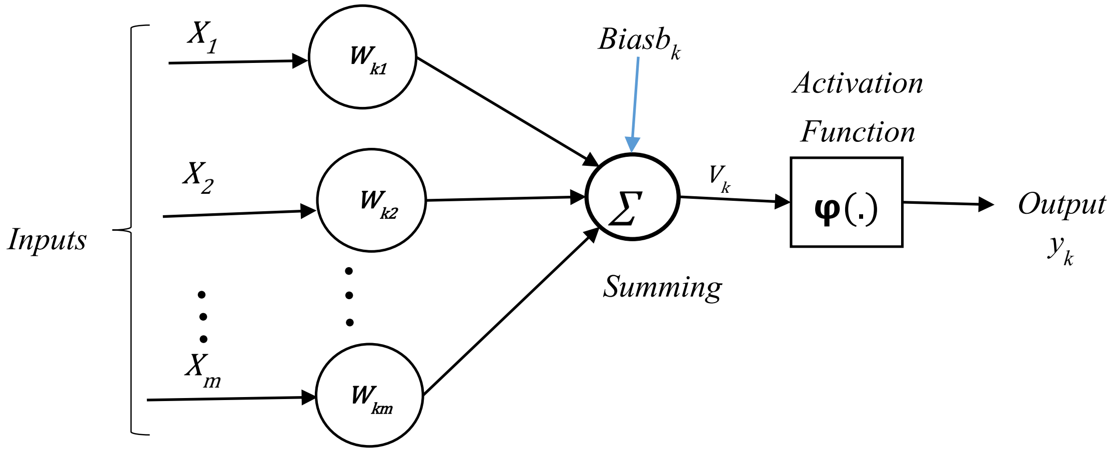


Figure 3-1 Nonlinear Model of a Neuron Labeled k [27].

The neural model of Figure 3- 1 also includes an externally applied bias, denoted by b_k. The bias b_k has the effect of increasing or decreasing the net input of the activation function, depending on whether it is positive or negative, respectively.

A single-input neuron is shown in Figure 3- 1. It consists of a processing unit which contains weights (w_{km}) and summing function followed by a transfer function **φ**(.). The summer output, referred to as the total synaptic input of the neuron, is given by the inner product of the input and weight vectors [28]. In mathematical terms, we may describe the neuron k depicted in Figure 3-1 by writing the pair of equations [27]:

$$\upsilon_k = \sum_{j=1}^{m} (w_{kj} x_{kj}) \qquad (3.1)$$

And

$$y_k = \varphi(\upsilon_k + b_k) \qquad (3.2)$$

Where:

$x_1, x_2..., x_k$ are the input signals;

$w_{k1}, w_{k2} ..., w_{kj}$ are the respective synaptic weights of neuron k;

u_k is the linear combiner output due to the input signals;

b_k is the bias, is the activation function; and

$\varphi(\upsilon)$ is activation function

y_k is the output signal of the neuron

Activation Function

The activation function, denoted by $\varphi(\upsilon)$, defines the output of a neuron in terms of the induced local field υ. There are two basic types of activation functions are [27]:

i. Threshold Function. For this type of activation function, we have

$$\varphi(\upsilon) = \begin{cases} 1, & \upsilon \geq 0 \\ 0, & \upsilon < 0 \end{cases} \tag{3.3}$$

Correspondingly, the output of neuron k employing such a threshold function is expressed as

$$y_k = \begin{cases} 1, & \upsilon_k \geq 0 \\ 0, & \upsilon_k < 0 \end{cases} \tag{3.4}$$

Where υ_k is the induces local field of neuron.

ii. Sigmoid Function. Whose graph is "S"-shaped, is by far the most common form of activation function used in the construction of neural networks. It is defined as a strictly increasing function that exhibits a graceful balance between linear and nonlinear behavior. An example of the sigmoid function is the logistic function, defined by

$$\varphi(\upsilon) = \frac{1}{1+\exp(-a\upsilon)} \tag{3.5}$$

In the limit, as the slope parameter approaches infinity, the sigmoid function becomes simply a threshold function.

3.1.2. Neural Network Architecture

The manner in which the neurons of a neural network are structured is closely linked with the learning algorithm used to learn the network. We may therefore speak of training algorithms (rules) used in the construction of neural networks as being structured. Architecture of the Neural Network consists Input layer, Hidden layer and Output layer [29].

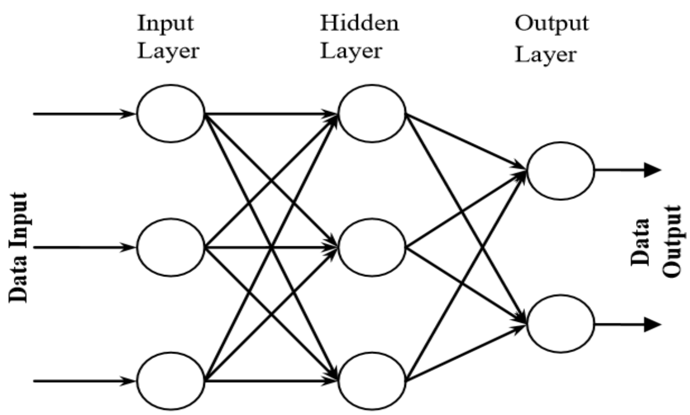


Figure 3-2 Multilayer Neural Network [29].

Input Layer: The Input Layer is a layer which connects with the external environment. Input layer presents a pattern to neural network. Once a pattern is obtained to the input layer, the output layer will produce another pattern. It also represents the condition for which purpose we are training the neural network [29].

Output Layer: The number of output neurons should be directly related to the type of the work that the neural network is to perform The Output layer of the neural network is what actually presents a pattern to the external environment [29].

Hidden Layer: The Hidden layer of the neural network is the layer between input and output layer. Activation function put on hidden layer if it is available. Hidden layer consists hidden nodes. Hidden nodes or neurons are the neurons that are neither in the input layer nor the output layer [29].

3.1.3. Type of Neural Network

In general, there are three basically different classes of network types [27]:

Single-Layer Feedforward Networks: It is the simplest form of a layered network, it has an input layer of source nodes that projects directly onto an output layer of neurons, but not vice versa.

Multilayer Feedforward Network: The second class of a feedforward neural network differentiates itself by the presence of one or more hidden layers, whose computation nodes are correspondingly called hidden neurons or units. The term "hidden" represents the type of neurons in the neural network that are not not seen directlyfrom either the input or output of the network. The purpose of hidden neurons is to intervene between the external input and the network output in some useful form.

Recurrent Network: It distinguishes itself from a feedforward neural network in that it has at least one feedback loop. For example, a recurrent network may contains a single layer of neurons with each neuron feeding its output signal back to the inputs of all the other neurons.

3.2. Data Pre-proccessing

This section illustrates, the major steps involved in data preprocessing, namely, data cleaning, data integration, data reduction, and data transformation.

Data Cleaning: Real-world data have a tendency to to be incomplete, noisy, and inconsistent. Data cleaning procedures attempt to fill in missing values, smooth out noise while detecting outliers, and correct irregularities in the data [30].

Data Integration: It helps to reduce or avoid redundancies and inconsistencies in the resulting dataset to improve the accuracy and speed of the subsequent data mining process. Some redundancies can be detected by correlation analysis. we can evaluate the correlation between two attributes, A and B, by computing the correlation coefficient (also known as **Pearson's** product moment coefficient) [30] This is

$$\mathbf{r}_{A,B} = \sum_{n=1}^{n}(a_i b_i) - (n\bar{A}\bar{B}))/(n\,\sigma_A\sigma_B) \quad (3.6)$$

Where

n is the number of tuples

a_i and b_i are the respective value of A and B in tuple i

$\bar{A}$ and $\bar{B}$ are respective mean value of A and B

σ_A and σ_B are respective standard deviation of A and B

Note that $-1 \leq r_{A,B} \leq +1$. r is greater than 0, then A and B are positively correlated, meaning that the value of A increased as the value of B increase. The higher the value, the stronger the correlation. Hence, a higher value may shows that A (or B) may be removed as a redundancy. If the resulting value is equivalent to 0, then A and B are not dependent and there is no correlation between them. If the resulting value is less than 0, then A and B have opposite relationship, where the value of one attribute is increase as the value of the other attribute decrease. This means this attribute discourages the other.

Data Transformation: It is one of apre-processing step to be considered, to have more efficient data mining process. One of the data transformation method is data normalization. Spefcifically, min-max normalization [30]. It does a linear transformation on the original data. Assume that min_A and max_A are the minimum and maximum values of an attribute, A. Min-max normalization maps a value, v_i, of A to v_i^- in the range [new min_A, new max_A] by computing [30].

$$V^- = \frac{(v_i - Min_A)}{(Max_A - Min_A)} \quad (3.7)$$

Min-max normalization preserves the relationships among the original data values.

3.3. Validation Metrics

We know that the models are used for estimation. For appropriate estimation, it is important to check first the capability of these models. So, to check its performance, performce metrics are required. This thesis used R Squared and MSE (Mean Square error) performance metrics to check the model is valid or not.

R-squared: It is a statistical measure of how close the data are to the fitted regression line, it is also known as the coefficient of determination, or the coefficient of multiple determination for multiple regression [31]. High values of R-Squared represent a strong relationship between actual and model output while low values mean that developed estimated model is not appropriate for required purpose. The value of R is between 0 and 1, that 0 means no relationship between sample data and 1 mean exact linear relationship. To calculate R Squared we used the following formula:

$$R^2 = \frac{SSR}{SST} \tag{3.8}$$

In this formula respectively SST (Sum of Squares of Total) and SSR (Sum of Squares of Regression) are the total sums of the squares and measures how far the data are from the mean and the sum of squares of errors and measures how far the data are from the model's predicted values.

$$SST = (Actual\ Value - \ Mean\ Actual\ Value)^2 \tag{3.9}$$

$$SSR = (Estimated\ Value - \ Mean\ Actual\ Value)^2 \tag{3.10}$$

According to our goal which is estimate QoE or MOS value, we calculated SST and SSR based on actual and model output.

MSE (Mean Squared Error): It measures the average of the squares of the errors, that is, the average squared difference between the actual value and estimated value [31]. The mean squared error tells us how close a regression line is to a set of points. It does this by taking the distances from the points to the regression line (these distances are the "errors") and squaring them, the squaring is necessary to remove any negative signs.

The measure of mean squared error requires a target of estimation along with an estimator which is said to be the function of the given data. The unit of MSE is the same as the unit of measurement for the quantity which is being estimated [31]. For calculate MSE we used the following formula:

$$\mathrm{MSE} = 1/\mathrm{n} \sum_{\mathrm{i}=1}^{\mathrm{n}} (\mathrm{A_t} - \mathrm{F_t}) \quad (3.11)$$

Where A_t indicates the actual number, F_t indicates the estimation number and n indicates number of observations.

The smaller the means squared error, the closer you are to finding the line of best fit but this number is related to the range of your values. Depending on your data, it may be impossible to get lesser value for the mean squared error.

Chapter 4 QoE Modeling Methodology

This chapter presents a model that estimate the QoE on LTE web browsing service using the neural network approach. It consists the system model, data collection, data pre-processing and neural network approach to model web QoE.

4.1. System Model

Figure 4-1 is the system model for estimating the web browsing QoE from QoS value. In broad sense this system model has two blocks, the first one is learning the neural network and the second one is testing the neural network. In learning the neural network, QoS parameter from service monitoring tool and target MOS value from existing model were input for the developed QoE model to train the neural network using supervised learning process. But in testing the neural network, the QOS parameters from the different situation of learning sets i.e. the type of data that are different from the data given for training the neural network. The testing process is performed after the neural network trained using the QoS parameters as input and MOS as a target value. The output of the testing the QoE model is estimated QoE value in terms of MOS.

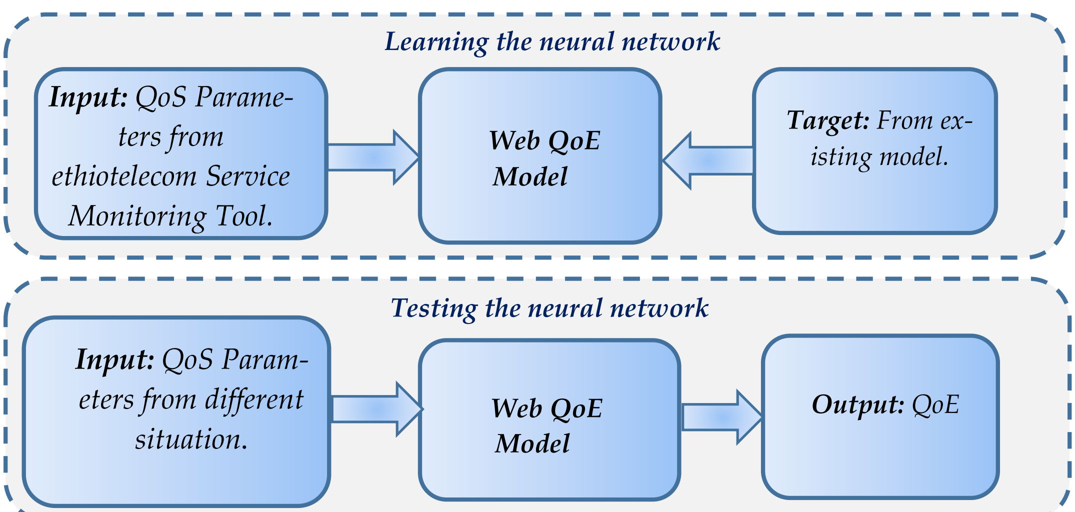


Figure 4-1 LTE Web Browsing QoE Estimation System Model.

4.2. Neural Network Tool

In fitting problem, we want a neural network to map between a data set of numeric input and a set of numeric targets. The neural network fitting tool will assistance us to select data, construct and train a network, and evaluate its performance using mean square error and regression analysis. A two-layer feed-forward network with sigmoid hidden neurons and linear output neurons, can fit multidimensional mapping problems arbitrarily well, given stable data and adequate number of neuron in its hidden layers. The network will have trained with Leverage-Marquardt backpropagation algorithm. This algorithm typically needs more memory but less time. Training automatically stop when generalization stops learning, as indicated by an increase in the mean square error of the validation samples [32].

The proposed work used this tool to estimate web browsing QoE from the selected input web browsing QoS parameters. All selected QoS parameters (the page response delay, page content browsing delay and page download throughput) are the input values for the neural network. The MOS was the target value for the process of learning in neural network.

4.3. Data Collection

4.3.1. Input Data

All Key Quality Indicator (KQI) parameters for web browsing services are measured and stored by using service monitoring tools (from Huawei Smart Care). The amount of data collected is for eight months' length and 5049 in number. Those are QoS parameters, that affects the web browsing QoE in different weights. Those QoS parameters are Page Resopnse Delay (PRD), Page Response Success Rate (PRSR), Page Content Browsing Delay (PCBD) and Page Download Throughput (PDTH).

4.3.2. Customer Survey

The survey was conducted in Addis Ababa city, Ethiopia. Data was collected among Sixzone: Central Addis Ababa Zone (CAAZ), East Addis Ababa Zone (EAAZ), North Addis Ababa (NAAZ), South Addis Ababa Zone (SAAZ), West Addis Ababa Zone (WAAZ), South West Addis Ababa Zone (SWAAZ) in Addis Ababa city. Two key information were collected from the prepared questionnaire. The first information was the QoS value for the web browsing service, that was collected by using the installed application on customer premises. The installed application, used to collect the QoS value is Website Performance Tester (WPT). The second information collected from the questionnaire was customer perception on the LTE web browsing service for their respective QoS values. Both informations were collected to select the best mapping function among the existing mapping function.

Sampling Procedures

Population

The population from which the sample was drawn for this survey, involved subscribers who are subscribed LTE data services in Addis Ababa. The sample frame was a subscriber from six zones namely: CAAZ, EAAZ, NAAZ, SAAZ, WAAZ and SWAAZ in Addis Ababa City.

Determining Sample Size

Ethio telecom has a total of 400,000 LTE subscribers. In order to get the sample size, the researcher adopted the formula [33] [34].

$$S= \frac{N}{(1+N(e)^2)} \tag{4.1}$$

Where S - sample size, N – Total customer, e – detection error expressed in to percentage (5% - 10%).

And thus Total customer (N) is 400,000, Detection error (e) = 8%, any value for e from 5% - 10% is acceptable.

$$S = \frac{400000}{(1+400000(0.08)^2)} = 156 \text{ Respondents}$$

That means, the respondents were not required to be less than 156. In order to make the research practicable, the sample size of 177 subscribers from six Addis Ababa zones was involved in this survey. The 177 respondents were selected by using random sampling.

Questionnaires

The questionnaires are provided in an electronic way through email, which allowed a flexible data collection mechanism. In this survey we have taken around 200 LTE subscribers, the outlier were filtered from the total involved users. Outliers were defined based on the user feedbacks. If a subscriber satisfies any one of the below conditions, then the subscriber is called outlier.

- If the selected user perception score by the subscriber is more than one options, then the subscriber is called as outlier.
- If the demographic information is incomplete, then the subscriber is called as outlier.

We have a total of 23 outliers out of that 7 outliers are due to scoring more than one option and the others are due to filling incomplete demographic information. The collected web browsing QoS data are shown in Figure 4-2. The figure displays all the available web browsing QoS parameters. All web browsing QoS values that are measured using the installed application WPT are filled in the table prepared on the questionnaire. Each parameter is defined as follows:

Figure 4-2 Website Performance Tester Measurement Results.

Total Load Time: It indicates how long it takes for a page to fully load in the browser after a user clicks a link or makes a request [35].

Start Render Time: The Start Render time is the start point in time that something was displayed to the screen. Before this point in time the user was bignning at a blank page. Start rendering time does not necessarily mean the user saw the page content, it could just be something as simple as a background color but it is the first sign of something happening for the user [36]. According to the definition given above, start rendering time is the same as the page response delay in the system side.

Page Content Browsing Time is the time it takes to download the content after the server respond the request. (Total load time minus Start render time).

Byte in: This is the amount of data that the browser had to download in order to load the page. It is also usually referred to as the "Page Size". From here we can determine the Download throughput [37]. The page download Throughput is calculated using the formula discussed on (2). So, we have a measurement for page response delay, Page content browsing delay and Page download throughput the customer side.

Survey Result

The survey is takes place by 177 LTE subscriber, the test was performed by different kind of subscribers, they are living and working in different place. Also, It was performed at different time with fair gender distribution (99 Male,78 Female), and they have different age and educational back ground. Table 4-1 depicts the summary of personal profile on the subscribers that perform the survey.

Table 4-1 Summary of Participant Profile.

Location	period						Gender		Age					Educational back ground					
	Early morning	Morning	Lunch time	After Noon	Evening	Night	Male	Female	Below 15	15 - 24	25 - 54	54 - 64	65+	MA/ MSC and above	BA/ BSC	University or College	Diploma or Certificate	Preparatory	Grade 10 and below
CAAZ	4	8	7	9	4	2	25	9	0	2	29	3	0	10	18	4	2	0	0
EAAZ	7	4	4	8	12	8	19	24	0	7	28	5	0	3	36	7	0	0	0
NAAZ	2	3	2	11	2	3	9	13	0	7	10	5	0	3	13	6	0	0	0
SAAZ	5	3	3	7	3	7	18	10	0	3	21	4	0	1	24	1	2	0	0
WAAZ	7	2	2	5	2	5	10	8	0	4	16	4	0	2	14	6	2	0	0
SWAAZ	7	5	2	8	2	2	18	9	0	8	21	0	0	1	18	4	0	0	0
Total	**32**	**25**	**20**	**48**	**25**	**27**	**99**	**73**	**0**	**31**	**125**	**21**	**0**	**20**	**123**	**28**	**6**	**0**	**0**
Total Prticipant = 177																			

After the end of experiment, the participant filled a questionnaire where the QoE of the service is graded. Table 4-2 is sample data that shows the personal profile for each participant, web browsing QoS parameter value and their respective user perception in terms of MOS.

Table 4-2 Sample of Survey Result Including Each Participant Profile.

No	Total Load Time (s)	Rendering Time (s)	Page content browsing time (s)	Page down load through put(Kbps)	Time Period	Location	Perception	Gender	Age	Educational back ground
1	5.667	1.833	3.834	3367.079291	Morning	EAAZ	3	Male	25 - 54	Bachler's degree
2	4.816	2.10925	2.70675	3603.471913	Lunch tim	SAAZ	3	Male	25 - 54	Bachler's degree
3	5.845	1.70725	4.13775	2568.652045	Early Mor	SWAAZ	2.5	Male	25 - 54	Bachler's degree
4	5.905	1.668	4.237	8884.602313	evenning	SWAAZ	2.5	Male	25 - 54	Bachler's degree
5	5.238	1.9697	3.2683	3686.207508	Morning	SWAAZ	2.75	Female	25 - 54	Bachler's degree
6	4.407	2.2895	2.1175	4319.435183	Morning	SWAAZ	3	Female	25 - 54	Bachler's degree
7	4.384	1.255	3.129	1901.091211	Morning	SWAAZ	3	Female	25 - 54	Bachler's degree
8	6.552	2.0953	4.4567	3316.603406	Lunch tim	SWAAZ	2.5	Male	25 - 54	Bachler's degree
9	3.934	1.474	2.46	11061.36667	After noo	SWAAZ	3.5	Female	15 - 24	University or college student
10	5.712	2.6945	3.0175	3684.671947	After noo	SWAAZ	3	Female	15 - 24	University or college student

4.3.3. Target Value

The proposed work used the supervised learning process to train the neural network. So, the input and target value are required to train the neural network. For this thesis, the target values are obtained from the existing mapping function, which takes into account the total download time of a web page.

The user perception from customer survey is the reference value. In terms of accuracy, the subjective method to determine the QoE is better than objective method. To measure the closeness of the mapping functions with that of the survey QoE results, the performance metrics i.e, the Mean Square Error (MSE) was calculated for each mapping function. Table 4-3 shows different existing mapping functions and their closeness with the real customer survey result using their MSE.

Table 4-3 Web browsing QoE Mapping Functions [9][4][38][12].

Name of Mapping function	Existing Web QoE functional fitting models	Mean Square Error (MSE)
Logarithmic	MOS = -1.426*ln(T) + 4.469	1.202599
Linear	MOS = -0.318*T + 4.469	0.742978
Exponential	MOS = 4.836*exp(- 0.15*T)	1.069522
Lorentzian	$MOS = 5 - \frac{578}{1+\left(11.77+\frac{22.61}{T}\right)^\wedge 2}$	0.328652

As Table 4-3 illustrated, Lorentzian function model output has minimum mean square error value, compare to others mapping functions. It has a mean square value of 0.328652. So, Lorenzian function model is selected, to obtained the target value for the proposed thesis work. The selected mapping function that used to get the target value, takes into account the page load time or the time that the web page takes to downloaded. The selected mapping function in [12] is given by (4.2).

$$MOS = 5 - \frac{578}{1+(11.77+22.61/d)^\wedge 2} \quad (4.2)$$

where d is the page load time in second.

The page load time parameter is measured in time series of one-hour interval, corresponding each one to one MOS measurement. Once the QoE is determined based on the selected mapping function, the next step is pre-processing to select the input parameters and to provide the proper input data for neural network.

4.4. Data Pre-proccessing

The estimation quality is meaningfully affected by the data preprocessing. Some of the data preprocessing steps that were performed in this thesis are data cleaning, parameter selection and data normalization.

4.4.1. Data Cleaning

Some of the data that are collected from the System have missed values. Table 4-4 showed, the missed values for the page download throughput parameter. One of the technique to fill the missed value is filling the missed value with most probable values. The total amount of data that have missed values are 25 in number. So, all of the missed value are filled by the values that are frequently occurred.

Table 4-4 Sample Missed Data.

Current period	Page Response Success Rate(%)	Page Response Delay(ms)	Page Browsing Success Rate(%)	Page Browsing Delay(ms)	Page Download Throughput(Kbps)
2019-02-06 10:00	100.00	631	80.00	1757	--
2019-02-15 11:00	100.00	611	100.00	1029	--
2019-02-15 12:00	50.00	1039	50.00	2502	--

4.4.2. Parameter Selection

Correlation is performed, to select the most influential parameters that affects the web browsing service. The selection process is performed by correlation between each parameter with respective MOS and the correlation between one parameter with another parameter. If the correlation between input parameters with MOS is high, we can say that the parameter is highly influencing the MOS. Also if the correlation between each parameter are high, their influence on the MOS is the same or those parameters are not independent. So, we should take only one of the parameter to avoid redundancy.

Table 4-5 Correlation Result Between Each Parameters.

	PRSR	PRD	PBSR	PCBD	PDTH	MOS
PRSR	100%	-7.4%	92%	1.5%	-6.1%	3.3%
PRD	-7.4%	100%	-2.6%	-3.13%	-4.8%	-31.4%
PBSR	92%	-2.6%	100%	-8.9%	-5.9%	7.7%
PCBD	1.5%	-3.13%	-8.9%	100%	23.7%	-80.5%
PDTH	-6.1%	-4.8%	-5.9%	23.7%	100%	27.5%
MOS	3.3%	-31.4%	7.7%	-80.5%	27.5%	100%

So, the Pearson Correlation between each parameters and the respective MOS was calculated. The parameters with a higher correlation with MOS were selected. Table 4-5 shows, the relationship among the collected five parameters as well as their correlation with MOS. PRSR and PBSR are not considered as an input, since the correlation between the measured MOS and these parameters are 3.3% and 7.7% respectively. According to [39], PRSR and PCBD have no correlation with MOS because of lower value of correlation coefficients. So, only the three parameters (PRD, PCBD and PDTH) are selected as an input for the model. The negative sign (-) indicates that the parameters are negatively correlated or they have opposite relationship.

4.4.3. Normalizing the Data

One of the very important data preprocessing step is normalizing the input data. Before applying the input data to the neural network for the training, the data should be normalized, to fall within a smaller range, such as −1 to 1, or 0 to 1. In this thesis the normalization is performed by Min-Max normalization, that means the all value are scaled with in 0 and 1. Table 4-6 and 4-7 depicts the values of the selected parameters before and after normalization respectively.

Table 4-6 Sample Data Before Normalization.

Current period	Page Browsing Success Rate(%)	Page Browsing Delay(ms)	Page Download Throughput(Kbps)
2018-11-27 00:00	48.92	5245	1108.48
2018-11-27 01:00	45.48	5240	1125.97

Table 4-7 Sample Data After Normalization.

Current period	Page Browsing Success Rate(%)	Page Browsing Delay(ms)	Page Download Throughput(Kbps)
2018-11-27 00:00	0.544214694	0.287966805	0.148236446
2018-11-27 01:00	0.493767415	0.287275242	0.151648643

4.5. Learning and Testing the Neural Network

The artificial neural network is learned using the parameters page response delay, Page content browsing delay and Page download throughput as input and MOS value as target. Table 4-8 illustrates, the sample data consisting of three QoS web browsing parameters and QoE value in terms of MOS. These QoS parameters are to be presented as input data and the MOS values are assigned as target that define the desired output.

Table 4-8 QOS Parameters and MOS Values in LTE Networks.

Page Response Delay(s)	Page Content Browsing delay(s)	Page Download Throughput(Kbps)	MOS
0.942	4.303	1108.48	2.75
0.899	3.572	1397.50	3
1.35	4.777	1072.62	2.5
0.93	3.818	1199.00	3

The total amount of data that are collected from ethio telcom service monitoring tool (Huawei Smart Care) for training the neural network was 5049 in numbers. To load those data to the MATLAB software, use the input and target option in the selected data. The input and the target data was written in the matrix form below.

Input Values:

$$\begin{bmatrix} 0.899 & 3.572 & 1397.50 \\ 1.35 & 4.777 & 1072.62 \\ 0.93 & 3.818 & 1199.00 \\ 1.068 & 8.487 & 309.58 \end{bmatrix}$$

Output Values:

$$\begin{bmatrix} 2.75 & 2.75 & 3 & 2.5 \end{bmatrix}$$

After loaded all the data, the data set is divided randomly in to three kind of sample data as indicated in Figure 4- 3: 70% for training, 15% for validation and 15% for test. The training samples are used for training the neural network to develop the web browsing QoE estimation model. During training, the network is adjusted according to its error. The validation sample data were used to measure network generalization or to determine the stopping point of the training process. The testing sample data were used for measuring the performance of the model. It has no effect on the training.

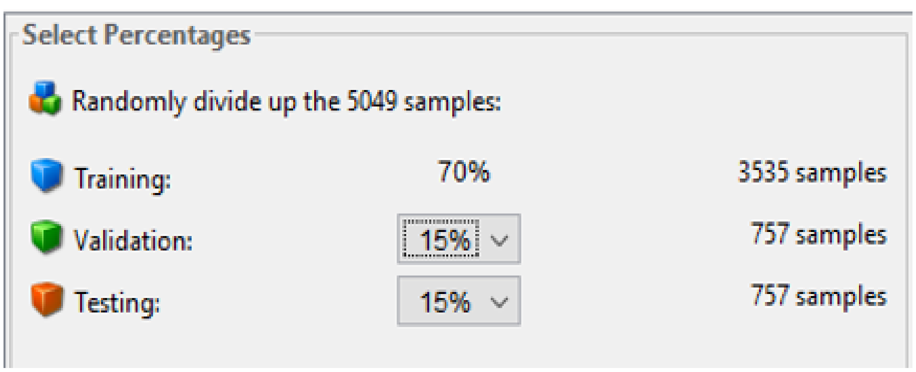


Figure 4-3 Percentage of Data Samples.

4.5.1. Network Architecture

One of the model issue that affect the performance of the estimation model is its network architecture. To design the appropriate architecture, the selection of the number of layers, the number of nodes in each layer and the type of network should be correctly chosen.

The number of inputs is usually transparent and reasonably easy to choose. Since, number of input parameters for this thesis were three due to that input node number was set to 3 to estimate the QoE values. Also the number of output nodes is relatively easy to specify as it is directly related to the problem under study. So, the output node set to 1. But, the selection of number of hidden layers and hidden nodes are not easier like input and output node selection.

4.5.2. Number of Hidden Layer and Node Selection

The other important issue for the construction of neural network architecture are selection of hidden layer and hidden nodes. Mostly, a single hidden layer can approximate any complex nonlinear function with any desired accuracy. But, a single layer needs a number of hidden nodes. There are two types of hidden nodes effect, over fitting and under fitting. Overfitting occurs when unnecessary more neurons are present in the network and its effect is deteriorating the network generalization ability [40].

The most common approach to define the number of hidden nodes is try and error method. This method divides in to two approaches, these are forward approach and backward approach. These approach begin by selecting a small or large number of neuron respectively, then train and test the neural network. After that gradually increase or decrease the number of hidden neurons and again train and test the Neural Network. Repeat the above process until training and testing improved [29]. Based on all the knowledge explained above, this thesis selected 1 hidden layer and as a sample 10 neuron in the hidden layer to design the neural network. The hidden neurons were selected by using a forward hidden neuron selection approach. Train and test the NN by observing the performance metrics when hidden neuron number increases. The overall neural network architecture is illustrated on Figure 4-4.

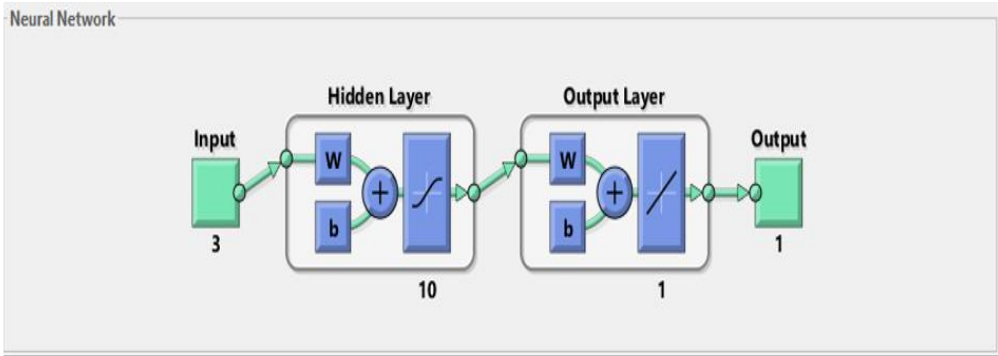


Figure 4-4 Two-Layer-Forward Neural Network.

The other point that are carefully considered to design the neural network is the selection of the activation function. For this research, the hidden layer used sigmoidal activation function and the output layer used the linear type of activation function, as indicated on Figure 4-4.

For this thesis, the algorithm to train the neural network is the Levenberg Marquand. This method is more efficient nonlinear optimization methods and are used in most optimization packages. Their faster convergence, robustness, and the ability to find good local minima make them attractive in ANN training [29]. This algorithm requires more memory and less time. Training automatically stop when generalization stops learning as indicated by an increase in the mean square error of the validation sample.

Chapter 5 Results and Discussion

This chapter presents the results on the performance assessment of the neural network model and the impact of each QoS metrics on the web browsing QoE. Also, presents detail discussion on each results.

5.1. Training and Testing

To train the neural network, the ratio of dataset among training, validation and test are 70%, 15% and 15% respectively. Figure 5-1 shows, the total amount of data set and the three data samples. The total amount of data that are entered to the neural network are 5049 samples. The amount of data samples for training are 3535, for validation that used to decide the stopping point of the training process are 757 data samples and the rest 757 sample data are for testing the developed model.

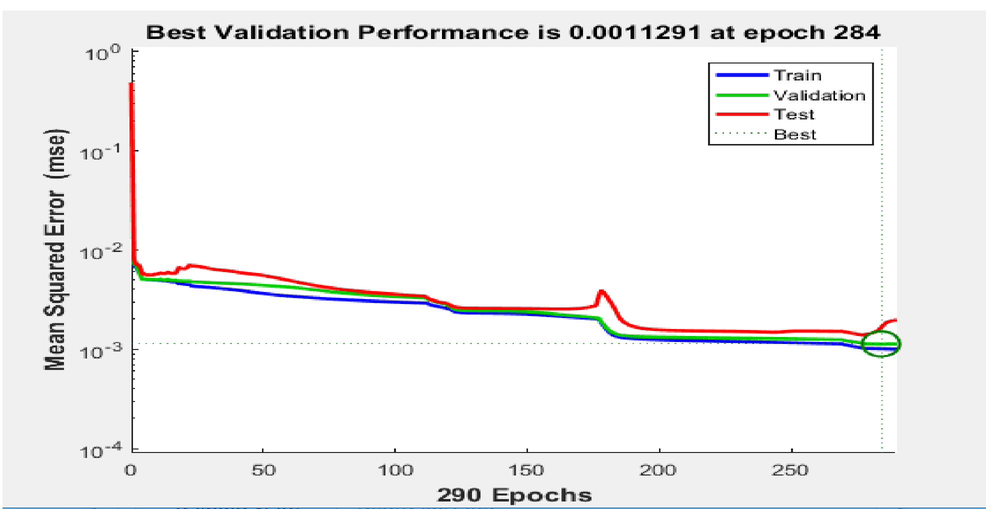


Figure 5-1 Neural Network Training Performance Graph.

This graph shows us the relation between the measured mean square error and epoch or number of iteration to learn the given training data that are entered to the neural network. From the graph, we see, where the best validation performance is attained. The value is indicated at the point marked by circle or the intersection of the two broken lines i.e. at

MSE value of 0.00113 and at epoch 284. This tells the training stops after the mean square error begins to increase.

As stated earlier, a training and a test data are normally required for constructing a Neural Network estimator. The training data is used for model development and the test data is implemented for evaluating the estimating capacity of the model. For proper estimation, it is essential to check the capability of the Neural Network. To check the neural network capacity, performance metrics MSE, R-Square are used.

The regression plot is used to validate the network performance. It shows the network outputs with respect to targets for test sets. For a best fit, the data should fall along a 45-degree line, where the network estimated output values are equal to the target values.

Table 5-1 Correlation Result by Different Hidden Layer Neuron Numbers.

NO	**Number of Neuron**	**R Value**	**Mathematical Relationship**
1	*2*	*0.87961*	*Output = Target*0.85+0.4*
2	*4*	*0.91834*	*Output = Target*0.81+0.53*
3	*6*	*0.89532*	*Output = Target*0.77+0.65*
4	*8*	*0.9351*	*Output = Target*0.86+0.39*
5	*10*	*0.97185*	*Output = Target*0.97+0.095*
6	*12*	*0.84388*	*Output = Target*0.98+0.052*
7	*14*	*0.90175*	*Output = Target*0.84+0.44*
8	*16*	*0.96054*	*Output = Target*0.9+0.27*
9	*18*	*0.92279*	*Output = Target*0.8+0.54*

In this thesis, the neural network is tested using all test samples and the outcome with correlation factor or regression value (R) for different hidden neuron numbers are presented in Table 5- 1 below. The column represented by mathematical relationship, illustrates the relationship between the estimated output and the actual (target) values. As the table indicates, the estimator model is performed best at hidden layer neuron number is 10. Due that this research used 10 hidden layer neurons for constructing the neural network to get best model performance. Figure 5- 2 depicts the model performance result at the selected hidden layer neuron numbers.

Results

	Samples	MSE	R
Training:	3535	1.00958e-3	9.83354e-1
Validation:	757	1.12906e-3	9.77790e-1
Testing:	757	1.65259e-3	9.71847e-1

Figure 5-2 Neural Network Model Training Performance Result.

Also Figure 5-3 is regression plot to show the model output (estimated value) with respect to target for test set. The correlation coefficient or R value and the other information obtained from the regression plot is the mathematical relationship between the actual and the estimated MOS.

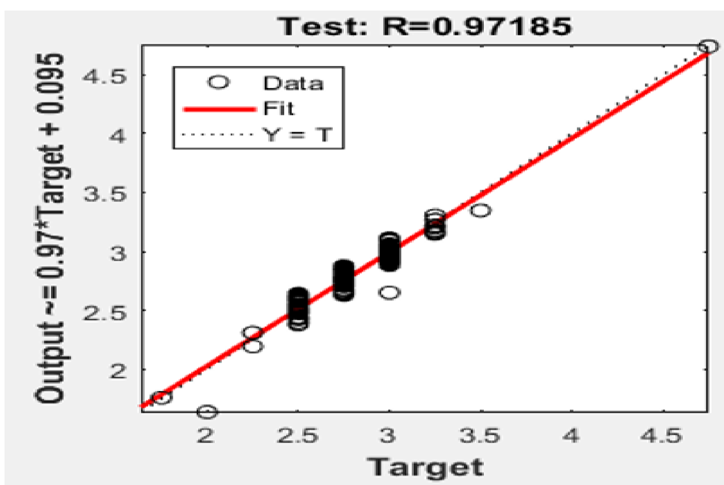


Figure 5-3 Scatter Plot for Model Output and Actual Value

Generally, the proposed model has a correlation coefficient of 0.97185 or 97% correlated and mean square error of 0.002. As the result indicates, the estimated and measured QoE values is highly correlated. And the error between them is very low.

So, this model can be used for estimating the QoE for the mobile operator to get the benefit of objective measurement, instead of using subjective QoE measurement.

5.2. Impact of Individual Parameters

This section presents the correlation results and their discussion on the impact of individual web browsing QoS parameters on the user perception.

The pearson correlation analysis is conducted on each of the available five web browsing QoS parameters with MOS values. The correlation coefficients are calculated according to (3.6) and find the coefficient values as listed in Table 4-5. Based on the pearson correlation results, among the available five parameters, three of them are selected as in put. The correlation results showed, the relationship between the user perception with each web browsing QoS parameters that used to determine their influence on the QoE or user perception on the web browsing service. Generally, this section indicated, the individual web browsing QoS parameters impact on the user perception.

PRD Vs QoE

Let us first consider the first type of feature (i.e., page response delay). this parameter indicates, the amount of time a subscriber delays before the web page information starts to display in the title bar of the browser after the subscriber types a URL in the address bar. The page response time includes the network latency of sending your request to the web server, the amount of time the web server spent processing and generating a response, and amount of time it needed to send the first byte of that response back from the server to your browser.

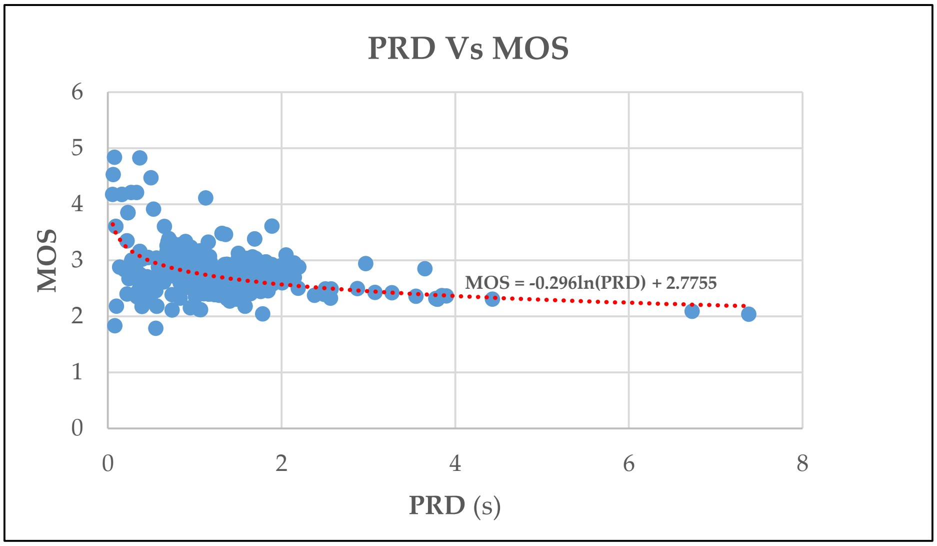


Figure 5-4 The Relationship Between PRD and MOS.

As shown in Table 4-5, we can see that the correlation coefficient of page response delay with MOS is -0.314. It implying that the page response delay has correlation with QoE, but their correlations is weak. The negative sign shows, PRD has a negative impact on the QoE. As indicated in Figure 5-4, for web browsing, the QoE decrease with the logarithm of the page response delay or searching time of the web pages.

PCBD Vs QoE

Concerning to the second type (i.e., page content browsing delay), it is depicted in Table 4-5. The correlation coefficient of page content browsing delay with MOS is -0.805, which is relatively close to 1, implying that PCBD is strongly correlated with MOS. This parameter strongly affects the web browsing QoE. Also it has a negative impact on QoE like page response delay. Figure 5-5 depicts, the relationship between the page content browsing delay and the QoE. Web browsing QoE decrease with the logarithm of the page content browsing delay or loading time of the page.

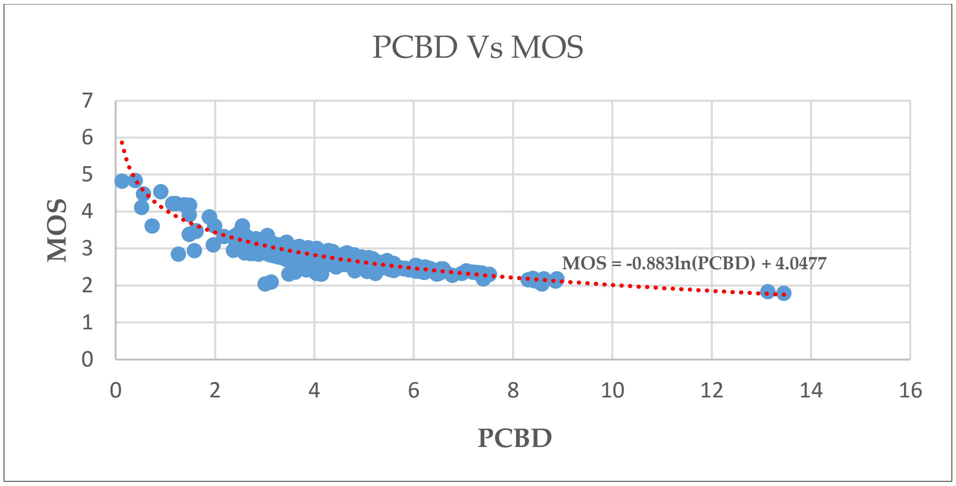


Figure 5-5 The Relationship Between PDTH and MOS.

PDTH Vs QoE

The third feature is page download throughput. This parameter is defined also in section 4.3.1 as follow, PDTH indicates the speed at which a web page is cached to a subscriber's mobile equipment after the subscriber writes a Uniform Resource Locator (URL) or re-freshes a web page. as Table 4-5 illustrates, the page download throughput has weak correlation with QoE (i.e., 0.275), implying that it has less influence on web browsing QoE. Figure 5-6 shows, the relationship between PDTH and MOS. The parameter has a linear relationship with QoE, and it has also a positive impact on user perception.

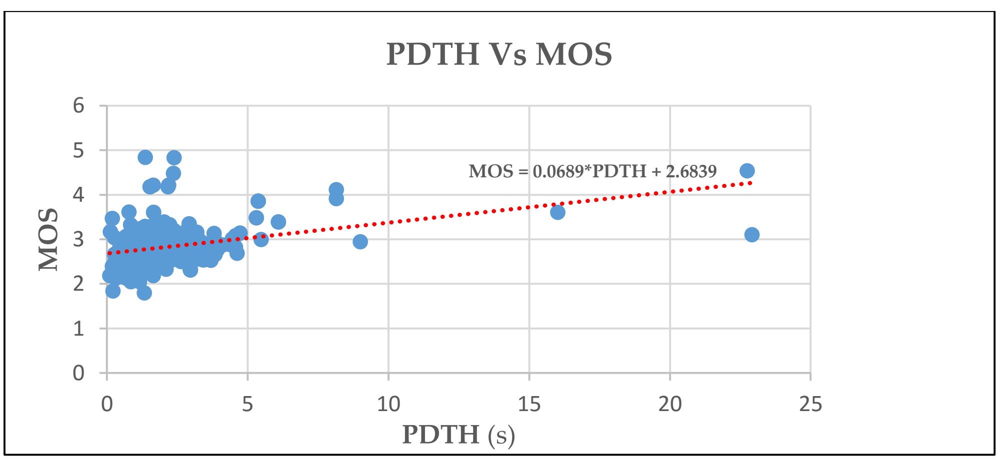


Figure 5-6 The Relationship Between PDTH and MOS.

Chapter 6 Conclusion and Future Work

This chapter presents two sections. The first section presents a conclusion that summarized the thesis work. And the second section describes recommendation or future work that are not addressed in this thesis.

6.1. Conclusion

The proposed goal for this thesis was to develop a model that estimate QoE by considering the QoS metrics. Taking this into account, web browsing QoE model in LTE network is developed using neural network approach, more specifically, artificial neural network algorithm.

The QoS data from ethiotelecom service monitoring tool (HUAWEI Smart Care) are using as inputs. also MOS values determined by an existing mapping function, that estimate the perceived quality for LTE web browsing service using webpage down load time are using as output, to develop the web QoE model. Webpage download time is one of the most essential factor that decide the Web browsing QoE.

Analyzing the collected input data is performed, to define the parameters having the highest influence on the web browsing QoE. Only the parameters that are more correlated with the web browsing QoE are selecting as inputs for the model. According to the Pearson correlation result, three parameters are selecting among the available five parameters. The selected parameters were Page response delay (in s), Page content browsing delay (in s) and Page download throughput (in kbps).

The Neural network approach is using to develop web browsing QoE model, to estimate web browsing QoE from the selected QoS parameters. The web browsing QoE model takes PRD, PCBD and PDTH as input parameters. As discussed in Chapter 4, The input data set is divided randomly in to three kind of sample data as follows: 70% for training, 15% for validation and 15% for test. The testing sample data were used for measuring the performance ability of the neural network model. The proposed work used R square and MSE as a performance metrics to check the model performance.

The model performs a QoE estimation with a R square value of 97.2% and a MSE of 0.002. The R square and MSE are all measured by comparing the estimated QoE with the measured or target QoE values. As the result indicated, the estimated and measured QoE values is highly correlated. And the error between them is very low. So, this model can be used for estimating the web browsing QoE for the mobile operator to get the benefit of objective measurement, instead of using subjective QoE measurement. And, it can be used by mobile network operators to prioritize the improvement of network factors that most influence web QoE.

6.2. Future Work

- This thesis proposes a specific QoE model for only web browsing service. For future work, new model applied to different services can developed for each technology, that estimates the QoE for all available services.
- The proposed model used only delay and throughput as input. But, the model can be developed using other network performance parameters, to have better accuracy.
- The spatio-temporal distribution of the user perception on web browsing should be studied.

References

[1] B. B. Haile, D. A. Bulti, and B. M. Zerihun, “On the Relevance of Capacity Enhancing 5G Technologies for Ethiopia,” no. June, 2017.

[2] E.telecom, “ethio telecom 2017 / 18 Business Plan,” no. February, 2017.

[3] HUAWEI, “Low Level Design Documentation for eUTRAN (LTE) Mobile Network in Addis Ababa V9.” .

[4] V. Cristina, “Identifying Quality of Experience (QoE) in 3G / 4G Radio Networks based on Quality of Service (QoS) Metrics,” no. November, 2017.

[5] Z. Zhen, “The effect of mobile cellular network performance and contextual factors on smartphone users ’ satisfaction A study on QoE evaluation for YouTube video,” 2015.

[6] European Telecommunications Standards Institute, “Etsi tr 102 643,” vol. 2, pp. 1–37, 2010.

[7] T. R. Team, *Customer Satisfaction Survey on Mobile Services : National Level Residents Presented to : Ethio telecom*. 2017.

[8] V. A. Siris, “Mobile Quality of Experience : Recent Advances and Challenges,” pp. 425–430, 2014.

[9] V. Pedras and M. Sousa, “A No-Reference User Centric QoE Model for Voice and Web Browsing based on 3G / 4G Radio Measurements,” *2018 IEEE Wirel. Commun. Netw. Conf.*, pp. 1–6, 2018.

[10] P. Anchuen, P. Uthansakul, and M. Uthansakul, “QOE Model in Cellular Networks Based on QOS Measurements Using Neural Network Approach,” 2016.

[11] A. Balachandran *et al.*, “Modeling web quality-of-experience on cellular networks,” *Proc. Annu. Int. Conf. Mob. Comput. Networking, MOBICOM*, no. September, pp. 213–224, 2014.

[12] A. Singh *et al.*, “Enhancing quality of experience (QoE) assessment models for web traffic,” *Lect. Notes Inst. Comput. Sci. Soc. Telecommun. Eng. LNICST*, vol. 125 LNICST, pp. 202–215, 2013.

[13] C. Kappler, *UMTS Networks and Beyond*. 2008.

[14] M. B. S.Sesia, I.Toufik, *LTE - The UMTS Long Term Evolution*, 2nd-. France: A John Wiley & Sons, Ltd., 2011.

[15] R. Kreher, *LTE Signaling,Troubleshooting , and Optimization*. Germany: Wiley & Son, Ltd, 2011.

[16] D. Soldani, M. Li, and R. Cuny, *QoS and QoE Management in UMTS Cellular Systems*. 2006.

[17] M. Services and L. T. E. Coverage, "Assessment of Quality of Service," no. June, 2017.

[18] T. S. S. of ITU, "ITU-T G.1030," *Estim. end-to-end Perform. IP networks data Appl.*, 2014.

[19] P. Paul and G. Quraishy, "Impact of HTTP Object Load Time on Web Browsing QoE," *Electr. Eng.*, no. August, 2013.

[20] HUAWEI, *SEQ Analyst Solution Product Documentation*, vol. 05. 2016.

[21] B. Hoffman, "How Website Speed Actually Impacts Search Ranking." [Online]. Available: https://moz.com/blog/how-website-speed-actually-impacts-search-ranking. [Accessed: 30-Jul-2019].

[22] European Telecommunications Standards Institute, "Definition of Quality of Service parameters and their computation," *ETSI TS 102 250-2*, vol. 1, pp. 1–252, 2017.

[23] M.Rukh Fida, "Understanding Mobile Network Quality and Infrastructure with User-Side Measurements," Edinburgh, 2018.

[24] D. Neves, D. Hora, D. E. L. U. Pierre, E. T. Marie, and C. De, "Predicting home Wi-Fi QoE from passive measurements on commodity access point," 2018.

[25] International Telecommunication Union, "P.800: Methods for subjective determination of transmission quality," *ITU-T Recomm.*, vol. 800, 1996.

[26] E. Vesterinen and V. President, "Content • QoS – QoE framework for data services," no. November, 2014.

[27] S.Haykin, *Neural Networks and Learning Machine*, 3rd ed. PEARSON.

[28] S. Benkachcha, E. H. H, P. Km, and R. El, "Seasonal Time Series Forecasting Models

based on Artificial Neural Network," vol. 116, no. 20, pp. 9–14, 2015.

[29] F. S. Panchal and M. Panchal, "Review on Methods of Selecting Number of Hidden Nodes in Artificial Neural Network," *Int. J. Comput. Sci. Mob. Comput.*, vol. 3, no. 11, pp. 455–464, 2014.

[30] J. pi. J.Han, M.Kamber, *Data Mining Concept and Techniques*, Third. USA: Morgan Kaufmann, 2012.

[31] P. D. I. Torino, "Machine Learning for Network-Based Prediction of Web browsing QoE," no. May, pp. 1–55, 2018.

[32] M. H. Beale, M. T. Hagan, and H. B. Demuth, *Neural Network Toolbox ™ 7 User ' s Guide*. 2010.

[33] A. B. Mtaho and F. R. Ishengoma, "Factors Affecting QoS in Tanzania Cellular Networks," *Int. J. Comput. Sci. Netw. Solut.*, vol. 2, no. 4, pp. 29–36, 2014.

[34] J. Joskow and T. Yamane, *Statistics, an Introductory Analysis.*, 2nd ed., vol. 60, no. 310. NEW YORK: HARPER & ROW, 1965.

[35] Webpagetest, "Webpagetest Documentation Metrics." [Online]. Available: https://sites.google.com/a/webpagetest.org/docs/using-webpagetest/metrics. [Accessed: 05-Jul-2019].

[36] Webpagetest, "Webpagetest Documentation Quick Start Guide." [Online]. Available: https://sites.google.com/a/webpagetest.org/docs/using-webpagetest/quick-start-quide. [Accessed: 10-Jul-2019].

[37] A. Savoia, "Web Page Response Time," *STQE*, no. August, pp. 48–53, 2001.

[38] P. Ameigeiras, J. J. Ramos-Munoz, J. Navarro-Ortiz, P. Mogensen, and J. M. Lopez-Soler, "QoE oriented cross-layer design of a resource allocation algorithm in beyond 3G systems," *Comput. Commun.*, vol. 33, no. 5, pp. 571–582, 2010.

[39] M. C. Madden, M. Friedrnan, L. L. Keyes, H. S. Koren, and G. R. Burleson, "Scatterplots and Correlation," *Inhal. Toxicol.*, vol. 3, no. 1, pp. 73–90, 1991.

[40] G. Zhang, B. Eddy Patuwo, and M. Y. Hu, "Forecasting with artificial neural networks: The state of the art," *Int. J. Forecast.*, vol. 14, no. 1, pp. 35–62, 1998.

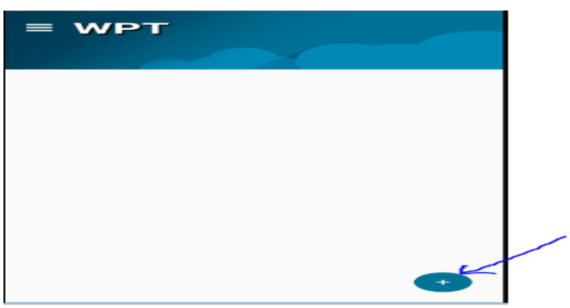
Appendix A

Questionnaire

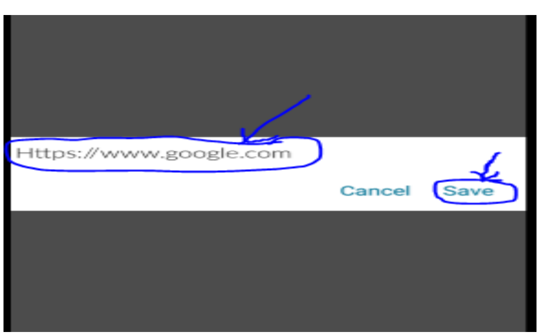
<u>Instruction</u> to use the **<u>Website Performance Tester (WPT)</u>** <u>application.</u>

<u>**Step 1**</u>. Download and install **WPT** application from play store or apple store.

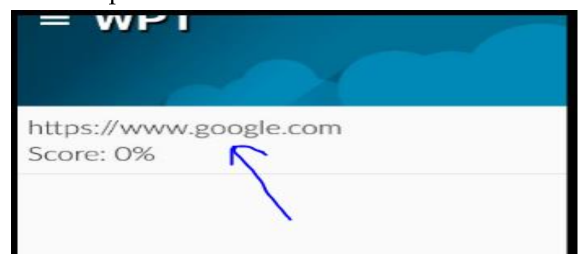
<u>**Step 2.**</u> Make your network selection is 4G (LTE).
<u>**Step 3**</u>. Enable the cellular data connection.
<u>**Step 4.**</u> Run the application



<u>**Step 5**</u>. Write the required website name on the provided space and save it



<u>**Step 6.**</u> Click the save website name to start the website performance measurement.



<u>**Step 7.**</u> Finally, the measurement result will be displayed as followed.

<u>**Step 8.**</u> Record the values of **Total load time**, **first byte received**, **started rendering** and **Byte in** to respond **Q1.**

<u>**Step 9.**</u> Repeat the above procedures to perform tests for different websites.

Part-I Impact of the parameter on the service

Q1. Please fill the website name and the value for each parameter. (Read from **WPT** application.)

No	Website Name	Total load time	First byte received	Started rendering	Byte in
1					
2					
3					

Q2. Choose the time period in which you get the above measurement result.

☐ Early Morning (5:00AM - 8:30AM ☐ Morning (8:30AM - 12:00PM)
☐ Lunch time (12:00PM - 1:30PM) ☐ Afternoon (1:30PM - 6:00PM
☐ Evening (6:00PM – 9:00PM) ☐ Night (9:00PM - 10:00PM)

Q3. Choose the location in which you get the above measurement result.

☐ Central AA ☐ East AA ☐ North AA ☐ West AA ☐ South AA
☐ South West AA

Part-II User perception about the service.

Please put "X" mark in the box to choose the response that best presents your level of perception on the web browsing service.

Note that *(**5** = Very satisfied, **4** = Somewhat satisfied, **3** = Neither satisfied nor dissatisfied, **2** = Somewhat dissatisfied, **1** = Very dissatisfied).*

Q4. According to your response on Q1, how do you feel about its page loading time?

☐ 5 ☐ 4.75 ☐ 4.5 ☐ 4.0 ☐ 3.75 ☐ 3.5 ☐ 3.0
☐ 2.75 ☐ 2.5 ☐ 2.0 ☐ 1.75 ☐ 1.5 ☐ 1.0

Part-III Demographic Details

Q5. Your gender is

☐ Male ☐ Female

Q6. Your Age

☐ Under 15 ☐ 15 - 24 ☐ 25 - 54 ☐ 55 - 64 ☐ 65+

Q7. Your educational back ground is

☐ Master's degree and above ☐ Bachler's degree ☐ University or College Student
☐ Diploma or Certificate ☐ Grade 11 - 12 (Preparatory) ☐ Grade 10 and Below

Q8. Any other comments
