DOKUZ EYLÜL UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

THE VOICE QUALITY IMPROVEMENT AND PERFORMANCE MANAGEMENT FOR SMARTPHONES

by

Şefika Esra ULUTAŞ

June, 2019

İZMİR

THE VOICE QUALITY IMPROVEMENT AND PERFORMANCE MANAGEMENT FOR SMARTPHONES

A Thesis Submitted to

the Graduate School of Natural and Applied Sciences of Dokuz Eylül University

In Partial Fulfillment of the Requirements for the Degree of Master of Science in Electrical and Electronics Engineering

by

Şefika Esra ULUTAŞ

June, 2019

İZMİR

M.Sc THESIS EXAMINATION RESULT FORM

We have read the thesis entitled **"THE VOICE QUALITY IMPROVEMENT AND PERFORMANCE MANAGEMENT FOR SMARTPHONES"** completed by **ŞEFİKA ESRA ULUTAŞ** under supervision of **ASSIST. PROF. DR. REYAT YILMAZ** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Assist. Prof. Dr. Reyat YILMAZ

Supervisor

Prof. Dr. Alp Kut

(Jury Member)

Prof. Dr. Mustafa Gündüzalp

(Jury Member)

Prof. Dr. Kadriye ERTEKİN

Director

Graduate School of Natural and Applied Sciences

ACKNOWLEDGEMENT

I would like to express my gratitude to my supervisor Assist. Prof. Dr. Reyat YILMAZ for the continuous support, guidance and and patience in every stage of my thesis. It would not have been possible to write this thesis without his support.

Also importantly, this thesis would not have been possible without the patience and everlasting support of my family. I would like to send all my love to my father Mustafa ULUTAŞ and my mother Emine ULUTAŞ for their endless support and understanding during my education.

I wish to thank my colleagues and friends Kemal ATAKAN, Ali AYTEKİN and Saffet İlkay ULUSOY who always give me motivation and encouragement. They never left me alone when I need motivation and help.

This work has been supported by The Scientific and Technological Research Council of Turkey (TUBITAK) under Grant No. Teydeb-3160452. The project name is “Development and performance management of sound quality of Vestel new generation intelligent mobile phones”. In addition to this, this project is supported by Vestel Elektronik Sanayi ve Ticaret A.Ş.

Şefika Esra ULUTAŞ

THE VOICE QUALITY IMPROVEMENT AND PERFORMANCE MANAGEMENT FOR SMARTPHONES

ABSTRACT

In today's global World, voice transmission has become an effective form of the communication. The expectation of the better sound quality is increased with the increasing importance of voice transmission. In order to measure sound quality, there are two basic methods: subjective and objective approach.

The system which is introduced in this thesis is used to measure and improve the sound quality of the mobile-based device according to mean opinion score scale. The measurements were made according to audio codec bit rate and signal strength at the audio bandwidths which are narrowband, wideband and super wideband. This study is performed according to ITU-T Recommendation P.863.

The results show that eighty percent, ninety percent and seventy seven point seventy seven percent sound quality improvement at downlink; ninety one point sixty six percent, eighty eight point eighty eight percent and eighty percent at uplink signal direction are achieved respectively for 2G, 3G and 4.5G mobile network.

Keywords: Smartphone audio transmission, audio quality measurement methods, POLQA, ITU-T Recommendation P.863, audio quality improvement

AKILLI TELEFONLARIN SES KALİTESİNİN İYİLEŞTİRİLMESİ VE PERFORMANS YÖNETİMİ

ÖZ

Günümüz küresel dünyasında, ses iletimi etkili bir iletişim biçimi haline gelmiştir. Ses iletiminin önemi arttıkça daha iyi ses kalitesine olan beklenti de artmıştır. Konuşma kalitesini ölçmek için subjektif ve objektif olarak iki temel yöntem vardır.

Bu tezde, mobil tabanlı cihazın ses kalitesini ölçmek ve iyileştirmek için bir sistem geliştirilmiştir. Ölçümler; dar bant, geniş bant ve süper geniş bant olan ses bant genişliklerinde ses kodek bit hızına ve sinyal gücüne göre yapılmıştır. Bu çalışma ITU-T Tavsiye P.863'e göre gerçekleştirilmiştir.

Sonuçlar, 2G, 3G ve 4.5G mobil ağlarında, gelen sesin ses kalitesinin iyileştirilmesi için sırasıyla yüzde seksen, yüzde doksan ve yüzde yetmiş yedi tam yüzde yetmiş yedi; giden sesin iyileştirilmesi için sırasıyla yüzde doksan bir tam yüzde altmış altı, yüzde seksen sekiz tam yüzde seksen sekiz ve yüzde seksen iyileşme elde edildiğini gösterir.

Anahtar kelimeler: Akıllı telefon ses iletimi, ses kalitesi ölçüm yöntemleri, POLQA, ITU-T Tavsiye P.863, ses kalitesinin iyileştirilmesi

CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER ONE

INTRODUCTION

Voice communication with mobile-based devices over telecommunication networks is now part of everyday life. Voice can be transmitted or received from point to another with the mobile-based device by using a circuit switched or packet switched network with the developing telecommunication generation (Sharma, 2013). The importance of sound quality has increased and the expectation for better sound quality has increased in line with the expansion of voice communication and the development of telecommunication generations.

The voice quality is defined as goodness level on perception of talking normally. Various subjective and objective methods of measurement are available to assess speech quality. Subjective listening tests are the most reliable way to obtain accurate measurement of the user's voice quality perception and have good results in terms of correlation with actual speech quality but they are very expensive and time consuming (Kondo, 2012). As a long time, subjective measurement method have been the only method of assessing sound quality. Because such tests are generally not expensive or practical, objective measurement methods based on the psycho-acoustic behavior of the human ear have been developed. An objective measurement method uses a computerized model that measures sound quality much cheaper than subjective methods and performs well in terms of correlation with subjective measurements. P series International Telecommunication Union (ITU) standards are widely used to measure speech quality objectively.

Perceptual objective listening quality assessment that is briefly called ITU-T Recommendation P.863 or POLQA, is standardized by International Telecommunication Union - Telecommunication Standardization Sector (ITU-T) and is last published standard of objective measurement method. The most commonly used objective speech quality evaluation method is ITU-T Recommendation P.863. This method is very suitable for evaluating the quality of speech on 2G, 3G or 4.5G mobile network, because it supports the voice transmission of packet switch and circuit switch. Voice is transmitted or received over LTE with packet switch in 4.5G mobile network. Voice is transmited or received over 2G or 3G with circuit switch (ITU-T P.863, 2014).

Until now, there have been several studies for measurement, evaluation or improvement of voice quality. However, no study has been conducted that fulfills these three criteria. A work by Kondo (2012), sound quality evaluation methods and standards have been mentioned and sound quality measurement methods have been evaluated. However, there is no study on measuring and improving sound quality. Ebem (2009) validated Perceptual Objective Listening Quality Assessment (POLQA) benchmark, which is the new ITU-T benchmarking for objective measurement of speech quality, on the tonal language Igbo – a language spoken in south eastern-Nigeria. The results obtained by using POLQA standard objectively are validated by subjective methods. Goudarzi (2008) researched that the speech quality in 3G mobile networks is evaluated by setting up a testbed platform based on Asterisk open source PBX to mediate between 3G mobile network and quality measurement equipment with ITU-T P.862. However, no improvement studies were made for sound quality. In another work by Tüysüz (2007), improving voice quality in Voice over IP and its factors as well as network impairments and their causes in a converged telephony and internet protocol (IP) network, all from the points of view of the network quality, is explained. In the study of Assem (2013), it was evaluated as subjective and objective evaluation of audio and video sent via internet protocol and improvement studies were done.

The thesis is organised into 5 chapters as follows:

Chapter 2 provides a summary of the mobile network technologies, in which the voice transmission is provided, namely the telecommunications generation. This chapter handled the features of 4G technologies, the first generation, from the 1G to the latest generation.

Chapter 3 presents a summary of the current literature on sound quality measurement techniques related to this research. The main ideas and the basic principles of subjective and objective sound quality measurement are presented. In addition to this, the methods and standards used in sound quality measurements are discussed in detail in this section.

Chapter 4 decribes on the technical aspects of ITU-T Recommendation P.863 that is Perceptual Objective Listening Quality Assessment (POLQA), the objective measurement method selected to determine the sound quality. The characteristics of the reference signal which should be used in the measurements are mentioned. It is discussed how to obtain the mean opinion score that is briefly called MOS with using POLQA algorithm.

Chapter 5 is the most important section of this thesis since it presents the introduced system and its performance. The detailed specifications and necessary equipment of the system for performing sound quality tests using the ITU-T recommendation P.863 standard are presented. The sound quality is measured in two different ways that are according to audio codec bit rate variables and signal strength variables. According to the measurement results, improvement studies were performed in conditions where sound quality is not satisfied. The results of the improvement studies were analyzed and evaluated.

Finally, Chapter 6 is the conclusion part and it summarizes and concludes the thesis.

CHAPTER TWO

GENERATIONS IN TELECOMMUNICATION

The cellular networks are evolving through several generations which largerly specifies the type of services and the data transfer speeds class of technologies. Mobile technologies have experience 4.5G generations of technology evolution, namely from 1G to 4.5G. Current research is contining with 5G technology. The evoluation to 5G is shown Figure 2.1 from past to present. Currently 5G technology is not officially used.

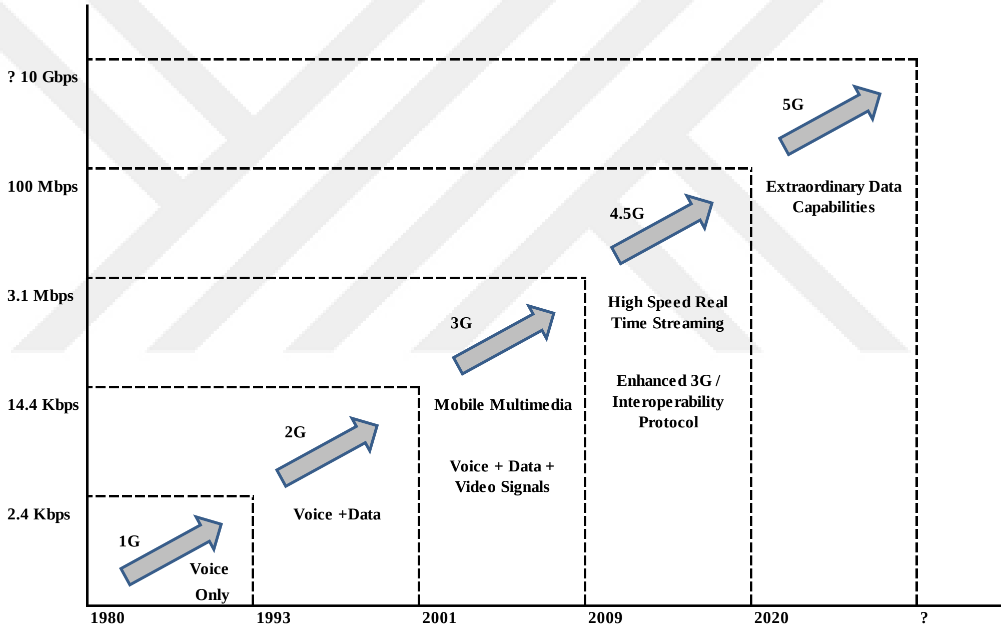

Figure 2.1 Evaluation to generation of 5G technology

2.1 First Generation Technology

First generation mobile network appeared in 1980 under the symbol 1G. The location of first commercialization is USA (Seo, 2013). It is the analog telecommunications standards which are the use of multiple cell sites and the ability to transfer calls from one site to the next as the user travelled between cells during the conversation. It is transmitted in analogue made and is able to handle only voice traffic.

Advantage of this technology is simple network elements. On the other hand, disadvantages of it are limited capacity, not secure, poor battery life, large phone size, poor audio quality and background interference.

Some features are listed below about 1G (Lopa & Vora, 2015):
Technology: Advanced Mobile Phone System (AMPS), Nordic Mobile Telephones (NMT) and Total Access Communication System (TACS)
Network Switching Type: Circuit Switching
Speed (Data Rates): 2.4 kbps
Supports: Voice
Internet Service: No internet
Bandwidth: Analog
Operating Frequency: 800 MHz
Band (Frequency) Type: Narrowband
Carrier Frequency: 30 kHz
Applications: Voice calls
Features: Voice only

2.2 Second Generation Technology

Second generation mobile network appeared in 1993 under the symbol 2G. The location of first commercialization is Finland. (Seo, 2013). It is based on digital transmission that offers several benefits over analog transmission. For encoding or decoding the digital signal, codec is used. The some benefits are listed below.

- Digital traffic channels
- Encryption-preventing eavesdropping
- Error detection and correction-giving clear voice reception
- Channel access- allow channels to be dynamically shared by a number of users.

The 2G mobile network supports voice and data services. Voice is the priority, data has very slow speed. This data services starting with Short Message Service (SMS) plain text based messages. In addition to this, the data services enable the various mobile phone networks to provide picture messages and Multimedia Message Service (MMS) (Gawas, 2015).

Advantages of this technology are that digital signals require consume less battery power, so it help mobile batteries to last long. Digital coding improves the voice clarity and reduces noise in the line. Digital encryption has provided secrecy and safety to the data and voice calls. On the other hand, the disadvantages of it are that network range and slow data rates.

Some features are listed below about 2G (Lopa & Vora, 2015):

Technology: Interim Standard 95 (IS-95) was the first ever CDMA - based digital cellular technology and Global System for Mobile Communications (GSM)

Network Switching Type: Circuit Switching for Voice and Packet Switching for Data

Speed (Data Rates): 14.4 kbps

Supports: Voice and Data

Internet Service: Narrowband

Bandwidth: 25 MHz

Operating Frequency: 900 MHz and 1800 MHz for GSM and 800 MHz for CDMA

Band (Frequency) Type: Narrowband

Carrier Frequency: 200 kHz

Applications: Voice calls, short messages, browsing(partial)

Features: Multiple users on single channel

2.3 Third Generation Technology

Third generation mobile network appeared in 2001 under the symbol 3G. The location of first commercialization is Japan. 3G was developed to overcome the drawbacks of 2G. It has better data/internet speed. But, voice is still the priority on

3G network. It supports voice calling, mobile TV and mobile internet access. The main difference is that distunguishes 3G technology from 2G technology is the use of packet switching rather than circuit switching for data transmission (Mehta et al., 2014). An evolution of this technology begun to implemented, namely High-Speed Packet Access (HSPA) that has high speed data and new features. In HSPA, voice and data are used simultaneously. That means, you can access the data even if you are on voice call.

Advantages of this technology are providing faster communication than 2G, high security and international roaming. On the other hand, disadvantages of it are high power consumption, low network coverage and high cost of spectrum licence.

<u>Some features are listed below about 3G (Lopa & Vora, 2015):</u>

Technology: International Mobile Telecommunications-2000 (IMT2000 is developed by the International Telecommunications Union (ITU). It provides worldwide mobile broadband media services via a single global frequency band. Wideband Code Division Multiplexing Access (WCDMA)

Network Switching Type: Packet Switching expect for Air Interface

Speed (Data Rates): 3.1 Mbps

Supports: Voice and Data

Internet Service: Broadband

Bandwidth: 25 MHz

Operating Frequency: 2100 MHz

Band (Frequency) Type: Wideband

Carrier Frequency: 5 MHz

Applications: Video conferencing, Mobile TV and Global Positioning System

Features: Multimedia features and Video Call

2.4 Fourth Generation Technology

Fourth generation mobile network appeared in 2009 under the symbol 4G. The location of first commercialization is South Korea. Data is the priority in 4G network. Data speed is very high, so it supports good quality of voice calling, video calling and video streaming services (Kumar et al., 2010). The main difference between 4G and 3G is using an all - IP network with elimination of circuit switching. The 4G is giving in a treatment of voice calls just like any other type of streaming audio media utilizing packet switching. The expectation for the 4G technology is basically the high quality audio or video streaming over end to end Internet Protocol. Over time, 4G was developed. The name after this development is LTE that means Long Term Evolution. It provides very high data speed. The data is end to end digitally encrypted. (Chen et al., 2007).

Advantages of this technology are data speed, high speed handoffs, multi-input and multi-output technology (MIMO) and global mobility. On the other hand, disadvantages of it that are hard to implement and complicated hardware required (Mishra, 2004).

<u>Some features are listed below about 4G (Lopa & Vora, 2015):</u>

Technology: Long Term Evoluation (LTE) and Worldwide Interoperability for Microwave Access (WiMAX)

Network Switching Type: Packet Switching

Speed (Data Rates): 100 Mbps

Supports: Voice and Data

Bandwidth: 100 MHz

Operating Frequency: 850 MHz and 1800 MHz

Band (Frequency) Type: Super Wideband

Carrier Frequency: 15 MHz

Applications: High speed applications, mobile TV, wearable devices

Features: High speed data, real time streaming

CHAPTER THREE

MEASUREMENT OF SPEECH QUALITY

Speech quality is defined as the grade of goodness in the perception of speaking normally; speech intelligibility is how well or clearly one can understand what is being said. Essentially quality is the gap between what you expect and what you really get. People's perceptions, cultural attitudes, preferences and expectations are added to evaluation for producing a valid quality scale. The quality perception is guided by experience and expectation. Speech quality in telecommunication is an important factor for the success of a product and the success of communication. The high quality of speech ensures that the user's effort is low for speech intelligibility.

Speech quality can be measured in two ways: subjective and objective measurements (Pourmand, 2013). The speech quality measurement methods are shown Figure 3.1 according to types of the measurement. Subjective tests includes that evaluating the quality of speech according to a group of listeners. The subjective measurement methods are such as ITU-T Recommendation P.800, P.805, P.835 and ITU-R BS.1534-1 (Kondo, 2012). The objective tests take measurements from sound samples and apply an algorithm to the data to obtain the perceived sound quality of the user. There are two methods of measuring the objective measurement methods: intrusive and non-intrusive. The intrusive methods are such as ITU-T P.861, Perceptual Analysis Measurement System (PAMS), Measuring Normalizing Blocks (MNB), ITU-T P.862, ITU-T P.863. The non-intrusive methods are such as E-model and ITU-T P.563 (Quackenbush et al., 1988).

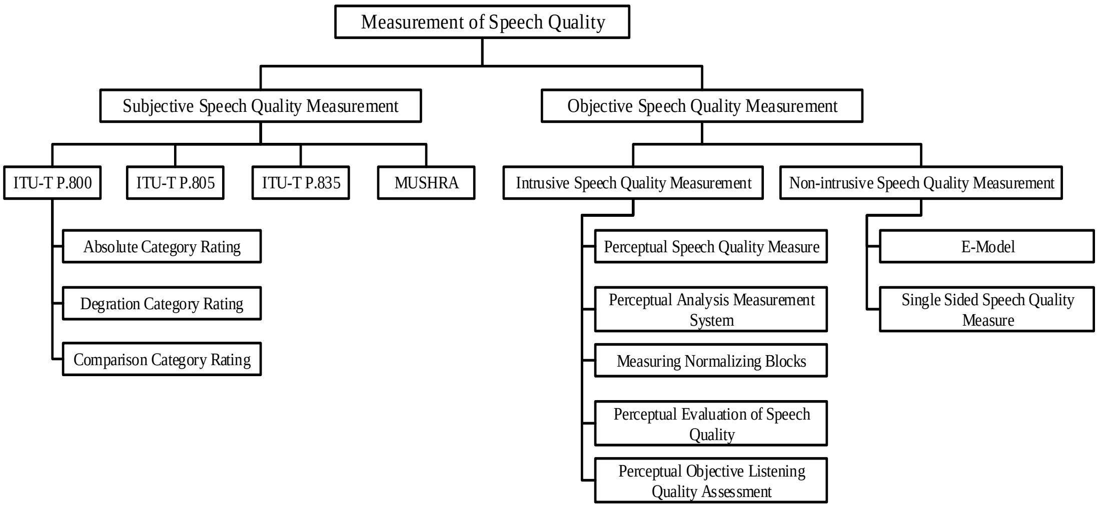

Figure 3.1 Measurement methods of speech quality

Although the validity of the subjective measures is high, it is expensive and time consuming. For these reasons, objective measures are used which are less costly, less time-consuming, do not need human listeners and measure speech quality from clean and distorted signals.

3.1 Subjective Speech Quality Measurement

People are the end users of telecommunication systems, so subjective listeners are not only the best quality measure, but also the true quality measure. The subjective measurement of speech quality is based on a perceptual assessment of speaking by people who listen to live or recorded speech and subjectively evaluate it according to a pre-determined scale (Côté, 2011). Test participants should be provide necessary conditions, even if they are randomly selected. For example, they should have knowledge of telecommunication systems and quality assessment. The purpose of subjective assessment methods; determining people's preferences, accurately reflecting real-life scenarios, getting accurate results, and minimizing costs.

The benefits of the subjective measurement method are listed below (Rango et al., 2006).

- More reliable than objective measurements.
- Tests evaluate the subjective quality of the voice or evaluate how it is perceived by the end user.
- All kinds of degraded are added to the measurement
- It gives definite results.
- It is useful for evaluating an audio system with any kind of speech or music.

The disadvantages of the subjective measurement method are listed below (Rango et al., 2006).

- It consumes more time and money.
- It is difficult or impossible to repeat the ambient conditions during the test.

Subjective measures have two groups of listeners that divided into average listeners and expert listeners. The average listeners are uneducated and inexperienced about speech evaluation. For this reason, tests with average listeners are less accurate than expert listeners. But it is the situation that best corresponds to the real life scenario. A large number of listeners are required per measurement. The number of listeners should be between 50 and 200 people. This is also expensive and time consuming. The expert listeners have experience and training about listening. The one to one match for a real life scenario is low. As a result of this, the subjective test with the expert listeners is more consistent and reliable results with fewer listeners. The mean that fewer listeners is that maximum 10 people. This is more cheaper and faster than the subjective measurement with average people (ITU-T Recommendation P.830, 1996).

3.1.1 ITU-T Recommendation P.800

One of the subjective measurement methods is the ITU-T Recommendation P.800 that is "Methods for subjective determinations of transmission quality". The ITU-T 800 describes the necessary procedure for the subjective evaluation of speech distorted by different factors such as environmental noise, transmission error, echo etc. in telecommunication systems. In addition to this, the speech material and the features of the speakers and listeners are detailed in recommendation. The listener should be inexperienced and the speech material should be selected randomly from simple, meaningful, short sentences selected from non-technical current literature or newspapers, and easy to understand (ITU-T Recommendation P.800, 1996).

The ITU-T Recommendation P.800 has three subcategories that are Absolute Category Rating (ACR), Degration Category Rating (DCR) and Comparison Category Rating (CCR).

3.1.1.1 Absolute Category Rating (ACR)

In the Absolute Category Rating that is briefly ACR, the evaluation is made without comparison to a non-degraded reference signal that is original signal because the listener can not access the other side original voice in the real life telecommunication scenario. Therefore, the signal quality is evaluated with an absolute quality according to a scale that is a 5-point scale (Möller, 2000). The listener evaluates the speech quality using this scale. The scale is shown at Table 3.1.

Table 3.1 Absolute Category Rating scale (ITU-T Recommendation P.800, 1996)

Quality of Speech	Score
Excellent	5
Good	4
Fair	3
Poor	2
Bad	1

The evaluations are collected in a listener pool, and the arithmetic average of these evaluations is the overall result of the test. This overall result is known as Mean Opinion Score or MOS (ITU-T Recommendation P.800.1, 2003). In the ACR test, a test pool should contain 16 or more people for obtaining a safe result, and should be tested in a quiet environment under controlled conditions. Generally, as the number of listeners increases, more stable scores are achieved.

Harvard Sentences: List 1 (Veen, 1969)

1. The birch canoe slid on the smooth planks.
2. Glue the sheet to the dark blue background.
3. It’s easy to tell the depith of a well.
4. These days a chicken leg is a are dish.
5. Rice is often served in round bowls.
6. The juice of lemans makes fine punch.
7. The box was thrown beside the parked truck.
8. The logs were fed chopped corn and gorbage.
9. For hours of steady work faced us.
10. Large size in stockings is hard to sell.

3.1.1.2 Degration Category Rating (DCR)

Degration Category Rating that is briefly DCR, is used to evaluate high quality sound samples. This measurement also known as the double-stimulus impair scale that is shortly DSIS. It is performed using two samples: the examples as A and B signals. The sample A represents the processed high quality reference signal, and the sample B represents the degraded signal.The listeners evaluate the quality impairment by the reference to the degraded speech signal by listening in pair (AB) or repeated pair (ABAB) that is pronounced on the same speaker (Rango et al., 2006). This evaluation is expressed by the Degraded Mean Opinion score, which is a 5-point evaluation scale. This score is also called Degraded MOS or DMOS. The Degraded Mean Opinion scale is shown in Table 3.2.

Table 3.2 Degration Category Rating scale (ITU-T Recommendation P.800, 1996)

Score	The degration is…
5	inaudible
4	audible but not annoying
3	slightly annoying
2	annoying
1	very annoying

3.1.1.3 Comparison Category Rating (CCR)

Comporison Category Rating that is briefly CCR, is used for evaluating high quality sound samples such as DCR. This measurement also known as the double-stimulus comporison scale that is shortly DSCS. It is performed using two samples like DCR: the examples as A and B signals. The processed and unprocessed samples are selected randomly in the CCR measurement method and the quality of first sample is compared to the second sample. The comparison is expressed by the Comporion Mean Opinion Score, which is a scale of 7. This score is aslo called CMOS. The Comporion Mean Opinion Score scale is shown in Table 3.3.

Table 3.3 Comporison Category Rating scale (ITU-T Recommendation P.800, 1996)

Score	Quality of the second compared to the quality of the first …
3	Much better
2	Better
1	Slightly better
0	About the same
-1	Slightly worse
-2	Worse
-3	Much worse

3.1.2 ITU-T Recommendation P.805

The name of the ITU-T Recommendation P.805 is "Subjective evoluation of conversational quality". Purpose of this that measuring the effect of distortion such as delay, echo and packet loss in speech quality during speech. In this test; expert, experienced or naive listeners are used. The subjects speak through a communication chain in two separate soundproof rooms as in Figure 3.2 (ITU-T Recommendation P.805, 2007).

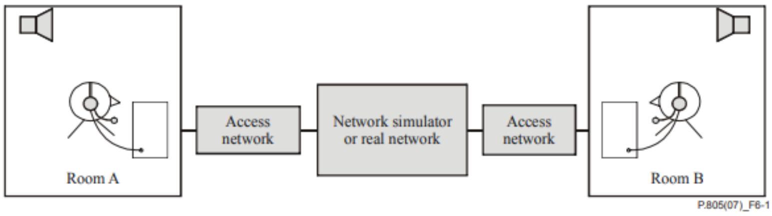

Figure 3.2 Test facilities for ITU-T Recommendation P.805 (ITU-T Recommendation P.805, 2007)

At the end of the talk in this communication chain; the subjects are asked to make an evaluation by answering the following questions that is shown Table 3.4, Table 3.5 and Table 3.6.

Table 3.4 Sample question 1 (ITU-T Recommendation P.805, 2007)

Question 1: What is your opinion of the connection you have just been using?	
Evaluation	**Score**
Excellent	5
Good	4
Fair	3
Poor	2
Bad	1

Table 3.5 Sample question 2 (ITU-T Recommendation P.805, 2007)

Question 2: Did you or your partner have any difficulity in talking or hearing over the connection?	
Evaluation	**Score**
Yes	1
No	2

Table 3.6 Sample question 3 (ITU-T Recommendation P.805, 2007)

Question 3: How well did you understand what the other person was telling you?	
Evaluation	**Score**
No loss of understanding	5
Minimal loss of understanding	4
Moderate loss of understanding	3
Considerable loss of understanding	2
Severe loss of understanding	1

3.1.3 ITU-T Recommendation P.835

The name of the ITU-T Recommendation P.835 is "Subjective test methodlogy for evaluating speech communication systems that include noise suppression algorithm". The ITU-T Recommenration P.835 describes a subjective test procedure for evaluating degraded speech quality, in particular with noise reduction algorithms. Typically, it is used to evaluate communication systems that contain noise suppression technologies. In the evaluation procedure, the listeners is determined the quality of speech by using separate rating scales for "Speech", "Backgorund Noise" and "Speech + Backgorund Noise (Overall noise)" (ITU-T Recommendation P.835, 2003). The test structure is shown in Figure 3.3.

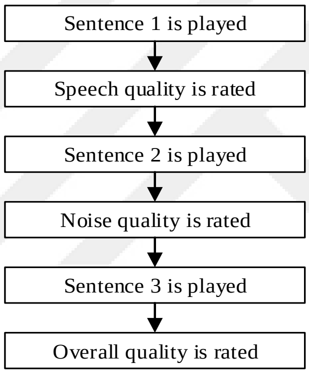

Figure 3.3 Test structure of ITU-T Recommendation P.835

The scales used in the "Speech", "Noise" and "Overall" grades evaluated in ITU-T P.835 are given in Table 3.7, Table 3.8 and Table 3.9 (Pourmand, 2013). If the speech signal is evaluated, attending ONLY to the SPEECH SIGNAL, the category which best describes the sample you just heard, is selected.If the noise is evaluated, attending ONLY to the BACKGROUND, the category which best describes the sample you just heard, is selected. If the overall signal is evaluated, the category which best describes the sample you just heard for purposes of everyday speech communication, is selected.

Table 3.7 Signal Rating for ITU-T Recommendation P.835 (ITU-T Recommendation P.835, 2003)

The SPEECH SIGNAL in this sample was	
Evaluation	**Score**
Not Distorted	5
Slightly Distorted	4
Somewhat Distorted	3
Fairly Distorted	2
Very Distorted	1

Table 3.8 Noise Rating for ITU-T Recommendation P.835 (ITU-T Recommendation P.835, 2003)

The BACKGROUND in this sample was	
Evaluation	**Score**
Not Noticeable	5
Slightly Noticeable	4
Noticeable but Not Intrusive	3
Somewhat Intrusive	2
Very Intrusive	1

Table 3.9 Overall Rating for ITU-T Recommendation P.835 (ITU-T Recommendation P.835, 2003)

The OVERALL SPEECH SAMPLE was	
Evaluation	**Score**
Excellent	5
Good	4
Fair	3
Poor	2
Bad	1

3.1.4 Recommendation ITU-R BS.1534-1

The name of Recommendation ITU-R BS.1534-1 is "Multiple Stimuli with Hidden Reference and Anchor", is MUSHRA shortly. The MUSHRA methodology is the most appropriate method for evaluating similar distortions with reference signal with using expert listeners. It is used to evaluate the sound quality of signals that have intermediate quality. The listeners listen to different processed samples of the same sound signal. These samples include a reference that is known by listeners, a hidden reference, and many anchors. The hidden reference signal is an unaltered copy of the original. The anchors are the modified versions of the original. In addition, these samples that include anchors and references, should be less than 15, and the perceptual difference between the samples should not be too small. The generated samples are presented to listeners randomly in a multi-scale user interface (Pulkki & Karjalainen). The listeners evaluate the audio quality with a scale defined between 0 and 100 for all samples except the known reference. This scale is given Table 3.10. The scale that is defined between 0 and 100, is divided into 5 equal intervals. The overall score is given as the average of all listeners (Recommendation ITU-R BS.1534-1, 2003).

Table 3.10 Recommendation ITU-R BS.1534-1 scale (Recommendation ITU-R BS.1534-1, 2003)

Quality of Speech	Score
Excellent	5
Good	4
Fair	3
Poor	2
Bad	1

3.2 Objective Speech Quality Measurement

Subjective quality measurement methods give the most accurate results to evaluate speech, but these measurement methods have disadvantages as we have already mentioned before. These disadvantages have been the main reason for the development of objective measures. Objective quality measurement methods evaluate speech quality based on physical parameters obtained from the system or speech signal being tested. The aim is to achieve the similar result as subjective tests with healthy listeners (Bäckström, 2017). Generally, objective speech quality is measured using some mathematical formulas with original speech and degraded speech signal. This measurement does not require human listeners, so it costs less and takes less time. In general, objective measurements are used to obtain rough estimates of speech quality (Rix et al., 2006).

The benefits of the objective measurement method are listed below (Benesty et al., 2008).

- Faster and cheaper than subjective methods
- Measurements that will last for months with subjective measurements can

be done with objective methods within minutes or hours, so the objective measurements are easy to repeat.

- If the objective measurement is make twice time in the same reference and degraded signal, it will give the same result.
- It can provide delay estimates between the reference signal and the degraded signal.
- No human need for these measurements.

The disadvantage of the objective measurement method is that it does not give exact results like the subjective measurements, it just predicts the behavior of the subjective listeners.

The objective measurements are classified in two groups; intrusive and non-intrusive measurement. The basis of this classification is based on the inputs needed for measurement. Intrusive and non-intrusive types of quality measurement are given Figure 8. Intrusive quality measurement requires access to the original and degraded speech signal. Non-intrusive quality measurement only makes quality estimates with the degraded signal. The intrusive measurement method is stronger for the quality assessment because it uses both the reference and the degraded signal (Kráčala, 2016).

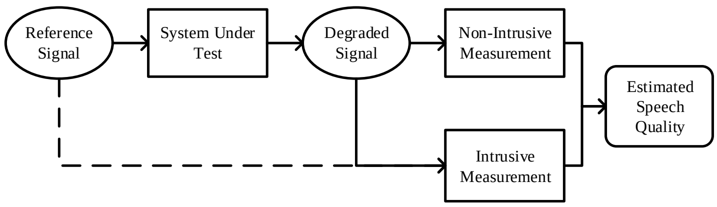

Figure 3.4 Intrusive and non-intrusive types of quality measurement

3.2.1 Intrusive Speech Quality Measurement

The intrusive speech quality measurement system typically uses two input signals that are a reference or original signal and a degraded signal from the network or the system under test (Cotanis, 2010). This measurement is also known as double-ended measurement methods since two input signals are used. These tests are insufficient for operations such as traffic monitoring, while the perceived end-to-end speech quality is very suitable method.

The intrusive model is the the best objective speech quality measurement method that is the perceptual domain measures based on the human voice perception model. This model is shown in Figure 9.The Perceptual Speech Quality Measure (PSQM), Perceptual Analysis Measurement System (PAMS), Perceptual Evaluation of Speech Quality (PESQ) and Perceptual Objective Listening Quality Assessment (POLQA) measures standardized for speech quality in communication systems by ITU-T represent perceptual domain measures (Sun, 2004).

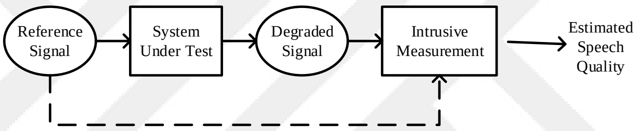

Figure 3.5 Intrusive speech quality model

3.2.1.1 Perceptual Speech Quality Measure (PSQM)

Perceptual Speech Quality Measure that is briefly PSQM, was developed by John G.Breerends in 1993 at KPN Research. Then the PSQM model was approved as the ITU-T standard and named "ITU-T P.861: Objective Quality Measurement of Telephone-Band Speech Codecs" in 1996. It consists of a mathematical algorithm that objectively evaluates the sound quality of speech codecs in the 300-3400 Hz sound band (ITU-T Recommendation P.861, 1996). The idea behind the PSQM model is to imitate the perception of the listeners as they are in real life. This method evaluates the difference between a degraded signal and a reference signal in a telephone system. The PSQM model consists of 3 separate blocks that are pre-processing, perceptual modelling and cognitive modelling, are shown in Figure 3.6 (Opticom, 2000).

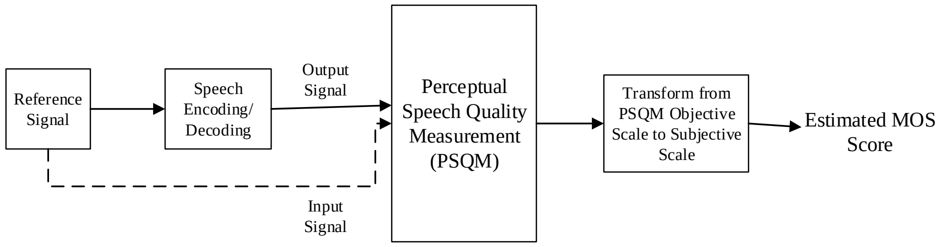

Figure 3.6 Perceptual Speech Quality Measure Model (ITU-T Recommendation P.861, 1996)

Perceptual Speech Quality Measure Model:

- Pre-processing: The input and output signals are checked for alignment. If these signals are not aligned, they will be aligned to a synchronous time zone.
- Perceptual modelling: The signal representation of the input and output signals is converted to obtain a real human perception signal with using a mathematical model.
- Cognitive modelling: The perceived quality is measured by comparing the input and output signals, and the noise impairment is calculated.

The block diagram of the PSQM algorithm is shown in Figure 3.7. The PSQM algorithm gives a PSQM value which is a criterion of signal quality degradation, which ranges from 0 to 6.5. The value of 0 means that it has no degradation. The value of 6.5 means that it has the highest degradation. These results can be converted into a MOS value, which is regarded as a measure of the perceived quality in a numerical scale ranging from 1 to 5. The value of 1 means that it is an unacceptable, and has low level quality. The value of 5 means that it has high sound quality.

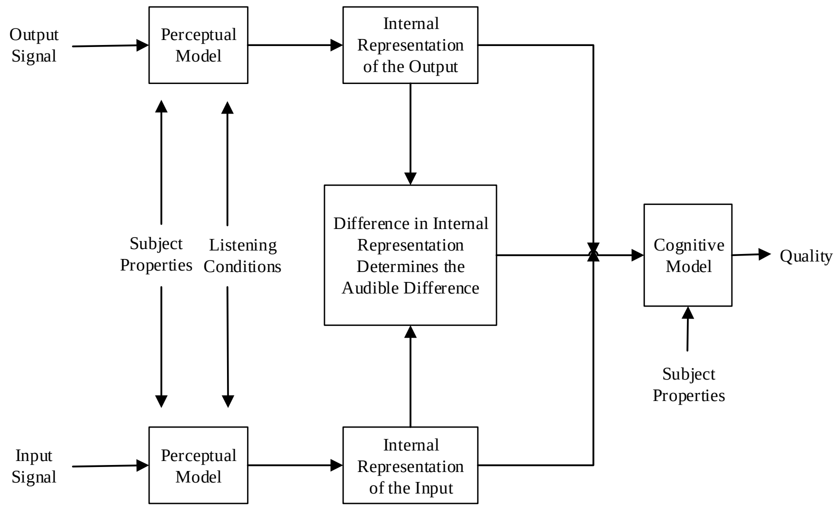

Figure 3.7 Block Diagram of Perceptual Speech Quality Measure (ITU-T Recommendation P.861, 1996)

3.2.1.2 Perceptual Analysis Measurement System (PAMS)

The Perceptual Analysis Measurement System that is briefly PAMS, was developed in 1998 by PsyTechnics that is an investigative group within the British Telecommunication (Rix & Hollier, 2000). For this reason, it is not an international standard defined by ITU-T. This speech quality algorithm is designed for mouth-ear evaluation of phone bandwidth transmission systems. Purpose of PAMS is that it measures the sound quality objectively by considering the time clipping, packet loss, delay and the degradation caused by codec usage in a system (Mohamed, 2003). This objective test uses a model based on human perception factors to measure the perceived speech clarity of the output signal when compared to the input signal. The model is shown in Figure 3.8.

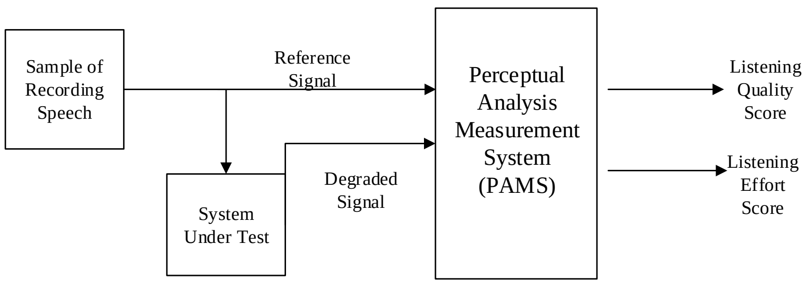

Figure 3.8 Perceptual Analysis Measurement System Model (Rix et al., 2006)

Perceptual Analysis Measurement System Model (Picovici & Nelson, 2009):

- Pre-processing: The reference and the degraded signal are aligned at the same time for compensate for the delay.
- Auditory transform: It uses a perceptual filter bank to model time/frequency analysis, frequency detection and simultaneous masking performed by the human auditory system.
- Error Parametrization: The difference between the reference and the degraded signal is evaluated to detect errors. The Listening Quality Score and Listening Effort Score corresponding to the ACR scale are produced. The listening quality is the most commonly used measure in network assessment and is recommended for general usage. The listening effort is more suitable for use with high level of degradation.

3.2.1.3 Measuring Normalizing Blocks (MNB)

Measuring Normalizing Blocks that is briefly MNB, was developed in 1997 by the US department of Commende. A simple perceptual transformation was developed for following the PSQM idea proposed by John G.Breerends. MNB is a perceptual analysis that calculates multiple time and frequency measures based on the assumption of differences of opinion obtained by listeners about short and long term spectrum deviations. MNB is an alternative intrusive quality measurement for PSQM quality measurement and it is published by ITU-T as ITU-T P.861 Rec.3 in 1998.

Two types of calculations to produce quality estimates are used in MNB algorithm: time measuring normalizing blocks (TMNB) and frequency measuring normalizing blocks (FMNB). Each of TMNB blocks integrates on frequency scales and measures differences over time intervals. Each of FMNB blocks integrates on time intervals and measures differences over frequency scales. The MNB model is shown in Figure 3.9 (Voran, 1997).

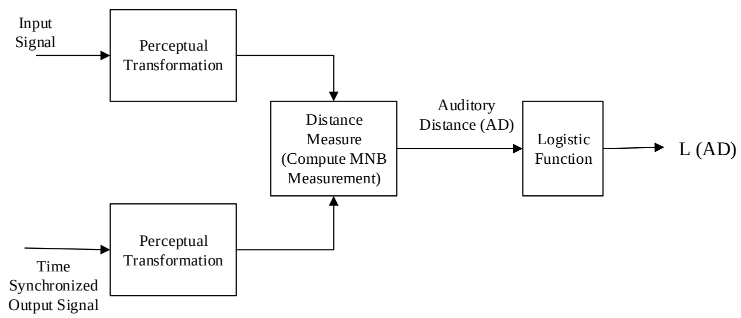

Figure 3.9 Measuring Normalizing Blocks Model (Liu et al., 2006)

Measuring Normalizing Blocks Model (Liu et al., 2006):

- Time sequential reference and output signals are added to the model. The signals are aligned in the frequency domain of the model. Frames that is below a certain threshold or zero-power frequency component are removed from the resulting frame sequence.
- The signal is sent as an input to the FMNB block and then sent to the TMNB block for the task of presenting a parameter set to be used on the signal side to receive different measurements.
- The algorithm combines the two MNB signals that are FMBN and TMBN on the output with a single value. The name of this value is Auditory Distance that is briefly AD. AD represents quality measurement on the basis of two signals comparisons. Then this value is mapped on the quality scale to obtain a subjective quality estimate.

3.2.1.4 Perceptual Evaluation of Speech Quality (PESQ)

Perceptual Evaluation of Speech Quality that is an intrusive method, was developed in 2001 for the evaluation of objective speech quality in telecommunication systems. It has been standardized by ITU-T and it has been named P.862: "An objective method for end-to-end speech quality assessment (PESQ): Narrowband telephone networks and speech codecs". It is based on comparing the transmitted audio signal with the reference signal like other intrusive methods (ITU-T Recommendation P.862, 2001). The purpose of the PESQ method is imitating voice perceptions of people. It is designed to estimate the outcome of ITU-T P.800 the Absolute Category Rating test and the result is given as mean opinion score (MOS) on a scale between 1 (bad) and 5 (excellent). The model of PESQ is shown in Figure 3.10 (Gambhir, 2009).

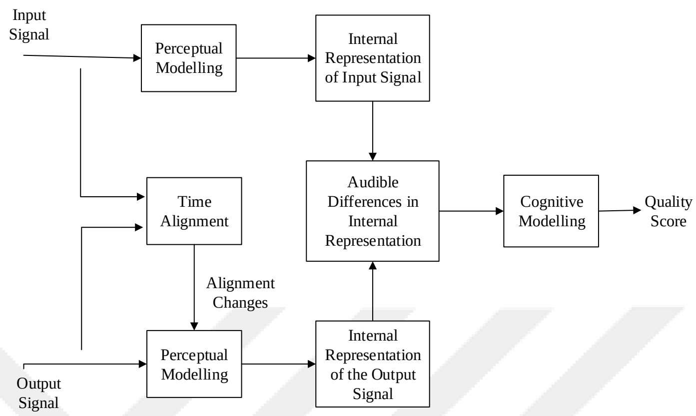

Figure 3.10 Perceptual Evaluation of Speech Quality Model (Breerends, 2001)

Perceptual Evaluation of Speech Quality Model:

- Signalling pre-processing: Includes the frequency and temporal alignment of the input signals. The temporal alignment module includes a multi-stage normalization of the delay of the test and reference signals.
- Perceptual modelling: Includes filtering the signal by matching the time and frequency domain in the bandwidth of the telephone network in order not to affect the PESQ measurement.
- Cognitive modelling: The difference between the reference signal and the degraded signal is calculated for each time-frequency frame. The difference in the positive direction indicates the presence of noise, while the negative difference indicates the presence of minimum noise. The values obtained in this noise calculation are combined to provide an estimate of the MOS score.

ITU-T P.862 standardized input signals were defined in the narrowband range of 40-3100 Hz in 2001. This recommendation was further expanded to support broadband telephone systems in the frequency range of 50 to 7000 Hz in 2005 and was standardized as P.862.2 (ITU-T Recommendation P.862.2, 2007).

3.2.1.5 Perceptual Objective Listening Quality Assessment (POLQA)

At the beginning of 2011 with the development of improved audio coding techniques, the third generation of the objective measurement algorithm is that "Perceptual Objective Listening Quality Assessment (POLQA)", was standardized by ITU-T as P.863 (ITU-T Recommendation P.863, 2011). This measurement is the latest ITU-T recommendation in series P.860 of objective perceptual quality measures for speech. The POLQA algorithm compares the reference signal and the degraded signal that is pass through the communication system that is coding, decoding and RF components to measure quality. It was designed similar with PESQ. The algorithm output is an estimate of the perceived quality of the signal that is degraded by the person in a subjective listening test. It is designed to estimate the outcome of ITU-T P.800: Absolute Category Rating (ACR) test like PESQ and the result is given as the mean opinion score (MOS) on a scale between 1 (bad) and 5 (excellent). The PESQ algorithm that is standardized in 2001 with ITU-T P.862, focuses on the determination of narrowband sound quality. The narrowband is range between 300 Hz and 3400 Hz. In 2005, broadband support was added to the standard with P.862.2. The broadband is range between 50 Hz and 7000 Hz. POLQA was developed to estimate the result of subjective tests in the super wideband mode. The super wideband is range between 50 Hz and 14000 Hz. The ITU-T Recommendation P.863 is one of the most important innovations compared to Recommendation P.862, a process called "idealization" that compensates for the quality difference in recording the original speech signals (Cotanis, 2011). This process suppress low sound levels in the recording naturally and sets the tone of the original recording to a globally acceptable timbre. The POLQA model is shown in Figure 3.11 (Gerlach, 2012).

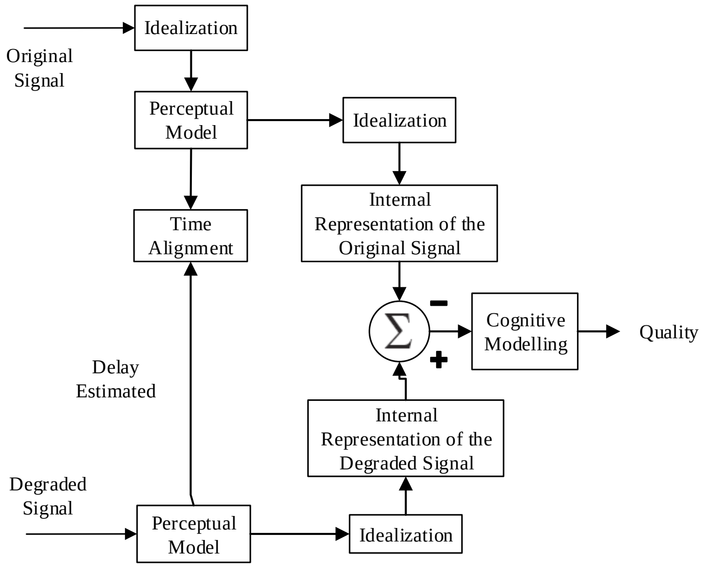

Figure 3.11 Simplified block diagram of the POLQA algorithm (ITU-T Recommendation P.863, 2011)

In the POLQA, firstly the reference signal is time synchronized with the degraded signal. For this reason, both signals are divided into time periods called frames. In the POLQA, firstly the reference signal is time synchronized with the degraded signal. For this reason, both signals are divided into time periods called frames. The delay between each part of the reference signal and the degraded signal is calculated. Then both signals are converted into an internal representation of the human hearing system similar to the PESQ model. In addition to the PESQ algorithm, a small amount of noise in the reference signal is detectable and this process can measure a signal of low quality (Ghimire, 2012). This represents the idealization process that people usually experience during quality decisions. Finally, the estimated quality value that is the MOS score, is calculated with using cognitive model. The obtained MOS value that is mean opinion score, is the most common metric used in evaluating the quality of voice call. At the same time, the MOS value is an estimate of the subjective evaluation of the users. MOS is an internationally accepted metric. MOS is expressed as a number value from 1 to 5, with 1 representing the worst and 5 representing the best. The evaluation scale is shown in Table 3.11.

Table 3.11 The ITU-T MOS listening quality scale (ITU-T Recommendation P.800.1, 2003)

User Opinion	MOS (ITU Scaled)
Very Satisfied	4.3 - 5.0
Satisfied	4.0 - 4.3
Some Users Satisfied	3.5 - 4.0
Many Users Dissatisfied	3.1 - 3.5
Nearly All Users Dissatisfied	2.6 - 3.1
Not Recommended	1.0 - 2.6

The last of the subjective sound quality evaluation methods that ITU-T P.863: POLQA algorithm standardized in 2010, is mentioned in detail in the next chapter that is Chapter 3. The ITU-T Recommendation Evolution for sound quality testing is also shown in Figure 3.12 (Slavata, 2017). The application of POLQA algorithm is mentioned in detail in Chapter 4.

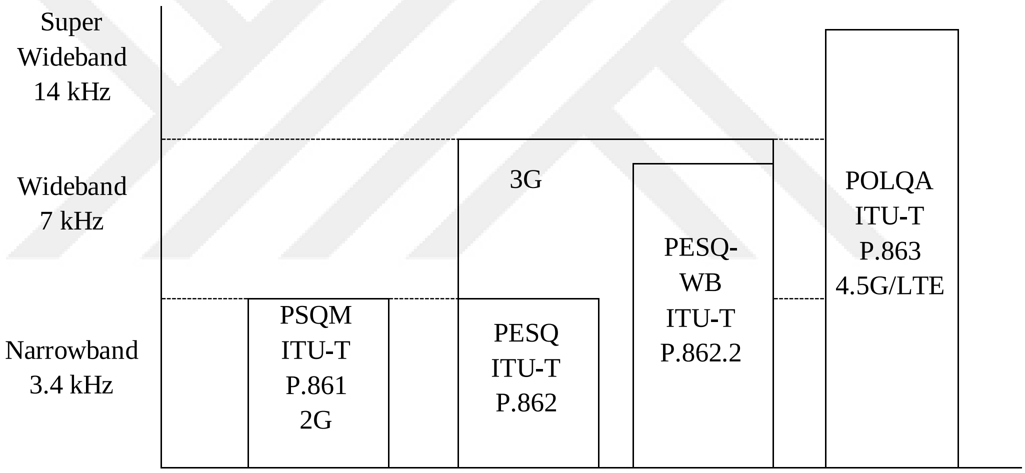

Figure 3.12 Evaluation of ITU-T Recommendation for speech quality measure

3.2.2 Non-Intrusive Speech Quality Measurement

Non-intrusive models evaluate the speech quality based on the signal processed by the system. This model is also known as single-ended or output-based measurement methods because it uses only a degraded speech signal. The non-intrusive model is shown in Figure 3.13 (Benesty et al., 2008). This method has been developed to perform real-time traffic measurements in telecommunication systems. Unlike Intrusive methods; this measurement result is evaluated without knowing the reference signal. Only the degraded speech signal is used to predict speech quality.

Estimation of non-intrusive speech quality mostly models of human speech perception. ITU-T P.563 is standardized as a recommendation for non-intrusive speech quality evaluation by ITU-T in 2004.

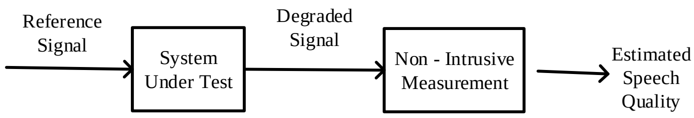

Figure 3.13 Non-intrusive speech quality measurement

Intrusive methods are stronger than non-intrusive methods for quality assessment because they use both reference and processed signals. Non-intrusive methods can be used in many places such as voice and satellite communication over IP network where reference signal can not be accessed and intrusive methods can not be applied.As a result, they are based on the processing of the speech signal using the human perception model and these algorithms are recommended when there are a large number of samples. Single Sided Speech Quality Measure (3SQM) standardized by ITU-T and E-model are typical non-intrusive models (Côté, 2011).

3.2.2.1 E-Model

The European Telecommunications Standards Institute(ETSI) Computation Model that is briefly E-model, was developed in 1998 by ETSI working group that is selected by ITU-T. The E-model is the most popular objective measurement method and a non-intrusive model because it does not need the original speech signal. In particular, the E-model is detailed in the Recommendation G.107.

The E-model is based on a concept founded by J. Allnatt more than 20 years ago, the fundamental principle is that "Psychological factors on the psychological scale are additive" It is based on the assumption that is degradations because of transmission parameters which have an additional effect on speech quality (ITU-T Recommendation G.107, 2000). Although it was originally a measurement tool that is developed for network planning, it is now used to predict sound quality nonintrusively. The E-model takes into account of various factors that affect speech quality and calculates a Rating Factor that is briefly R-Factor, ranges from 0-100. The R-Factor is a rough estimate of the quality expected when the network is implemented as planned (Vaighan, 2011). Depending on this model, the sound quality classes are given in Table 3.12.

Table 3.12 Speech quality classes according to E-model (ITU-T Recommendation G.107, 2000)

R-Value Range	**100-90**	**90-80**	**80-70**	**70-60**	**60-50**
User's Satisfaction	Very Satisfied	Satisfied	Some Users Dissatisfied	Many Users Dissatisfied	Nearly All Users Dissatisfied

R factor is calculated by the (3.1) equation (ITU-T Recommendation G.107, 2000):

$$R= R_o-I_s-I_d-I_e+A \tag{3.1}$$

R_o: represents that the signal / noise ratio that is including noise generated by the circuit and background noise.

I_s: represents that simultaneous degradations with speech

I_d: represents that degradations caused by delays

I_e: represents that the effects of equipment failure (decoder etc.)

A: represents that advantage factor (0 for landline, 10 for GSM etc.)

After the R factor has been calculated, the final step of the E-model is to map the R factor to an equivalent MOS value via the equation in ITU-T Rec.G.107. The (3.2) equation can be used for this mapping.

$$MOS = \begin{cases} 1 & for\ R<=0 \\ 1+0.035R+R\ (R-60)\ (100-R)\ 7*10\ (-6) & for\ 0<R<100 \\ 4.5 & for\ R=>100 \end{cases} \tag{3.2}$$

3.2.2.2 Single Sided Speech Quality Measure(3SQM)

Single sided speech quality that is briefly 3SQM, is a non-intrusive method developed for sound quality. For this reason, the measurement model analyzes the output of the degraded speech signal from the device or network being tested. It was standardized by the ITU under the title of "ITU-T P.563: Single ended method for objective speech quality assessment in narrowband telephony applications" in 2004 (ITU-T Recommendation P.563, 2004). The model of the algorithm is shown in Figure 3.14.

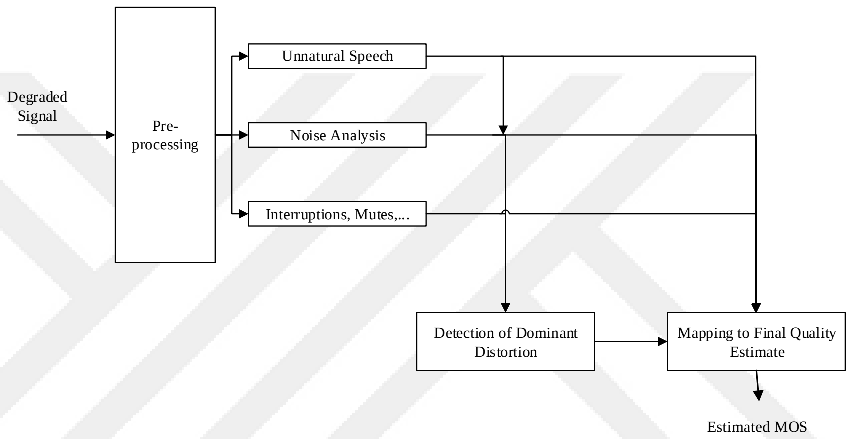

Figure 3.14 Single Sided Speech Quality Measure model (Opticom, 2004)

In order to be able to evaluate the audio signal quality correctly, it is necessary to pre-process it in the first step.

Pre-processing steps:

- IRS receive filtering: It simulates a standard handset used in laboratories for subjective listening tests.
- Speech level setting
- Using Voice Activity Detection that is briefly VAD to separate voice and silent sections.

After pre-processing, the transmitted audio signal is analyzed according to 3 different techniques. The first method involves the sensitive LPC model of the human vocal tract that tries to pronounce the degraded sample signal. For example, damage that can occur during signal transfer is determined. The sum of the obtained degradations is used to compute the MOS value. In the second approach, it tries to reconstruct the reference signal from the degraded signal using correction functions. The missing parts are recalculated, the sample is filtered and further regulated. This purified sample is compared with the degraded signal to obtain an MOS value. In the third method of evaluation that is the last method, typical parameters are calculated such as signal-to-noise ratio (SNR), time clipping, sound transmission characteristic type disorder in the sample for processing computer signal. Then the value range of the parameters is used to calculate the MOS value. Finally, the farthest of these 3 estimates of the MOS value is ignored, and the arithmetic mean of the remaining 2 MOS values is the MOS value which the sound quality is to be evaluated according to the 3SQM algorithm (Goudarzi et al., 2009). The MOS value allows the audio quality to be defined between 1 (bad) and 5 (excellent).

CHAPTER FOUR

PERCEPTUAL LISTENING QUALITY ASSESSMENT (POLQA)

In recent years, the objective speech quality measurement methods have been developed using a perceptual measurement approach. These perceptual based algorithms mimic the behavior of a person evaluating the quality of an audio in a listening test. The Perceptual Objective Listening Quality Assessment (POLQA) that is standardized by ITU-T Recommendation P.863 in 2011 for objective speech quality assessment, is used to estimate the overall speech quality perceived by the user. This technique is used for both of narrowband and super wideband. The narrowband covers the frequency band of 300-3400 Hz. The super wideband covers the frequency band of 50-14000 Hz. The audio quality is evaluated by the ITU-T P.863 algorithm using the Absolute Category Rating (ACR) with a 5-point opinion rating as defined in the ITU-T Recommendation P.800 standard. This scale is explained in Section 3.1.1.1. Also the scale summary table is given in Table 3.1. The POLQA algorithm is a full reference model that works by comparing the reference signal with the degraded signal. The full reference model for telecommunications networks is shown in Figure 4.1.

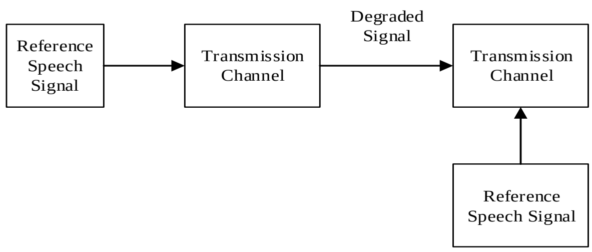

Figure 4.1 The full reference model

The basic view of the POLQA algorithm is shown in Figure 4.2. There are three main function blocks in this algorithm. These blocks are listed below and the basic layout of these major components is shown in Figure 4.3 (Ebem, 2009).

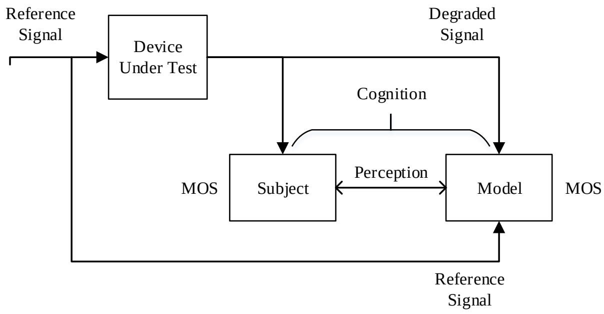

Figure 4.2 Basic overview of POLQA algorithm (ITU-T Recommendation P.863, 2011)

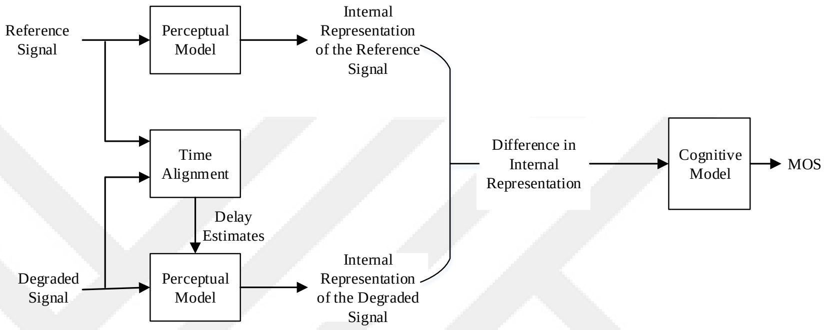

Figure 4.3 Basic scheme of the main components of POLQA (ITU-T Recommendation P.863, 2011)

The overview of the POLQA algorithm, the reference signal to should be used during testing and the main components of the algorithm that are the temporal alignment, the perceptual model and the cognitive model, are defined in Chapter 4.

4.1 General Overview of POLQA Algorithm

The algorithm needs two signals as referance and degraded signal, because, it is an intrusive measurement method. These signals are waveforms that are containing 16 PCM samples. The general overview of the POLQA algorithm is shown in Figure 4.4. In the first stage of the algorithm, both of the reference and the degraded signal are divided into small time frames and the delay of the reference signal with respect to the degraded signal is calculated using these time frames, and the sampling rate of these two signals is estimated with respect to each other. When the correct delay is determined and the sampling rate differences are compensated, the signals and delay information are transferred to the cognitive model. In the Cognitive model, the value of the quality measure that is the MOS value, is determined.

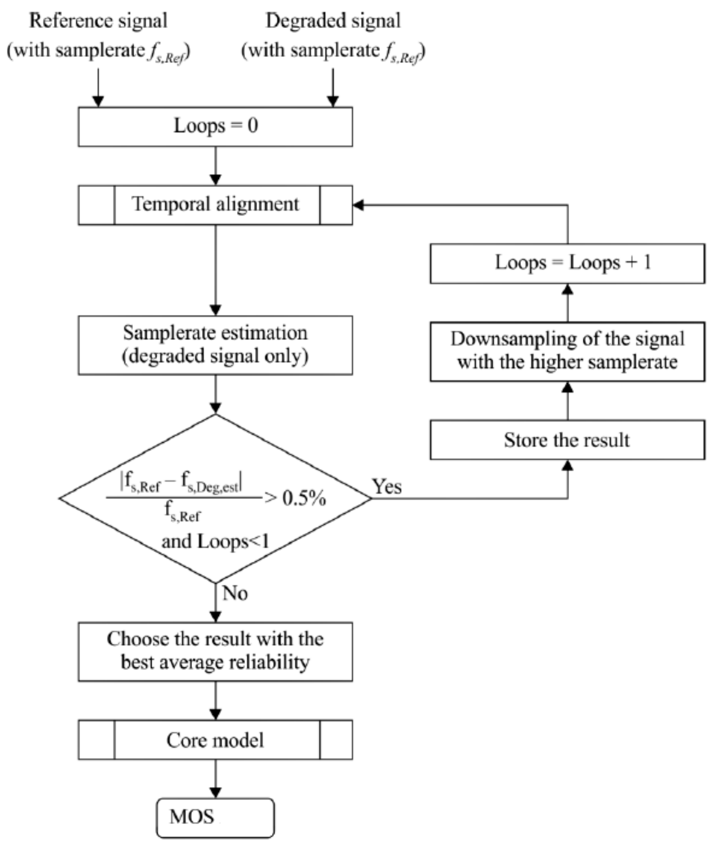

Figure 4.4 General overview of the POLQA algorithm (ITU-T Recommendation P.863, 2011)

4.2 Reference Signals for Audio Quality Measurement

The reference signal to be used in the measurements should have the following features (ITU-T Recommendation P.863, 2011):

- At least 3 active conversations
- At least 1 second silence between active talk periods
- There should be no more than 6 seconds of active talk
- The total length of the test sample, including silences, should not exceed 12 seconds
- Active speech level must be -26 dBoV

The "dB overload" that is briefly called dBoV, is the level value in the speech reference signal and this level refers to a full-scale square wave signal. A full-scale sine wave signal has a RMS level of -3 dBoV.

- Encoded with 16-bit linear PCM
- Noise level should be less than -80 dBoV
- Conversation silences should not be zero digitally but should contain a noise signal that is less than -80 dBoV level.

As is known, the POLQA algorithm supports two different bandwidths that are Narrowband and Super-wideband. For narrowband communication mode, reference signals should be limited between 50 Hz and 3800 Hz bandwidth and sampled at 8 kHz. For the super wideband communication mode, the reference signals should be limited between 50 Hz and 14000 Hz bandwidth and sampled at 48 kHz.

4.3 Temporal Alignment

Temporal alignment is to limit the search intervals as much as possible and gradually sensitize the delay estimate. This algorithm aligns the reference and degraded signals, allowing the signals in the same speech sections of the core model process to be compared. Both the reference and the degraded signal are converted to the time-frequency domain using FFT frames depending on the signal length sampling rate.The examples of the signal length sampling rate are 1024 for 48 kHz, 512 for 16 kHz and 256 for 8 kHz. A delay value is calculated for each frame obtained.

Basic concepts of temporal alignment (Breerends et al., 2013):

- The distribution of signals to the corresponding parts in equal lengths and the delay for each frame must be calculated separately.
- If possible for each part of the degraded signal, the corresponding part of the original signal is searched.

- Since each search of a very long segment of the signal requires high computational power and time, the delay of each frame is updated.

The temporal alignment algorithm is shown in blocks in Figure 4.5. These blocks are respectively Bandpass filter, Pre-alignment, Coarse alignment, Fine alignment and Section combination.

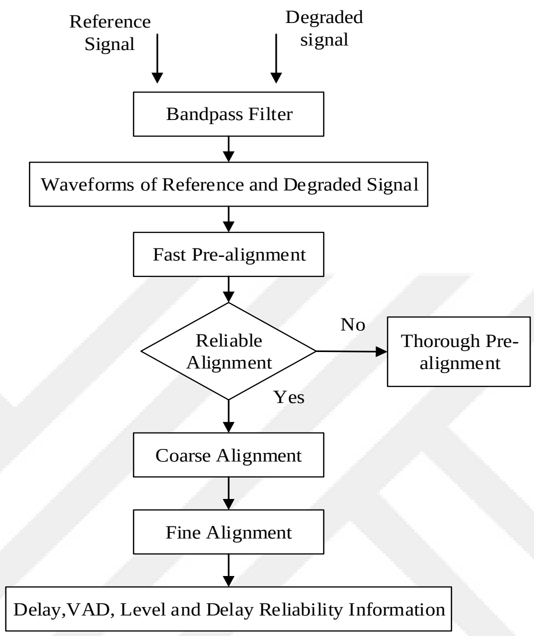

Figure 4.5 The major blocks of Temporal Alignment algorithm

Firstly, the reference and degraded signal are filtered by a bandpass filter. This filtering depends on whether the signal is in narrowband or super-wideband mode. In the pre-alignment section, the active speech parts of the signals are found, the initial delay estimate at the beginning of each signal frame is calculated, and a search range is determined to determine the delay of each frame. In the coarse alignment section, the delay values for each frame are gradually sensitized using the back tracking algorithm. In the fine alignment section, the exact delay value of each frame is determined. Finally, all frames with approximately the same delay are combined and called a “section combination” (Ghimire, 2012).

The features of the temporal alignment process are listed below.

- There is no static delay limit.

- The algorithm is designed to process the delay length changes up to 300 ms.
- The value of delay may change for each frame.
- Small differences of up to 2% at sampling rate can be compensated for in the time synchronization block.Larger differences are compensated for excepting of temporal alignment.
- Temporal alignment also works for very noisy signals.
- No problems are observed with variable level signals.

4.3.1 Fast Pre-Alignment

In cases where the delay in a degraded signal is fixed or constant per frame, the ITU-T Recommendation P.863 algorithm includes a pre-alignment method that allows for fast alignment of signals. The fast pre-alignment module works at the beginning of the temporary alignment process as shown in Figure 4.5. Sequential cross-correlation is used to separate the fragmented speech signal into fixed delayed segments, which allow to limit the delayed search range. If less than 75% of the active talk is matched to prevent misaligned signals during fast pre-alignment step, or the constant delays per frame change continuously throughout the signal, the algorithm is guided to thorough pre-alignment.

Fast pre-alignment consists of the following steps:

1. The average energy distribution of the speech is highly skewed towards low frequencies and a bandpass filter that has frequency range between 700 Hz and 3000 Hz, is used to avoid excessive deviation of the calculated cross correlations relative to these frequencies. After filtering, the signals are scaled to -26 dBoV active speech level (ASL).
2. The average speech level "speechLev" and the average noise level "noiseLev" are calculated.
3. The active talk threshold "thr" is calculated using the following formula:

   Thr = min((-26+3*max(noiseLev_ref, noiseLev_deg))/4, -29)

   This threshold value will be used to exclude signal frames that have insufficient signal/noise ratio (SNR).

4. The cross correlation of the signal envelopes and the delay estimate are calculated. Each envelope frame indicates the signal level in a 180 ms section of the signal. The active talk threshold value is briefly thr that we calculated in the previous step, is subtracted from the envelope value. In the parts where active speech is present, this value is positive. The envelope values for inactive speech that give negative data are equal to zero. Then, the general delay value is obtained from the maximum cross correlation value of the processed envelopes.
5. Sliding window power normalization is used to reduce the effect of level changes in the cross correlation of the reference and degraded signal. Normalization calculates the RMS values of a sliding window signal that is length of 26.625 ms and the obtained RMS value is normalized with the sample at the center of the window position. The signal before the sliding window power normalization and the signal after the normalization that are applied, are shown in Figure 4.6. After normalization, active speech has an almost flat signal envelope.

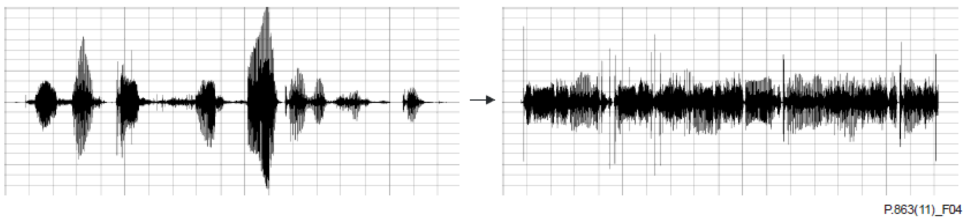

Figure 4.6 Sliding window power normalization (ITU-T Recommendation P.863, 2011)

The parts of the reference signal are matched to the related parts in the degraded signal. The matching is performed in a loop and the aim of this loop is that creating a reference list for the degraded signal fragments. The operating of the loop is shown in Figure 4.7. The location, length, delay, assigned, unsure and missing informations of the signal are kept for each signal on the signal fragments.

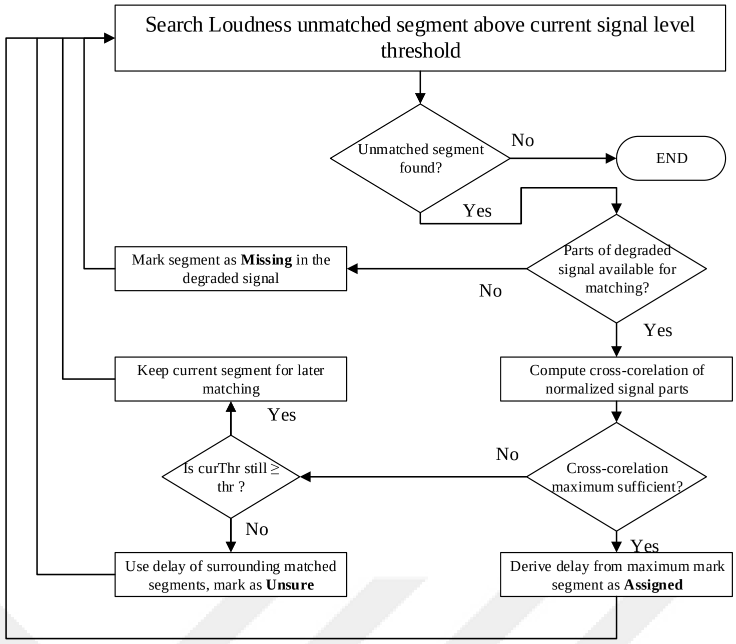

Figure 4.7 Flowchart of the segment-wise matching (ITU-T Recommendation P.863, 2011)

The parts of the reference signal are matched to the related parts in the degraded signal. The matching is performed in a loop and the aim of this loop is that creating a reference list for the degraded signal fragments. The operating of the loop is shown in Figure 4.7. The location, length, delay, assigned, unsure and missing informations of the signal are kept for each signal on the signal fragments.

The value of curThr in the loop must be greater than thr value. The (4.1) formula is used to calculate "curThr".

$$curThr= max(thr+1,\ 0.4xspeechLev\_ref+0.6xnoiseLev\_ref) \tag{4.1}$$

If the segment of degraded signal in the matching phase is smaller than the current segment length in the reference signal, the matching is performed for the current length only, and those with unmatched excess segment length are signals marked "missing". If the selected maximum absolute value in the cross correlation vector is sufficiently high and the segment length is at least 64ms, the current segment is marked as "assigned" and the delay is listed. If the sections marked as "Unsure" in the signal are equal to or less than the curThr value thr value, the predetermined estimated delay is stored and the current section is marked as "unsure".

4.3.2 Thorough Pre-alignment

In degraded signal, thorough pre-alignment defines the positions that from paused to active talk at the conversations. These locations are called reparse points. For each reparse point, all active speech part information starting from a reparse point that is called reparse section, is computed. The overview of thorough pre-alignment is shown in Figure 4.8 (Breerends et al., 2013).

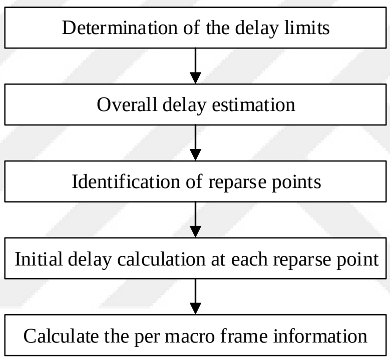

Figure 4.8 The overview of thorough pre-alignment

4.3.3 Coarse Alignment

In the coarse alignment section, the delay values are sensitized by incrementally increasing the frame delay value determined in the pre-alignment. This operation is done by dividing each signal into smaller parts. The subdivisions of the signal that are obtained by dividing the signal into smaller segments, are called "feature frames". These frames are equal distance. The resulting vectors are called feature vectors. All feature vectors are generated for both active sections of the reference and the degraded signal. These features use energy and fractal dimension per feature frame. In addition to this, a correlation matrix is created for each feature. The generated matrix is organized as correlation vectors per macro frame. Correlation vectors are the pre-alignment result includes all possible time delays within the estimated look-up range obtained. The occured matrix structure is Ncv x Nmf. The Ncv definition is the number of possible delays that can be tested in each correlation vector. The Nmf definition is the number of frames. Then the maximum correlation for the macro frame is determined. The position of the maximum correlation is equal to the optimum correction of the delay per frame to obtain a better match between the reference signal and the degraded signal.

Speech signals are usually periodic and when a packet loss occurs, an incorrect delay can give a better correlation than correct delay. For this reason, a backtrace algorithm similar to the Viterbi algorithm is used to find the correct correlation value. The key of the backtrace algorithm is recursion. The aim here is to find the way to give the best general probability to the mathematical, without having to calculate all possible combinations (ITU-T Recommendation P.863, 2011).

4.3.4 Fine Alignment

In the fine alignment section, the exact delay value of each macro frame is determined. Fine alignment works very much like coarse alignment. The difference between the two alignments is that the recursion are not used in fine alignment. Because the search range is largely limited by the previous alignment steps.

4.3.5 Section Combination

The result of the temporal alignment step is that the start and end points of each segment of the degraded signal are matched with the corresponding reference signal segment. The matching of these parts is called "Section Combination". Parts of the degraded signal that do not correspond to the reference signal, that is, parts that are added to the degraded signal, and parts of the reference signal that are missing in the degraded signal are also marked. The signal part in Figure 4.9 is an example that is temporally aligned. In Figure 4.9, the upper graphical reference signal segment and the lower graph shows the degraded signal segment.

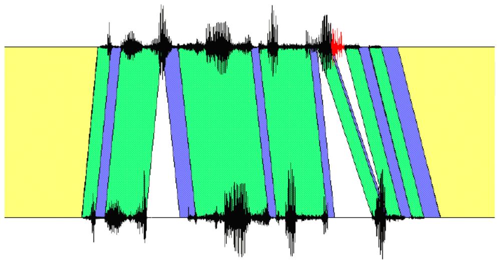

Figure 4.9 Example of an aligned pair of reference and degraded signal

4.4 Perceptual Model

The basic approach in the POLQA is to map the reference and degraded signal on an internal representation using the human perception model. A cognitive model is used to estimate the perceived speech quality of the degraded signal. The perceptual model used in the POLQA consists of three main parts.

In the first part, the internal representation of the reference and degraded signal is calculated as can be shown in Figure 13. Then both signals are scaled. As the degraded signal is scaled to the playback level, the reference signal is scaled to a predefined constant optimal level of 73 dB SPL. After the scaling process is over, the signals are converted to time-frequency domain with 50% overlapping FFT frames using Hann window. If small pitch shifts are found that cause the impaired frequency axis at the degraded signal, they should be compensated. This compensation process is shown in Figure 4.10. After the frequency compensation, the three quality indices in the POLQA algorithm are calculated (ITU-T Recommendation P.863, 2011). These quality indicators are frequency response distortions (FREQ), additive noise (NOISE), and room reverberations (REVERB). In addition to this, the internal representation of the reference and degraded signal is calculated in terms of pitch-loudness-time in the first part of the perceptual model. In the next process, the masking is done in both of time and frequency domain to determine the perceived sound intensity for each frequency component. This operation is done for each frame in the reference and degraded speech signal. After the masking process, pitch density representations are converted to pitch loudness density representations. The low noise level in the reference signal is suppressed and the constant background noise in the degraded signal is partially removed for optimization. As a result of these operations; four different variants of the internal representations of the reference and the degraded signal are calculated. Two of these variables focus on normal degradation and high degradation of the system under test. The other two variables focus on normal and loud audio distortions for added disturbances (Breerends, 2013).

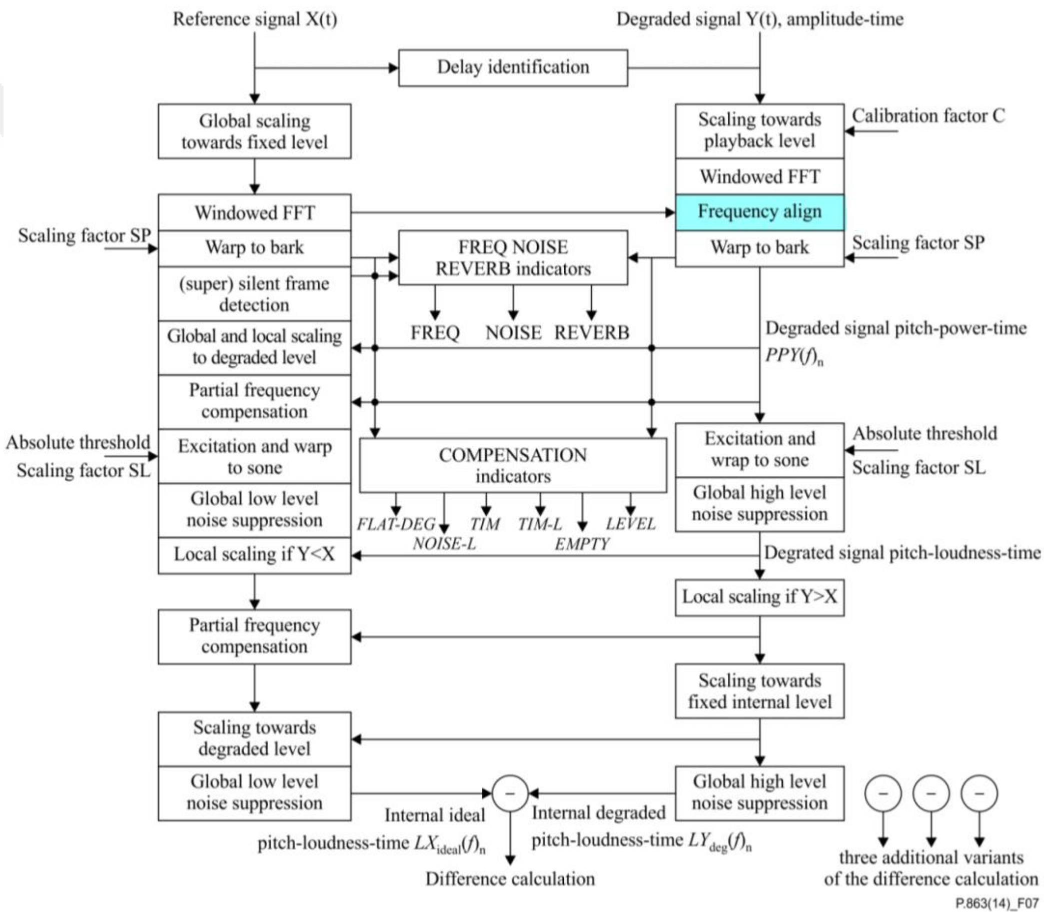

Figure 4.10 Overview of the first part of the ITU-T P.863 (ITU-T Recommendation P.863, 2011)

The calculation of the distortion intensity and the added distortion intensity by using the variables calculated in the first part of the algorithm as the input that constitutes the second main part. An overview of the second main section of the perceived model is shown in Figure 4.11. The intensity of impairment is a display for the perceivability of impairments, because cognitive effects are not yet considered. The cognitive model takes into account the cognitive aspects that are important for people to assess the quality of what they perceive. The disturbance density that is the measure of perceivability, is transformed into an annoyance measure. The conversion is carried out by compensating for the distortion intensity in the following condition.

- Significant level variations
- Frame repretitions
- Timbre
- Spectral flatness
- Noise switching during speech pauses
- Delay variations
- Strong variations of the disturbance density over time
- Loudness jumps

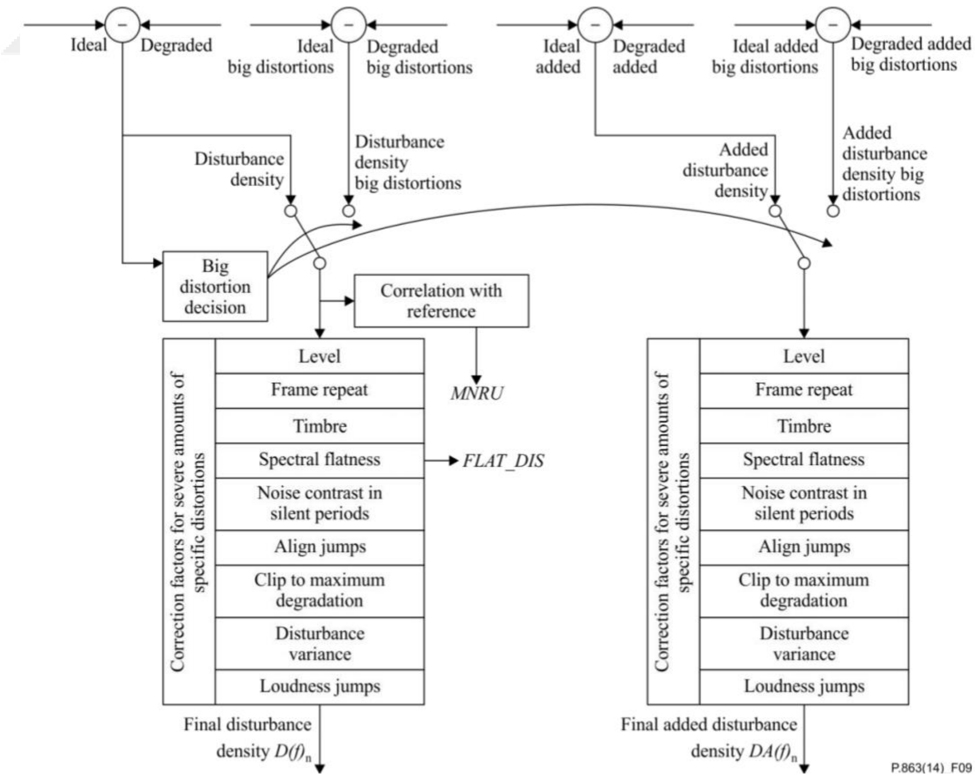

Figure 4.11 Overview of the second part of the ITU-T P.863 (ITU-T Recommendation P.863, 2011)

Two more indicators that are LEVEL and FLATNESS, are calculated during the compensation operation. The LEVEL indicator indicates significant deviations from the optimal listening level. The FLATNESS indicator quantifies the effect of the loudness frequency field peaks in the disturbance loudness function.

The disturbance density values obtained as the result of the second main part of the algorithm are collected with the pitch, spurt and time in the third part of the POLQA algorithm. An overview of the third main section of the perceived model is shown in Figure 4.12. The added disturbance density is compensated using the REVERB and NOISE indications for loud reverberations and loud additional noise.The two disturbance densities are combined with the FREQ indicator to obtain a MOS-like that is intermediate quality indicator. The indicators of LEVEL and FLATNESS and the MOS-like intermediate indicator obtained in the previous step are used in order to obtain a raw POLQA score. The obtained MOS value that is mean opinion score, is the most common metric used in evaluating the quality of voice call. At the same time, the MOS value is an estimate of the subjective evaluation of the users. MOS is an internationally accepted metric. MOS is expressed as a number value from 1 to 5, with 1 representing the worst and 5 representing the best. The evaluation scale is shown in Table 3.11.

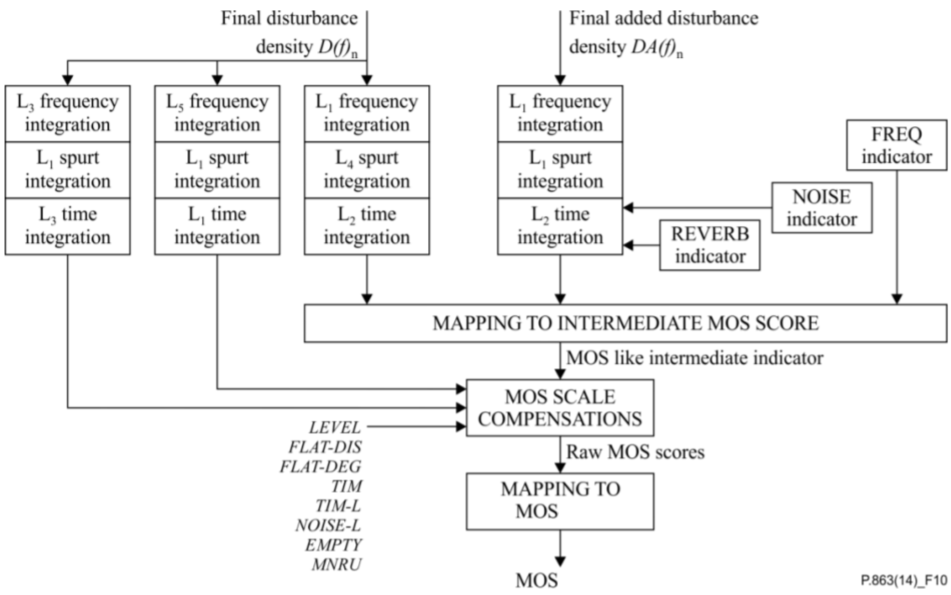

Figure 4.12 Overview of the third part of the ITU-T P.863 (ITU-T Recommendation P.863, 2011)

CHAPTER FIVE

AUDIO QUALITY MEASUREMENT AND IMPROVEMENT

This chapter consists of two main parts. The first major section is the measurement of the sound quality of mobile-based devices objectively using the POLQA algorithm described in Chapter 4. The equipment required to perform these measurements and the provision of the interface settings necessary for the usage of these equipments in measurements is described. In the second main section, the quality scores obtained by using POLQA algorithm are evaluated and the sound codecs which do not have sufficient sound quality are tried to improve sound quality with audio calibration tool.

5.1 Audio Quality Measurement with POLQA Algorithm

In order to test the sound quality of mobile based devices using POLQA algorithm, some hardware equipments are needed firstly. These hardware equipments are listed below and their connection to each other is shown in Figure 5.1. For mobile-based devices, sound quality can be evaluated in three conditions that are handset, handsfree and headset, but an anechoic chamber is required for accurate measurement of handset and handsfree sound quality. The purpose of using the chamber is to take an accurate measurement without being affected by environmental factors. The measurements will be make at normal room conditions and will only be make at the headset condition because the anechoic chamber does not needed for sound quality measurement in the case of headphones.

The hardware equipment for sound quality measurement:

- Rohde&Schwarz CMW500 Communication tester that is connected to the network via RF. This equipment provides 2G, 3G and 4.5G infrastructure and establishes a 2G, 3G and 4.5G calls.
- Rohde&Schwarz UPV Audio Analyzer that is using for for performing POLQA and PESQ measurements of downlink and uplink audio signals.
- Application PC equipped with a LAN interface with CMW500.

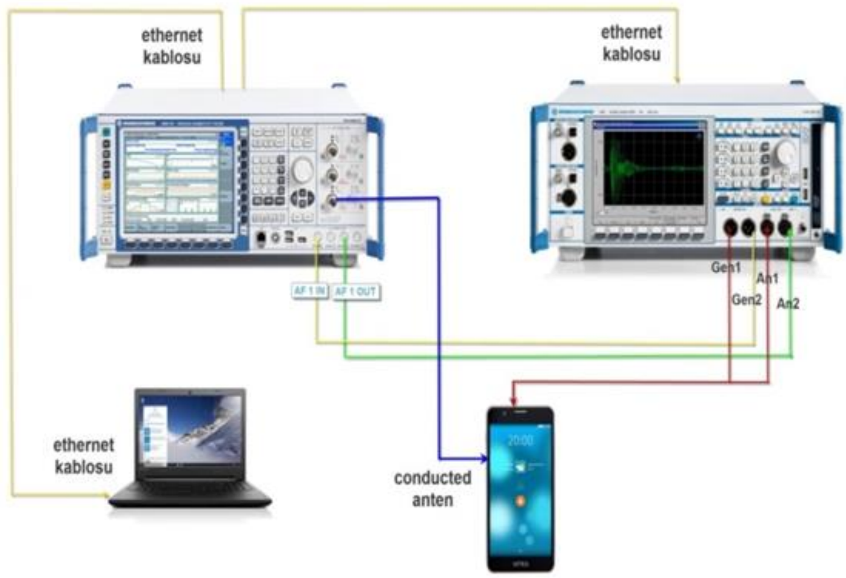

Figure 5.1 Test setup with CMW500 Communication Tester and Audio Analyzer

The sound quality with POLQA algorithm is classified according to the transmission direction of the signal. The transmission direction of the signal is divided into two parts that are uplink and downlink. “Uplink” is also called “sending”, seen from the mobile phone’s point of view and “Downlink” is also called “receiving”.

5.1.1 Audio Quality Measurement at Downlink Signal Direction

In the downlink measurements, the audio analyzer generates an audio test signal that is fed from the audio interface line input. This audio test signal is Harvard sentences that include all of the speech frequencies. This analog signal is converted to a digital signal and sent to the CMW500 communication tester via the LAN. In the CMW500 communication tester, the 2G, 3G or 4.5G test signals are then transmitted to the mobile based device to be tested via conducted RF connection. It is decoded here and converted to an analog signal on the headphone plug. This degraded audio signal is sent to the audio analyzer input and the quality of the sound is calculated using this degraded audio signal and the original signal with POLQA algorithm in the audio analyzer (Rohde & Schwarz, 2015). The downlink sound quality measurement with POLQA is shown in Figure 5.2.

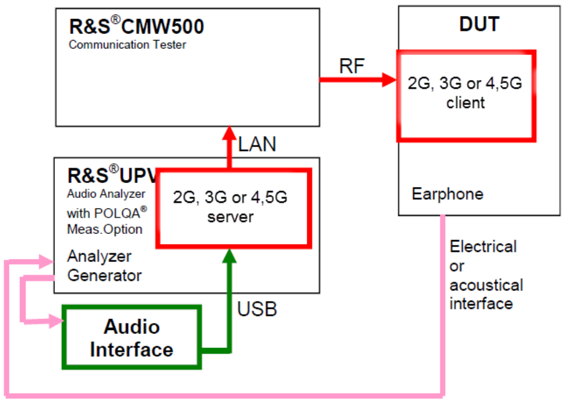

Figure 5.2 The downlink audio quality measurement with POLQA

5.1.2 Audio Quality Measurement at Uplink Signal Direction

In uplink measurements, a reference speech signal fed by the audio analyzer is sent to the microphone of the mobile-based test device. The signal of 2G, 3G or 4.5G is encoded into voice packets and modulated to RF carrier. The audio packets sent are demodulated in the CMW500 Communication tester and the decoded packets are sent to the audio analyzer via the LAN (Rohde & Schwarz, 2017). In the audio analyzer, the quality of the sound is calculated using this degraded audio signal and the original signal with POLQA algorithm (Rohde & Schwarz, 2015). The uplink sound quality measurement with POLQA is shown in Figure 5.3.

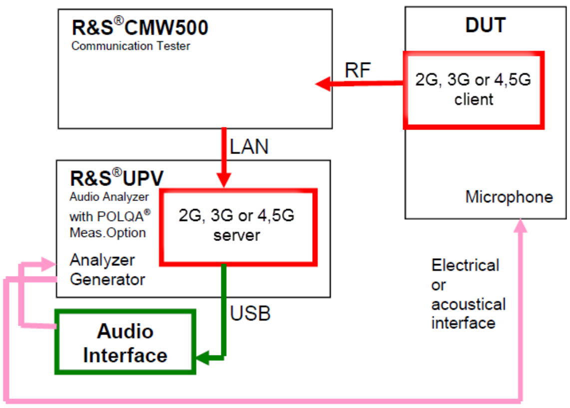

Figure 5.3 The uplink audio quality measurement with POLQA

5.1.3 *Wideband Radio Communication Tester Configuration*

After the installation of the test setup for the POLQA sound quality measurements is completed, the interface settings of the equipment in the measurement are required to be set. Since the last network technology is 4.5G, the measurement interface settings are used VoLTE that is voice over LTE. The mobile-based device used in the measurements supports five bands that are band 1, 3, 7, 8 and 20 in the 4.5G network. In order to be exemplary, it is described how to perform test interface adjustments in the 4.5G band 7 network.

Firstly, the mobile-based device to be tested must be connected to the CMW500 Communication Tester. In order to provide this connection, the back cover of the device is opened and the main antenna output on main board of the device is connected to the CMW500 Communication Tester with RF conducted cable. The connection between the device and the CMW500 Communication Tester is shown in Figure 5.4. For communication between the device and the CMW500, the special simcard for the CMW500 Communication Tester is inserted into the device and then the device is powered on. Once the connection is made and the simcard is inserted in the device, further adjustments will be made with the interface of CMW500 Communication Tester (Gerlach, 2012).

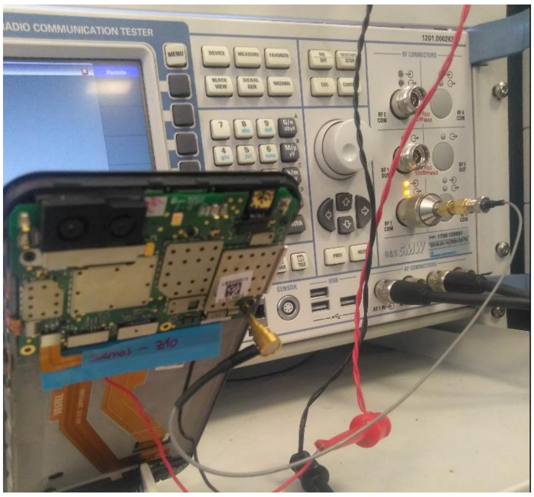

Figure 5.4 The connection between mobile device and CMW500 Communication Tester (Personal archive, 2019)

For further adjustments, the interface of CMW500 Communication Tester and its buttons are used. Firstly, the signal is generated with the help of equipment button that name is "Signal Gen". After pressing the button of "Signal Gen", the equipment interface shows the menu as shown in Figure 5.5. "LTE signaling" is activated with the navigation key.

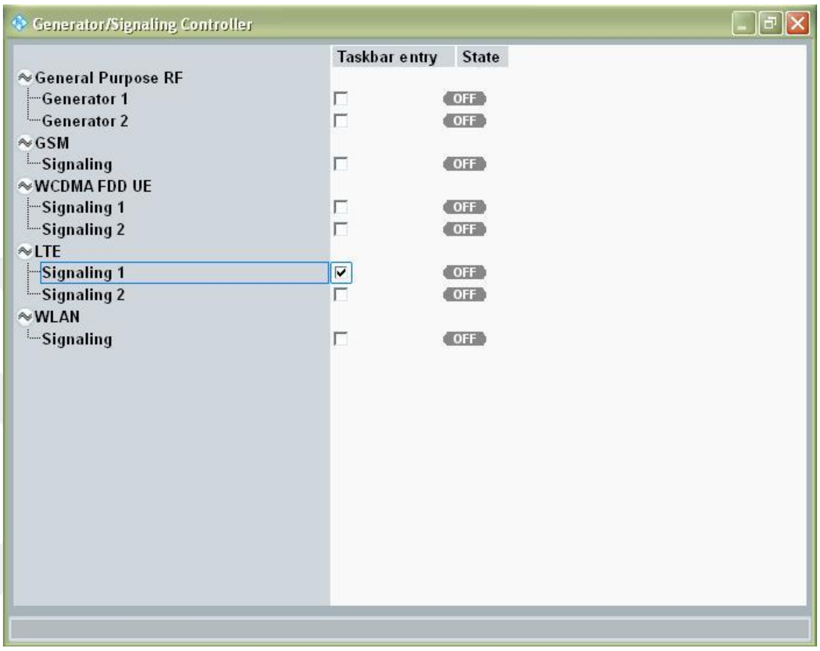

Figure 5.5 Activating the LTE signal with communication tester

After activating the LTE signal; variables such as the "Operating Band" and "Cell Bandwidth" to be used in measurement are set in the interface shown in Figure 5.6. Since the LTE band to be tested is Band 7, the Band 7 is selected as shown in Figure 7. The variable of "Channel" is set according to signal direction. The downlink variable is set 2850 Ch and the uplink variable is set 20850 Ch. The variable of "Frequency" is set 2630 MHz for downlink and 2510 MHz for uplink. "Cell bandwidth" is set to 10 Mhz. The reference signal (RS) energy per resource element (EPRE) specified as RS-EPRE in the interface is set to -79.8 dBm /15kHz (Rohde & Schwarz, 2017).

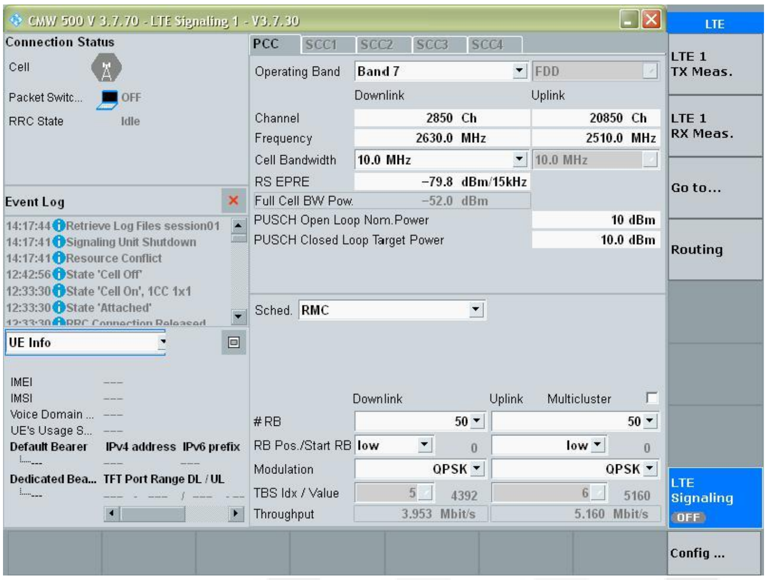

Figure 5.6 LTE signaling menu

At CMW500 Wideband Radio Communication Tester, the button of "CONFIG" is pressed to open the menu of "LTE SIGNALLING CONFIGURATION" that is shown in Figure 5.7. Since the data or voice can be transferred with the Packet Switch in LTE or 4.5G calls, the option of "Connection Type: Data Application" is selected as the shown in Figure 5.7 (Rohde & Schwarz, 2017). The IP version and APN information of the mobile based device must be set as shown in Figure 5.7, because 4.5G is telecommunication system via Internet Protocol that is briefly IP.

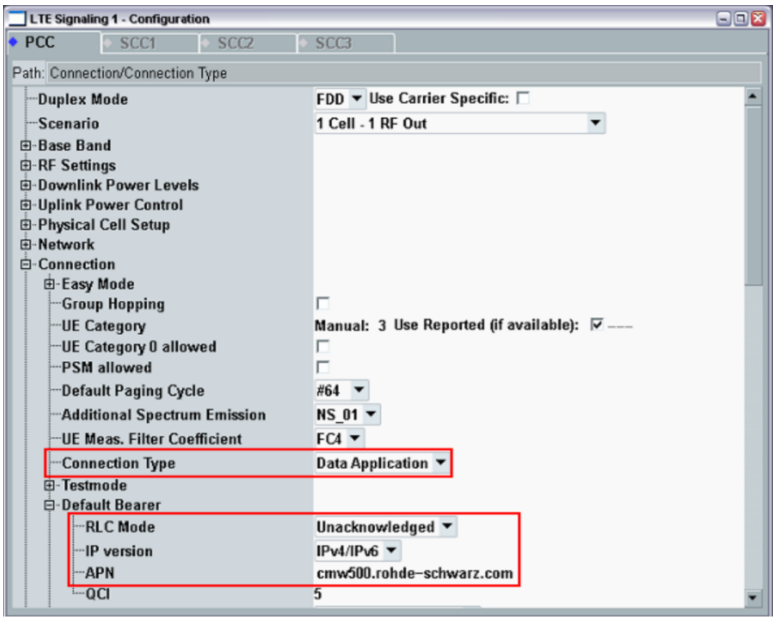

Figure 5.7 LTE signaling configuration

After the process of LTE signal configuration is completed, the button of "Measure" is pressed on the CMW500 Communication Tester in order to activate audio measurement and the checkbox of “Audio -> Measurements 1” is enable as shown in Figure 5.8 (Rohde & Schwarz, 2015).

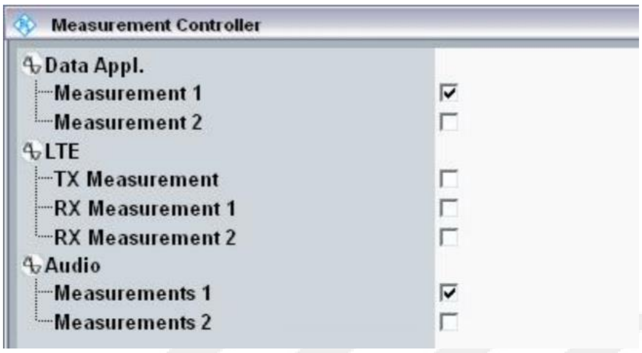

Figure 5.8 Enable audio measurement

After the audio measurement is enabled, the LTE signal is sent to the mobile based device and the network connection of the device is activated. The option of "LTE signaling" that is shown in Figure 5.6 at the right bottom, is switched on with using the button of "ON/OFF". After "LTE signaling" is enabled, the variable of "Packet Switched" is changed from "Off" to "Attached" and "RCC State" is changed from "Idle" to "Connected" at the menu of connection status as shown in Figure 5.9.

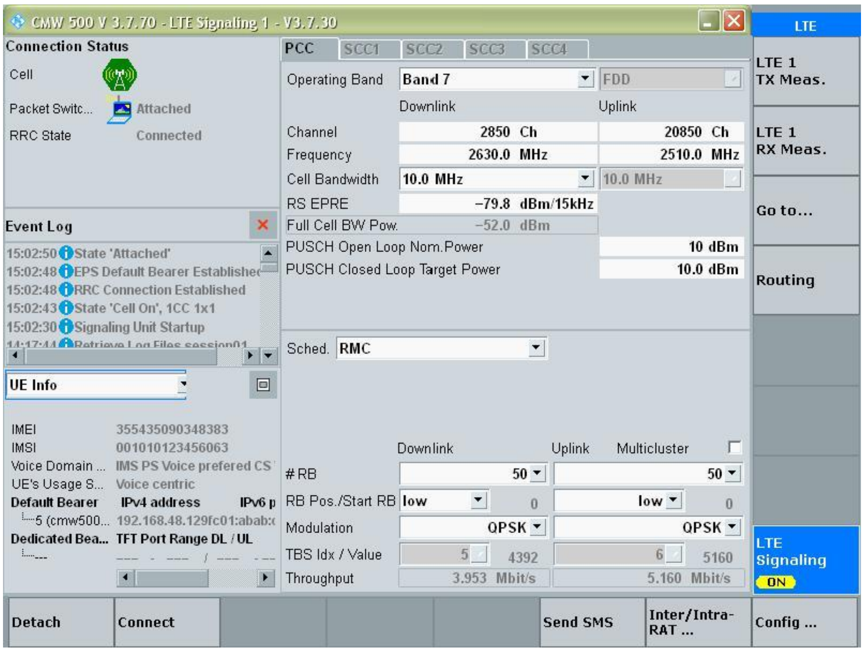

Figure 5.9 LTE signaling state

Adjustments for LTE Band 7 have made the device suitable for making VoLTE calls and the device can be used to measure uplink/downlink audio quality with POLQA depending on the ITU-T P.863 standard. The option of "Connect" is selected as shown in Figure 5.9 and a VoLTE call is sent to the mobile-based device. And then the calling is answered from the ringing device. The audio file specified as Harvard Sentences, which is used as reference sound in the standard, is sent or received via UPV Audio Analyzer depending on the measurement direction. The path that the incoming and outgoing sound to the mobile device follows in the device, is shown in Figure 5.10. As seen in Figure 5.10, the sound from the microphone or the sound from the antenna goes to the vocoder. Vocoder is a hardware circuit that converts the spoken word that is analog signal into a digital code. It can also do the opposite, so it converts the digital code into an analog signal. The method of encoding and decoding the signals in digital media is called a codec. The main purpose of the coding process is; to represent the conversation using the minimum number of bits. The efficient digital representation of the speech signal enables both the transmission of the signal and the efficiency of the bandwidth during signal transmission. In order to make a voice call, the number of bits per second, which must be transmitted or received, is expressed as the codec bit rate and is expressed in kilobits per second that is briefly kbps. The network provider determines the codec bit rate to be sent, but the mobile based device cannot determine the codec bit rate. The following parameters are taken into account when codec bit rate determining.

- Provided bandwidth
- Provided signal level
- Quality of the provided channel (noise level)

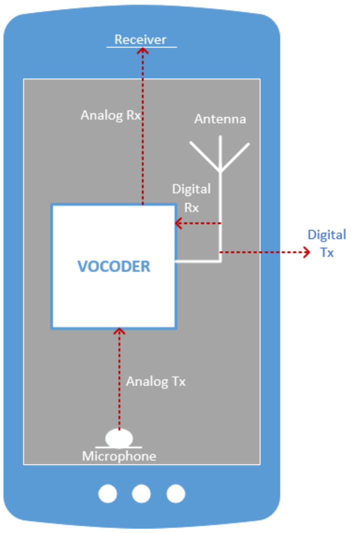

Figure 5.10 Phone audio transmission

Since the bit rate to be sent will be determined by the network provider according to the above parameters, measurements are performed in different codec bit rates supported by 2G, 3G and 4.5G networks in order to measure the best and worst conditions. The sound quality will be adversely affected as the number of bits received or sent per second decreases as the Codec bit rate decreases. In addition to this, the performance of the constant codec bit rate is measured with poor, fair and good signal strength and the effect of the signal strength on the sound quality is evaluated. The measurements are made by preferring the most widely used networks in Turkey. These networks can be listed as GSM900 for 2G, WCDMA Band 1 for 3G and LTE Band 1 for 4.5G mobile network. The audio codec bit rate values according to network bands are shown in Table 5.1. The results of the measurement are shared in Section 5.1.4. According to the results of the measurement, the improvement trials will be made for the failures of the codec bit rate that has insufficient sound quality. The improvement trials and the results of the measurements carried out after the improvement trials will be explained in Section 5.2.

Table 5.1 Audio codec bit rates according to network

Audio Codec Bit Rate Variables (kbps)		
2G	**3G**	**4.5G**
12.20	12.20	23.85
10.20	10.20	23.05
7.95	7.95	19.85
7.40	7.40	18.25
6.70	6.70	15.85
5.90	5.90	14.25
5.15	5.15	12.65
4.75	4.75	8.85
	23.85	6.60
	23.05	12.20
	19.85	10.20
	18.25	7.95
	15.85	7.40
	14.25	6.70
	12.65	5.90
	8.85	5.15
	6.60	4.75

5.1.4 Audio Quality Test Results with Current ACDB

The purpose of this test is the objective test with using ITU-T P.863 standard that similar result with Absolute Category Rating which is used in subjective tests. ITU-T MOS listening quality scale can be used when evaluating the MOS values obtained as a result of this test. This scale is shown in Chapter 3 as Table 3.11. According to this scale, the sound signals having 3.5 MOS values have a sound quality that can satisfy most users, so the target MOS value is 3.5. The audio quality measurements are taken in two different conditions that are tests based on codec bit rate variables and tests based on signal strength variables. Firstly, the measurements are taken at a constant signal strength in order to evaluate the results according to the variation of the audio codec bit rate without affecting the signal strength. In the second evaluation, the measurements are taken at a constant audio codec bit rate in order to evaluate the results according to the variation of the signal strength.

5.1.4.1 Tests results based on Audio Codec Bit Rate Variables

The measurements are performed at a constant signal strength in order to evaluate the variation of the audio codec bit rate without affecting the signal strength. The selected constant signal strength for the measurement is -85 dBm. The decibel-milliwatts that is briefly dBm, is unit of level used to indicate that a power ratio is expressed in decibels with reference to one milliwatt (Bigelow et al., 2001). The audio quality test results of 14 different codec bit rate values in 2G network at the downlink signal direction are shown in Table 5.2. For instance, in the downlink signal direction of 2G network, the MOS value, POLQA result graph and the waveform of the transmitted signal at 12.20 kbps audio codec bit rate are shown in Figure 5.11.

Table 5.2 2G downlink audio quality test result with current audio parameters

Mobile Network	Signal Direction	Speech Coding System	Codec Bit Rate (kbps)	Measured MOS Value	Target Value (MOS)	Result
2G	Downlink	Full Rate	12.2	3.65	>3.5	Pass
2G	Downlink	Full Rate	10.2	3.62	>3.5	Pass
2G	Downlink	Full Rate	7.95	3.60	>3.5	Pass
2G	Downlink	Full Rate	7.4	3.56	>3.5	Pass
2G	Downlink	Full Rate	6.7	**3.45**	>3.5	Fail
2G	Downlink	Full Rate	5.9	**3.40**	>3.5	Fail
2G	Downlink	Full Rate	5.15	**3.34**	>3.5	Fail
2G	Downlink	Full Rate	4.75	**3.28**	>3.5	Fail
2G	Downlink	Half Rate	7.95	**3.42**	>3.5	Fail
2G	Downlink	Half Rate	7.4	**3.43**	>3.5	Fail
2G	Downlink	Half Rate	6.7	**3.45**	>3.5	Fail
2G	Downlink	Half Rate	5.9	**3.21**	>3.5	Fail
2G	Downlink	Half Rate	5.15	**3.01**	>3.5	Fail
2G	Downlink	Half Rate	4.75	**2.96**	>3.5	Fail

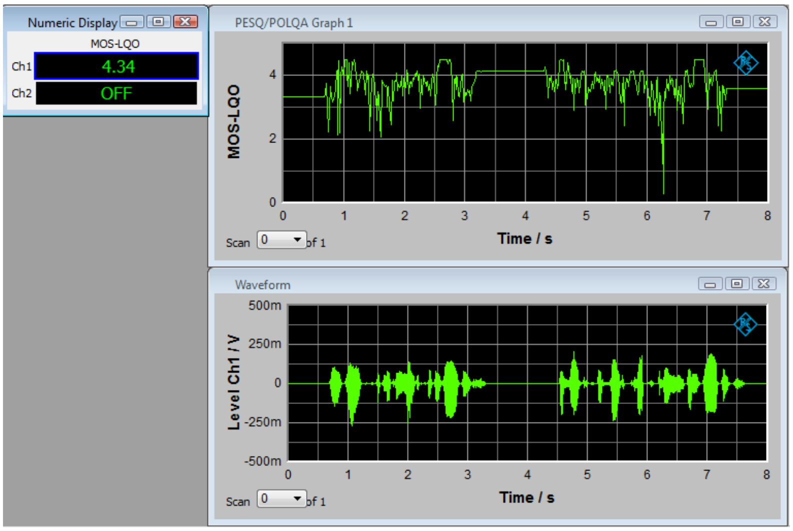

Figure 5.11 Graphical Test Result for 2G downlink signal direction at 12.20 kbps

The audio quality test results of 14 different codec bit rate values in the 2G network band at the uplink signal direction are shown in Table 5.3.

Table 5.3 2G uplink audio quality test result with current audio parameters

Mobile Network	Signal Direction	Speech Coding System	Codec Bit Rate (kbps)	Measured MOS Value	Target Value (MOS)	Result
2G	Uplink	Full Rate	43508	3,62	>3,5	Pass
2G	Uplink	Full Rate	43506	3,6	>3,5	Pass
2G	Uplink	Full Rate	34881	**3,35**	>3,5	Fail
2G	Uplink	Full Rate	43562	**3,11**	>3,5	Fail
2G	Uplink	Full Rate	43652	**3,06**	>3,5	Fail
2G	Uplink	Full Rate	43713	**3,1**	>3,5	Fail
2G	Uplink	Full Rate	42125	**2,98**	>3,5	Fail
2G	Uplink	Full Rate	27485	**2,78**	>3,5	Fail
2G	Uplink	Half Rate	34881	**3,15**	>3,5	Fail
2G	Uplink	Half Rate	43562	**3,13**	>3,5	Fail
2G	Uplink	Half Rate	43652	**2,85**	>3,5	Fail
2G	Uplink	Half Rate	43713	**2,58**	>3,5	Fail
2G	Uplink	Half Rate	42125	**2,73**	>3,5	Fail
2G	Uplink	Half Rate	27485	**2,97**	>3,5	Fail

The audio quality test result of the downlink and uplink signal directions in the 2G network are shown in Figure 5.12.

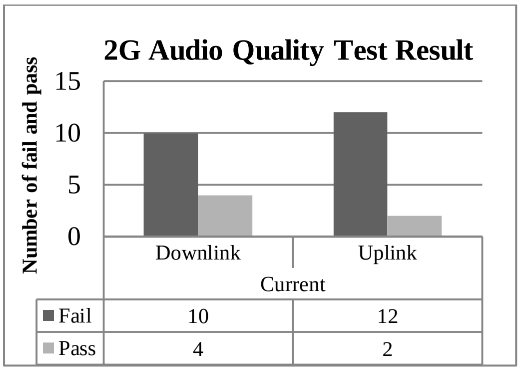

Figure 5.12 Audio quality test result of 2G mobile network according to audio codec bit rate variables

The audio quality test results of 17 different codec bit rate values in the 3G network at the downlink signal direction are shown in Table 5.4.

Table 5.4 3G downlink audio quality test result with current audio parameters

Mobile Network	Signal Direction	Codec Bit Rate (kbps)	Measured MOS Value	Target Value (MOS)	Result
3G	Downlink	12.2	3.72	>3.5	Pass
3G	Downlink	10.2	3.63	>3.5	Pass
3G	Downlink	7.95	**3.45**	>3.5	Fail
3G	Downlink	7.4	**3.41**	>3.5	Fail
3G	Downlink	6.7	**3.34**	>3.5	Fail
3G	Downlink	5.9	**3.21**	>3.5	Fail
3G	Downlink	5.15	**3.15**	>3.5	Fail
3G	Downlink	4.75	**3.12**	>3.5	Fail
3G	Downlink	23.85	3.85	>3.5	Pass
3G	Downlink	23.05	3.82	>3.5	Pass
3G	Downlink	19.85	3.78	>3.5	Pass
3G	Downlink	18.25	3.69	>3.5	Pass
3G	Downlink	15.85	3.57	>3.5	Pass
3G	Downlink	14.25	**3.39**	>3.5	Fail
3G	Downlink	12.65	**3.23**	>3.5	Fail
3G	Downlink	8.85	**3.07**	>3.5	Fail
3G	Downlink	6.6	**2.87**	>3.5	Fail

The audio quality test results of 17 different codec bit rate values in the 3G network at the uplink signal direction are shown in Table 5.5.

Table 5.5 3G uplink audio quality test result with current audio parameters

Mobile Network	Signal Direction	Codec Bit Rate (kbps)	Measured MOS Value	Target Value (MOS)	Result
3G	Uplink	12.2	375	>3.5	Pass
3G	Uplink	10.2	3.69	>3.5	Pass
3G	Uplink	7.95	**3.34**	>3.5	Fail
3G	Uplink	7.4	**3.39**	>3.5	Fail
3G	Uplink	6.7	**3.25**	>3.5	Fail
3G	Uplink	5.9	**3.19**	>3.5	Fail
3G	Uplink	5.15	**3.11**	>3.5	Fail
3G	Uplink	4.75	**3.03**	>3.5	Fail
3G	Uplink	23.85	3.88	>3.5	Pass
3G	Uplink	23.05	3.81	>3.5	Pass
3G	Uplink	19.85	3.75	>3.5	Pass
3G	Uplink	18.25	3.66	>3.5	Pass
3G	Uplink	15.85	3.52	>3.5	Pass
3G	Uplink	14.25	3.55	>3.5	Pass
3G	Uplink	12.65	**3.21**	>3.5	Fail
3G	Uplink	8.85	**3.02**	>3.5	Fail
3G	Uplink	6.6	**2.91**	>3.5	Fail

The audio quality test result of the downlink and uplink signal directions in the 3G network are shown in Figure 5.13.

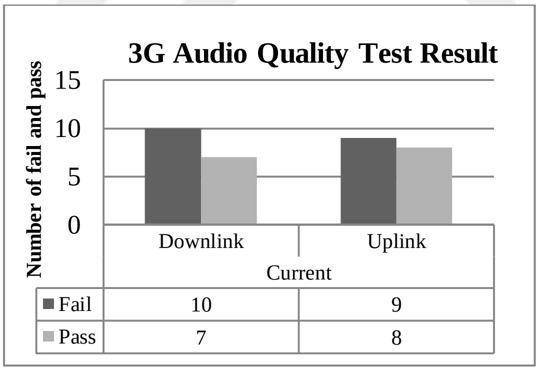

Figure 5.13 Audio quality test result of 3G mobile network according to audio codec bit rate variables

The audio quality test results of 17 different codec bit rate values in the 4.5G network at the downlink signal direction are shown in Table 5.6.

Table 5.6 4.5G downlink audio quality test result with current audio parameters

Mobile Network	Signal Direction	Codec Bit Rate (kbps)	Measured MOS Value	Target Value (MOS)	Result
4.5G	Downlink	23.85	3.93	>3.5	Pass
4.5G	Downlink	23.05	3.89	>3.5	Pass
4.5G	Downlink	19.85	3.81	>3.5	Pass
4.5G	Downlink	18.25	3.77	>3.5	Pass
4.5G	Downlink	15.85	3.72	>3.5	Pass
4.5G	Downlink	14.25	3.68	>3.5	Pass
4.5G	Downlink	12.65	3.63	>3.5	Pass
4.5G	Downlink	8.85	3.54	>3.5	Pass
4.5G	Downlink	6.6	**3.4**	>3.5	Fail
4.5G	Downlink	12.2	**3.33**	>3.5	Fail
4.5G	Downlink	10.2	**3.35**	>3.5	Fail
4.5G	Downlink	7.95	**3.21**	>3.5	Fail
4.5G	Downlink	7.4	**3.15**	>3.5	Fail
4.5G	Downlink	6.7	**3.06**	>3.5	Fail
4.5G	Downlink	5.9	**2.93**	>3.5	Fail
4.5G	Downlink	5.15	**2.82**	>3.5	Fail
4.5G	Downlink	4.75	**2.71**	>3.5	Fail

The audio quality test results of 17 different codec bit rate values in the 4.5G network at the uplink signal direction are shown in Table 5.7.

Table 5.7 4.5G uplink audio quality test result with current audio parameters

Mobile Network	Signal Direction	Codec Bit Rate (kbps)	Measured MOS Value	Target Value (MOS)	Result
4.5G	Uplink	23.85	3.95	>3.5	Pass
4.5G	Uplink	23.05	3.91	>3.5	Pass
4.5G	Uplink	19.85	3.85	>3.5	Pass
4.5G	Uplink	18.25	3.79	>3.5	Pass
4.5G	Uplink	15.85	3.73	>3.5	Pass
4.5G	Uplink	14.25	3.66	>3.5	Pass
4.5G	Uplink	12.65	3.58	>3.5	Pass
4.5G	Uplink	8.85	**3.47**	>3.5	Fail
4.5G	Uplink	6.6	**3.41**	>3.5	Fail
4.5G	Uplink	12.2	**3.38**	>3.5	Fail
4.5G	Uplink	10.2	**3.31**	>3.5	Fail
4.5G	Uplink	7.95	**3.26**	>3.5	Fail
4.5G	Uplink	7.4	**3.12**	>3.5	Fail
4.5G	Uplink	6.7	**3.01**	>3.5	Fail
4.5G	Uplink	5.9	**2.89**	>3.5	Fail
4.5G	Uplink	5.15	**2.79**	>3.5	Fail
4.5G	Uplink	4.75	**2.65**	>3.5	Fail

The audio quality test result of the downlink and uplink signal directions in the 4.5G network are shown in Figure 5.14.

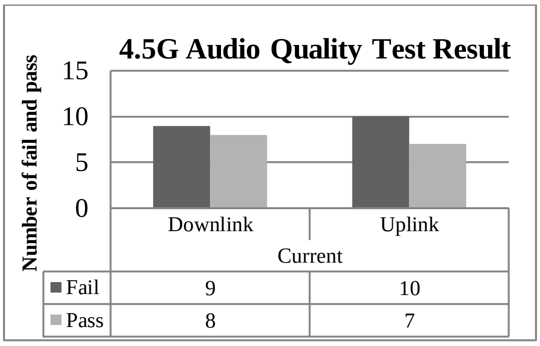

Figure 5.14 Audio quality test result of 4.5G mobile network according to audio codec bit rate variables

According to the audio quality test results on 2G, 3G and 4.5G networks, the audio quality of the device is insufficient for many codec bit rate values. Improvements are tried on the audio tuning side to improve the audio quality. These studies are described in Section 5.2.

5.1.4.2 Test results based on Signal Strength Variables

Measurements are performed at a constant audio codec bit rate in order to evaluate the variation of the signal strength without being affected by the audio codec bit rate. The chosen constant audio codec bit rate for measurement is 12.20 kbps that is common in 2G, 3G and 4.5G network bands. The measurement are performed at signal strength of poor, fair and good on 2G, 3G and 4.5G network bands. The signal strength is defined by the Received Signal Strength Indicator value that is briefly RSSI. The RSSI is a negative value and the signal near 0 is stronger. The values and identification of the RSSI according to the network bands are shown in Table 5.8 (Mobile Signal Strength Recommendations, 2018).

Table 5.8 Signal strength indicator value according to mobile network

Network	RSSI	Signal strength	Description
2G	>= -70 dBm	Excellent	Strong signal with maximum data speeds
	-70 dBm to -85 dBm	Good	Strong signal with good data speeds
	-86 dBm to -100 dBm	Fair	Fair but useful, fast and reliable data speeds may be attained, but marginal data with drop-outs is possible
	< -100 dBm	Poor	Performance will drop drastically
	-110 dBm	No signal	Disconnection
3G	>= -70 dBm	Excellent	Strong signal with maximum data speeds
	-70 dBm to -85 dBm	Good	Strong signal with good data speeds
	-86 dBm to -100 dBm	Fair	Fair but useful, fast and reliable data speeds may be attained, but marginal data with drop-outs is possible
	< -100 dBm	Poor	Performance will drop drastically
	-110 dBm	No signal	Disconnection
4,5G	> -65 dBm	Excellent	Strong signal with maximum data speeds
	-65 dBm to -75 dBm	Good	Strong signal with good data speeds
	-75 dBm to -85 dBm	Fair	Fair but useful, fast and reliable data speeds may be attained, but marginal data with drop-outs is possible
	-85 dBm to -95 dBm	Poor	Performance will drop drastically
	<= -95 dBm	No signal	Disconnection

The poor, fair and good signals are determined in 2G, 3G and 4.5G network bands for using in measurement at considering the signal strength values in Table 8. The determined signal strength values for measurement are shown in Table 5.9.

Table 5.9 The determined signal strength for measurement

	Signal Strength (dBm)		
	Good	**Fair**	**Poor**
2G	-75	-90	-105
3G	-75	-90	-105
4.5G	-70	-80	-90

The test results of the current audio calibration database at the signal strength values that is specified in Table 5.9 at the 2G, 3G and 4.5G mobile networks are shown in Table 5.10. The test results are also shown graphically in Figure 5.15. According to test results; the change in the signal strength are found to affect the sound quality. It is seen that as the signal level improved, the sound quality also increased.

Tablo 5.10 Test results according to signal strength variable with current audio parameters

Current acdb	Good Signal		Fair Signal		Poor Signal	
	Downlink	**Uplink**	**Downlink**	**Uplink**	**Downlink**	**Uplink**
2G	3.85	3.83	3.65	3.62	3.46	3.4
3G	3.9	3.91	3.72	3.75	3.58	3.52
4.5G	3.79	3.75	3.33	3.38	3.12	3.15

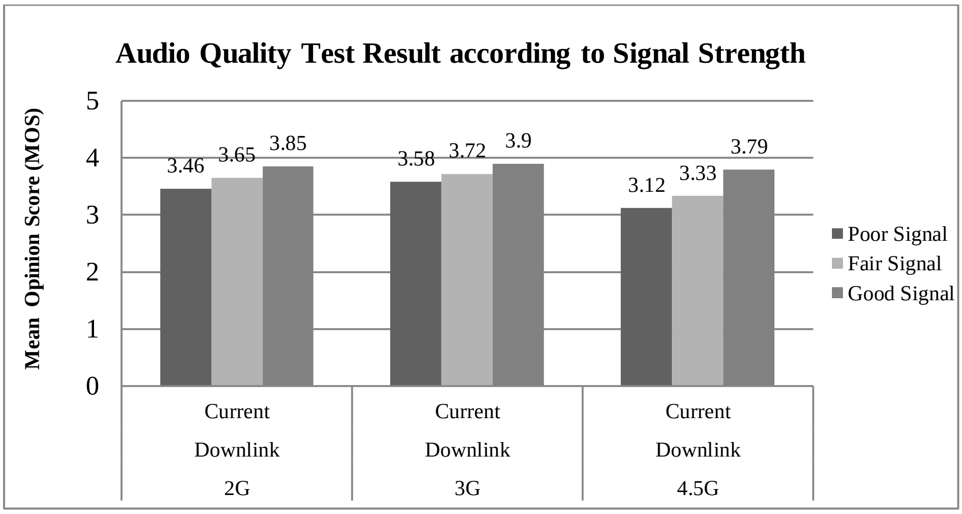

Figure 5.15 Audio Quality Test Result according to Signal Strength

5.2 Audio Quality Improvement Studies

The audio related parameters of the mobile based devices are kept in the audio calibration database that is called acdb file. If the audio parameters in the device are wanted to change in order to improving the audio quality, the audio calibration tool of the mobile based device's chipset manufacturer can be used to make changes. According to the measurement results, this audio calibration tool is used to improve the insufficient audio quality at the audio codec bit rate values. The interface of the audio calibration tool is shown in Figure 5.16.

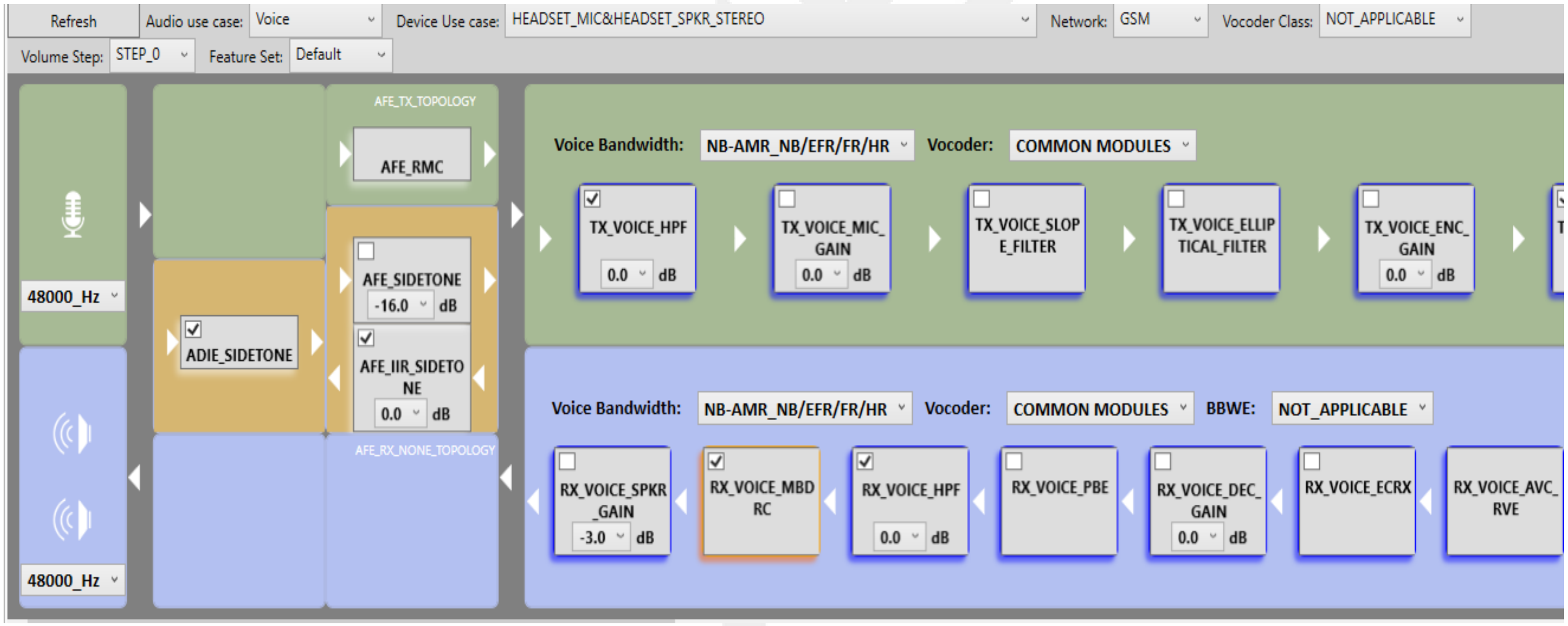

Figure 5.16 The interface of the audio calibration tool

The parameters of downlink and uplink that can be changed in the tool interface are shown in Table 5.11 and Table 5.12. After the parameter changes in the direction of the uplink and downlink signal direction through the audio calibration tool, a new acdb file is created by means of the tool. The new generated acdb file is uploaded in the mobile based device to improve the audio quality of the device.

Table 5.11 T_x path parameters for downlink

Tx Path Parameters	**Variables**
High Pass Filter (HPF)	Enable/Disable
Microphone Gain (MIC_GAIN)	Changing dB values
Slope Filter	Enable/Disable
Elliptical Filter	Enable/Disable
Encoder Gain (ENC_GAIN)	Changing dB values
Microphone Infinite Impulse Response (IIR_MIC1)	Enable/Disable
Single Microphone Echo Cancellation and Noise Suppression (SMECNS_V2)	Enable/Disable
Finite Impulse Response (FIR)	Enable/Disable
Infinite Impulse Response (IIR)	Enable/Disable
Adaptive Input Gain (AIG)	Enable/Disable
Dynamic Range Control (DRC)	Enable/Disable
Volume (VOL)	Changing dB values

Table 5.12 R_x path parameters for uplink

RX Path Parameters	**Variables**
Receive Gain (RX_GAIN)	Changing dB values
Far-end Noise Suppression (FNS)	Enable/Disable
Wide Volume V2 (WV_2)	Enable/Disable
WideVolume (WV)	Enable/Disable
Adaptive Input Gain (AIG)	Enable/Disable
Dynamic Range Control (DRC)	Enable/Disable
Finite Impulse Response (FIR)	Enable/Disable
Infinite Impulse Response (IIR)	Enable/Disable
Volume (VOL)	Changing dB values
Automatic Voice Control/Receive Voice Enhancement (AVC_RVE)	Enable/Disable
Echo Cancellar Receive (ECRX)	Enable/Disable
Decoder Gain (DEC_GAIN)	Changing dB values
Psychoacoustic Bass Enhancement (PBE)	Enable/Disable
High Pass Filter (HPF)	Enable/Disable
Multiband Dynamic Range Control (MBDRC)	Enable/Disable
Speaker Gain (SPKR_GAIN)	Changing dB values

In order to provide improvement in uplink and downlink signal direction, changes are made in T_x and R_x path parameters.

5.2.1 Improvement studies according to Audio Codec Bit Rate Variables

5.2.1.1 Improvement Studies for Downlink Signal Direction

In order to improve the audio quality in the downlink signal direction, parameters in Table 5.11 are used. In Table 5.13, the parameters in the current audio parameters of the mobile-based device are shown. In the first improvement trial, 4 parameters marked with yellow in Table 5.13 are changed. The changed parameters are listed below.

- Microphone gain is activated and 1dB gain is achieved.
- Encoder Gain (ENC_GAIN) is enabled and 1dB gain is achieved.
- Adaptive Input Gain (AIG) is activated.
- Volume (VOL) is increased by 2dB.

Table 5.13 1st improvement trial vs current audio parameters parameters for downlink

Tx Path Parameters	Variables	Current Parameters	1st Improvement Trial	Current Parameters	1st Improvement Trial	Current Parameters	1st Improvement Trial
		2G	2G	3G	3G	4,5G	4,5G
High Pass Filter(HPF)	Enable/Disable	ON	ON	ON	ON	ON	ON
Microphone Gain(MIC_GAIN)	Changing dB values	OFF-0dB	ON-1dB	OFF-0dB	ON-1dB	OFF-0dB	ON-1dB
Slope Filter	Enable/Disable	OFF	OFF	OFF	OFF	OFF	OFF
Elliptical Filter	Enable/Disable	OFF	OFF	OFF	OFF	OFF	OFF
Encoder Gain(ENC_GAIN)	Changing dB values	OFF-0dB	ON-1dB	OFF-0dB	ON-1dB	OFF-0dB	ON-1dB
Microphone Infinite Impulse Response(IIR_MIC1)	Enable/Disable	ON-6dB	ON-6dB	ON-6dB	ON-6dB	ON-6dB	ON-6dB
Single Microphone Echo Cancellation and Noise Suppression(SMECNS_V2)	Enable/Disable	ON	ON	ON	ON	ON	ON
Finite Impulse Response(FIR)	Enable/Disable	OFF	OFF	Deactive	Deactive	Deactive	Deactive
Infinite Impulse Response(IIR)	Enable/Disable	OFF	OFF	Deactive	Deactive	Deactive	Deactive
Adaptive Input Gain(AIG)	Enable/Disable	OFF	ON	OFF	ON	OFF	ON
Dynamic Range Control(DRC)	Enable/Disable	ON	ON	Deactive	Deactive	Deactive	Deactive
Volume(VOL)	Changing dB values	0 dB	2 dB	0 dB	2 dB	0 dB	2 dB

The measurement of audio quality are performed on 2G, 3G and 4.5G mobile network after the first improvement trial. The results of the first improvement trial are presented in Table 5.14 for 2G network. Table 5.15 for 3G network and Table 5.16 for 4.5G network.

Table 5.14 2G downlink audio quality test result after 1st improvement trial

Mobile Network	Signal Direction	Speech Coding System	Codec Bit Rate (kbps)	Measured MOS Value	Target Value (MOS)	Result
2G	Downlink	Full Rate	12.2	3.72	>3.5	Pass
2G	Downlink	Full Rate	10.2	3.75	>3.5	Pass
2G	Downlink	Full Rate	7.95	3.76	>3.5	Pass
2G	Downlink	Full Rate	7.4	3.68	>3.5	Pass
2G	Downlink	Full Rate	6.7	3.60	>3.5	Pass
2G	Downlink	Full Rate	5.9	3.55	>3.5	Pass
2G	Downlink	Full Rate	5.15	**3.38**	>3.5	Fail
2G	Downlink	Full Rate	4.75	**3.32**	>3.5	Fail
2G	Downlink	Half Rate	7.95	3.73	>3.5	Pass
2G	Downlink	Half Rate	7.4	3.75	>3.5	Pass
2G	Downlink	Half Rate	6.7	**3.46**	>3.5	Fail
2G	Downlink	Half Rate	5.9	**3.34**	>3.5	Fail
2G	Downlink	Half Rate	5.15	**3.25**	>3.5	Fail
2G	Downlink	Half Rate	4.75	**3.12**	>3.5	Fail

Table 5.15 3G downlink audio quality test result after 1st improvement trial

Mobile Network	Signal Direction	Codec Bit Rate (kbps)	Measured MOS Value	Target Value (MOS)	Result
3G	Downlink	12.2	3.83	>3.5	Pass
3G	Downlink	10.2	3.76	>3.5	Pass
3G	Downlink	7.95	3.56	>3.5	Pass
3G	Downlink	7.4	3.52	>3.5	Pass
3G	Downlink	6.7	**3.42**	>3.5	Fail
3G	Downlink	5.9	**3.32**	>3.5	Fail
3G	Downlink	5.15	**3.33**	>3.5	Fail
3G	Downlink	4.75	**3.24**	>3.5	Fail
3G	Downlink	23.85	3.85	>3.5	Pass
3G	Downlink	23.05	3.91	>3.5	Pass
3G	Downlink	19.85	3.88	>3.5	Pass
3G	Downlink	18.25	3.79	>3.5	Pass
3G	Downlink	15.85	3.71	>3.5	Pass
3G	Downlink	14.25	3.58	>3.5	Pass
3G	Downlink	12.65	3.52	>3.5	Pass
3G	Downlink	8.85	**3.21**	>3.5	Fail
3G	Downlink	6.6	**3.16**	>3.5	Fail

Table 5.16 4.5G downlink audio quality test result after 1st improvement trial

Mobile Network	Signal Direction	Codec Bit Rate (kbps)	Measured MOS Value	Target Value (MOS)	Result
4.5G	Downlink	23.85	3.95	>3.5	Pass
4.5G	Downlink	23.05	3.94	>3.5	Pass
4.5G	Downlink	19.85	3.85	>3.5	Pass
4.5G	Downlink	18.25	3.8	>3.5	Pass
4.5G	Downlink	15.85	3.75	>3.5	Pass
4.5G	Downlink	14.25	3.72	>3.5	Pass
4.5G	Downlink	12.65	3.68	>3.5	Pass
4.5G	Downlink	8.85	3.61	>3.5	Pass
4.5G	Downlink	6.6	3.58	>3.5	Pass
4.5G	Downlink	12.2	3.52	>3.5	Pass
4.5G	Downlink	10.2	**3.37**	>3.5	Fail
4.5G	Downlink	7.95	**3.25**	>3.5	Fail
4.5G	Downlink	7.4	**3.18**	>3.5	Fail
4.5G	Downlink	6.7	**3.09**	>3.5	Fail
4.5G	Downlink	5.9	**2.97**	>3.5	Fail
4.5G	Downlink	5.15	**2.85**	>3.5	Fail
4.5G	Downlink	4.75	**2.79**	>3.5	Fail

The general evaluation of the first improvement trial at the downlink signal direction in the 2G, 3G and 4.5G network bands is shown in Figure 5.17. The measurement results with the current audio paremeters show that the sound quality of downlink is not satisfied at 10 of fourteen audio codec bit rate variables, 10 and 9 of seventeen audio codec bit rate variables respectively for 2G, 3G and 4.5G mobile network. After the first improvement trial, the sound quality of downlink is not satisfied at 6 of fourteen audio codec bit rate variables, 6 and 7 of seventeen audio codec bit rate variables respectively for 2G, 3G and 4.5G mobile network.

If the audio quality of 2G, 3G and 4.5G network is generally respectively evaluated after the first improvement trial, the sound quality of downlink is not satisfied at 42.85%, 35.29% and 41.17% of the total audio codec variables according to the MOS scale. A second improvement trial should be attempted.

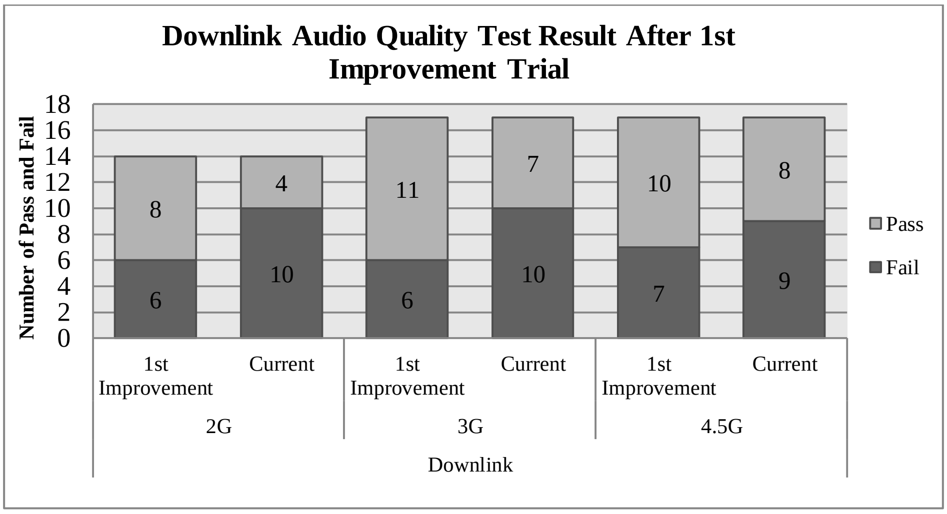

Figure 5.17 Downlink audio quality test result after 1st improvement trial

As the first improvement trial is not sufficient, a second improvement trial is made. In the second improvement trial, 5 parameters and filters marked with yellow in Table 5.17 are changed. The changed parameters are listed below.

- Microphone gain is activated and 2dB gain is achieved.
- Elliptical filter has been activated.
- Encoder Gain (ENC_GAIN) is enabled and 2dB gain is achieved.
- Adaptive Input Gain (AIG) is activated.
- Volume (VOL) is increased by 3dB.

Table 5.17 2nd improvement trial vs current ACDB parameters for downlink

Tx Path Parameters	Variables	Current Parameters	2nd Improvement Trial	Current Parameters	2nd Improvement Trial	Current Parameters	2nd Improvement Trial
		2G	2G	3G	3G	4,5G	4,5G
High Pass Filter(HPF)	Enable/Disable	ON	ON	ON	ON	ON	ON
Microphone Gain(MIC_GAIN)	Changing dB values	OFF-0dB	ON-2dB	OFF-0dB	ON-2dB	OFF-0dB	ON-2dB
Slope Filter	Enable/Disable	OFF	OFF	OFF	OFF	OFF	OFF
Elliptical Filter	Enable/Disable	OFF	ON	OFF	ON	OFF	ON
Encoder Gain(ENC_GAIN)	Changing dB values	OFF-0dB	ON-2dB	OFF-0dB	ON-2dB	OFF-0dB	ON-2dB
Microphone Infinite Impulse Response(IIR_MIC1)	Enable/Disable	ON-6dB	ON-6dB	ON-6dB	ON-6dB	ON-6dB	ON-6dB
Single Microphone Echo Cancellation and Noise Suppression(SMECNS_V2)	Enable/Disable	ON	ON	ON	ON	ON	ON
Finite Impulse Response(FIR)	Enable/Disable	OFF	OFF	Deactive	Deactive	Deactive	Deactive
Infinite Impulse Response(IIR)	Enable/Disable	OFF	OFF	Deactive	Deactive	Deactive	Deactive
Adaptive Input Gain(AIG)	Enable/Disable	OFF	ON	OFF	ON	OFF	ON
Dynamic Range Control(DRC)	Enable/Disable	ON	ON	Deactive	Deactive	Deactive	Deactive
Volume(VOL)	Changing dB values	0 dB	3 dB	0 dB	3 dB	0 dB	3 dB

The measurements of audio quality are performed on 2G, 3G and 4.5G network bands after the second improvement trial. The results of the second improvement trial are presented in Table 5.18 for 2G network, Table 5.19 for 3G network and Table 5.20 for 4.5G network.

Table 5.18 2G downlink audio quality test result after 2nd improvement trial

Mobile Network	Signal Direction	Speech Coding System	Codec Bit Rate (kbps)	Measured MOS Value	Target Value (MOS)	Result
2G	Downlink	Full Rate	12.2	3.85	>3.5	Pass
2G	Downlink	Full Rate	10.2	3.88	>3.5	Pass
2G	Downlink	Full Rate	7.95	3.90	>3.5	Pass
2G	Downlink	Full Rate	7.4	3.86	>3.5	Pass
2G	Downlink	Full Rate	6.7	3.75	>3.5	Pass
2G	Downlink	Full Rate	5.9	3.72	>3.5	Pass
2G	Downlink	Full Rate	5.15	**3.59**	>3.5	Pass
2G	Downlink	Full Rate	4.75	**3.41**	>3.5	Fail
2G	Downlink	Half Rate	7.95	3.75	>3.5	Pass
2G	Downlink	Half Rate	7.4	3.81	>3.5	Pass
2G	Downlink	Half Rate	6.7	**3.62**	>3.5	Pass
2G	Downlink	Half Rate	5.9	**3.59**	>3.5	Pass
2G	Downlink	Half Rate	5.15	**3.42**	>3.5	Fail
2G	Downlink	Half Rate	4.75	**3.38**	>3.5	Fail

Table 5.19 3G downlink audio quality test result after 2nd improvement trial

Mobile Network	Signal Direction	Codec Bit Rate (kbps)	Measured MOS Value	Target Value (MOS)	Result
3G	Downlink	12.2	3.97	>3.5	Pass
3G	Downlink	10.2	3.88	>3.5	Pass
3G	Downlink	7.95	3.73	>3.5	Pass
3G	Downlink	7.4	3.69	>3.5	Pass
3G	Downlink	6.7	3.61	>3.5	Pass
3G	Downlink	5.9	3.56	>3.5	Pass
3G	Downlink	5.15	**3.41**	>3.5	Fail
3G	Downlink	4.75	**3.3**	>3.5	Fail
3G	Downlink	23.85	4.03	>3.5	Pass
3G	Downlink	23.05	4.01	>3.5	Pass
3G	Downlink	19.85	3.93	>3.5	Pass
3G	Downlink	18.25	3.84	>3.5	Pass
3G	Downlink	15.85	3.8	>3.5	Pass
3G	Downlink	14.25	3.67	>3.5	Pass
3G	Downlink	12.65	3.62	>3.5	Pass
3G	Downlink	8.85	**3.42**	>3.5	Fail
3G	Downlink	6.6	**3.32**	>3.5	Fail

Table 5.20 4.5G downlink audio quality test result after 2nd improvement trial

Mobile Network	Signal Direction	Codec Bit Rate (kbps)	Measured MOS Value	Target Value (MOS)	Result
4.5G	Downlink	23.85	4.09	>3.5	Pass
4.5G	Downlink	23.05	4.03	>3.5	Pass
4.5G	Downlink	19.85	4.01	>3.5	Pass
4.5G	Downlink	18.25	3.98	>3.5	Pass
4.5G	Downlink	15.85	3.92	>3.5	Pass
4.5G	Downlink	14.25	3.82	>3.5	Pass
4.5G	Downlink	12.65	3.79	>3.5	Pass
4.5G	Downlink	8.85	3.76	>3.5	Pass
4.5G	Downlink	6.6	3.69	>3.5	Pass
4.5G	Downlink	12.2	3.61	>3.5	Pass
4.5G	Downlink	10.2	3.55	>3.5	Pass
4.5G	Downlink	7.95	3.52	>3.5	Pass
4.5G	Downlink	7.4	**3.33**	>3.5	Fail
4.5G	Downlink	6.7	**3.25**	>3.5	Fail
4.5G	Downlink	5.9	**3.14**	>3.5	Fail
4.5G	Downlink	5.15	**3.09**	>3.5	Fail
4.5G	Downlink	4.75	**2.99**	>3.5	Fail

The general evaluation of the second improvement trial at the downlink signal direction in the 2G, 3G and 4.5G network bands is shown in Figure 5.18. The measurement results with the current audio paremeters show that the sound quality of downlink is not satisfied at 10 of fourteen audio codec bit rate variables, 10 and 9 of seventeen audio codec bit rate variables respectively for 2G, 3G and 4.5G mobile network. After the second improvement trial, the sound quality of downlink is not satisfied at 3 of fourteen audio codec bit rate variables, 4 and 5 of seventeen audio codec bit rate variables respectively for 2G, 3G and 4.5G mobile network.

If the audio quality of 2G, 3G and 4.5G network is generally respectively evaluated after the second improvement trial, the sound quality of downlink is not satisfied at 21.42%, 23.52% and 29.41% of the total audio codec variables according to the MOS scale. A third improvement trial should be attempted.

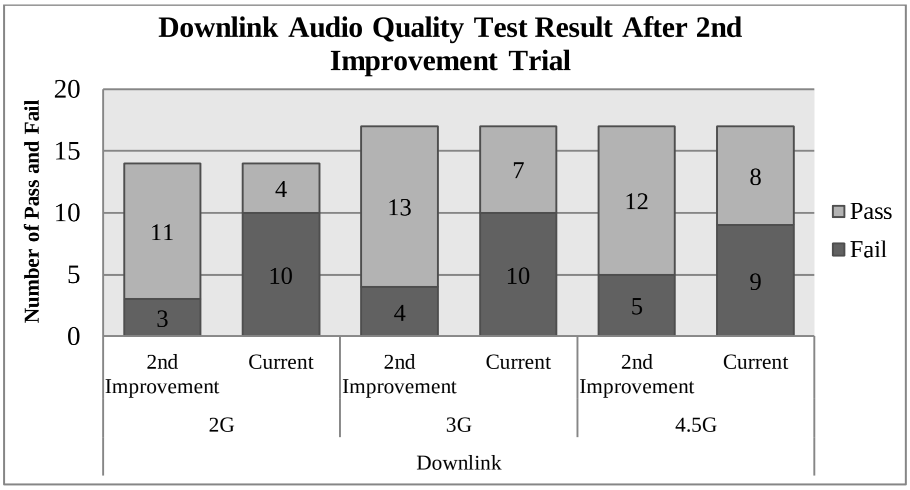

Figure 5.18 Downlink audio quality test result after 2nd improvement trial

As the first and second improvement trials are not sufficient, a third improvement trial is made. In the third improvement trial, 6 parameters and filters marked with yellow in Table 5.21 are changed. The changed parameters are listed below.

- Microphone gain is enabled and 3dB gain is achieved.
- The slope filter is activated.
- Encoder Gain (ENC_GAIN) is enabled and 3dB gain is achieved.
- Finite impulse response (FIR) is activated for GSM band.
- Adaptive Input Gain (AIG) is activated.
- Volume (VOL) is increased by 3dB.

Table 5.21 3rd improvement trial vs current ACDB parameters for downlink

Tx Path Parameters	Variables	Current Parameters	3rd Improvement Trial	Current Parameters	3rd Improvement Trial	Current Parameters	3rd Improvement Trial
		2G	2G	3G	3G	4,5G	4,5G
High Pass Filter(HPF)	Enable/Disable	ON	ON	ON	ON	ON	ON
Microphone Gain(MIC_GAIN)	Changing dB values	OFF-0dB	ON-3dB	OFF-0dB	ON-2dB	OFF-0dB	ON-2dB
Slope Filter	Enable/Disable	OFF	ON	OFF	ON	OFF	ON
Elliptical Filter	Enable/Disable	OFF	OFF	OFF	OFF	OFF	OFF
Encoder Gain(ENC_GAIN)	Changing dB values	OFF-0dB	ON-3dB	OFF-0dB	ON-3dB	OFF-0dB	ON-3dB
Microphone Infinite Impulse Response(IIR_MIC1)	Enable/Disable	ON-6dB	ON-6dB	ON-6dB	ON-6dB	ON-6dB	ON-6dB
Single Microphone Echo Cancellation and Noise Suppression(SMECNS_V2)	Enable/Disable	ON	ON	ON	ON	ON	ON
Finite Impulse Response(FIR)	Enable/Disable	OFF	ON	Deactive	Deactive	Deactive	Deactive
Infinite Impulse Response(IIR)	Enable/Disable	OFF	OFF	Deactive	Deactive	Deactive	Deactive
Adaptive Input Gain(AIG)	Enable/Disable	OFF	ON	OFF	ON	OFF	ON
Dynamic Range Control(DRC)	Enable/Disable	ON	ON	Deactive	Deactive	Deactive	Deactive
Volume(VOL)	Changing dB values	0 dB	3 dB	0 dB	3 dB	0 dB	3 dB

The measurements of audio quality are performed on 2G, 3G and 4.5G network bands after the third improvement trial. The results of the third improvement trial are presented in Table 5.22 for 2G network, Table 5.23 for 3G network and Table 5.24 for 4.5G network.

Table 5.22 2G downlink audio quality test result after 3rd improvement trial

Mobile Network	Signal Direction	Speech Coding System	Codec Bit Rate (kbps)	Measured MOS Value	Target Value (MOS)	Result
2G	Downlink	Full Rate	12.2	4.21	>35	Pass
2G	Downlink	Full Rate	10.2	4.10	>3.5	Pass
2G	Downlink	Full Rate	7.95	4.07	>3.5	Pass
2G	Downlink	Full Rate	7.4	4.11	>3.5	Pass
2G	Downlink	Full Rate	6.7	3.99	>3.5	Pass
2G	Downlink	Full Rate	5.9	3.84	>3.5	Pass
2G	Downlink	Full Rate	5.15	3.81	>3.5	Pass
2G	Downlink	Full Rate	4.75	3.75	>3.5	Pass
2G	Downlink	Half Rate	7.95	3.98	>3.5	Pass
2G	Downlink	Half Rate	7.4	3.91	>3.5	Pass
2G	Downlink	Half Rate	6.7	3.85	>3.5	Pass
2G	Downlink	Half Rate	5.9	3.76	>3.5	Pass
2G	Downlink	Half Rate	5.15	**3.44**	>3.5	Fail
2G	Downlink	Half Rate	4.75	**3.41**	>3.5	Fail

Table 5.23 3G downlink audio quality test result after 3rd improvement trial

Mobile Network	Signal Direction	Codec Bit Rate (kbps)	Measured MOS Value	Target Value (MOS)	Result
3G	Downlink	12.2	4.02	>3.5	Pass
3G	Downlink	10.2	3.92	>3.5	Pass
3G	Downlink	7.95	3.85	>3.5	Pass
3G	Downlink	7.4	3.77	>3.5	Pass
3G	Downlink	6.7	3.71	>3.5	Pass
3G	Downlink	5.9	3.69	>3.5	Pass
3G	Downlink	5.15	3.58	>3.5	Pass
3G	Downlink	4.75	**3.35**	>3.5	Fail
3G	Downlink	23.85	4.21	>3.5	Pass
3G	Downlink	23.05	4.15	>3.5	Pass
3G	Downlink	19.85	4.04	>3.5	Pass
3G	Downlink	18.25	4.01	>3.5	Pass
3G	Downlink	15.85	3.99	>3.5	Pass
3G	Downlink	14.25	3.75	>3.5	Pass
3G	Downlink	12.65	3.76	>3.5	Pass
3G	Downlink	8.85	3.61	>3.5	Pass
3G	Downlink	6.6	3.56	>3.5	Pass

Table 5.24 4.5G downlink audio quality test result after 3rd improvement trial

Mobile Network	Signal Direction	Codec Bit Rate (kbps)	Measured MOS Value	Target Value (MOS)	Result
4.5G	Downlink	23.85	4.32	>3.5	Pass
4.5G	Downlink	23.05	4.27	>3.5	Pass
4.5G	Downlink	19.85	4.23	>3.5	Pass
4.5G	Downlink	18.25	4.15	>3.5	Pass
4.5G	Downlink	15.85	4.09	>3.5	Pass
4.5G	Downlink	14.25	4.01	>3.5	Pass
4.5G	Downlink	12.65	3.95	>3.5	Pass
4.5G	Downlink	8.85	3.9	>3.5	Pass
4.5G	Downlink	6.6	3.85	>3.5	Pass
4.5G	Downlink	12.2	3.79	>3.5	Pass
4.5G	Downlink	10.2	3.72	>3.5	Pass
4.5G	Downlink	7.95	3.66	>3.5	Pass
4.5G	Downlink	7.4	3.61	>3.5	Pass
4.5G	Downlink	6.7	3.56	>3.5	Pass
4.5G	Downlink	5.9	3.52	>3.5	Pass
4.5G	Downlink	5.15	**3.15**	>3.5	Fail
4.5G	Downlink	4.75	**3.07**	>3.5	Fail

The general evaluation of the third improvement trial at the downlink signal direction in the 2G, 3G and 4.5G network bands are shown in Figure 5.19. The measurement results with the current audio paremeters show that the sound quality of downlink is not satisfied at 10 of fourteen audio codec bit rate variables, 10 and 9 of seventeen audio codec bit rate variables respectively for 2G, 3G and 4.5G mobile network. After the third improvement trial, the sound quality of downlink is not satisfied at 2 of fourteen audio codec bit rate variables, 1 and 2 of seventeen audio codec bit rate variables respectively for 2G, 3G and 4.5G mobile network.

If the audio quality of 2G, 3G and 4.5G network is generally respectively evaluated after the third improvement trial, the sound quality of downlink is satisfied at 85.17%, 94.11% and 88.23% of the total audio codec variables by most users.

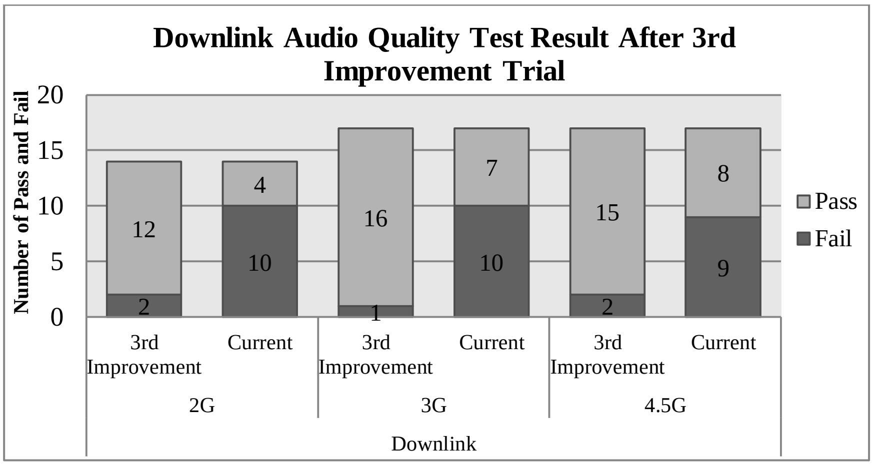

Figure 5.19 Downlink audio quality test result after 3rd improvement trial

5.2.1.2 Improvement Studies for Uplink Signal Direction

In order to improve the audio quality in the uplink signal direction, parameters and filters in Table 5.12 are used. In Table 5.25, the parameters and filters in the current acdb file of the mobile-based device are shown. In the first improvement trial, 4 parameters and filters marked with yellow in Table 5.25 are changed. The changed parameters are listed below.

- Receive Gain (RX_GAIN) is enabled and 1dB gain is achieved.
- Adaptive Input Gain (AIG) is activated for GSM band.
- Volume (VOL) is increased by 2dB.
- Decoder Gain (DEC_GAIN) is enabled and 1dB gain is achieved.

Table 5.25 1st Improvement Trial vs Current ACDB Parameters for Uplink

Rx Path Parameters	Variables	Current Parameters	1st Improvement Trial	Current Parameters	1st Improvement Trial	Current Parameters	1st Improvement Trial
		2G	2G	WCDMA	WCDMA	LTE	LTE
Receive Gain(RX_GAIN)	Changing dB values	OFF-0dB	ON-1dB	OFF-0dB	ON-1dB	OFF-0dB	ON-1dB
Far-end Noise Suppression(FNS)	Enable/Disable	ON	ON	Deactive	Deactive	Deactive	Deactive
Wide Volume V2(WV_2)	Enable/Disable	Deactive	Deactive	OFF	OFF	OFF	OFF
WideVolume(WV)	Enable/Disable	Deactive	Deactive	OFF	OFF	OFF	OFF
Adaptive Input Gain(AIG)	Enable/Disable	OFF	ON	Deactive	Deactive	Deactive	Deactive
Dynamic Range Control(DRC)	Enable/Disable	ON	ON	Deactive	Deactive	Deactive	Deactive
Finite Impulse Response(FIR)	Enable/Disable	OFF	OFF	Deactive	Deactive	Deactive	Deactive
Infinite Impulse Response(IIR)	Enable/Disable	ON	ON	Deactive	Deactive	Deactive	Deactive
Volume(VOL)	Changing dB values	0 dB	2 dB	0 dB	2 dB	0 dB	2 dB
Automatic Voice Control/Receive Voice Enhancement(AVC_RVE)	Enable/Disable	OFF	OFF	OFF	OFF	OFF	OFF
Echo Cancellar Receive(ECRX)	Enable/Disable	OFF	OFF	OFF	OFF	OFF	OFF
Decoder Gain(DEC_GAIN)	Changing dB values	OFF-0dB	ON-1dB	OFF-0dB	ON-1dB	OFF-0dB	ON-1dB
Psychoacoustic Bass Enhancement(PBE)	Enable/Disable	OFF	OFF	OFF	OFF	OFF	OFF
High Pass Filter(HPF)	Enable/Disable	OFF	OFF	OFF	OFF	OFF	OFF
Multiband Dynamic Range Control(MBDRC)	Enable/Disable	OFF	OFF	Deactive	Deactive	Deactive	Deactive
Speaker Gain(SPKR_GAIN)	Changing dB values	OFF-0dB	OFF-0dB	OFF-0dB	OFF-0dB	OFF-0dB	OFF-0dB

The measurements of audio quality are performed on 2G, 3G and 4.5G network bands after the first improvement trial. The results of the first improvement trial are presented in Table 5.26 for 2G network, Table 5.27 for 3G network and Table 5.28 for 4.5G network.

Table 5.26 2G uplink audio quality test result after 1st improvement trial

Mobile Network	Signal Direction	Speech Coding System	Codec Bit Rate (kbps)	Measured MOS Value	Target Value (MOS)	Result
2G	Uplink	Full Rate	12.2	3.71	>3.5	Pass
2G	Uplink	Full Rate	10.2	3.75	>3.5	Pass
2G	Uplink	Full Rate	7.95	3.68	>3.5	Pass
2G	Uplink	Full Rate	7.4	3.65	>3.5	Pass
2G	Uplink	Full Rate	6.7	3.58	>3.5	Pass
2G	Uplink	Full Rate	5.9	**3.15**	>3.5	Fail
2G	Uplink	Full Rate	5.15	**3.11**	>3.5	Fail
2G	Uplink	Full Rate	4.75	**3.02**	>3.5	Fail
2G	Uplink	Half Rate	7.95	3.56	>3.5	Pass
2G	Uplink	Half Rate	7.4	**3.21**	>3.5	Fail
2G	Uplink	Half Rate	6.7	**3.15**	>3.5	Fail
2G	Uplink	Half Rate	5.9	**2.87**	>3.5	Fail
2G	Uplink	Half Rate	5.15	**2.98**	>3.5	Fail
2G	Uplink	Half Rate	4.75	**3.14**	>3.5	Fail

Table 5.27 3G Uplink Audio Quality Test Result after 1st Improvement Trial

Mobile Network	Signal Direction	Codec Bit Rate (kbps)	Measured MOS Value	Target Value (MOS)	Result
3G	Uplink	12.2	3.81	>3.5	Pass
3G	Uplink	10.2	3.77	>3.5	Pass
3G	Uplink	7.95	3.51	>3.5	Pass
3G	Uplink	7.4	3.55	>3.5	Pass
3G	Uplink	6.7	**3.34**	>3.5	Fail
3G	Uplink	5.9	**3.31**	>3.5	Fail
3G	Uplink	5.15	**3.25**	>3.5	Fail
3G	Uplink	4.75	**3.19**	>3.5	Fail
3G	Uplink	23.85	3.95	>3.5	Pass
3G	Uplink	23.05	3.93	>3.5	Pass
3G	Uplink	19.85	3.86	>3.5	Pass
3G	Uplink	18.25	3.8	>3.5	Pass
3G	Uplink	15.85	3.75	>3.5	Pass
3G	Uplink	14.25	3.63	>3.5	Pass
3G	Uplink	12.65	3.54	>3.5	Pass
3G	Uplink	8.85	**3.25**	>3.5	Fail
3G	Uplink	6.6	**3.19**	>3.5	Fail

.

Table 5.28 4.5G Uplink Audio Quality Test Result after 1st Improvement Trial

Mobile Network	Signal Direction	Codec Bit Rate (kbps)	Measured MOS Value	Target Value (MOS)	Result
4.5G	Uplink	23.85	3.99	>3.5	Pass
4.5G	Uplink	23.05	3.96	>3.5	Pass
4.5G	Uplink	19.85	3.88	>3.5	Pass
4.5G	Uplink	18.25	3.83	>3.5	Pass
4.5G	Uplink	15.85	3.81	>3.5	Pass
4.5G	Uplink	14.25	3.75	>3.5	Pass
4.5G	Uplink	12.65	3.71	>3.5	Pass
4.5G	Uplink	8.85	3.63	>3.5	Pass
4.5G	Uplink	6.6	3.55	>3.5	Pass
4.5G	Uplink	12.2	3.51	>3.5	Pass
4.5G	Uplink	10.2	**3.42**	>3.5	Fail
4.5G	Uplink	7.95	**3.3**	>3.5	Fail
4.5G	Uplink	7.4	**3.22**	>3.5	Fail
4.5G	Uplink	6.7	**3.14**	>3.5	Fail
4.5G	Uplink	5.9	**3.04**	>3.5	Fail
4.5G	Uplink	5.15	**2.93**	>3.5	Fail
4.5G	Uplink	4.75	**2.82**	>3.5	Fail

The general evaluation of the first improvement trial at the uplink signal direction in the 2G, 3G and 4.5G network bands is shown in Figure 5.20. The measurement results with the current audio paremeters show that the sound quality of uplink is not satisfied at 12 of fourteen audio codec bit rate variables, 9 and 10 of seventeen audio codec bit rate variables respectively for 2G, 3G and 4.5G mobile network. After the first improvement trial, the sound quality of uplink is not satisfied at 8 of fourteen audio codec bit rate variables, 6 and 7 of seventeen audio codec bit rate variables respectively for 2G, 3G and 4.5G mobile network.

If the audio quality of 2G, 3G and 4.5G network is generally respectively evaluated after the first improvement trial, the sound quality of uplink does not satisfied at 47.05%, 35.29% and 41.17% of the total audio codec variables according to the MOS scale. A second improvement trial should be attempted.

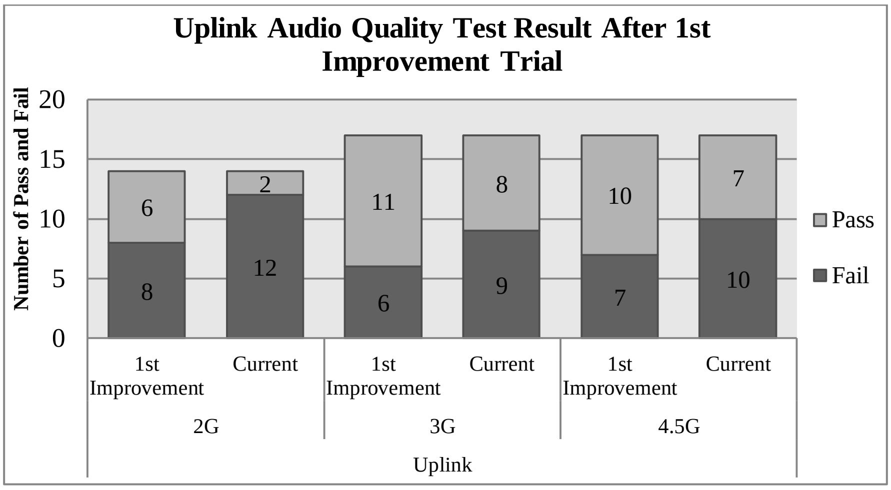

Figure 5.20 Uplink audio quality test result after 1st improvement trial

As the first improvement trial is not sufficient, a second improvement trial is made. In the second improvement trial, 6 parameters and filters marked with yellow in Table 5.29 are changed. The changed parameters are listed below.

- Receive Gain (RX_GAIN) is enabled and 2dB gain is achieved.
- Wide Volume is active for 3G and 4.5G bands.
- Adaptive Input Gain (AIG) is enabled for 2G band.
- Volume (VOL) is increased by 2dB.
- Decoder Gain (DEC_GAIN) is enabled and 1dB gain is achieved.
- High Pass Filter (HPF) is activated.

Table 5.29 2nd Improvement Trial vs Current ACDB Parameters for Uplink

Rx Path Parameters	Variables	Current Parameters	2nd Improvement Trial	Current Parameters	2nd Improvement Trial	Current Parameters	2nd Improvement Trial
		2G	2G	3G	3G	4,5G	4,5G
Receive Gain(RX_GAIN)	Changing dB values	OFF-0dB	ON-2dB	OFF-0dB	ON-2dB	OFF-0dB	ON-2dB
Far-end Noise Suppression(FNS)	Enable/Disable	ON	ON	Deactive	Deactive	Deactive	Deactive
Wide Volume V2(WV_2)	Enable/Disable	Deactive	Deactive	OFF	OFF	OFF	OFF
WideVolume(WV)	Enable/Disable	Deactive	Deactive	OFF	ON	OFF	ON
Adaptive Input Gain(AIG)	Enable/Disable	OFF	ON	Deactive	Deactive	Deactive	Deactive
Dynamic Range Control(DRC)	Enable/Disable	ON	ON	Deactive	Deactive	Deactive	Deactive
Finite Impulse Response(FIR)	Enable/Disable	OFF	OFF	Deactive	Deactive	Deactive	Deactive
Infinite Impulse Response(IIR)	Enable/Disable	ON	ON	Deactive	Deactive	Deactive	Deactive
Volume(VOL)	Changing dB values	0 dB	2 dB	0 dB	2 dB	0 dB	2 dB
Automatic Voice Control/Receive Voice Enhancement(AVC_RVE)	Enable/Disable	OFF	OFF	OFF	OFF	OFF	OFF
Echo Cancellar Receive(ECRX)	Enable/Disable	OFF	OFF	OFF	OFF	OFF	OFF
Decoder Gain(DEC_GAIN)	Changing dB values	OFF-0dB	ON-1dB	OFF-0dB	ON-1dB	OFF-0dB	ON-1dB
Psychoacoustic Bass Enhancement(PBE)	Enable/Disable	OFF	OFF	OFF	OFF	OFF	OFF
High Pass Filter(HPF)	Enable/Disable	OFF	ON	OFF	ON	OFF	ON
Multiband Dynamic Range Control(MBDRC)	Enable/Disable	OFF	OFF	Deactive	Deactive	Deactive	Deactive
Speaker Gain(SPKR_GAIN)	Changing dB values	OFF-0dB	OFF-0dB	OFF-0dB	OFF-0dB	OFF-0dB	OFF-0dB

The measurements of audio quality are performed on 2G, 3G and 4.5G network bands after the second improvement trial. The results of the second improvement trial are presented in Table 5.30 for 2G network, Table 5.31 for 3G network and Table 5.32 for 4.5G network.

Table 5.30 2G uplink audio quality test result after 2nd improvement trial

Mobile Network	Signal Direction	Speech Coding System	Codec Bit Rate (kbps)	Measured MOS Value	Target Value (MOS)	Result
2G	Uplink	Full Rate	12.2	3.93	>3.5	Pass
2G	Uplink	Full Rate	10.2	3.86	>3.5	Pass
2G	Uplink	Full Rate	7,.95	3.79	>3.5	Pass
2G	Uplink	Full Rate	7.4	3.75	>3.5	Pass
2G	Uplink	Full Rate	6.7	3.72	>3.5	Pass
2G	Uplink	Full Rate	5.9	3.67	>3.5	Pass
2G	Uplink	Full Rate	5.15	3.60	>3.5	Pass
2G	Uplink	Full Rate	4.75	**3.34**	>3.5	Fail
2G	Uplink	Half Rate	7.95	3.69	>3.5	Pass
2G	Uplink	Half Rate	7.4	3.63	>3.5	Pass
2G	Uplink	Half Rate	6.7	3.58	>3.5	Pass
2G	Uplink	Half Rate	5.9	**3.11**	>3.5	Fail
2G	Uplink	Half Rate	5.15	**3.01**	>3.5	Fail
2G	Uplink	Half Rate	4.75	**3.36**	>3.5	Fail

Table 5.31 3G uplink audio quality test result after 2nd improvement trial

Mobile Network	Signal Direction	Codec Bit Rate(kbps)	Measured MOS Value	Target Value (MOS)	Result
3G	Uplink	12.2	3.99	>3.5	Pass
3G	Uplink	10.2	3.91	>3.5	Pass
3G	Uplink	7.95	3.75	>3.5	Pass
3G	Uplink	7.4	3.62	>3.5	Pass
3G	Uplink	6.7	3.59	>3.5	Pass
3G	Uplink	5.9	3,57	>3.5	Pass
3G	Uplink	5.15	**3.37**	>3.5	Fail
3G	Uplink	4.75	**3.29**	>3.5	Fail
3G	Uplink	23.85	4.01	>3.5	Pass
3G	Uplink	23.05	4.04	>3.5	Pass
3G	Uplink	19.85	3.94	>3.5	Pass
3G	Uplink	18.25	3.91	>3.5	Pass
3G	Uplink	15.85	3.87	>3.5	Pass
3G	Uplink	14.25	3.72	>3.5	Pass
3G	Uplink	12.65	3.67	>3.5	Pass
3G	Uplink	8.85	**3.39**	>3.5	Fail
3G	Uplink	6.6	**3.26**	>3.5	Fail

Table 5.32 4.5G uplink audio quality test result after 2nd improvement trial

Mobile Network	Signal Direction	Codec Bit Rate(kbps)	Measured MOS Value	Target Value (MOS)	Result
4.5G	Uplink	23.85	4.18	>3.5	Pass
4.5G	Uplink	23.05	4.11	>3.5	Pass
4.5G	Uplink	19.85	4.05	>3.5	Pass
4.5G	Uplink	18.25	4.02	>3.5	Pass
4.5G	Uplink	15.85	3.95	>3.5	Pass
4.5G	Uplink	14.25	3.85	>3.5	Pass
4.5G	Uplink	12.65	3.82	>3.5	Pass
4.5G	Uplink	8.85	3.77	>3.5	Pass
4.5G	Uplink	6.6	3.72	>3.5	Pass
4.5G	Uplink	12.2	3.63	>3.5	Pass
4.5G	Uplink	10.2	3.59	>3.5	Pass
4.5G	Uplink	7.95	3.55	>3.5	Pass
4.5G	Uplink	7.4	**3.38**	>3.5	Fail
4.5G	Uplink	6.7	**3.31**	>3.5	Fail
4.5G	Uplink	5.9	**3.19**	>3.5	Fail
4.5G	Uplink	5.15	**3.13**	>3.5	Fail
4.5G	Uplink	4.75	**3.01**	>3.5	Fail

The general evaluation of the second improvement trial at the uplink signal direction in the 2G, 3G and 4.5G network bands are shown in Figure 5.21. The measurement results with the current audio paremeters show that the sound quality of uplink is not satisfied at 12 of fourteen audio codec bit rate variables, 9 and 10 of seventeen audio codec bit rate variables respectively for 2G, 3G and 4.5G mobile network. After the second improvement trial, the sound quality of uplink is not satisfied at 4 of fourteen audio codec bit rate variables, 4 and 5 of seventeen audio codec bit rate variables respectively for 2G, 3G and 4.5G mobile network.

If the audio quality of 2G, 3G and 4.5G network is generally respectively evaluated after the second improvement trial, the sound quality of uplink does not satisfied at 28.57%, 23.52% and 29.41% of the total audio codec variables according to the MOS scale. A third improvement trial should be attempted.

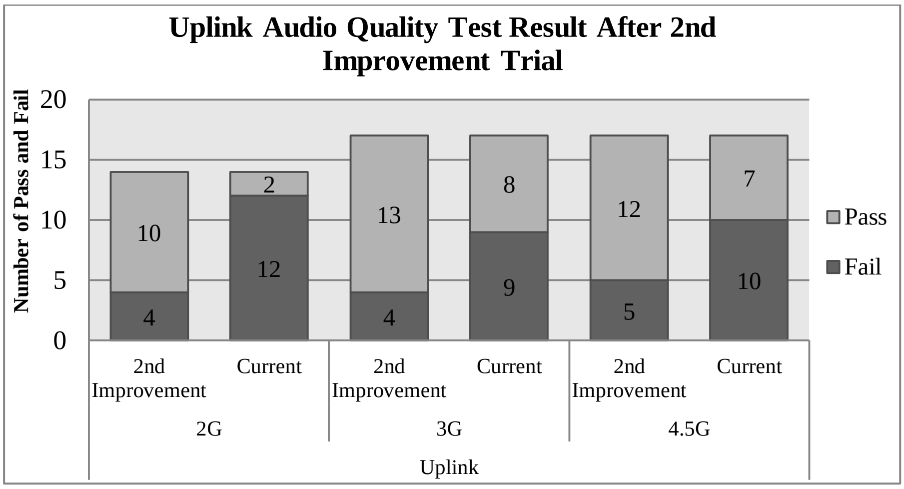

Figure 5.21 Uplink audio quality test result after 2nd improvement trial

As the first and second improvement trials are not sufficient, a third improvement trial is made. In the third improvement trial, 8 parameters and filters marked with yellow in Table 5.33 are changed. The changed parameters are listed below.

- Receive Gain (RX_GAIN) is enabled and 3dB gain is achieved.
- Wide Volume is active for WCDMA and LTE bands.
- Adaptive Input Gain (AIG) is activated for GSM band.
- Volume (VOL) is increased by 3dB.
- Echo Canceles Receive (ECRX) is activated.
- Decoder Gain (DEC_GAIN) is activated and 2dB gain is achieved.
- High Pass Filter (HPF) is activated.
- Speaker Gain (SPKR_GAIN) is activated to gain 2dB.

Table 5.33 3rd Improvement Trial vs Current ACDB Parameters for Uplink

Rx Path Parameters	Variables	Current Parameters	3rd Improvement Trial	Current Parameters	3rd Improvement Trial	Current Parameters	3rd Improvement Trial
		2G	2G	3G	3G	4,5G	4,5G
Receive Gain(RX_GAIN)	Changing dB values	OFF-0dB	ON-3dB	OFF-0dB	ON-3dB	OFF-0dB	ON-3dB
Far-end Noise Suppression(FNS)	Enable/Disable	ON	ON	Deactive	Deactive	Deactive	Deactive
Wide Volume V2(WV_2)	Enable/Disable	Deactive	Deactive	OFF	OFF	OFF	OFF
WideVolume(WV)	Enable/Disable	Deactive	Deactive	OFF	ON	OFF	ON
Adaptive Input Gain(AIG)	Enable/Disable	OFF	ON	Deactive	Deactive	Deactive	Deactive
Dynamic Range Control(DRC)	Enable/Disable	ON	ON	Deactive	Deactive	Deactive	Deactive
Finite Impulse Response(FIR)	Enable/Disable	OFF	OFF	Deactive	Deactive	Deactive	Deactive
Infinite Impulse Response(IIR)	Enable/Disable	ON	ON	Deactive	Deactive	Deactive	Deactive
Volume(VOL)	Changing dB values	0 dB	3 dB	0 dB	3 dB	0 dB	3 dB
Automatic Voice Control/Receive Voice Enhancement(AVC_RVE)	Enable/Disable	OFF	OFF	OFF	OFF	OFF	OFF
Echo Cancellar Receive(ECRX)	Enable/Disable	OFF	ON	OFF	ON	OFF	ON
Decoder Gain(DEC_GAIN)	Changing dB values	OFF-0dB	ON-2dB	OFF-0dB	ON-2dB	OFF-0dB	ON-2dB
Psychoacoustic Bass Enhancement(PBE)	Enable/Disable	OFF	OFF	OFF	OFF	OFF	OFF
High Pass Filter(HPF)	Enable/Disable	OFF	ON	OFF	ON	OFF	ON
Multiband Dynamic Range Control(MBDRC)	Enable/Disable	OFF	OFF	Deactive	Deactive	Deactive	Deactive
Speaker Gain(SPKR_GAIN)	Changing dB values	OFF-0dB	ON-2dB	OFF-0dB	ON-2dB	OFF-0dB	ON-2dB

The measurements of audio quality are performed on 2G, 3G and 4.5G network bands after the third improvement trial. The results of the third improvement trial are presented in Table 5.34 for 2G network, Table 5.35 for 3G network and Table 5.36 for 4.5G network.

Table 5.34 2G uplink audio quality test result after 3rd improvement trial

Mobile Network	Signal Direction	Speech Coding System	Codec Bit Rate(kbps)	Measured MOS Value	Target Value (MOS)	Result
2G	Uplink	Full Rate	12.2	4.25	>3.5	Pass
2G	Uplink	Full Rate	10.2	4.12	>3.5	Pass
2G	Uplink	Full Rate	7.95	4.11	>3.5	Pass
2G	Uplink	Full Rate	7.4	4.03	>3.5	Pass
2G	Uplink	Full Rate	6.7	3.99	>3.5	Pass
2G	Uplink	Full Rate	5.9	3.86	>3.5	Pass
2G	Uplink	Full Rate	5.15	3.77	>3.5	Pass
2G	Uplink	Full Rate	4.75	3.66	>3.5	Pass
2G	Uplink	Half Rate	7.95	3.91	>3.5	Pass
2G	Uplink	Half Rate	7.4	3.88	>3.5	Pass
2G	Uplink	Half Rate	6.7	3.81	>3.5	Pass
2G	Uplink	Half Rate	5.9	3.72	>3.5	Pass
2G	Uplink	Half Rate	5.15	3.61	>3.5	Pass
2G	Uplink	Half Rate	4.75	**3.42**	>3.5	Fail

Table 5.35 3G uplink audio quality test result after 3rd improvement trial

Mobile Network	Signal Direction	Codec Bit Rate(kbps)	Measured MOS Value	Target Value (MOS)	Result
3G	Uplink	12.2	4.11	>3.5	Pass
3G	Uplink	10.2	4.02	>3.5	Pass
3G	Uplink	7.95	3.96	>3.5	Pass
3G	Uplink	7.4	3.81	>3.5	Pass
3G	Uplink	6.7	3.76	>3.5	Pass
3G	Uplink	5.9	3.66	>3.5	Pass
3G	Uplink	5.15	3.61	>3.5	Pass
3G	Uplink	4.75	**3.41**	>3.5	Fail
3G	Uplink	23.85	4.26	>3.5	Pass
3G	Uplink	23.05	4.17	>3.5	Pass
3G	Uplink	19.85	4.07	>3.5	Pass
3G	Uplink	18.25	4.03	>3.5	Pass
3G	Uplink	15.85	3.96	>3.5	Pass
3G	Uplink	14.25	3.79	>3.5	Pass
3G	Uplink	12.65	3.78	>3.5	Pass
3G	Uplink	8.85	3.64	>3.5	Pass
3G	Uplink	6.6	3.59	>3.5	Pass

Table 5.36 4.5G uplink audio quality test result after 3rd improvement trial

Mobile Network	Signal Direction	Codec Bit Rate(kbps)	Measured MOS Value	Target Value (MOS)	Result
4.5G	Uplink	23.85	4.4	>3.5	Pass
4.5G	Uplink	23.05	4.33	>3.5	Pass
4.5G	Uplink	19.85	4.29	>3.5	Pass
4.5G	Uplink	18.25	4.21	>3.5	Pass
4.5G	Uplink	15.85	4.14	>3.5	Pass
4.5G	Uplink	14.25	4.04	>3.5	Pass
4.5G	Uplink	12.65	3.98	>3.5	Pass
4.5G	Uplink	8.85	3.92	>3.5	Pass
4.5G	Uplink	6.6	3.83	>3.5	Pass
4.5G	Uplink	12.2	3.84	>3.5	Pass
4.5G	Uplink	10.2	3.77	>3.5	Pass
4.5G	Uplink	7.95	3.7	>3.5	Pass
4.5G	Uplink	7.4	3.65	>3.5	Pass
4.5G	Uplink	6.7	3.59	>3.5	Pass
4.5G	Uplink	5.9	3.55	>3.5	Pass
4.5G	Uplink	5.15	**3.21**	>3.5	Fail
4.5G	Uplink	4.75	**3.1**	>3.5	Fail

The general evaluation of the third improvement trial at the uplink signal direction in the 2G, 3G and 4.5G network bands are shown in Figure 5.22. The measurement results with the current audio paremeters show that the sound quality of uplink is not satisfied at 12 of fourteen audio codec bit rate variables, 9 and 10 of seventeen audio codec bit rate variables respectively for 2G, 3G and 4.5G mobile network. After the third improvement trial, the sound quality of uplink is not satisfied at 1 of fourteen audio codec bit rate variables, 1 and 2 of seventeen audio codec bit rate variables respectively for 2G, 3G and 4.5G mobile network.

If the audio quality of 2G, 3G and 4.5G network is generally respectively evaluated after the third improvement trial, the sound quality of uplink is satisfied at 92.85%, 94.11% and 88.23% of the total audio codec variables by most users.

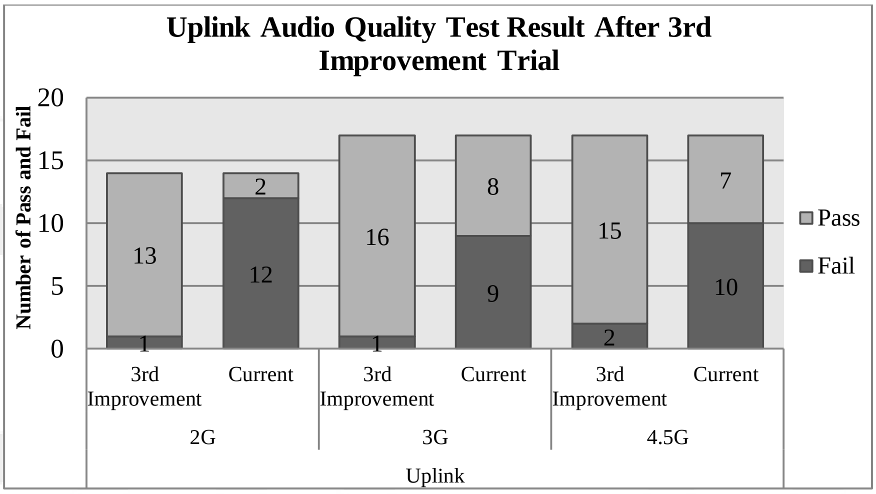

Figure 5.22 Uplink audio quality test result after 3rd improvement trial

5.2.2 Improvement Studies According to Signal Strength Variables

Improvement studies have been completed according to audio codec bit rate and they have been detailed in Section 5.2.1. In the measurement, the third improvement trial that is used at Section 5.2.1, has best result about audio quality for the downlink and uplink signal direction, so it is also used for signal strength variable measurement. At this improvement trial, the changed parameters and filters are shown in Table 5.21 for the downlink signal direction and Table 5.33 for the uplink signal direction. The audio calibration database of current and the 3rd improvement trial are used to test at the signal strength of poor, fair and good on 2G, 3G and 4.5G network.

The general evaluation of the results that include current and improvement trial is shown in Figure 5.23 for downlink signal direction and Figure 5.24 for the uplink signal direction.

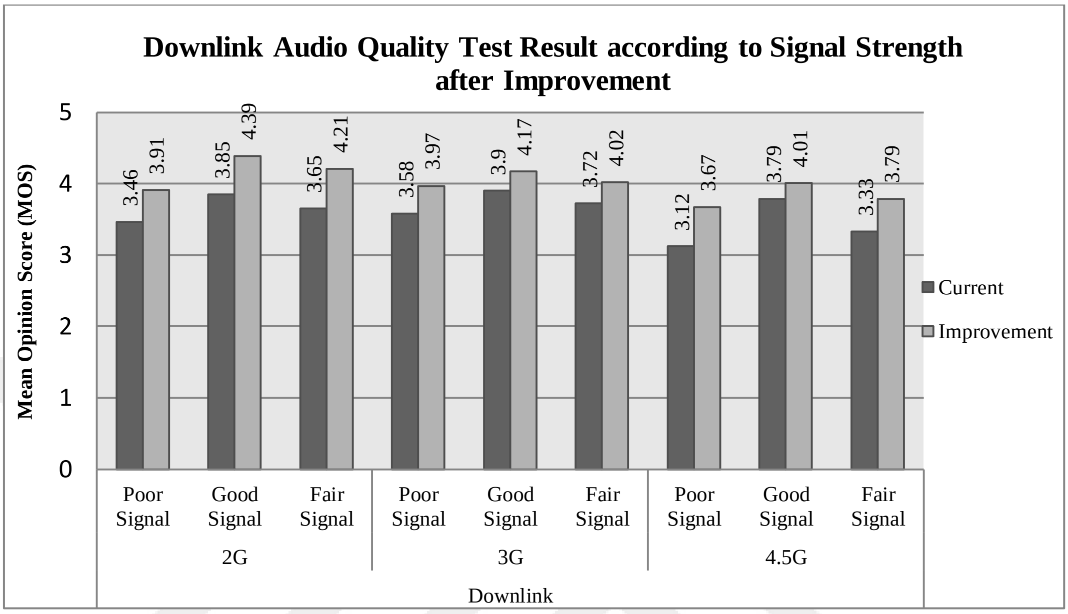

Figure 5.23 Downlink audio quality test result according to signal strength after improvement

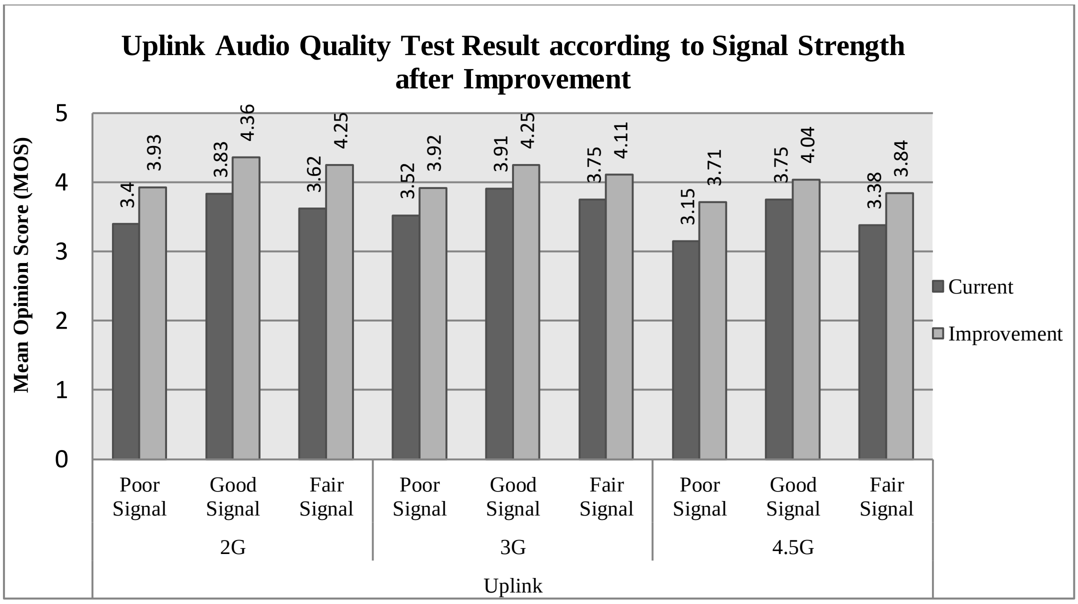

Figure 5.24 Uplink audio quality test result according to signal strength after improvement

CHAPTER SIX

CONCLUSION

The main purpose of this thesis is analyzing the sound quality of the mobile-based devices with Android operating system by measuring according to ITU-T Recommendation P.863 and improving the sound quality according to mean opinion score scale that is briefly called MOS scale. The system which is introduced in this thesis is used to measure and improve the sound quality of the mobile-based device. The measurements were made according to audio codec bit rate and signal strength at the audio bandwidths which are narrowband, wideband and super wideband. In addition to this, the sound quality measurements of downlink and uplink are needed to be handled separately. The audio related parameters of the mobile-based devices are kept in the audio calibration database that is called acdb.

There are fourteen, seventeen and seventeen audio codec bit rate variables for both of downlink and uplink signal direction respectively at 2G, 3G and 4.5G mobile network. In the beginning, the measurement results with the current audio paremeters show that the sound quality of downlink does not satisfied at 10 of fourteen audio codec bit rate variables, 10 and 9 of seventeen audio codec bit rate variables; the sound quality of uplink does not satisfied at 12 of fourteen audio codec bit rate variables, 9 and 10 of seventeen audio codec bit rate variables respectively at 2G, 3G and 4.5G mobile network.

An audio calibration tool of the mobile based device chipset manufacturer is used for the sound quality improvement trials. In 2G, 3G and 4.5G mobile network, respectively four, five and six parameters are changed using the audio calibration tool in order to improve the sound quality of downlink signal direction; respectively four, six and eight parameters are changed using the audio calibration tool in order to improve the sound quality of downlink signal direction.

The sound quality of downlink is not satisfied at two, one and two variables; the sound quality of uplink is not satisfied at one, one and two variables respectively for 2G, 3G and 4.5G mobile network after improvement studies. If the results of current and improvement audio paremeters are comparised, 80%, 90% and 77.77% improvement are achieved for the sound quality of downlink; 91.66%, 88.88% and 80% improvement are achieved for the sound quality of uplink respectively at 2G, 3G and 4.5G mobile network.

Future work: In this thesis, the recommendation P.863 is used to measure the sound quality of transmitted or received sounds objectively at narrowband, wideband and super-wideband. Audio bandwidth in broadcast and communication systems have four different bandwidths that are including narrowband, wideband, super-wideband and fullband. Network providers and mobile-based device manufacturers have now begin to support fullband audio transmission, but there is no objective standard to assess the sound quality of the sound transmitted or received through the fullband. In future work, the improvement studies of audio quality can be made by evaluating the sound quality of the sound transmitted or received through the full-band according to the subjective measurement standards. In addition to this, if a standard is recommended to measure the sound quality of the transmitted sound at the full band in later times, the sound quality can be measured objectively at full band audio transmission and the improvement studies can be made for better sound quality.

REFERENCES

Analog Devices (2012). *Isolated precision half-bridge driver.* Retrieved November 23, 2014, from http://www.analog.com/media/en/technical-documentation/data-sheets/ADuM3223{_}4223.pdf.

Assem, H. (2013). *Assessing and improving the VVoIP call quality.* M.Sc. Thesis, National University of Ireland Maynooth, Republic of Ireland.

Breerends, J. G., Rix, A. W., Hollier, M. P., & Hekstra, A. P. (2001). Perceptual evaluation of speech quality (PESQ) - a new method for speech quality assessment of telephone networks and codecs. *IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings, 2,* 2-5.

Breerends, J. G., Schmidmer, C., Berger, J., Obermann, M., Ullmann, R., Pomy, J., & Keyhl, M. (2013). Perceptual objective listening quality assessment (POLQA), The third generation ITU-T standard for end-to-end speech quality measurement part I – Temporal alignment. *Journal of the Audio Engineering Society, 61,* 1-6.

Benesty, J., Sondhi, M. M., & Huang, Y. (2008). *Springer handbook of speech processing.* Heidelberg: Springer Science & Business Media.

Berger, J. (2011). *An introduction to measuring quality and MOS testing* [White Paper]. Germany.

Bigelow, S. J., Carr, J. J., & Winder, S. (2001). *Understanding telephone electronics.* United States of America: Buttenvorth-Heinemann.

Chen, H. H., Guizani M., & Mohr, W. (2007). Evolution toward 4G wireless networking. *IEEE Network,* 4-5.

Cotanis, I. (2010). *Voice service quality evaluation techniques and the new technology POLQA* [White Paper]. Retrieved November 3, 2018, from https://pdfs.semanticscholar.org/24d8/c05d0d329ae2c41439afcdfd56ba225697aa.pdf

Côté, N. (2011). *Integral and diagnostic intrusive prediction of speech quality.* Heidelberg: Springer Science & Business Media.

Ebem, D. U. (2009). *Validating perceptual objective listening quality assessment methods on the tonal language Igbo.* M.Sc. Thesis, University of Nigeria, Nigeria.

Gawas, A. U. (2015). An overview on evolution of mobile wireless communication networks: 1G-6G. *International Journal on Recent and Innovation Trends in Computing and Communication, 3,* 3130-3133.

Gambhir, N. M. (2009). *Objective measurement of speech quality in VoIP over wireless LAN during handoff.* M.Sc. Thesis, San Jose State University, California.

Gerlach, O. (2012). *Next - generation (3G/4G) voice quality testing with POLQA* [White Paper]. Retrieved March 23, 2018, from https://cdn.rohde-schwarz.com/pws/dl_downloads/dl_application/application_notes/1ma202/1MA202_1e_3G4G_voice_quality_testing_POLQA.pdf

Ghimire, S. (2012). *Speech intelligibility measurement on the basis of ITU-T recommendation P.863.* M.Sc. Thesis, Halmstad University, Sweden.

Goudarzi, M., Sun, L., & Ifeachor, E. (2009). *PESQ and 3SQM measurement of voice quality over live 3G networks.* Retrieved January 13, 2019, from https://pdfs.semanticscholar.org/e1ee/348f48a4ab3294eff53c2401cfa921dfaf34.pdf

Goudarzi, M. (2008). *Evaluation of voice quality in 3G mobile networks.* M.Sc. Thesis, University of Plymouth, England.

ITU-T Recommendation P.830 (1996). *Subjective performance assessment of telephone - band and wideband digital codecs.* Geneva.

ITU-T Recommendation P.800 (1996). *Methods for subjective determination of transmission quality.* Geneva.

ITU-T Recommendation P.800.1 (2003). *Mean opinion score (MOS) terminology.* Geneva.

ITU-T Recommendation P.805 (2007). *Subjective evaluation of conversational quality.* Geneva.

ITU-T Recommendation P.835 (2003). *Subjective test methodology for evaluating speech communication systems that include noise suppression algorithm.* Geneva.

ITU-T Recommendation P.861 (1996). *Objective quality measurement of telephone band (300 - 3400 Hz) speech codecs.* Geneva.

ITU-T Recommendation P.862 (2001). *Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs.* Geneva.

ITU-T Recommendation P.863 (2011). *Perceptual objective listening quality assessment.* Geneva.

ITU-T Recommendation P.862.2 (2007). *Wideband extension to recommendation P.862 for the assessment of wideband telephone networks and speech codecs.* Geneva.

ITU-T Recommendation G.107 (2000). *The E-model: a computational model for use in transmission planning.* Geneva.

ITU-T Recommendation P.563 (2004). *Single-ended method for objective speech quality assessment in narrow-band telephony applications.* Geneva.

Khurram, M. (2002). The perceptual analysis measurement system for end-to-end network speech quality. *IEEE International Multi Topic Conference,* 88-94.

Kondo, K. (2012). *Subjective quality measurement of speech - Its evaluation, Estimation and Applications.* Heidelberg: Springer Science & Business Media.

Kráčala, M. (2016). *Měření hlasové kvality technologie VoLTE/VoWiFi.* M.Sc. Thesis, Brno University of Technology, Czech republic.

Kumar, A., Liu, Y., & Sengupta, J. (2010). Evolution of mobile wireless communication networks: 1G to 4G. *International Journal of Electronics & Communication Technology, 1,* 68-72.

Lopa, M., & Vora, J. (2015). Evaluation of mobile generation technology: 1G to 5G and review of upcoming wireless technology 5G. *International Journal of Modern Trends in Engineering and Research (IJMTER), 2,* 281-290.

Liu, W. M., Mason, J. S. D, Evans, N. W. D., & Jellyman, K. A. (2006). An assessment of automatic speech recognition as speech intelligibility estimation in the context of additive noise. *The Ninth International Conference on Spoken Language Processing,* 2166-2169.

Mehta, H., Patel, D., Joshi, B., & Modi, H. (2014). 0G to 5G mobile technology: a survey. *Journal of Basic and Applied Engineering Research, 1,* 56-60.

Mishra, A. K. (2004). *Fundamentals of cellular network planning and optimization, 2G/2.5G/3G evolution of 4G.* USA: John Wiley & Sons Inc.

Mobile signal strength recommendations. (2018, November 20). In Teltonika. Retrieved February 20, 2019, from https://wiki.teltonika.lt/index.php= Mobile_Signal_Strength_Recommendations

Mohamed, S. A. (2003). E*valuation automatique de la qualité des flux multimédias en temps réel: une approche par réseaux de neurones.* Ph.D. Thesis, Versailles Saint-Quentin-en-Yvelines University, France.

Möller, S. (2000). *Assessment and prediction of speech quality in telecommunications.* United States: Springer.

Opticom GmbH (2000). *State of the art voice quality testing* [White Paper]. Germany.

Opticom GmbH (2004). *3SQM™ advanced non-intrusive voice quality testing* [White Paper]. Germany.

Picovici, D., & Nelson, J. (2009). *Handbook of research on wireless multimedia:*

quality of service and solutions. United States of America: IGI Global.

Pourmand, N. (2013). *Objective and subjective evaluation of wideband speech quality.* Ph.D. Thesis, The University of Western Ontario, London.

Pulkki, V., & Karjalainen, M. (2015). *Communication acoustics: an introduction to speech, audio and psychoacoustics.* USA: John Wiley & Sons Inc.

Quackenbush, S. R., Barnwell, T.P., & Clement, M. A. (1988). *Objective measures Of speech quality.* United States of America: Prentice Hall.

Rango, D. F., Tropea, M., Fazio, P., & Marano, S. (2006). Overview on VoIP: subjective and objective measurement methods. *International Journal of Computer Science and Network Security, 6*, 140-153.

Recommendation ITU-R BS.1534-1 (2003). *Method for the subjective assessment of intermediate quality level of coding systems.* Geneva.

Rix, A. W., Beerends, J. G., Kim, D. S., Kroon, P., & Ghitza, O. (2006). Objective assessment of speech and audio quality - technology and applications. *IEEE Transactions on Audio, Speech, and Language Processing, 14* (6) , 1890-1901.

Rix, A. W., & Hollier, M. P. (2000). The perceptual analysis measurement system for robust end-to-end speech quality assessment. *IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings, 3,* 1515–1518.

Rohde & Schwarz GmbH & Co (2015). *UPV audio analyzer operating manual.* Germany.

Rohde & Schwarz GmbH & Co (2017). *CMW500 wideband radio communication tester user manual.* Germany.

Seo, D. (2013). *Evolution and standardization of mobile communications technology.* United States of America: IGI Global.

Sharma, P. (2013). Evolution of mobile wireless communication networks - 1G to 5G as well as future prospective of next generation communication network. *International Journal of Computer Science and Mobile Computing, 2,* 47-53.

Slavata, O. (2017). *Impact of IP chanel parameters on the final quality of the transferred voice.* Ph.D. Thesis, Czech Technical University, Prague.

Sun, L. (2004). *Speech quality prediction for voive over internet protocol networks.* Ph.D. Thesis, University of Plymouth, England.

Tüysüz, M. F. (2007). *Improving voice quality in VOIP.* M.Sc. Thesis, Dokuz Eylül University, İzmir.

Vaighan, F. B. (2011). *VOIP voice quality measurement by network traffic analysis.* M.Sc. Thesis, Dokuz Eylül University, İzmir.

Veen, F. V. (1969 IEEE recommended practice for speech quality measurements. *IEEE Audio and Electroacoustics Group*, *12*, 225-246.

Voran, S. (1999). Objective estimation of perceived speech quality. *IEEE Transactions on Speech and Audio Processing,* 371 - 382.

Voran, S. (1997). Estimation of perceived speech quality using measuring normalizing blocks. *IEEE Speech Coding Workshop,* 83-84.