QUALITY OF EXPERIENCE FOR VOICE OVER INTERNET PROTOCOL (VoIP)

Asmaa Ali Jaish 1, Basim K. J. Al-Shammarri 1

Abstract
Today, the service of Voice over Internet Protocol (VoIP) is one of the most widely used around the globe in many fields, especially with video conferencing applications. Communication stakeholders aim to keep old customers happy and attract new customers. There was a need for some kind of service quality measures. Metrics are used to quantify the quality of transmission link which are known as Quality of Service (QoS). Whereas; user satisfaction level is identified by using Quality of Experience QoE. This paper aims to compare the user satisfaction level of the VoIP service by using different coding schemes at the application level of the User Equipment (UE). The measured QoS parameters over the heterogeneous transmission link were: IP packet delay, IP Packet jitter, IP packet loss, MAC layer delay, and PHY layer throughput in the proposed network. Additionally, simulation model uses the E-Model to assess QoE level using the Mean Opinion Score (MOS) metric in predicting VoIP call quality. The results during 30 seconds of the voice call simulation time were as follow: Delay was 0.57 mSec in G.729A, and 0.25 mSec in G.711. Throughput was 290 Kbits/sec in G.729A, and 70 Kbits/sec in G.711. Jitter was 0.02 uSec G.729A, and 0 Sec in G.711. MOS value was 3.76 G.729A, and 3.08 in G.711.

Keywords: Voice over Internet Protocol (VoIP), Quality of Service (QoS), Quality of Experience (QoE), Mean opinion score (MOS), R-factor, Delay, Throughput, OPNET.

1. INTRODUCTION
Voice over Internet Protocol is a service that many users of packet switching communication networks employ for personal, business, or other purposes. VoIP services with varying features and costs are offered globally by several telecom firms. VoIP is advantageous to individuals as well as organizations. VoIP benefits service providers by opening up a new, lucrative revenue stream for them. VoIP offers consumers a glimpse into the future of communication by enabling low-cost, faster, more reliable, and secure connections for conversations across long distances or on tiny networks like local area networks or wide area networks [1]. In order to determine the effectiveness of packet switching communication network services, Quality of Service (QoS) criteria are deemed insufficient [2]. Voice awareness and human perception are two significant aspects that were overlooked in the QoS analysis. In order to gauge how the End User (EU) perceives the used communication
services from the EU point of view, the International Telecommunication Union (ITU) has created a measure known as Quality of Experience (QoE). To estimate, monitor, and evaluate the quality level of the service in various scenarios and across a wide range of network types and topologies that employ wired or wireless technologies, VoIP service quality evaluation methodologies and metrics are necessary [3]. Both service providers and clients will utilize the results of an evaluation to determine what sort of services they are utilizing and what advantages they are receiving. Customers should be aware of the areas where the service is lacking so that a quick fix may be found to improve the quality of the offered service.

With billions of people having instant access to the internet, the usage of networks has grown and evolved into a need of daily life, particularly the internet. With millions of customers using packet-switching communication networks to make voice calls, VoIP has grown to be a significant service. The majority of experts or consumers seek to communicate the level of quality they encounter in a service that enables comparisons between various algorithms, services, technologies, and protocols. The opinion score is a novel statistic developed by international organizations like ITU. It is described as the "value on a predefined scale that a subject assign to his opinion of the performance of a system" [4].

The main work in this article is to construct a heterogeneous communication link and enable VoIP voice calls using the OPNET modeler simulator. Additionally, in every modeler simulation run, this VoIP service attempts to employ two different vocoder variations. Furthermore, the QoE of these VoIP conversations is determined using MOS. What’s more, network QoS metrics are monitored in the above-mentioned simulated transmission link.

The rest of this paper is organized as follows, the related work is given in section Two. In addition, an overview of the VoIP idea and its definitions was presented in section Three. Whereas, in section four, the problems that presented as challenges and limitations in the basis of this work were presented. Section Five the research methodologies were discussed presented by the QoE measure’s metric, which is MOS. In section six, the proposal model was presented, including the formulation of the model and the development of the solution algorithm. Additionally, results and discussion are presented in section Seven. Finally, conclusions are presented in the conclusion section.

2. Related work

This part of paper summarizes the work done by various researchers. In this context the article [5] presents the study's conclusions about the quality of voice data transmission in IP packet networks. It looks into the mechanisms that make it possible to assess the data transmission quality in packet telephony. Potential transmission quality levels, suitable quality metrics, and limit values enforcing acceptable voice data transmission quality were discussed and made clear in order to be included in standardization bodies' recommendations.

A packet switching network model was developed and tested to take VoIP architecture supporting various audio codecs used for voice compression into consideration. Studies on transmission systems using the audio codecs G.711, G.723, G.726, G.728 and G.729 have been done. It was shown that for delay-sensitive communication that fluctuates above and beyond its nominal rate, some codecs are more beneficial than others, allowing VoIP to be sent with improved quality.

Researchers in Scientific Research [6] presented a survey of the gradual development and use of reliability techniques to obtain high voice quality in a VoIP network. They stressed that it is necessary for the data packet to reach the destination from its source with a high level of reliability. It is very important to analyse link failure, packet loss, delay and jitter during data communication. The components involved in communicating data must be reliable. An innovative decentralized identity identification method for VoIP networks is suggested in the article [7]. The suggested method uses the fault tolerance, transparency, and immutability inherent in the blockchain platform to safeguard user privacy and enable secure communication in VoIP apps. The experiment's findings demonstrate that the suggested scheme is an efficient and cost-effective method of managing communication operations when employed as a decentralized identity authentication system in VoIP. Additionally, end-to-end secure calling achieves an average time delay of between 30% and 70% less than current blockchain-based authentication techniques. The suggested party authentication mechanism is ten times quicker than the TLS process. Moreover, the proposed system has reasonable latency for the communication process when compared to simple, fast and less secure alternatives via SIP authentication. The researchers in the work [8] provided a thorough investigation of the impacts of QoS on QoE using experimental methods and correlation testing approach rather than examining each parameter separately. To integrate the impacts of QoS parameters with temporal or geographical features on QoE, a method is put forward. To forecast non-intrusive speech quality in various network contexts, machine learning regression techniques with QoS impairment, noise impairment, and echo impairment are incorporated. Several experiments' outcomes demonstrate that these models may produce predictions that are reasonably accurate. The work demonstrated how IP network environments, noise attenuation, and echo affect the quality and reliability of VoIP traffic and provided QoS parameter requirements for a VoIP application operating at the desired environment quality level. It also improved the accuracy of the assessment of
VoIP QoE using QoS parameters. An investigation for jitter and packet loss for VoIP traffic is performed in [9] and a modelling for packet loss is presented, one of the QoS criteria. in an article [10] The packet-loss mechanisms in User Datagram Protocol (UDP) connections are modeled using a Gilbert-gamma topology. State duration modeling using gamma distributions is introduced by the suggested structure. The suggested topology significantly reduces the subjective assessment of voice quality conveyed over IP networks by up to 70% and increases the chance of observable packet-loss processes when compared to the standard Gilbert model. It is simple to apply the results to other real-time applications, such streaming audio and video. While the researchers in [11] suggested a solution to the packet transmission delay and proposed an E-model for the Thai context. They defended the correctness and dependability of the E-model. The findings were offered as the most recent support for language detection in VoIP systems by the authors. However, only Thai voice samples were taken into account, and this bias aspect needs more research. Furthermore, the researcher study in [12] presented an assessment of a company’s trustworthiness VoIP. The developed method identified two VoIP design elements: matrices for measuring enterprise VoIP services and VoIP design standards. Researchers examined VoIP’s QoS performance on converged MPLS networks in the study, that were presented in reference [13]. VoIP cannot be transmitted in the network architecture unless traffic is gradually monitored and managed due to the extreme dynamic nature of IP systems. The Internet Service Provider (ISP) network infrastructure, which consists of converged networks carrying many VoIP instances, was studied in order to give novel configuration management. This is accomplished by guaranteeing a guaranteed QoS limit at the acceptance level under unfavourable circumstances by employing path reflectors and path redistribution at the micro level.

3. VoIP application and service

VoIP technology was a breakthrough in the internet world because it enables users to make sound calls over IP networks and makes it simple to send data packets over a Public Switched Telephone Network (PSTN) until they reach their destination using the best path offered by the PSTN. VoIP and its features to the extent necessary to comprehend how VoIP functions and how to enhance its functionality by creating two distinct simulation models using modeler. Any VoIP service provider will be concerned about both QoS and QoE impacts [14].

VoIP is a real time service that uses a Real Time Protocol (RTP) combined with Real Time Control Protocol (RTCP) and UDP packet transmission.

3.1. Real Time Protocol (RTP)

A Real-time Transport Protocol is used to send audio and video data over IP networks. RTP is utilized in systems for communication and entertainment that include streaming media, including web-based push-to-talk functions, video conferencing apps, television services, and telephone services. One of the technological pillars of VoIP is RTP, which is frequently used in tandem with a signalling protocol like the Session Initiation Protocol (SIP) to create connections across the network. A network standard was developed for the purpose of sending audio or video data in a manner that is optimal for the reliable transmission of live data. Internet telephony, Voice over Internet Protocol (VoIP), and video telecommunication all make use of it. It is also possible to utilize it for one-to-many conferences as well as one-on-one calls (also known as unicast) (multicast) [15].

3.2. Real Time Control Protocol (RTCP)

The Real-Time Control Protocol collaborates with the RTP, in order to give feedback on the performance or quality of the RTP stream. In a VoIP network, RTCP is an optional companion protocol to RTP. It is employed to keep an eye on streaming media's QoS and transmission [16]. All participants in an RTP session get periodic RTCP packets, which offer feedback on the effectiveness of the data distribution being carried out via RTP. This function is connected to the flow control and congestion management functions of the network and is an essential component of RTP's position as a transport protocol. Although the feedback reports from RTCP only indicate that problems are occurring and not where they are, they can still be utilized as a tool to pinpoint the specific locations of problems. An administrator is able to determine potential weak spots in the operation of the network by analysing the information included in RTCP feedback reports. This data is generated by the various media gateways in the network.

3.3. User Datagram Protocol (UDP)

The User Datagram Protocol, sometimes known as UDP, is the simplest communication protocol that can be used over the Transport Layer of the TCP/IP protocol suite. It requires only a minimal amount of communication mechanism to be carried out. It is commonly believed that UDP is an unstable transport protocol; however, it makes use of IP services, which provide a best effort delivery method [1]. In UDP, the receiver is not required to
produce an acknowledgement of the packet that it has received, and the sender is not required to wait for any
acknowledgement of the packet that it has delivered.
The VoIP service have some weak points considering the security side, this service is weak against some kind of
attacks such as that aims to prevent the service from working by flooding the server with requests, this attack is
called Denial of Services (DOS) [17]. In addition, another kind of attacks that aims to steal the packets running in
the network from a third party, this attack called Man-in-the-middle attack [18]. This kind of attack will probably
lead to stopping the service or listening and stealing the user's private data. These all kinds of methods and
techniques are called the protocol suite of VOIP. VOIP service design benefits from the existence of some
infrastructure such as the Internet Protocol (IP), in order to use it as a transporting channel to transport the voices
from the source to the destination in the form of digital packets. So, any network in the world that uses the IP can
run the VoIP service such as the Internet, Intranets and Local Area Networks (LAN). To perform the service,
whether it needs to change the form of the recorded signals and then transmit the signals over the targeted network.
To do so the voice signal is digitized, compressed and converted to IP packets and then transmitted over the IP
network. The quality of the sessions and connections are estimated using a method from the RTCP. The VoIP
uses both of protocols from the network layer called the TCP and UPD [19].
The VoIP is categorized into multiple types depending on how using it and what the devices are running the service. VoIP Signalling DOS attack, VoIP Media DOS attack and Man-in-the-middle attack. These attacks may
cause the service to be stopped, and some of them are security issues that aim to steal user data or listen to user’s
private calls. Some work has been done to make VoIP operate as it should be for all users using different networks
and devices, so the need for standard protocol has emerged to a standard that is called the protocol suite of VoIP.
The H.323 standard protocol suite has a set of standards. For example, the standard protocol G.711 (64 kbps
channel) gives the minimum requirement needed for audio transfer over the IP. For the protocols G.723 (both of
the channels 5.3 and 6.3 kbps), G.728 standard (16 kbps channel) and the standard protocol G.722 (48, 56, and 64
kbps channels), G.729 (8 kbps channel) [20, 21].
The H245 control protocol is only employed during the connection-establishing phase of multimedia application
communication [22], when the devices' capabilities, audio encoding techniques, and appropriate media channels
are discussed and agreed upon. A technique from the RTCP is used to assess the quality of the sessions and
connections. Both TCP and UDP are used for data communications and VoIP control. Numerous forms of VoIP
exist, based on how it is used and what kind of equipment is providing the service.

4. CHALLENGES and LIMITATIONS

As is currently known, human recorded speech cannot be directly transferred over an IP network as packets are
transported in a digital format. As a result, an additional procedure is required to prepare the voice for transmission
over an IP network. The analogue signal is being sampled or digitalized throughout this procedure. So, a reliable
method for producing the desired signal for VoIP is required. One of these techniques is known as pulse code
modulation, or PCM. It is the most widely used technique in the sampling industry for converting analogue to
digital and vice versa. The Nyquist theorem, which states that sampling frequency should only be equal to the rate
of the highest frequency in the signal, must be followed while utilizing PCM [23].
The primary cause of VoIP's occasionally poorer voice quality than typical performance or VoIP's lower quality
among users can be attributed to technical factors including network speed, type, and device configuration.
Additionally, some networks that utilize improperly designed or mishandled security mechanisms, such firewalls,
PAT, and NAT, may have a severe impact on VoIP. Despite this, the VoIP business is still expanding, and service
providers are fiercely competing to offer the highest level of client pleasure.
VoIP service performance can be impacted by a variety of issues, but packet loss is the most significant or frequent
issue. Ignoring this issue might lead to too much difficulties for the network designer. One of the most challenging
and frequent issues that any network may experience is packet loss. The issue and its solutions are covered in a
number of research. The VoIP service is highly susceptible to packet loss inside the network of transfers. These
lost packets are caused by using bad or old network equipment, and lead to bad performance of the VoIP service
and any other service that affects running on the network. The more networks have collided packets, the more
delay will affect the VoIP. Also, packet loss affects the quality of voice calls [24]. So, building the communication
networks using the best available performance and design is needed. A study to design the network and simulate
it using the relevant programs has been done in the article, in order to test the network properly, the best network
devices that are now on the market, and equipment to construct a network that has as few packet collisions as
possible. Furthermore, the more packets that are transmitted at any one moment owing to an increase in network
user density might result in more lost packets. If one has a thorough understanding of the planned network, using
high-speed channels to tackle the issue will benefit a large number of people. Implementing a fire wall and access
list on the network to stop improper usage of its resources is another concept worth discussing.
5. MEAN OPINION SCORE (MOS)

The Mean Opinion Score (MOS), which has been around for over a decade in the current communications sector, provides a significant numerical value about the quality of the service. Using MOS, the service provider businesses may assess the general quality of voice calls. The MOS scale has values between 1 and 5. A service level of 1 denotes poor service, whereas a score of 5 denotes the highest possible degree of service. Since the ITU created the MOS, many businesses and organizations have used it. The MOS metric is frequently used in IP networks that employ VoIP services to assure the quality of speech transmission. It offers a statistic that may quantify VoIP service level deterioration and performance. VoIP has become a more popular service among UEs. To guarantee client satisfaction as a QoE level, MOS scoring is crucial [25].

Another key and important metric is the R-factor, which ranges from 1 to 100. This metric can be mapped from MOS and vice versa. It could use the R-factor metric for calculating the performance of VoIP services as well. The R-Factor value and MOS level provide a quick access about the QoE level of the VoIP service.

MOS can be found and calculated using the R-factor via the following formulas:

\[ R = 93 - \left( \frac{\text{effectiveLatency}}{\text{factor} \times \text{latency based}} \right) \]

\[ R = R - (\text{losspacket} \times \text{impact}) \]

\[ \text{MOS} = \left( (R - 60)(100 - R) \times 0.000007R \right) + 0.035R + 1 \]

Where R is the individual ratings.

6. Proposed Simulation Model

This section presents the simulation model that have been performed and discussed to obtain results from the simulation process. First of all, it demonstrates the used networks in the simulation work as presented in Figure 1, a heterogeneous transmission link between two devices UE1 and UE2, access point, switch, and voice call server.

Through this network, five parameters will be measured after introducing a voice call between UE1 and UE2 using (PCM G.729A codec in the first scenario and PCM G.711 codec in the second scenario). These parameters are IP Layer QoS parameters, they are IP packets (Jitter, Delay, Throughput, Traffic sent, and Traffic received).
6.1. **G711 Voice Codec:**

Narrowband codec G.711, sometimes referred to as PCM, provides uncompressed audio quality with extremely accurate voice reproduction. Although each call uses 64 kbps, G.711 may require a lot of bandwidth if call volumes are high. The ITU first released this codec in 1972, and is one of the oldest ones.

6.2. **G.729A Voice Codec**

G.729 uses only 8 kbps per call and provides strong compression and excellent bandwidth usage. It is a licensed codec, thus customers who purchase G.729-compatible devices inadvertently pay to use the codec. However, if the call is compressed and decompressed more than once, as can happen if a call is routed over various networks and is consequently transcoded more than once.

6.3. **The differences between G.711 and G.729A**

The main differences between G711 and G729A are presented in Table 1 below.

<table>
<thead>
<tr>
<th>Codec</th>
<th>Algorithm</th>
<th>Codec Delay (ms)</th>
<th>Bit Rate (Kbps)</th>
<th>Packets Per Second</th>
<th>IP Packet Size (bytes)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>G.711</td>
<td>pulse-code modulation (PCM)</td>
<td>0.375</td>
<td>64</td>
<td>100</td>
<td>120</td>
<td>Delivers precise speech transmission. Very low processor requirements</td>
</tr>
<tr>
<td>G.729A</td>
<td>Conjugate Structure – Algebraic Code Excited Linear Prediction (ACELP)</td>
<td>35</td>
<td>8</td>
<td>100</td>
<td>50</td>
<td>Excellent bandwidth utilization. Error tolerant</td>
</tr>
</tbody>
</table>

6.4. **Network Model setting**

6.4.1. **Application Configuration:**

1- PCM quality speech:
   - Encoder scheme: G.711
   - Voice frame per packet: 1
   - Type of service: Interactive Voice (6)
   - Coding rate (bits/sec): 64Kbps

2- IP Telephony:
   - Encoder scheme: G.729A
   - Voice frame per packet: 1
   - Type of service: Best efforts (0)
   - Coding rate: 8 Kbps

6.4.2. **Profile configuration:**

Start time offset (seconds): uniform (5,10)
Duration (seconds): End of profile
Simulation parameters
Start time: 0 sec
7. Results and Discussion

7.1. Voice Jitter

When two consecutive packages leave the source node with timestamps t1 and t2 and run at the destination node at time t3 and t4, then:

\[ \text{Jitter} = (t4 - t3) - (t2 - t1) \]

A negative jitter indicates that the time difference between packets at the destination node was less than that at the source node. As shown in Figure 2, the jitter in G.729A changes more than the jitter in G.711 especially at the beginning of the simulation run time.

7.2. Delay

Network delay is the ratio of the time the sender node gives the packet to RTP to the time the recipient gets the packet from RTP. The corresponding features in the audio application configuration cause compression and decompression delays. These are combined to give the total delay that the UE experienced at the application level.

7.3. Link Throughput:

A link's throughput is a metric that indicates how much data it can effectively transfer in a specific time period. It considers all of the traffic that was sent across the network. Only data packets that were delivered from the source and successfully received at the destination are considered in the application throughput.
The theoretical maximum rate of link data transfer, known as bandwidth or channel capacity, which is always higher than throughput, it is commonly measured in bits per second.

![Fig. 4 Link Throughput](image)

As shown in Figure 4, throughput of G.711 is lower than of G.729A, and both of them keep the constant values of throughput after 10 secs of beginning the simulation run time.

7.4. Traffic Sent and Received (Packets per second)

They are important metrics for measuring the performance of a communication network. They provide information about the amount of data being transmitted and received over the network. Link throughput takes into account the entire traffic that was sent through the link, while application throughput only takes into account those data packets that were successfully received at the destination. Throughput is the amount of data moved successfully from one place to another in a given time period and is typically measured in bits per second.

It is the statistic, which is defined as the average number of packets per second forwarded to FTP application by the transport layers in the network and it is known as Traffic Received (packets/sec)

![Fig. 5 Traffic sent](image)
Figures 5 and 6 shows that traffic send and received in G.711 are higher than traffic send and received in G.729A.

7.5. MOS

Mean Opinion Score is a metric used to measure the perceived quality of voice transmission in VoIP calls. MOS is a quality measure that has been used in telephony for decades as a way to assess human users’ opinion of call quality.

In the scenario, the delay of the PCM (G.729A) is less than the delay of PCM (G.711).

Finally, in Figure 7, the obtained results for MOS have compared between the MOS of the G.729A and G.711. The G.729A has a higher MOS than the G.711, which indicates that using a VOIP over a C.729A telephone network is better than G.711.

8. Conclusions

The QoE for VoIP is a critical aspect of ensuring customer satisfaction and network stability, as well as, good QoS parameters is vital to the functionality and effectiveness of a VoIP application. The use of high bandwidth transmission media, combined with high quality of voice, makes VoIP a flexible alternative for speech transmission. The presented work in this paper, aimed at assessing the QoE for voice applications, demonstrates that the MOS score serves as a valuable diagnostic tool for identifying and addressing quality issues. The MOS score reports on three different metrics: listening quality, transmission quality, and conversational quality. The heterogeneous transmission link in simulated network (wired and wireless) was built using Modeler and results were obtained in which the QoS parameters (voice jitter, delay, throughput, traffic sent and traffic received) were calculated, which leads to measuring the QoE. During the simulation run of the proposed scenario, delay values are noticed were too small, due to the number of few nodes in the network. Whereas more nodes will increase
network delay. There are differences in QoS parameters between using G.729A codec and G.711 codec in VoIP application. It is noted that the value of delay in the G.729A codec is higher than its value in the G.711 codec. The results during 30 seconds of the voice call simulation run time were as follows: Delay was 0.57 mSec in G.729A, and 0.25 mSec in G.711. The throughput was 290 Kbits/sec in G.729A, and 70 Kbits/sec in G.711. The jitter was 0.02 uSec G.729A, and 0 Sec in G.711. The MOS value was 3.76 G.729A, and 3.08 in G.711.

REFERENCES